DAGSTUHL
REPORTS

Volume 3, Issue 10, October 2013

ISSN 2192-5283

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

Publication date
February, 2014

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

Digital Object Identifier: 10.4230/DagRep.3.10.i

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

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

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

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

www.dagstuhl.de/dagrep

Report from Dagstuhl Seminar 13411

Deduction and Arithmetic

Edited by

Nikolaj Bjørner[1], Reiner Hähnle[2], Tobias Nipkow[3], and Christoph Weidenbach[4]

1 **Microsoft Research, Redmond, US,** nbjorner@microsoft.com
2 **Technische Universität Darmstadt, DE,** haehnle@cs.tu-darmstadt.de
3 **Technische Universität München, DE,** nipkow@in.tum.de
4 **Max-Planck-Institut für Informatik, Saarbrücken, DE,** weidenbach@mpii.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13411 "Deduction and Arithmetic". The aim of this seminar was to bring together researchers working in deduction and fields related to arithmetic constraint solving. Current research in deduction can be categorized in three main strands: SMT solvers, automated first-order provers, and interactive provers. Although dealing with arithmetic has been in focus of all three for some years, there is still need of much better support of arithmetic. Reasong about arithmetic will stay at the center of attention in all three main approaches to automated deduction during the coming five to ten years. The seminar was an important event for the subcommunities involved that made it possible to communicate with each other so as to avoid duplicate effort and to exploit synergies. It succeeded also in identifying a number of important trends and open problems.

Seminar 6.–11. October, 2013 – www.dagstuhl.de/13411
1998 ACM Subject Classification D.2.4 Software/Program Verification, F.3.1 Specifying and Verifying and Reasoning about Programs, F.4.1 Mathematical Logic, G.1.2 Approximation, G.1.3 Numerical Linear Algebra, G.1.6 Optimization, I.2.3 Deduction and Theorem Proving
Keywords and phrases Automated Deduction, Program Verification, Arithmetic Constraint Solving
Digital Object Identifier 10.4230/DagRep.3.10.1
Edited in cooperation with Jasmin Christian Blanchette

1 Executive Summary

Nikolaj Bjørner
Reiner Hähnle
Tobias Nipkow
Christoph Weidenbach

Arithmetic plays a fundamental role in deduction. Logical constraints over arithmetical properties occur frequently in classical theorems in mathematics, as well as in program analysis and verification. The first automatic theorem prover was an implementation of Presburger Arithmetic in 1954. With the availability of powerful predicate calculus proof procedures some years later, arithmetic would be relegated to the sidelines. Interest in arithmetic revived in the 1980s with the advent of powerful interactive theorem provers that needed and supported arithmetic for their applications. The need for efficient computer

aided deduction with support for arithmetic in the area of program analysis and verification recently gave birth to a new technology, so called SMT solvers.

Thus we have three strands of automated deduction: SMT solvers, automated first-order provers, and interactive provers in need of (more) arithmetic.

SMT SMT (satisfiability modulo theories) solvers distinguish themselves by integrating built-in support for a combination of theories, including prominently the theory of arithmetic. Most often handling arithmetic formulas in isolation is not sufficient. Applications typically use a non-disjoint combination of arithmetic and other theory reasoning. SMT solvers nowadays handle quantifier-free arithmetic well, but are not directly equipped to solve arithmetical formulas with quantifiers. Recent progress on building in quantifier-elimination procedures for linear and non-linear arithmetic have made practical integration of such richer arithmetic deduction viable.

ATP Research in first-order logic theorem proving used to concentrate on efficient calculi in general and the integration of equational theories in particular. It is obvious that further integration of "richer" arithmetic theories into first-order logic should be done by rather a combination approach than an integration approach. One major challenge of combining first-order logic calculi with arithmetic procedures is that of compactness/completeness and termination. While Boolean combinations of ground atoms, as they are considered by SMT solvers typically do not cause trouble with respect to those challenges, combining first-order clauses with an arithmetic theory can never result in a compact/complete/terminating calculus, in general. The actual combination typically requires the solution of purely arithmetic problems in order to establish valid inferences and simplifications. These problems are of a specific nature in that the form of the arithmetic formulas and the way they need to be tested require specific variants of the known arithmetic procedures.

ITP Interactive theorem provers initially came with built-in decision procedures for quantifier-free linear arithmetic. More foundational systems then developed new techniques to implement these decision procedures by reducing them to pure logic, trading efficiency for guaranteed correctness. Aspects of arithmetic reasoning are present in deductive software verification systems: interactive systems combine a number of automatic arithmetic reasoning methods and control them with heuristics that are specific for verification. A challenging application of interactive proof and arithmetic is the Flyspeck project, an effort to formalize Hales's proof of the Kepler conjecture in an interactive theorem prover.

The Dagstuhl seminar was a timely event that brought together experts in the above subareas of deduction, and in reasoning about arithmetic, to exchange experiences and insights. The research questions pursued and answered included:

- Which arithmetic problems are best solved with which approach?
- How to handle very complex numeric representations such as the IEEE floating-point standard with a high degree of automation?
- Arithmetic in combination with other theories results easily in languages with a very complex decision problem—how can a high degree of automation be obtained nevertheless?
- How can SMT-based reasoning be combined with model-based reasoning?
- What is the best way to incorporate arithmetic simplification available in computer algebra systems into deductive frameworks?
- How can the specific structure of arithmetic problems generated by deduction systems be exploited?

In addition to the technical contributions, the seminar participants attempted in an open discussion session to identify the major trends and open questions around Deduction and Arithmetic. The outcome of this discussion is summarized in section 3.

2 Table of Contents

3 Trends and Open Questions in Deduction and Arithmetic

All seminar participants; notes taken by Tjark Weber, edited and completed by Reiner Hähnle

A major trend is the growing interest in and partial support of *non-linear arithmetic*, including not only multiplication, but even transcendental functions, and ranging over floating-point data types.

Support for non-linear arithmetic and floating-point data types is particularly important for verification of embedded systems and hybrid systems.

Current tool support for non-linear arithmetic and floating-point data types is immature for the time being, but the community expects considerable progress and activity in this area.

At the same time, the problems around *linear arithmetic with quantifiers* are far from being solved satisfactorily either and work on this front continues as well.

Model-based and *model-guided* approaches are prominent and are perceived as crucial for progress, in particular, to help tackling the partially open problem of how to handle *quantifiers*.

Formal specification of code that uses floating-point types is essentially an unsolved problem. At the moment, generic precision assertions using intervals or deltas are employed, but how to come up with functional specifications is unclear. It was widely agreed that is in general infeasible to write specifications completely by hand, they should be generated at least semi-automatically. The specification problem is compounded by the lack the of semantic abstractions that would make specifications interchangeable.

It was agreed that currently there is a *certification gap* in the tool chain when SMT solvers are involved due to the lack of a common proof format for them. It is desirable to continue work on a common proof framework for SMT solvers, such as the Logical Framework with Side Conditions (LFSC).

4 Overview of Talks

4.1 Approximate decidability and verification of hybrid systems

Jeremy Avigad (Carnegie Mellon University, US)

Joint work of Sicun Gao; Avigad, Jeremy; Clarke, Edmund
Main reference S. Gao, J. Avigad, E. M. Clarke, "Delta-complete decision procedures for satisfiability over the reals," in Proc. of the 27th Annual IEEE Symposium on Logic in Computer Science (LICS'12), pp. 305–314, IEEE, 2012.
URL http://dx.doi.org/10.1109/LICS.2012.41

Control systems that combine analog and discrete components are now ubiquitous. Methods of verifying discrete components are by now well understood, but the task of modeling an analog component, whose evolution over time is often described as the solution to a system of differential equations, poses new challenges. In particular, issues of decidability and complexity limit the reach of what can be done purely symbolically. I will discuss a general framework, based on the methods of computable analysis, for reasoning about specifications

using numerical approximations. I will discuss the notion of a "δ-decision procedure", which can be used to evaluate a verification claim. A positive answer provides a guarantee that the system description meets its specification. A negative answer indicates that either the system is unsafe, or that a small perturbation would render it so. I will show that the set of first-order bounded sentences involving computable functions on the real numbers is delta-decidable. I will also briefly discuss complexity considerations, and an implementation, "dreal", due to Gao, Soonho Kong, and Clarke.

References

1 Sicun Gao, Jeremy Avigad, and Edmund Clarke.,"Delta-decidability over the reals." In *Logic in Computer Science (LICS)*, 305–314, 2012.

2 Sicun Gao, Jeremy Avigad, and Edmund Clarke, "Delta-complete decision procedures for satisfiability over the reals." In Bernard Gramlich et al., eds., *International Joint Conference on Automated Reasoning (IJCAR)*, 286–300, 2012.

3 Sicun Gao, Soonho Kong, and Edmund Clarke, "dReal: An SMT Solver for Nonlinear Theories of Reals." In *International Conference on Automated Deduction (CADE)*, 2013.

4.2 Hierarchic superposition with weak abstraction and the Beagle theorem prover

Peter Baumgartner (NICTA, Canberra, AU)

Joint work of Baumgartner, Peter; Waldmann, Uwe
Main reference P. Baumgartner, U. Waldmann, "Hierarchic superposition with weak abstraction," in Proc. of the 24th Int'l Conf. on Automated Deduction (CADE'13), LNCS, Vol. 7898, pp. 39–57, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-38574-2_3

Many applications of automated deduction require reasoning in first-order logic modulo background theories, in particular some form of integer arithmetic. A major unsolved research challenge is to design theorem provers that are "reasonably complete" even in the presence of free function symbols ranging into a background theory sort. The earlier hierarchic superposition calculus of Bachmair, Ganzinger, and Waldmann [1] already supports such symbols, but, not optimally. We have devised a new calculus, hierarchic superposition with weak abstraction, which rectifies this situation by introducing a novel form of clause abstraction, a core component in the hierarchic superposition calculus for transforming clauses into a form needed for internal operation [2]. Additionally, it includes a definition rule that is generally useful to find refutations more often, and, specifically, gives completeness for the clause logic fragment where all background-sorted terms are ground.

The talk provides an overview of the calculus, its implementation in the Beagle theorem prover and experiments with it.

References

1 Leo Bachmair, Harald Ganzinger, and Uwe Waldmann. Refutational theorem proving for hierarchic first-order theories. *Appl. Algebra Eng. Commun. Comput*, 5:193–212, 1994.

2 Peter Baumgartner and Uwe Waldmann. Hierarchic superposition with weak abstraction. In Maria Paola Bonacina, editor, *CADE-24 – The 24th International Conference on Automated Deduction*, volume 7898 of *Lecture Notes in Artificial Intelligence*, pages 39–57. Springer, 2013.

4.3 Basic Hilbert basis

Nikolaj Bjørner (Microsoft Research, Redmond, US)

Joint work of Bjørner, Nikolaj; Voronkov, Andrei

We present a new algorithm for computing Hilbert bases and present an efficient implementation using tailored data-structures. A profound feature of the algorithm is that it is also a constructive proof that *every* basis vector of a Hilbert Basis requires at a polynomial number of bits to represent. This re-establishes that Integer Linear Programming is in NP, and it shows that all minimal basis solutions are small.

4.4 Arithmetic in Sledgehammer

Jasmin Christian Blanchette (Technische Universität München, DE)

Joint work of Blanchette, Jasmin Christian; Böhme, Sascha; Paulson, Lawrence C.
Main reference J. C. Blanchette, S. Böhme, L. C. Paulson, "Extending Sledgehammer with SMT Solvers", J. Autom. Reasoning 51(1):109–128, 2013.
URL http://dx.doi.org/10.1007/s10817-013-9278-5

Sledgehammer is a subsystem of Isabelle/HOL that integrates automatic theorem provers (ATPs). It uses, among others, the first-order resolution provers E, SPASS, Vampire and the SMT solvers CVC3, Yices, and Z3 as backends. Interpreted arithmetic is used for the SMT solvers, but the impact is not as great as one might have expected. This talk attempted to give a few explanations and suggested directions for future research.

4.5 Sound compilation of reals

Eva Darulova (EPFL, Lausanne, CH)

Joint work of Darulova, Eva; Kuncak, Viktor
Main reference E. Darulova, V. Kuncak, "Sound Compilation of Reals," in Proc. of the 41st Annual ACM SIGPLAN-SIGACT Symp. on Principles of Programming Languages (POPL'14), pp. 235–248, ACM, 2014.
URL http://dx.doi.org/10.1145/2535838.2535874

Writing accurate numerical software is hard because of many sources of unavoidable uncertainties, including finite numerical precision of implementations. We present a programming model where the user writes a program in a real-valued implementation and specification language that explicitly includes different types of uncertainties. We then present a compilation algorithm that generates a conventional implementation that is guaranteed to meet the desired precision with respect to real numbers. Our verification step generates verification conditions that treat different uncertainties in a unified way and encode reasoning about floating-point roundoff errors into reasoning about real numbers. We show that current state-of-the art SMT solvers do not scale well to solving such verification conditions. We propose a new procedure that combines exact SMT solving over reals with approximate and sound affine and interval arithmetic. We report on results from our initial prototype implementation.

4.6 Making invariants inductive

Stephan Falke (Karlsruhe Institute of Technology, DE)

Joint work of Falke, Stephan; Kapur, Deepak

While non-inductive invariants might be easier to write, inductive invariants are easier to prove. In this talk, I will present an algorithm that aims at strengthening a given potential invariant so it becomes inductive. I will discuss properties of the algorithm and pose open questions.

4.7 The poor man's way to integrate non-linear real arithmetic reasoning capabilities within SMT

Pascal Fontaine (LORIA, Nancy, FR)

Joint work of Déharbe, David; Fontaine, Pascal; Sturm, Thomas

In this talk, we report some very preliminary results on a cooperation between the veriT SMT solver and Redlog. Thanks to this cooperation, the language handled by veriT now accepts non-linear real arithmetic expressions. The technique is based on a model-based combination. The linear arithmetic and congruence closure decision procedures are used as simplifiers for the set of constraints given to Redlog.

4.8 Deduction and arithmetic – a functional marriage: iSAT, odeSAT, and related procedures

Martin Fränzle (Universität Oldenburg, DE)

Joint work of Fränzle, Martin; Eggers, Andreas; Herde, Christian; Teige, Tino; Becker, Bernd; Kupferschmid, Stefan; Scheibler, Karsten; Schubert, Tobias; Ratschan, Stefan; Ramdani, Nacim; Nedialkov, Nedlialko

Over the last decade, arithmetic SAT modulo theory (SMT) solving has found wide-spread application within diverse analysis tasks for hybrid discrete-continuous systems, cyber-physical systems, or software systems involving numerical computations. Traditional SMT approaches are, however, confined to decidable theories, i.e. small fragments of arithmetic. Combining ideas from abstract interpretation, namely to manipulate lattices of subsets of the value domain, with ideas from verified computer arithmetic, namely to use safe numeric interval enclosures, and with techniques from SMT solving, namely conflict-driven clause learning, this restriction can be relaxed in practice. While the resulting procedures, which represent a "functional marriage" between automated deduction and computer arithmetic, are necessarily incomplete, they can solve many intricate problems in practice: For complex-structured Boolean combinations of arithmetic constraints involving transcendental functions and thousands of variables, or for combinations of arithmetic constraints involving first-order relations defined by ordinary differential equations, they are able to rigorously prove

unsatisfiability, do provide strong hints (and sometimes even proofs) for satisfiability, and can compute Craig interpolants.

References

1 Martin Fränzle, Christian Herde, Stefan Ratschan, Tobias Schubert, and Tino Teige. Efficient solving of large non-linear arithmetic constraint systems with complex boolean structure. *Journal on Satisfiability, Boolean Modeling and Computation – Special Issue on SAT/CP Integration*, 1:209- -236, 2007.

2 Andreas Eggers, Martin Fränzle, and Christian Herde. SAT modulo ODE: A direct SAT approach to hybrid systems. In Sungdeok (Steve) Cha, Jin-Young Choi, Moonzoo Kim, Insup Lee, and Mahesh Viswanathan, editors, *Proceedings of the 6th International Symposium on Automated Technology for Verification and Analysis (ATVA'08)*, volume 5311 of Lecture Notes in Computer Science (LNCS), pages 171–185. Springer-Verlag, 2008.

4.9 Cooperation for better termination proving

Carsten Fuhs (University College London, GB)

Joint work of Brockschmidt, Marc; Cook, Byron; Fuhs, Carsten
Main reference M. Brockschmidt, B. Cook, C. Fuhs, "Better termination proving through cooperation," in Proc. of 25th Int'l Conf. on Computer Aided Verification (CAV'13), LNCS, Vol. 8044, pp. 413–429, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-39799-8_28

One of the difficulties of proving program termination is managing the subtle interplay between the finding of a termination argument and the finding of the argument's supporting invariant. We propose a new mechanism that facilitates better cooperation between these two types of reasoning. In an experimental evaluation we find that our new method leads to dramatic performance improvements.

References

1 Marc Brockschmidt, Byron Cook, and Carsten Fuhs. *Better termination proving through cooperation.* In *Proc. CAV '13*, volume 8044 of *LNCS*, pages 413–429, 2013.

4.10 Alternating runtime and size complexity analysis of integer programs

Jürgen Giesl (RWTH Aachen, DE)

We present a new modular approach to automatic complexity analysis. Based on a novel alternation between finding symbolic time bounds for program parts and using these to infer size bounds on program variables, we can restrict each analysis step to a small part of the program while maintaining a high level of precision. Extensive experiments with the implementation of our method demonstrate its performance and power in comparison with other tools. In particular, our method finds bounds for many programs whose complexity could not be analyzed by automatic tools before.

4.11 Simple interpolation for floating-point arithmetic with abstract CDCL

Alberto Griggio (Bruno Kessler Foundation, Trento, IT)

Joint work of Brain, Martin; D'Silva, Vijay; Griggio, Alberto; Haller, Leopold; Kroening, Daniel
Main reference M. Brain, V. D'Silva, A. Griggio, L. Haller, D. Kroening, "Interpolation-Based Verification of Floating-Point Programs with Abstract CDCL," in Proc. of the 20th Int'l Symp. on Static Analysis (SAS'13), LNCS, Vol. 7935, pp. 412–432, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-38856-9_22

One approach for SMT solvers to improve efficiency is to delegate reasoning to abstract domains. Solvers using abstract domains do not support interpolation and cannot be used for interpolation-based verification. We extend Abstract Conflict Driven Clause Learning (ACDCL) solvers with proof generation and interpolation. Our results lead to the first interpolation procedure for floating-point logic and subsequently, the first interpolation-based verifiers for programs with floating-point variables. We demonstrate the potential of this approach by verifying a number of programs which are challenging for current verification tools.

4.12 Model-constructing satisfiability calculus

Dejan Jovanović (SRI, Menlo Park, US)

Joint work of Jovanović, Dejan; Barrett, Clark; de Moura, Leonardo
Main reference D. Jovanović, C. Barrett, L. De Moura, "Design and implementation of the model-constructing satisfiability calculus." Presentation at 2013 Formal Methods in Computer-Aided Design Conference (FMCAD'13).
URL http://www.cs.utexas.edu/users/hunt/FMCAD/FMCAD13/papers/71-Model-Construction-SAT-Calculus.pdf

We present the Model Constructing Satisfiability (MCSat) calculus. MCSat calculus generalizes ideas found in CDCL-style propositional SAT solvers to SMT solvers, and provides a common framework where recent model-based procedures and techniques can be justified and combined. We describe how to incorporate support for linear real arithmetic and uninterpreted function symbols in the calculus. We report encouraging experimental results, where MCSat performs competitive with the state-of-the art SMT solvers without using pre-processing techniques and ad-hoc optimizations. The implementation is flexible, additional plugins can be easily added, and the code is freely available.

4.13 Practical exploitation of mixed integer programming solvers by SMT solvers

Tim A. King (New York University, US)

Simplex based methods used within SMT solvers become overwhelmed by problems that are efficiently solved by the Simplex solvers used within Mixed Integer Programming solvers.

This talk will describe practical usage of MIP solvers from within SMT solvers to extend SMT solvers to be able to handle such benchmarks. A preliminary implementation of this technique solves the more SMTLIB QFLRA benchmarks then other state-of-art SMT solvers.

4.14 Solving linear arithmetic by bound propagation

Konstantin Korovin (University of Manchester, GB)

Joint work of Korovin, Konstantin; Voronkov, Andrei
Main reference K. Korovin, A. Voronkov, "Solving Systems of Linear Inequalities by Bound Propagation," in Proc. of the 23rd Int'l Conf. on Automated Deduction, LNCS, Vol. 6803, pp. 369–383, Springer, 2011.
URL http://dx.doi.org/10.1007/978-3-642-22438-6_28

Bound propagation is a method for solving systems of linear inequalities using DPLL-style reasoning, shown to be complete in our previous work ([1], extending conflict resolution [2]). The bound propagation method adapts constraint propagation, dynamic variable ordering, lemma learning and backjumping. In this talk I overview the bound propagation method, discuss non-trivial issues such as termination, and present recent implementation and experimental results [3, 4]. Joint work with Ioan Dragan, Laura Kovács, Nestan Tsiskaridze and Andrei Voronkov.

References

1 K. Korovin and A. Voronkov Solving Systems of Linear Inequalities by Bound Propagation. *Proc. of the 23rd International Conference on Automated Deduction, (CADE 2011)*, 369–383, LNAI, vol 6803, Springer, 2011.
2 K. Korovin, N. Tsiskaridze and A. Voronkov. Conflict Resolution. *Proc. of the 15th International Conference on Principles and Practice of Constraint Programming (CP'09)*, 509–523, Lecture Notes in Computer Science, vol 5732, Springer, 2009.
3 K. Korovin, N. Tsiskaridze and A. Voronkov. Implementing Conflict Resolution. *Proc. of the 8th International Conference on Perspectives of Systems Informatics (PSI'2011)*, LNCS, vol 7162, Springer, 2012.
4 I. Dragan, K. Korovin, L. Kovacs and A. Voronkov. Bound Propagation for Arithmetic Reasoning in Vampire. Proc. of SYNASC'13, IEEE, 2013.

4.15 An SMT-based approach to memory access optimization problem

Marek Košta (Max-Planck-Institut für Informatik, Saarbrücken, DE)

Joint work of Košta, Marek; Karrenberg, Ralf; Sturm, Thomas
Main reference R. Karrenberg, M. Košta, T. Sturm, "Presburger Arithmetic in Memory Access Optimization for Data-Parallel Languages," in Proc. of the 9th Int'l Symp. on Frontiers of Combining Systems (FroCoS'13), LNCS, Vol. 8152, pp. 56–70, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-40885-4_5

To exploit capabilities of modern SIMD capable CPUs, state-of-the-art compilers use vectorization techniques. One of the main drawbacks of these techniques is that they serialize memory access operations. We begin with memory access optimization problem definition and describe how its solution can help to generate vectorized, i.e. more efficient code in presence of memory accesses. Then we will refer about our previous work, which uses SMT

solving to solve the memory access optimization problem. Next, we will describe our current work: Development of a system using Z3 library capable of fully-automatic solution of the memory access optimization problem. We will conclude with a few theoretical and practical questions motivated by our current work.

4.16 Acceleration for Petri nets

Jérôme Leroux (University of Bordeaux, FR)

Main reference J. Leroux, "Acceleration for Petri Nets," in Proc. of the 11th Int'l Symp. on Automated Technology for Verification and Analysis (ATVA'13), LNCS, Vol. 8172, pp. 1–4, Spinger, 2013.
URL http://dx.doi.org/10.1007/978-3-319-02444-8_1

The reachability problem for Petri nets is a central problem of net theory. The problem is known to be decidable by inductive invariants definable in the Presburger arithmetic. When the reachability set is definable in the Presburger arithmetic, the existence of such an inductive invariant is immediate. However, in this case, the computation of a Presburger formula denoting the reachability set is an open problem. Recently this problem got closed by proving that if the reachability set of a Petri net is definable in the Presburger arithmetic, then the Petri net is flatable, i.e. its reachability set can be obtained by runs labeled by words in a bounded language. As a direct consequence, classical algorithms based on acceleration techniques effectively compute a formula in the Presburger arithmetic denoting the reachability set. In this presentation, the framework of verification of infinite-state systems based on acceleration techniques is recalled. We also explain the completeness of this method on the computation of Presburger formulas denoting the reachability sets of Petri nets.

4.17 Automated (formal) proofs of summation identities

Assia Mahboubi (INRIA Saclay–Île-de-France–Orsay, FR)

Joint work of Mahboubi, Assia; Chyzak, Frédéric; Sibut-Pinote, Thomas; Tassi, Enrico

Among the most cited references in the mathematical literature are the handbooks of properties of special functions or sequences. These dictionaries consist of tables of results about functions arising naturally in applications like physics, number theory, economy etc. It happens that a large subset of these functions belongs to a well-behaved class of objects, which benefits from the theory of so-called holonomic systems. In particular, it is possible to increase dramatically the confidence in the results displayed in the handbooks by relying on algorithmic proofs and computer algebra systems. In this talk I will discuss the issues raised by the certification of these theorems by a proof assistant.

4.18 From ordinal interpretations to elementary interpretations

Aart Middeldorp (Universität Innsbruck, AT)

Joint work of Winkler, Sarah; Zankl, Harald; Middeldorp, Aart
Main reference S. Winkler, H. Zankl, A. Middeldorp, "Beyond Peano Arithmetic – Automatically Proving Termination of the Goodstein Sequence," in Proc. of the 24th Int'l Conf. on Rewriting Techniques and Applications (RTA'13), LIPIcs, Vol. 21, pp. 335–351, Dagstuhl Publishing, 2013.
URL http://dx.doi.org/10.4230/LIPIcs.RTA.2013.335

Kirby and Paris (1982) proved in a celebrated paper that a theorem of Goodstein (1944) cannot be established in Peano (1889) arithmetic. In a recent paper we presented an encoding of Goodstein's theorem as a termination problem of a finite rewrite system. Using a novel implementation of ordinal interpretations, we were able to automatically prove termination of this system, resulting in the first automatic termination proof for a system whose derivational complexity is not multiply recursive. After recapitulating this work, we discuss the challenges when trying to implement the elementary interpretations of Lescanne (1995).

4.19 Death by a thousand cuts (worst-case execution time by bounded model checking)

David Monniaux (VERIMAG, Grenoble, FR)

Joint work of Monniaux, David; Henry, Julien; Maïza, Claire; Caminha, Diego

The trace semantics of programs (with finite loop unrolling) can be compiled into first-order logic formulas, and thus bounded model checking maps to satisfiability testing for such formulas. Modern satisfiability modulo theory (SMT) solvers currently solve many large instances. Unfortunately, if certain non-functional properties such as timing and energy are naively encoded into the formula, the SMT problems become intractable since they contain "diamond formulas". We propose a general redundant encoding scheme for such properties, resulting in tractable SMT problems: we conjoin to the original problem some "cuts", which do not change the solution set but kill the complexity of the solving.

We illustrate this encoding with worst-case execution time (WCET) of loop-free programs. In real-time systems, it is often necessary to bound the WCET of program fragments; for instance, in safety-critical control systems, this time should be less than the period of the control loop. The conventional approach is to first run special static analyses which give for each basic block an upper bound for its WCET, taking into account information from the history of the computation (e.g. possible pipeline and cache states), then reassemble these local WCET into a global WCET through a path analysis. Unfortunately, this path analysis may take into account paths that are infeasible due to the semantics of the instructions, resulting in over-estimation of the WCET. We replace this path analysis by optimization modulo theory.

We experimented our analysis on benchmark and industrial programs, gaining as much as 53% improvement on the bound on WCET for a fly-by-wire program. On most of these programs, a naive encoding results in timeout, but with our encoding the analysis terminates within minutes.

4.20 Non-numerical permissions for concurrent reasoning

Wojciech Mostowski (University of Twente, NL)

We present a new, symbolic system for permission accounting that can be easily used with first-order logic based reasoning systems typically used in program verification. Permission accounting is a core building block in approaches to modular, thread-local reasoning about concurrent programs. Our permission system differs from the existing ones in that it provides a symbolic and global view of permissions, rather than value-based and thread-local one. That in turn enables (a) better understanding of permission tracking from the point of the view of the specifier, (b) specification of complex permission transfer scenarios, and (c) more efficient reasoning for the verification tools (in particular, no reasoning about rational numbers is required). Our system is based on the idea of "I-owe-you" chains of permission owners to track the history of permission transfers, and the idea of symbolic permission slicing to divide permissions between multiple owners/threads. Underneath, special lists with dedicated operations are used. We axiomatised our permission system and its vital properties in the KeY verifier as well as in PVS. KeY is a verification tool for Java programs based on first-order dynamic logic and our primary target to employ our permission system for reasoning about multi-threaded Java programs. Initial results with the verification of actual Java programs using our permission system and KeY are also discussed.

4.21 Quantifier elimination for dense and discrete linear orders

Tobias Nipkow (Technische Universität München, DE)

In earlier work I verified a number of quantifier elimination procedures for dense linear orders without endpoints in Isabelle [1]. In ongoing work I have extend these quantifier elimination procedures to discrete orderings and to orderings with endpoints (typically plus and minus infinity). It is known from the literature that these theories admit quantifier elimination, but my focus is on efficient procedures. It turns out that the method of interior points (some arbitrary point between a lower and upper bound, e.g. $(l + u)/2$) and of infinitesimals (due to Loos and Weispfenning) can be extended to endpoints and to discrete orders.

References

1 Tobias Nipkow. *Linear Quantifier Elimination*. In A. Armando, P. Baumgartner, and G. Dowek (eds.), *International Joint Conference on Automated Reasoning (IJCAR)*, LNCS 5195, pages 18–33, Springer, 2008.

4.22 Exact global optimization on demand

Grant Olney Passmore (University of Edinburgh, GB)

Joint work of Passmore, Grant Olney; de Moura, Leonardo

We present a method for exact global nonlinear optimization based on a real algebraic adaptation of the conflict-driven clause learning (CDCL) approach of modern SAT solving. This method allows polynomial objective functions to be constrained by real algebraic constraint systems with arbitrary boolean structure. Moreover, it can correctly determine when an objective function is unbounded, and can compute exact infima and suprema when they exist. The method requires computations over real closed fields containing infinitesimals (cf. [1]).

References

1 Leonardo de Moura and Grant Olney Passmore. *Computation in Real Closed Infinitesimal and Transcendental Extensions of the Rationals.* In Proceedings of the 24th International Conference on Automated Deduction (CADE) (2013)

4.23 Automating separation logic using SMT

Ruzica Piskac (Yale University, US)

Joint work of Piskac, Ruzica; Wies, Thomas; Zufferey, Damien
Main reference R. Piskac, T. Wies, D. Zufferey, "Automating Separation Logic Using SMT," in Proc. of the . 25th Int'l Conf. on Computer Aided Verification (CAV'13), LNCS, Vol. 8044, pp. 773–789, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-39799-8_54

Separation logic (SL) has gained widespread popularity because of its ability to succinctly express complex invariants of a program's heap configurations. Several specialized provers have been developed for decidable SL fragments. However, these provers cannot be easily extended or combined with solvers for other theories that are important in program verification, e.g., linear arithmetic. In this talk, we present a reduction of decidable SL fragments to a decidable first-order theory that fits well into the satisfiability modulo theories (SMT) framework. We show how to use this reduction to automate satisfiability, entailment, frame inference, and abduction problems for separation logic using SMT solvers. Our approach provides a simple method of integrating separation logic into existing verification tools that provide SMT backends, and an elegant way of combining SL fragments with other decidable first-order theories.

4.24 Logic of hybrid games

André Platzer (Carnegie Mellon University, US)

Main reference A. Platzer, "A Complete Axiomatization of Differential Game Logic for Hybrid Games," School of Computer Science, Carnegie Mellon University, CMU-CS-13-100R, January 2013, extended in revised version from July 2013.
URL http://reports-archive.adm.cs.cmu.edu/anon/2013/CMU-CS-13-100R.pdf

Hybrid systems model cyber-physical systems as dynamical systems with interacting discrete transitions and continuous evolutions along differential equations. They arise frequently in many application domains, including aviation, automotive, railway, and robotics. This talk studies hybrid games, i.e. games on hybrid systems combining discrete and continuous dynamics. Unlike hybrid systems, hybrid games allow choices in the system dynamics to be resolved adversarially by different players with different objectives.

This talk describes how logic and formal verification can be lifted to hybrid games [2, 1]. The talk describes a logic for hybrid systems called differential game logic dGL. The logic dGL can be used to study the existence of winning strategies for hybrid games, i.e. ways of resolving the player's choices in some way so that he wins by achieving his objective for all choices of the opponent. Hybrid games are determined, i.e. one player has a winning strategy from each state, yet their winning regions may require transfinite closure ordinals. The logic dGL, nevertheless, has a sound and complete axiomatization relative to any expressive logic. Separating axioms are identified that distinguish hybrid games from hybrid systems. Finally, dGL is proved to be strictly more expressive than the corresponding logic of hybrid systems.

References

1 André Platzer. *A Complete Axiomatization of Differential Game Logic for Hybrid Games.* School of Computer Science, Carnegie Mellon University, CMU-CS-13-100, January 2013, extended in revised version from July 2013.
2 André Platzer. *Differential Game Logic for Hybrid Games.* School of Computer Science, Carnegie Mellon University, CMU-CS-12-105, March 2012.

4.25 Proving unsatisfiability in non-linear arithmetic by duality

Enric Rodríguez-Carbonell (UPC, BarcelonaTech, ES)

Joint work of Larraz, Daniel; Oliveras, Albert; Rodríguez-Carbonell, Enric; Rubio, Albert

Non-linear problems arise in many contexts, for instance, in the generation of invariants and ranking functions following the constraint-based program analysis approach. Solvers based on linearization by case analysis efficiently find solutions for satisfiable instances. On the other hand, unsatisfiability is difficult to prove with this method. In this talk we propose to prove a conjunction of non-linear atoms to be unsatisfiable by finding a Positivstellensatz refutation certificate, which is more amenable to satisfiability-goaled non-linear solvers.

4.26 Exploring interpolants

Philipp Rümmer (Uppsala University, SE)

Joint work of Rümmer, Philipp; Subotić, Pavle
Main reference Philipp Rümmer, Pavle Subotić. Exploring Interpolants. FMCAD 2013.

Craig Interpolation is a standard method to construct and refine abstractions in model checking. To obtain abstractions that are suitable for the verification of software programs or hardware designs, model checkers rely on theorem provers to find the right interpolants, or interpolants containing the right predicates, in a generally infinite lattice of interpolants for any given interpolation problem. We present a semantic and solver-independent framework for systematically exploring interpolant lattices, based on the notion of interpolation abstraction [1]. We discuss how interpolation abstractions can be constructed for a variety of logics, and how they can be exploited in the context of software model checking.

References

1 Philipp Rümmer, Pavle Subotić. *Exploring Interpolants*. Formal Methods in Computer-Aided Design (FMCAD). Portland, USA, 2013.

4.27 Some challenges in applied software bounded model checking

Carsten Sinz (Karlsruhe Institute of Technology, DE)

Joint work of Sinz, Carsten; Falke, Stephan; Merz, Florian

In this talk I will give a personal view on what challenges software analysis tools (like bounded model checkers) are currently facing, especially when applied to low-level system and control software. The intention of the talk is not so much to present finished work, but to discuss ideas for future research directions.

4.28 Hierarchical reasoning and model generation in the verification of hybrid systems

Viorica Sofronie-Stokkermans (Universität Koblenz-Landau, DE)

Main reference V. Sofronie-Stokkermans, "Hierarchical Reasoning and Model Generation for the Verification of Parametric Hybrid Systems," in Proc. of the 24th Int'l Conf. on Automated Deduction (CADE'13), LNCS, Vol. 7898, pp. 360–376, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-38574-2_25

We study possibilities of using hierarchical reasoning, quantifier elimination and model generation for the verification of parametric hybrid systems, where the parameters can be constants or functions. Our goal is to automatically provide guarantees that such systems satisfy certain safety or invariance conditions.

We first analyze the possibility of automatically generating such guarantees in the form of constraints on parameters, then show that we can also synthesise so-called criticality functions, typically used for proving stability and/or safety of hybrid systems.

We illustrate our methods on several examples. The results are presented in detail in [1].

References

1 Viorica Sofronie-Stokkermans. *Hierarchical Reasoning and Model Generation for the Verification of Parametric Hybrid Systems.* In Maria Paola Bonacida (ed.) Automated Deduction – CADE-24 – 24th International Conference on Automated Deduction, LNCS 7898, pages 360–376, Springer 2013.

4.29 Certified abstract completion

Christian Sternagel (JAIST, Nomi, JP)

Joint work of Hirokawa, Nao; Middeldorp, Aart; Sternagel, Christian

The textbook proof of soundness of the Knuth-Bendix completion procedure is rather involved. To make this important result more accessible, especially in a classroom situation, Hirokawa and Middeldorp recently found an alternative proof via peak-decreasingness (which is a weaker but also simpler variant of the decreasing diagrams method). I present an Isabelle/HOL formalization of this new proof and compare it to a formalization of the traditional proof which is part of IsaFoR (a library of formalized results about rewriting).

4.30 Partial certification for termination proofs and more

Rene Thiemann (Universität Innsbruck, AT)

Joint work of Thiemann, Rene; Kuknat, Christian; Sternagel, Christian

Since untrusted termination tools will always use techniques that have not been formally verified in some proof assistant, it is hard to fully certify the generated proofs. We present an approach which still allows to certify large parts of the proof, where we support both an online and an offline modus for certification.

We further report on initial steps towards the certification of complexity proofs where we focus on problems with respect to arithmetic.

4.31 Integrating SAT, QBF and SMT solvers with interactive theorem provers

Tjark Weber (Uppsala University, SE)

This talk describes integrations of various automated solvers (for SAT, QBF and SMT) with the interactive theorem provers Isabelle/HOL and HOL4. Our integrations demonstrate that LCF-style proof checking is feasible for those solvers. We thereby increase not only automation in interactive theorem proving, but also confidence in the correctness of solver results.

5 Program

Monday, 7 October 2013

09:00–09:30 Organizers
Welcome, Announcements, Introduction of Participants

09:30–10:00 Jürgen Giesl
Analyzing Runtime and Size Complexity of Integer Programs

10:00–10:30 André Platzer
Logic of Hybrid Games

11:00–11:30 Peter Baumgartner
Hierarchic Superposition With Weak Abstraction and the Beagle Theorem Prover

11:30–12:00 Wojciech Mostowski
Non-numerical Permissions for Concurrent Reasoning

14:00–14:30 Enric Rodríguez-Carbonell
Proving Unsatisfiability in Non-Linear Arithmetic by Duality

14:30–15:00 Wolfgang Ahrendt
Verifying (In)stability in Floating-point Programs by Increasing Precision, using SMT Solving

15:00–15:30 Carsten Fuhs
Cooperation for Better Termination Proving

16:00–16:30 Marek Košta
An SMT-Based Approach To Memory Access Optimization Problems

16:30–17:00 Christian Sternagel
Certified Abstract Completion

17:00–17:30 David Monniaux
Death by a Thousand Cuts

Tuesday, 8 October 2013

09:00–09:30 Aart Middeldorp
From Ordinal Interpretations to Elementary Interpretations

09:30–10:00 Pascal Fontaine
The Poor Man's Way to Integrate Non-Linear Real Arithmetic Reasoning Capabilities within SMT

10:00–10:30 Jeremy Avigad
Approximate Decidability and the Verification of Hybrid Systems

11:00–11:30 Viorica Sofronie-Stokkermans
Hierarchical Reasoning and Model Generation in Verification

11:30–12:00 Eva Darulova
Sound Compilation of Reals

14:00–14:30 Assia Mahboubi
Automated (Formal) Proofs of Summation Identities

14:30–15:00 Konstantin Korovin
Solving Linear Arithmetic by Bound Propagation

15:00–15:30 Arie Gurfinkel
Exploring Interpolants

16:00–16:30 Bernhard Beckert
On the Specification and Verification of Voting Schemes

16:30–17:00 Jérôme Leroux
Computing Vector Addition System Reachability Sets

17:00–17:30 Philipp Rümmer
Interpolation for Software Verification

Wednesday, 9 October 2013

09:00–09:30 Christoph Weidenbach
Hierarchic Superposition – Current Status and Future Steps

09:30–10:00 Dejan Jovanović
Model-Constructing Satisfiability Calculus

10:00–10:30 Nikolaj Bjørner
Basic Hilbert Basis

11:00–11:30 Tjark Weber
Integrating SAT, QBF and SMT Solvers with Interactive Theorem Provers

11:30–12:00 Jasmin Christian Blanchette
Arithmetic in Sledgehammer

Thursday, 10 October 2013

09:00–09:30 Leonardo de Moura
Arithmetic Procedures in the Model-Constructing Satisfiability Calculus

09:30–10:00 Tobias Nipkow
Linear Quantifier Elimination for Linear Orderings with Endpoints

10:00–10:30 Thomas Sturm
Satisfiability Checking for the Sciences

11:00–11:30 Stephan Falke
Making Invariants Inductive

11:30–12:00 Grant Passmore
Exact Global Optimization on Demand

14:00–14:30 Ruzica Piskac
Automating Separation Logic Using SMT

14:30–15:00 Carsten Sinz
Some Challenges in Applied Software Bounded Model Checking

16:00–16:30 Tim King
Floating Point Simplex Solvers for SMT

16:30–17:00 Alberto Griggio
Simple Interpolation for Floating-Point Arithmetic with Abstract CDCL

Friday, 10 October 2013

09:00–09:30 Martin Fränzle
Deduction and Arithmetic – a Functional Marriage (iSAT, odeSAT, and Related Procedures)

09:30–10:00 René Thiemann
Partial Certification for Termination Proofs and More

10:00–10:30 James J. Hunn
Realtime Java and Formal Methods: Use and Open Issues

11:00–12:00 Organizers
Wrap-Up Session

Participants

- Wolfgang Ahrendt
Chalmers UT – Göteborg, SE
- Jeremy Avigad
Carnegie Mellon University, US
- Peter Baumgartner
NICTA – Canberra, AU
- Bernhard Beckert
KIT – Karlsruhe Institute of Technology, DE
- Nikolaj Bjørner
Microsoft Res. – Redmond, US
- Jasmin Christian Blanchette
TU München, DE
- Richard Bubel
TU Darmstadt, DE
- Eva Darulova
EPFL – Lausanne, CH
- Leonardo de Moura
Microsoft Res. – Redmond, US
- Stephan Falke
KIT – Karlsruhe Institute of Technology, DE
- Pascal Fontaine
LORIA – Nancy, FR
- Martin Fränzle
Universität Oldenburg, DE
- Carsten Fuhs
University College London, GB
- Jürgen Giesl
RWTH Aachen, DE
- Alberto Griggio
Bruno Kessler Foundation – Trento, IT
- Arie Gurfinkel
Carnegie Mellon University, US
- Reiner Hähnle
TU Darmstadt, DE
- James J. Hunt
aicas GmbH – Karlsruhe, DE
- Dejan Jovanovic
SRI – Menlo Park, US
- Tim A. King
New York University, US
- Konstantin Korovin
University of Manchester, GB
- Marek Kosta
MPI für Informatik – Saarbrücken, DE
- Jerome Leroux
University of Bordeaux, FR
- Assia Mahboubi
INRIA Saclay – Île-de-France – Orsay, FR
- Aart Middeldorp
Universität Innsbruck, AT
- David Monniaux
VERIMAG – Grenoble, FR
- Wojciech Mostowski
University of Twente, NL
- Tobias Nipkow
TU München, DE
- Grant Olney Passmore
University of Cambridge & University of Edinburgh, GB
- Ruzica Piskac
Yale University, US
- André Platzer
Carnegie Mellon University, US
- Enric Rodríguez-Carbonell
UPC – BarcelonaTech, ES
- Philipp Rümmer
Uppsala University, SE
- Peter H. Schmitt
KIT – Karlsruhe Institute of Technology, DE
- Carsten Sinz
KIT – Karlsruhe Institute of Technology, DE
- Viorica Sofronie-Stokkermans
Universität Koblenz-Landau, DE
- Christian Sternagel
JAIST – Nomi, JP
- Thomas Sturm
MPI für Informatik – Saarbrücken, DE
- René Thiemann
Universität Innsbruck, AT
- Cesare Tinelli
Univ. of Iowa – Iowa City, US
- Tjark Weber
Uppsala University, SE
- Christoph Weidenbach
MPI für Informatik – Saarbrücken, DE

Report from Dagstuhl Seminar 13412

Genomic Privacy

Edited by

Kay Hamacher[1], Jean Pierre Hubaux[2], and Gene Tsudik[3]

1 **TU Darmstadt, DE,** hamacher@bio.tu-darmstadt.de
2 **EPFL – Lausanne, CH,** Jean-Pierre.Hubaux@epfl.ch
3 **University of California – Irvine, US,** gene.tsudik@uci.edu

Abstract

Recent advances in genomics prompt a formidable privacy challenge: As the price of a complete genome profile has plummeted to as low as 99 USD for genome-wide genotyping, wide-spread usage of genomic information is about to become reality. Substantial progress is expected in the near future in terms of improved diagnoses and better preventive medicine.

The impact of the increased availability of genomic information on privacy, however, is unprecedented, for obvious reasons: First, genetic conditions and the predisposition to specific diseases (such as Alzheimer's) can be revealed. Second, a person's genomic information leaks substantial information about his relatives. Third, complex privacy issues can arise if DNA analysis is used for criminal investigations, epidemiological research, and personalized medicine purposes.

This report documents the program and the outcomes of the Dagstuhl Seminar 13412 "Genomic Privacy". The goal of the seminar was to bring together leading researchers, from different areas of academia and industry. The seminar welcomed participants from computer science, bioinformatics, genetics, ethics and medical fields. Through a series of presentations, discussions, and working groups, the seminar attempted to provide a coherent picture of the field, which transcends the borders of disciplines. The participants discussed many aspects of genomic privacy and jointly identified the main requirements and the possible technical solutions for protecting genomic data.

Seminar 07.–09. October, 2013 – www.dagstuhl.de/13412
1998 ACM Subject Classification K.4.1 Public Policy Issues, D.4.6 Security and Protection, J.3 Life and Medical Sciences
Keywords and phrases Genomics, Genetics, Health Data, Privacy Protection, Differential Privacy, Privacy by Design, Information Security, Cryptography, Secure Computation
Digital Object Identifier 10.4230/DagRep.3.10.25
Edited in cooperation with Jean Louis Raisaro

1 Executive Summary

Jean Louis Raisaro

The Dagstuhl seminar 13412 "Genomic Privacy" was a short two-and-a-half-day seminar, the first one on this topic ever, which took place from October 6th to 9th, 2013. The aim was to bring together researchers, from various research areas related to genomic privacy, and to inspire them to exchange theoretical results, practical requirements and ethical and legal implications related to the protection of genomic data. The rise of personalized medicine on the background of available, individual genomic sequences is taken for granted

in the biomedical community. Impressive advances in genome sequencing have opened the way to a variety of revolutionary applications in modern healthcare. In particular, the increasing understanding of the human genome, and of its relation to diseases and its response to treatments brings promise of improvements in preventive and personalized healthcare. However, because of the genome's highly sensitive nature, this progress raises important privacy and ethical concerns that need to be addressed. Indeed, besides carrying information about a person's genetic condition and his predisposition to specific diseases, the genome also contains information about his relatives. The leakage of such information can open the door to a variety of abuses and threats not yet fully understood. During the seminar, these points were addressed in particular:

- Expression and Requirements: What should be protected? For how long? Against whom? Who should be liable? Who would manage cryptographic keys? Anonymity vs. cryptography?
- Privacy Mechanisms & Regulations: What privacy enhancing techniques can be envisioned specifically for genomic data? What if some people publish their genome online against the will of their relatives? Which ethical guidelines can be adopted from traditional privacy regulations?
- Medical Perspective: Would medical specialists accept to have only a partial view on genomic data? How are epidemiological studies and biobanks affected by legal and/or technical restrictions?
- Patient Perspective: What patient's involvement can be reasonably expected? 'Can a person's genomic information be outsourced to some cloud storage service?
- Economics: What are the economic implications of genomic privacy; of its neglect?

The seminar fully satisfied the expectations. All participants briefly self-introduced themselves. Some of them were invited by the organizers to give survey talks about their recent research on genomic privacy, thus facilitating and encouraging inter-disciplinary discussions during the morning sessions. There were talks focusing both on the definition of the requirements for the efficient and secure implementation of genomic medicine and on the possible solutions to be addressed. The afternoon sessions were devoted to working groups.

The first speaker Regine Kollek (University of Hamburg, GER), addressed the meaning and context of genomic data, focusing on some of the social and ethical aspects of genomics. She was followed by Brad Malin (Vanderbilt University, US), who provided a summary of the ways (both legal and technical) such data can be protected and raised the question about its worth or if there exists some other practical approach that guarantees flexible genomic data protection plans. Satoru Miyano (University of Tokyo, JP) gave an overview of the requirements in term of storage, computational power and security needed to make "clinical sequencing " become a reality. He described the ongoing program that has been playing a key role in the International Cancer Genome Consortium (ICGC) in Japan. The morning session ended up with a joint talk by Jacques Fellay (EPFL – Lausanne, CH) and Amalio Telenti (Lausanne University Hospital, CH) about the current and future usage of genomic information in clinical settings. They outlined the importance of defining new threat models, emphasizing that trust is essential in healthcare.

The second day was focused on the possible technical solutions that can be used to ensure genomic privacy. If, on one hand, there are computational expensive cryptographical approaches such as homomorphic encryption or secure multi-party computation that guarantee accuracy at the expense of flexibility and increasing complexity, on the other hand there are also statistical-based solutions such as differential privacy, which are less accurate but more

flexible and less expensive in terms of computational and complexity costs. The first speaker of the day, Andreas Peter (University of Twente, NL), described his ongoing work on how to securely outsource genomic sequences in a privacy-preserving way by relying on an oblivious RAM construction. The second talk, by Erman Ayday (EPFL – Lausanne, CH), provided an overview of the activities on genomic privacy in Lausanne. Ayday first focused on how to protect and evaluate genomic privacy in the clinical context, he then showed how to process in a privacy-preserving fashion raw genomic data; and finally he described how to quantify kin genomic privacy. The third speaker of the morning, Vitaly Shmatikov (University of Texas – Austin, US), discussed about how to conduct privacy-preserving exploration in Genome-Wide Association Studies (GWAS). He presented a set of privacy-preserving data mining algorithms that produce significantly accurate results while guaranteeing differential privacy. Finally, the second-day morning session was closed by Emiliano De Cristofaro's (University College London, GB) survey about how to begin to address privacy-respecting genomic tests by relying on privacy-enhancing techniques based on private set intersection operations.

The final day started with a talk by Xiaofeng Wang (Indiana University – Bloomington, US) about the privacy-preserving sharing and analysis of human genomic data. In particular, he described some techniques for secure outsourcing of genome analysis, and differentially-private pilot data release and data source selection. The remaining part of the morning was devoted to a general discussion about the seminar's outcomes. Due to the seminar and the multi-disciplinary interactions, it became clear that protection of simple genomic sequences is not enough for a full-privacy preserving approach. The organizers, together with the participants, agreed that this problem should be addressed in a sequel Dagstuhl-seminar. Hence, they set up a future work agenda in order to organize again such a fruitful gathering.

We thank Schloss Dagstuhl for the professional and inspiring atmosphere it provides. Such an intense research seminar is possible because Dagstuhl so perfectly meets all researchers' needs.

2 Table of Contents

3 Overview of Talks

3.1 Genomics & Privacy: More Trouble than It's Worth?

Bradley Malin (Vanderbilt University, US)

URL http://www.hiplab.org/

Genomic sequence data is increasingly subject to attacks that disclose the identity of the corresponding individual (i.e., identity disclosure) or certain information about the individual (i.e., attribute disclosure). This talk reviews why such data is interesting and useful and how such attacks have transpired and are escalating over time. It then provides a high-level summary of ways in which such data can be protected from such attacks, both from a sociological (e.g., policy and contract) and technical perspective (disclosure control and encryption). After doing so, this questions if these are the only options for data protection or if risk-based models, which are cognizant of monetary costs and (dis)incentives might be more appropriate options for crafting practical and flexible genomic data protection plans.

3.2 Personalized Genomic Medicine at Institute of Medical Science of University of Tokyo

Satoru Miyano (University of Tokyo, JP)

The cost of human genome sequencing has decreased significantly. Today integrative systems' understanding of personal cancer based on personal omics data is getting important issue. It has become a reality to implement "clinical sequencing" of personal cancer and germline genomes together with their transcriptomic information. At the IMSUT (Institute of Medical Science, The University of Tokyo), a team comprised of members from the Human Genome Center, the Advanced Clinical Research Center and the IMSUT Research Hospital is installing a program for personalized genomic medicine of cancer based on whole genome sequencing and integrative omics analysis enhanced with the supercomputer system at Human Genome Center (225T FLOPS at peak, 4.6PB storage). In this talk, we present our ongoing program which uses the supercomputer system that has been playing a key role in the ICGC project of Japan and some important cancer genome sequencing projects. By running this program, we will facilitate the information system for personalized genomic medicine and foster people on biomedical informatics who can clinically interpret whole genome sequencing and omics data.

3.3 Needs and Myths in Medical Genomics

Jacques Fellay (EPFL – Lausanne, CH)
Amalio Telenti (Lausanne University Hospital, CH)

Joint work of Fellay, Jacques; Telenti, Amalio

In this presentation we will present a view on genomic privacy from the perspective of clinicians and genomic researchers. We will describe the current and future usage of genomic information in clinical settings, in order to define the requirements for the efficient and secure implementation of genomic medicine. While contemporary medical genetics still mostly focuses on rare variants involved in Mendelian diseases or exceptional response to drugs, the toolbox of genomic medicine is about to massively expand to encompass neonatal and prenatal sequencing, oncogenomics, and complex trait genomics. In addition, direct-to-consumer genomics (DTC) is likely to profoundly alter the classic patient-doctor interaction. Threat models need to be defined, keeping in mind that trust must remain at the center of healthcare.

3.4 A Practical Approach to Securely Outsource Genomic Sequences

Andreas Peter (University of Twente, NL)

Joint work of Peter, Andreas; Biedermann, Sebastian; Karvelas, Nikolaos; Katzenbeisser, Stefan

Recent developments in genomics enable new business and research models to have many individuals' genomes as an essential driving force. It is foreseeable that this genomic data will be outsourced to a central service provider (e.g., the cloud), thus allowing different applications, such as personalized medicine, large-scale genomic research, and disease susceptibility.

In my presentation, I describe a novel mechanism that enables a patient to securely outsource her genomic data to the cloud while delegating to an investigator (e.g. a physician) the right to run certain algorithms (e.g., medical tests) on her data. The mechanism is privacy-preserving, meaning that the investigator only learns the result of his algorithm on the patient's genome, while the cloud learns nothing at all. Our protocol only requires the patient to be online for the delegation of rights, and it is thereafter completely non-interactive with the patient.

We achieve reasonable efficiency by dividing the patient's genome into small blocks (e.g., SNPs) on which we can run efficient, secure, multiparty protocols. Our main technical contribution is a new ORAM-like construction for hiding the access patterns to these blocks. To circumvent an undesirable heavy workload at the investigator's side, our ORAM allows for the outsourcing of its most involved step (the "reshuffle") to an independent proxy without loss of privacy.

3.5 Protecting and Quantifying Genomic Privacy

Erman Ayday (EPFL – Lausanne, CH)

Joint work of Ayday, Erman; Raisaro, Jean Louis; Humbert, Mathias; Huang, Zhicong; Hubaux, Jean-Pierre
URL http://lca.epfl.ch/projects/genomic-privacy/

Genomics is becoming the next significant challenge for privacy. The price of a complete genome profile has plummeted below 100 USD for genome-wide genotyping (i.e., the characterization of about one million common genetic variants), which is offered by a number of companies. This low cost of DNA sequencing will break the physician/patient connection and it can open the door to all kinds of abuse not yet fully understood.

Access to genomic data prompts some important privacy concerns: (i) Genetic diseases can be unveiled; (ii) the propensity to develop specific diseases (such as Alzheimer's) can be revealed; (iii) a volunteer accepting de facto to have his genomic code made public can leak substantial information about genomic data of his relatives (possibly against their will); and (iv) complex privacy issues can arise if DNA analysis is used for criminal investigations and insurance purposes. Such issues could lead to abuse, threats, and genetic discrimination.

In this talk, after discussing the threats and challenges of genome privacy, I will summarize our solutions for protecting the privacy of genomic data. In particular, I will focus on (i) protecting and evaluating genome privacy in medical tests and personalized medicine, (ii) privacy-preserving processing of raw genomic data, and (iii) the quantification of kin genomic privacy using information theoretical tools.

3.6 Privacy-Preserving Data Exploration in Genome-Wide Association Studies

Vitaly Shmatikov (University of Texas – Austin, US)

Joint work of Shmatikov, Vitaly; Johnson, Aaron
Main reference A. Johnson, V. Shmatikov, "Privacy-Preserving Data Exploration in Genome-Wide Association Studies," in Proc. of the 19th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD'13), pp. 1079–1087, ACM, 2013.
URL http://dx.doi.org/10.1145/2487575.2487687

Genome-wide association studies (GWAS) have become a popular method for analyzing sets of DNA sequences in order to discover the genetic bases of diseases. Unfortunately, statistics published as the result of GWAS can be used to identify individuals participating in the study. To prevent privacy breaches, even previously published results have been removed from public databases, impeding researchers' access to the data and hindering collaborative research. Existing techniques for privacy-preserving GWAS focus on answering specific questions, such as correlations between a given pair of SNPs (DNA sequence variations). This does not fit the typical GWAS process, where the analyst might not know in advance which SNPs to consider, which statistical tests to use, how many SNPs are significant for a given dataset, etc.

We present a set of practical, privacy-preserving data-mining algorithms for GWAS datasets. Our framework supports exploratory data analysis, where the analyst does not know a priori how many and which SNPs to consider. We develop privacy-preserving

algorithms for computing the number and location of SNPs that are significantly associated with the disease, the significance of any statistical test between a given SNP and the disease, any measure of correlation between SNPs, and the block structure of correlations. We evaluate our algorithms on real-world datasets and demonstrate that they produce significantly more accurate results than prior techniques while guaranteeing differential privacy.

3.7 Whole Genome Sequencing vs. Privacy: Efficient Cryptographic Protocols to the Rescue

Emiliano De Cristofaro (University College London, GB)

Joint work of De Cristofaro, Emiliano; Baldi, Pierre; Baronio, Roberta; Faber, Sky; Gasti, Paolo; Tsudik, Gene
URL http://emilianodc.com/

Recent advances in DNA sequencing technologies have put ubiquitous availability of whole human genomes within reach. It is no longer hard to imagine the day when everyone will have the means to obtain and store his own DNA sequence. The widespread and affordable availability of whole genomes immediately opens up important opportunities in a number of health-related fields. In particular, common genomic applications and tests performed in vitro today will soon be conducted computationally, using digitized genomes. New applications will be developed as genome-enabled medicine becomes increasingly preventive and personalized. However, the very same progress also amplifies worrisome privacy concerns, as a genome represents a treasure trove of highly personal and sensitive information. In this talk, we will give an overview of biomedical advances in genomics and discuss associated privacy, ethical, and security challenges. We begin to address privacy-respecting genomic tests by focusing on some important applications, such as, personalized medicine, paternity tests, ancestry testing, and genetic compatibility tests. After carefully analyzing these applications and their requirements, we propose a set of efficient privacy- enhancing techniques based on private set operations. This enables us to implement, in silico, some operations that are currently performed via in vitro methods, in a secure fashion. Our experimental results demonstrate that the proposed techniques are both feasible and practical today. Finally, we explore a few alternatives for securely storing human genomes and allowing authorized parties to run tests in such a way that only the required minimum amount of information is disclosed. We present an Android API framework geared for privacy-preserving genomic testing.

3.8 Privacy-Preserving Sharing and Analysis of Human Genomic Data

Xiaofeng Wang (Indiana University – Bloomington, US)

Main reference Y. Chen, B. Peng, X. Wang, H. Tang, "Large-Scale Privacy-Preserving Mappings of Human Genomic Sequences on Hybrid Clouds," in Proc. of the 19th Annual Network and Distributed System Security Symp. (NDSS'12), 18pp., The Internet Society, 2012.
URL http://www.internetsociety.org/large-scale-privacy-preserving-mapping-human-genomic-sequences-hybrid-clouds

This presentation will present a view on genomic privacy from the perspective of clinicians and genomic researchers. in order to define the requirements for the efficient and secure

implementation of genomic medicine, we will describe the current and future usage of genomic information in clinical settings. While contemporary medical genetics still mostly focuses on rare variants involved in Mendelian diseases or exceptional responses to drugs, the toolbox of genomic medicine is about to massively expand to encompass neonatal and prenatal sequencing, oncogenomics, and complex trait genomics. In addition, direct-to-consumer genomics (DTC) is likely to profoundly alter the classic patient-doctor interaction. Threat models need to be defined, keeping in mind that trust must remain at the center of healthcare.

In this talk, we review our research on privacy-preserving genomic data computing and privacy-preserving data dissemination, particularly the techniques for secure outsourcing genomic analysis, differentially- private pilot data release and data source selection.

4 Working Groups

4.1 Privacy and Genomic/Health Data – What Makes Genomic Data Special

Chair: *Bradley Malin (Vanderbilt University, TN, United States)*

This working group focused on the definitions of the benefits and harms of genomic data. Recently, it has been shown that the widespread and affordable availability of fully-sequences human genomes has created enormous opportunities for better preventive and personalized medicine. In particular, in this working group, it was observed how the main benefits of genomic data consist in tailored drug dosages and choices and in the risk assessment for diseases based on the genetic makeup.

However, because of genome's highly sensitive nature, the leakage of genomic information can pave the way to a variety of abuses and threats. For example, health insurance companies might obtain the genetic information of their customers and deny their services to people with a high susceptibility of developing a chronic disease, or employers could hire applicants based on their genetic features. Access to this information could engender genetic discrimination or discredit an organization or a country willing to promote a personal genome project. Hence, there is an impelling need to design new privacy-enhancing technologies (PET) that guarantee the protection of such a data by keeping in mind that the first step should be the identification of the data-flow model both in the clinical practice and in the bioinformatics research in order to understand which part of the process PETs should focus on.

4.2 Scientific Methods in Human Genomics and Bioinformatics that Set the Requirements for Crypto Solutions

Chair: *Satoru Miyano (University of Tokyo, Japan)*

This working group focused on the definitions of the requirements for cryptographic solutions. In particular during this session, it was observed that the main requirements should consist in the protection of germline whole genome sequences along with whole markers (SNPs, insertions, deletions, etc..) linked with clinical data and family histories. The solution to these requirements is often not unique, rather a trade-off between (i) performance, (ii) utilization and (iii) privacy. For example, statistical-based solutions such as differential privacy focus on privacy at the expense of utilization, whereas cryptographic solutions

prioritize the utilization at the cost of bad performances. However, technical solutions are often not sufficient for protecting genomic privacy, therefore legal and professional guidelines are certainly needed to govern how genomic information is transmitted, stored and processed by the different stakeholders.

4.3 Cryptographic Protocols for the Protection of Genomic Data

Chair: *Marina Blanton (University of Notre Dame, IN, United States)*

This working group focused on defining what are the best cryptographic protocols for the protection of genomic data. In particular, it was observed by all the participants that there is a gap between the algorithms used for the processing of genomic data and algorithms that privacy-preserving solutions implement. Existing cryptographic protocols for privacy-preserving DNA processing started with rather simple functions and, so far, they still do not mimic complex algorithms used in practice (e.g. susceptibility tests, genotype imputations, association studies, etc.). Furthermore, current solutions do not simultaneously achieve full functionality, security and speed. Hence, it is clear that designing new cryptographic protocols that satisfy all the requirements represents a rather complex challenge. A possible solution outlined during the working group consists in protecting genomic data through multiple mechanisms such as cryptographic means, secret sharing, access control and legal means; by focusing on more precise functionality more so than on speed, while keeping in mind that cryptography can mitigate the risk but not eliminate it. In addition, the longevity of genomic data calls for levels of security stricter than in some other domains, and choosing larger security parameters for encrypted data represents only one part of the solution.

4.4 Anonymization of Genomic Data

Chair: *Vitaly Shmatikov (University of Texas Austin, TX, United States)*

This working group focused on the anonymization of genomic data. With some background information, only 75 SNPs are sufficient for identifying an individual. Because DNA is relatively stable in time and a small subset of other data can be revealed, and as any biological sample contain all markers required for identification, genomic data should be considered different from any other high dimensional profile of an individual. The general discussion of the working group was about the following question: "Can we apply traditional anonymization techniques to anonymization of genomic data?". The following points were proposed as potential solutions.

- *Recommendation 1*: Use multilevel data anonymization. Anonymized data could be made public (after some obfuscation), whereas precise data can be made available with strict access control.
- *Recommendation 2*: Combine technical, social (e.g., economic incentives), and legal(e.g., boundaries and damages) controls.
- *Recommendation 3*: Audit (physical and virtual).
- *Recommendation 4*: Pilot testing with real data and target populations.

It is important, however, to be aware that in general anonymization breaks when some other information is known and therefore every new model that provides a better understanding of humane genome would increase the risk of re-identification. Again, there does not exist a unique solution, but technical and legal privacy guarantees are needed.

Participants

- Gergely Acs
INRIA Grenoble – Rhône-Alpes, FR
- Erman Ayday
EPFL – Lausanne, CH
- Marina Blanton
Univ. of Notre Dame, US
- Jurgi Camblong
Sophia Genetics SA – Lausanne, CH
- Emiliano De Cristofaro
University College London, GB
- Zekeriya Erkin
TU Delft, NL
- Sky Faber
University California – Irvine, US
- Jacques Fellay
EPFL – Lausanne, CH
- Kay Hamacher
TU Darmstadt, DE
- Urs Hengartner
University of Waterloo, CA
- Zhicong Huang
EPFL – Lausanne, CH
- Jean Pierre Hubaux
EPFL – Lausanne, CH
- Mathias Humbert
EPFL – Lausanne, CH
- Murat Kantarcioglu
The University of Texas at Dallas, US
- Florian Kerschbaum
SAP AG – Walldorf, DE
- Regine Kollek
Universität Hamburg, DE
- Klaus A. Kuhn
TU München – Klinikum Rechts der Isar, DE
- Inald Lagendijk
TU Delft, NL
- Bradley Malin
Vanderbilt University, US
- Srdan Marinovic
ETH Zürich, CH
- Satoru Miyano
University of Tokyo, JP
- Andrzej Mizera
University of Luxembourg, LU
- Muhammad Naveed
EPFL – Lausanne, CH
- Andreas Peter
University of Twente, NL
- Jean-Jacques Quisquater
University of Louvain, BE
- Jean-Louis Raisaro
EPFL – Lausanne, CH
- Roded Sharan
Tel Aviv University, IL
- Vitaly Shmatikov
University of Texas – Austin, US
- Amalio Telenti
University of Lausanne, CH
- Gene Tsudik
Univ. of California – Irvine, US
- Xiaofeng Wang
Indiana University – Bloomington, US

Report from Dagstuhl Seminar 13421

Algorithms for Optimization Problems in Planar Graphs

Edited by

Glencora Borradaile[1], Philip Klein[2], Dániel Marx[3], and Claire Mathieu[4]

1 **Oregon State University, US,** glencora@eecs.oregonstate.edu
2 **Brown University, US,** klein@brown.edu
3 **HU Berlin, DE,** dmarx@cs.bme.hu
4 **Brown University, US,** claire@cs.brown.edu

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13421 "Algorithms for Optimization Problems in Planar Graphs". The seminar was held from October 13 to October 18, 2013. This report contains abstracts for the recent developments in planar graph algorithms discussed during the seminar as well as summaries of open problems in this area of research.

Seminar 13.–18. October, 2013 – www.dagstuhl.de/13421
1998 ACM Subject Classification F.2 Analysis of Algorithms and Problem Complexity
Keywords and phrases Algorithms, planar graphs, theory, approximation, fixed-parameter tractable, network flow, network design, kernelization
Digital Object Identifier 10.4230/DagRep.3.10.36
Edited in cooperation with Kyle Fox

1 Executive Summary

Glencora Borradaile
Philip Klein
Dániel Marx
Claire Mathieu

Planar graphs, and more generally graphs embedded on surfaces, arise in applications such as road map navigation and logistics, computational topology, graph drawing, and image processing. There has recently been a growing interest in addressing combinatorial optimization problems using algorithms that exploit embeddings on surfaces to achieve provably higher-quality output or provably faster running times. New algorithmic techniques have been discovered that yield dramatic improvements over previously known results. In addition, results have been generalized to apply to other families of graphs: excluded-minor, bounded-genus and bounded-treewidth graphs.

This Dagstuhl seminar brought together researchers who have been working in these areas to present recent research results, consolidate and share understanding of the emerging basic techniques, and collaborate to move past the current barriers.

- **Polynomial-time solvable problems.** There is a long tradition of finding fast algorithms for poly-time problems in planar graphs. In 1956, the first paper on maximum st-flow addressed the case where the network is planar (and s and t are adjacent). In 1976, a

linear-time algorithm was given for minimum spanning trees in planar graphs. In 1979, the paper introducing generalized nested dissection gave a fast algorithm for shortest paths in planar graphs with positive and negative lengths. The past couple of decades has witnessed the discovery of fast algorithms for a wide range of polynomial-time problems in planar graphs: variants of max flow, multicommodity flow, variants of shortest paths, Gomory-Hu cut trees, global min-cut, girth, matching, and min-cost flow. It seems, however, there is a long way yet to go; for many promising problems, no planarity-exploiting algorithm is known or there is reason to believe faster algorithms can be obtained.

- **Approximation schemes.** Research on polynomial-time approximation schemes (PTAS) for optimization problems in planar graphs goes back to the pioneering work of Lipton and Tarjan (1977) and Baker (1983), who introduced linear-time algorithms for certain problems in which the constraints were quite local, e.g. maximum-weight independent set and minimum-weight dominating set. For many years, little progress was made on problems with non-local constraints. In the mid-nineties, polynomial-time approximation schemes were developed for the traveling-salesman problem (TSP) in planar graphs, but in these the degree of the polynomial running time depended on the desired accuracy. A decade later, a linear-time approximation scheme was found for TSP. Shortly afterwards, the first polynomial time approximation schemes were found for problems, e.g. Steiner tree, in which the solution was much smaller than the input graph. Since then approximation schemes have been found for several other problems in planar graphs, such as two-connected spanning subgraph, Steiner forest, survivable network design, k-terminal cut, and k-center. Important new techniques have emerged, but we still lack fast approximation schemes for many important problems (e.g. facility location). The area of approximation schemes for planar graphs is ripe for further exploration.
- **Fixed-parameter tractable algorithms.** Another way to cope with computational intractability of some planar graph problems is through the lens of fixed-parameter tractability. The theory of bidimensionality and algorithms exploiting tree decompositions of planar graphs give a general methodology of dealing with planar problems. One way to obtain fixed-parameter tractability results is to show that there is a polynomial-time preprocessing algorithm that creates a "problem kernel" by reducing the size of the instance such that it is bounded by a function of the parameter k. Research on kernelization for planar graph problems has been a very active topic recently, culminating in a meta-theorem that gives problem kernels for a wide range of problems (2009).

The scientific program of the seminar consisted of 24 talks. Five of these talks were longer (60-90 minute) tutorials overviewing the three main areas of the seminar: Jeff Erickson ("Flows in planar and surface graphs") and Christian Wulff-Nilsen ("Separators in planar graphs with applications") covered polynomial-time algorithms; Philip Klein ("Some techniques for approximation schemes on planar graphs") covered approximation schemes; and Dániel Marx ("The square-root phenomenon in planar graphs") and Daniel Lokshtanov ("Kernels for planar graph problems") covered fixed-parameter tractability. One of the main goals of the seminar was to encourage collaboration between the three communities, and these well-received tutorials were very helpful by introducing the basics of each of these topics. The rest of the talks were 25-minute presentations on recent research of the participants.

The time between lunch and the afternoon coffee break was left open for individual discussions and collaborations in small groups. Two open-problem sessions were organized (on Monday evening and Wednesday evening). Notes on the presented problems can be found in this report.

2 Table of Contents

3 Overview of Talks

3.1 A polynomial-time approximation scheme for planar multiway cut

MohammadHossein Bateni (Google – New York, US)

Joint work of Bateni, MohammadHossein; Hajiaghayi, Mohammad Taghi; Klein, Philip; Mathieu, Claire
Main reference M. Bateni, M. Hajiaghayi, P. N. Klein, C. Mathieu, "A Polynomial-time Approximation Scheme for Planar Multiway Cut," in Proc. of the 23rd Annual ACM-SIAM Symp. on Discrete Algorithms (SODA'12), pp. 639–655, SIAM, 2012.
URL http://dl.acm.org/citation.cfm?id=2095116.2095170

Given an undirected graph with edge lengths and a subset of nodes (called the terminals), the multiway cut (also called the multi-terminal cut) problem asks for a subset of edges, with minimum total length, whose removal disconnects each terminal from all others. The problem generalizes minimum s-t cut, but is NP-hard for planar graphs and APX-hard for general graphs [2]. In this paper, we present a PTAS for multiway cut on planar graphs.

This work has been published in SODA [1].

References

1 MohammadHossein Bateni, MohammadTaghi Hajiaghayi, Philip N. Klein and Claire Mathieu. *A polynomial-time approximation scheme for planar multiway cut*. In SODA. SIAM, Philadelphia, PA, USA, 2012.
2 E. Dahlhaus, D. S. Johnson, C. H. Papadimitriou , P. D. Seymour, M. Yannakakis. *The Complexity of Multiterminal Cuts*. SIAM Journal on Computing: 23, 864–894, 1994.

3.2 Triangulating planar graphs under constraints

Therese Biedl (University of Waterloo, CA)

Main reference T. Biedl, "On triangulating k-outerplanar graphs," arXiv:1310.1845v2 [cs.DM] , 2013.
URL http://arxiv.org/abs/1310.1845v2

Most planar graph drawing algorithms operate by first triangulating the planar graph. If we want to use this approach, but draw the planar graph with approximately the optimal height, then we'll need to be careful about how we triangulate. It is easy to see that this can be done without increasing the area-requirement (at least if we don't insist on triangulating the outerface): take an optimal drawing and triangulate it in the computational geometry sense. But how can we find the edges to add without knowing the optimal drawing? This remains open.

In this talk, I will present two results that are close, in that they maintain some graph parameters that are closely related to the optimal height of a planar drawing. Namely, I will show how to triangulate a planar graph such that the outer-planarity remains (roughly) the same, and I will show how to triangulate a planar graph such that the pathwidth remains (asymptotically) the same.

3.3 The ball cover problem

Glencora Borradaile (Oregon State University, US)

Joint work of Borradaile, Glencora; Chambers, Erin

A recent result of Chepoi, Estellon and Vaxes [Disc. Comp. Geom. '07] states that any planar graph of diameter at most $2R$ can be covered by a constant number of balls of size R; put another way, there are a constant-sized subset of vertices within which every other vertex is distance half the diameter. We generalize this result to graphs embedded on surfaces of fixed genus with a fixed number of apices, making progress toward the conjecture that graphs excluding a fixed minor can also be covered by a constant number of balls. To do so, we develop two tools which may be of independent interest. The first gives a bound on the density of graphs drawn on a surface of genus g having a limit on the number of pairwise-crossing edges. The second bounds the size of a non-contractible cycle in terms of the Euclidean norm of the degree sequence of a graph embedded on the surface.

3.4 Parameterized complexity of 1-planarity

Sergio Cabello (University of Ljubljana, SI)

Joint work of M. J. Bannister; Cabello, Sergio; D. Eppstein
Main reference M. J. Bannister, S. Cabello, D. Eppstein, "Parameterized complexity of 1-planarity," in Proc. of the 13th Int'l Symp. on Algorithms and Data Structures (WADS'13), LNCS, Vol. 8037, pp. 97–108, Springer, 2013; also available as pre-print as arXiv:1304.5591v1 [cs.DS].
URL http://dx.doi.org/10.1007/978-3-642-40104-6_9
URL http://arxiv.org/abs/1304.5591v1

We consider the problem of finding a 1-planar drawing for a general graph, where a 1-planar drawing is a drawing in which each edge participates in at most one crossing. Since this problem is known to be NP-hard we investigate the parameterized complexity of the problem with respect to the vertex cover number, tree-depth, and cyclomatic number. For these parameters we construct fixed-parameter tractable algorithms. However, the problem remains NP-complete for graphs of bounded bandwidth, pathwidth, or treewidth.

3.5 Multiple source shortest paths in embedded graphs

Erin Moriarty Wolf Chambers (St. Louis University, US)

Let G be a graph with non-negative edge lengths, embedded on an orientable surface of genus g, and let f be an arbitrary face of G. We describe an algorithm to preprocess the graph in $O(gn \log n)$ time, so that the shortest-path distance from any vertex on the boundary of f to any other vertex in G can be retrieved in $O(\log n)$ time. Our result directly generalizes the $O(n \log n)$-time algorithm of Klein [Proc. SODA 2005] for multiple-source shortest paths in planar graphs. Intuitively, our preprocessing algorithm maintains a shortest-path tree as its source point moves continuously around the boundary of f. As an application of our algorithm, we describe algorithms to compute a shortest non-contractible or non-separating cycle in G in $O(g^2 n \log n)$ time.

3.6 Catalan structures & embedded dynamic programming – a survey

Frederic Dorn (SINTEF – Trondheim, NO)

Joint work of Dorn, Frederic; Fomin, Fedor; Penninkx, Eelko; Bodlaender, Hans; Thilikos, Dimitrios
Main reference F. Dorn, E. Penninkx, H. L. Bodlaender, F. V. Fomin, "Efficient Exact Algorithms on Planar Graphs: Exploiting Sphere Cut Decompositions," Algorithmica 58(3):790–810, 2010.
URL http://dx.doi.org/10.1007/s00453-009-9296-1
Main reference F. Dorn, "Planar Subgraph Isomorphism Revisited," in Proc. of the 27th Int'l Symp. on Theoretical Aspects of Computer Science (STACS'10), LIPIcs, Vol. 5, pp. 263–274, Dagstuhl Publishing, 2010.
URL http://dx.doi.org/10.4230/LIPIcs.STACS.2010.2460

The topic of this talk is a survey on tools for solving NP-hard planar graph problems exactly. For basic problems such as Planar Vertex Cover a commonly used exact algorithm goes as follows: find a structure of the input graph with small separators which are used to do some Myhill-Nerode equivalence class type of dynamic programming. Planarity only plays a role for finding such structure, typically tree- or branch-decompositions, and for proving the bounded separator size. Sphere-cut decompositions represent a novel tool that allows one to do dynamic programming which explicitly exploits planarity. The small separators are connected by simple, closed curves in the plane and thereby obtaining a circular order of the separator vertices. For problems like Planar Longest Cycle, where the solution intersects the separator as a set non-crossing paths, one can bound the number of equivalent solutions by the Catalan numbers. Embedded dynamic programming goes one step further. For problems such as Planar Subgraph Isomorphism one looks at how a separator curve may intersect the pattern. The latter problem for patterns of size k and input graphs with n vertices may be solved in solved in time $2^{O(k)}n$.

3.7 Node-weighted network design in planar and minor-free graphs

Alina Ene (Princeton University, US and University of Warwick, GB)

Joint work of Chekuri, Chandra; Ene, Alina; Vakilian, Ali
Main reference C. Chekuri, A. Ene, A. Vakilian, "Node-weighted network design in planar and minor-closed families of graphs," in Proc. of the 39th Int'l Colloquium on Automata, Languages, and Programming (ICALP'12), Part I, LNCS, Vol. 7391, pp. 206–217, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-31594-7_18

We consider node-weighted network design problems in planar graphs. In particular, we focus on the survivable network design problem (SNDP). The input consists of a node-weighted undirected graph G and connectivity requirements $r(uv)$ for each pair of nodes uv. The goal is to find a minimum weight subgraph H of G such that, for each pair uv of nodes, H contains $r(uv)$ edge-disjoint paths between u and v. In this talk, we describe an $O(k)$-approximation algorithm for the problem when the graph is planar; here k is the maximum requirement of a pair. This improves the $O(k \log n)$-approximation known for node-weighted SNDP in general graphs [Nutov '10].

3.8 Minimum cuts and maximum flows in planar and surface graphs

Jeff Erickson (University of Illinois – Urbana Champaign, US)

This talk is a survey of the state of the art for the classical minimum cut and maximum flow problems for planar and surface-embedded graphs.

3.9 Online node-weighted Steiner forest in planar graphs and extensions

MohammadTaghi Hajiaghayi (University of Maryland, US)

Consider a graph $G = (V, E)$ with a weight value $w(v)$ associated with each vertex v. A demand is a pair of vertices (s, t). A subgraph H satisfies the demand if s and t are connected in H. In the (offline) node-weighted Steiner forest problem, given a set of demands the goal is to find the minimum-weight subgraph H which satisfies all demands. In the online variant, the demands arrive one by one and we need to satisfy each demand immediately; without knowing the future demands.

In the online variant of the problem, we give a randomized $O(\log^3(n))$-competitive algorithm. The competitive ratio is tight to a logarithmic factor. This result generalizes the recent result of Naor, Panigrahi, and Singh for the Steiner tree problem, thus answering one of their open problems. When restricted to planar graphs (and more generally graphs excluding a fixed minor) we give a deterministic primal-dual algorithm with a logarithmic competitive ratio which is tight to a constant factor.

3.10 Some techniques for approximation schemes in planar graphs

Philip N. Klein (Brown University, US)

An approximation scheme for an optimization problem gives a $(1 + \varepsilon)$-approximation algorithm for every $\varepsilon > 0$. I survey some techniques for obtaining approximation schemes for optimization problems in planar graph. I briefly illustrate Baker's framework in addressing vertex cover. I then turn to the traveling salesman problem (TSP) in planar graphs with edge-weights. I discuss a framework in which one can obtain a linear-time approximation scheme for TSP. The framework has been used to obtain approximation schemes for a variety of problems. The key step is computing a spanner designed for the specific optimization problem. I outline the spanner construction for TSP and for Steiner tree, and finish with a few words on the tools needed to extend the Steiner tree result to Steiner forest.

References

1 Brenda S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. Assoc. Comput. Mach.* 41:153–180, 1994.

2 Philip N. Klein. A linear-time approximation scheme for TSP in undirected planar graphs with edge-weights. *SIAM Journal on Computing* 37(6):1926–1952, 2008.
3 Glencora Borradaile, Philip N. Klein, and Claire Mathieu. An $O(n \log n)$ approximation scheme for Steiner tree in planar graphs. *ACM Trans. Algorithms* 5(3):article 31, 2009.
4 MohammadHossein Bateni, Mohammadtaghi Hajiaghayi, and Dániel Marx. Approximation schemes for Steiner forest on planar graphs and graphs of bounded treewidth. *J. ACM*, 58(5):21:1–21:37, October 2011.
5 David Eisenstat, Philip N. Klein, and Claire Mathieu. An efficient polynomial-time approximation scheme for Steiner forest in planar graphs. *Proceedings of the 23nd ACM-SIAM Symposium On Discrete Algorithms*, 626–638, 2012.

3.11 All-or-nothing multicommodity flow problem with bounded fractionality in planar graphs

Yusuke Kobayashi (University of Tokyo, JP)

Joint work of Kawarabayashi, Ken-ichi; Kobayashi, Yusuke
Main reference K.-I. Kawarabayashi, Y. Kobayashi, "All-or-nothing multicommodity flow problem with bounded fractionality in planar graphs," in Proc. of the 54th Annual IEEE Symp. on Foundations of Computer Science (FOCS'13), pp. 187–196, IEEE CS, 2013.
URL http://dx.doi.org/10.1109/FOCS.2013.28

We study the following all-or-nothing multicommodity flow problem in planar graphs.

Input: A graph G with n vertices and k pairs of vertices $(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)$ in G.

Find: A largest subset W of $\{1, \ldots, k\}$ such that for every i in W, we can send one unit of flow between s_i and t_i.

This problem is different from the well-known maximum edge-disjoint paths problem in that we do not require integral flows for the pairs. This problem is APX-hard even for trees, and a 2-approximation algorithm is known for trees. For general graphs, Chekuri et al. (STOC'04) give a poly-logarithmic factor approximation algorithm and show that a natural LP-relaxation has a poly-logarithmic integrality gap. This result is in contrast with the integrality gap $\Omega(\sqrt{n})$ for the maximum edge-disjoint paths problem.

Our main result considerably strengthens this result when an input graph is planar. Namely, for the all-or-nothing multicommodity flow problem in planar graphs, we give an $O(1)$-approximation algorithm and show that the integrality gap is $O(1)$. In particular, in polynomial time, we can find an index set W with $|W| = \Omega(\mathrm{OPT})$ and eight s_i-t_i paths for each $i \in W$ such that each edge is used at most eight times in these paths (with multiplicity), where OPT is the optimal value of the LP-relaxation of the all-or-nothing multicommodity flow problem.

Our result can be compared to the recent result by Séguin-Charbonneau and Shepherd (FOCS'11) who give an $O(1)$-approximation algorithm for the maximum edge-disjoint paths problem in planar graphs with congestion 2 (but not implied by this result).

3.12 Prize-collecting network design in planar graphs

Nitish Korula (Google – New York, US)

Joint work of Bateni, MohammadHossein; Chekuri, Chandra; Ene, Alina; Hajiaghayi, MohammadTaghi; Korula, Nitish; Marx, Dániel
Main reference M. Bateni, C. Chekuri, A. Ene, M. T. Hagjiaghayi, N. Korula, D. Marx, "Prize-collecting Steiner problems on planar graphs," in Proc. of the 22nd ACM-SIAM Symp. on Discrete Algorithms (SODA'11), pp. 1028–1049, SIAM, 2011.
URL http://dx.doi.org/10.1137/1.9781611973082.79

In this talk, based on work in [1], we describe reductions from Prize-Collecting Steiner TSP (PCTSP), Prize-Collecting Stroll (PCS), Prize-Collecting Steiner Tree (PCST), Prize-Collecting Steiner Forest (PCSF), and more generally Submodular Prize-Collecting Steiner Forest (SPCSF), on planar graphs (and also on bounded-genus graphs) to the corresponding problems on graphs of bounded treewidth. We show that for each of the mentioned problems, an α-approximation algorithm for the problem on graphs of bounded treewidth implies an $(\alpha + \epsilon)$-approximation algorithm for the problem on planar graphs (and also bounded-genus graphs), for any constant $\epsilon > 0$. PCS, PCTSP, and PCST can be solved exactly on graphs of bounded treewidth and hence we obtain a PTAS for these problems on planar graphs and bounded-genus graphs.

In contrast, we show that PCSF is APX-Hard on series-parallel graphs, which are planar graphs of treewidth at most 2. Apart from ruling out a PTAS for PCSF on planar graphs and bounded-treewidth graphs, this result is also interesting since it gives the first provable hardness separation between the approximability of a problem and its prize-collecting version.

References

1 MohammadHossein Bateni, Chandra Chekuri, Alina Ene, Mohammad Taghi Hagjiaghayi, Nitish Korula and Dániel Marx. Prize- collecting Steiner problems on planar graphs. *SODA*, 1028–1049, 2011.

3.13 Kernels for planar graph problems

Daniel Lokshtanov (University of Bergen, NO)

Joint work of Fomin, Fedor; Lokshtanov, Daniel; Saurabh, Saket
Main reference F. V. Fomin, D. Lokshtanov, S. Saurabh, "Efficient Computation of Representative Sets with Applications in Parameterized and Exact Algorithms," to appear in Proc. of the 25th Symp. on Discrete Algorithms (SODA'14), 2014.
URL http://www.ii.uib.no/~daniello/papers/EfficientRepSet.pdf

Bollobás' lemma and its generalization to matroids, due to Lovász, are classical results in extremal combinatorics. In this talk we will discuss algorithmic variants of these lemmas, and survey some recent applications in parameterized and exact algorithms.

3.14 The square-root phenomenon in planar graphs

Dániel Marx (Hungarian Academy of Sciences, HU)

Most of the classical NP-hard problems remain NP-hard when restricted to planar graphs, and only exponential-time algorithms are known for the exact solution of these planar problems. However, in many cases, the exponential-time algorithms on planar graphs are significantly faster than the algorithms for general graphs: for example, 3-Coloring can be solved in time $2^{O(\sqrt{n})}$ in an n-vertex planar graph, whereas only $2^{O(n)}$-time algorithms are known for general graphs. For various planar problems, we often see a square-root appearing in the running time of the best algorithms, e.g., the running time is often of the form $2^{O(\sqrt{n})}$, $n^{O(\sqrt{k})}$, or $2^{O(\sqrt{k})} \cdot n$. By now, we have a good understanding of why this square-root appears. On the algorithmic side, most of these algorithms rely on the notion of treewidth and its relation to grid minors in planar graphs (but sometimes this connection is not obvious and takes some work to exploit). On the lower bound side, under a complexity assumption called Exponential Time Hypothesis (ETH), we can show that these algorithms are essentially best possible, and therefore the square root has to appear in the running time.

3.15 Approximating k-center in planar graphs

Claire Mathieu (Brown University, US)

Joint work of Eisenstat, David; Klein, Philip; Mathieu, Claire
Main reference D. Eisenstat, P. N. Klein, C. Mathieu, "Approximating k-center in planar graphs," to appear in the Proc. of the 25th Symp. on Discrete Algorithms (SODA'14), 2014.
URL http://www.davideisenstat.com/cv/EisenstatKM14.pdf

We consider variants of the metric k-center problem. Imagine that you must choose locations for k rehouses in a city so as to minimize the maximum distance of a house from the nearest rehouse. An instance is specified by a graph with arbitrary nonnegative edge lengths, a set of vertices that can serve as rehouses (i.e., centers) and a set of vertices that represent houses. For general graphs, this problem is exactly equivalent to the metric k-center problem, which is APX-hard. We give a polynomial-time bicriteria approximation scheme when the input graph is a planar graph [1].

References

1 David Eisenstat, Philip N. Klein, and Claire Mathieu. *Approximating k-center in planar graphs.* In Proceedings of the Twenty-Fifth Symposium on Discrete Algorithms (SODA), January 2014. To be published.

3.16 Lower and upper bounds for long induced paths in 3-connected planar graphs

Tamara Mchedlidze (KIT – Karlsruhe Institute of Technology, DE)

Joint work of Mchedlidze, Tamara; Di Giacomo, Emilio; Liotta, Giuseppe
Main reference E. Di Giacomo, G. Liotta, T. Mchedlidze, "Lower and Upper Bounds for Long Induced Paths in 3-connected Planar Graphs," in Proc. of the 39th Int'l Workshop on Graph-Theoretic Concepts in Computer Science (WG'13), LNCS, Vol. 8165, pp. 213–224, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-45043-3_19

Let G be a 3-connected planar graph with n vertices and let $p(G)$ be the maximum number of vertices of an induced subgraph of G that is a path. We prove that $p(G) \geq \frac{\log n}{12 \log \log n}$. To demonstrate the tightness of this bound, we notice that the above inequality implies $p(G) \in \Omega((\log_2 n)^{1-\varepsilon})$, where ε is any positive constant smaller than 1, and describe an infinite family of 3-connected planar graphs for which $p(G) \in O(\log n)$.

As a byproduct of our research, we prove a result of independent interest: Every 3-connected planar graph with n vertices contains an induced subgraph that is outerplanar and connected and that contains at least $\sqrt[3]{n}$ vertices. The proofs in the paper are constructive and give rise to $O(n)$-time algorithms.

3.17 Large independent sets in triange-free planar graphs

Matthias Mnich (Universität des Saarlandes, DE)

Joint work of Dvorak, Zdenek; Mnich, Matthias
Main reference Z. Dvorak, M. Mnich, "Large Independent Sets in Triangle-Free Planar Graphs," arXiv:1311.2749v1 [cs.DM], 2013.
URL http://arxiv.org/abs/1311.2749v1

Every triangle-free planar graph on n vertices has an independent set of size at least $(n+1)/3$, and this lower bound is tight. We give an algorithm that, given a triangle-free planar graph G on n vertices and an integer $k >= 0$, decides whether G has an independent set of size at least $(n + k)/3$, in time $2^{O(\sqrt{k})}n$. Thus, the problem is fixed-parameter tractable when parameterized by k. Furthermore, as a corollary of the result used to prove the correctness of the algorithm, we show that there exists $\varepsilon > 0$ such that every planar graph of girth at least five on n vertices has an independent set of size at least $n/(3 - \varepsilon)$.

3.18 Multiple-source multiple-sink maximum flow in directed planar graphs in near-linear time

Shay Mozes (Interdisciplinary Center Herzliya, IL)

Joint work of Borradaile, Glencora; Klein, Philip; Mozes, Shay; Nussbaum, Yahav; Wulff-Nilsen, Christian
Main reference G. Borradaile, P. Klein, S. Mozes, Y. Nussbaum, C. Wulff-Nilsen, "Multiple-Source Multiple-Sink Maximum Flow in Directed Planar Graphs in Near-Linear Time," in Proc. of the IEEE 52nd Annual Symp. on Foundations of Computer Science (FOCS'11), pp. 170–179, IEEE, 2011.
URL http://dx.doi.org/10.1109/FOCS.2011.73

In this talk I describe an $O(n \log^3 n)$ algorithm that, given an n-node directed planar graph with arc capacities, a set of source nodes, and a set of sink nodes, finds a maximum flow from the sources to the sinks. I give an overview of the algorithm and go into the details of a procedure to redistribute flow among nodes of a cycle separator. The procedure is based on a representation of circulations via face potentials and efficiently computing shortest paths using an extension of Fakcharoenphol and Rao's fast Dijkstra on dense distance graphs.

3.19 Min-cost flow duality in planar networks

Yahav Nussbaum (Tel Aviv University, IL)

Joint work of Kaplan, Haim; Nussbaum, Yahav
Main reference H. Kaplan, Y. Nussbaum, "Min-Cost Flow Duality in Planar Networks," arXiv:1306.6728v1 [cs.DM] , 2013.
URL http://arxiv.org/abs/1306.6728v1

In this talk we will discuss the minimum-cost flow problem in planar graphs.

We begin with a minimum-cost flow problem in a planar graph and modify the problem using the following two transformations. First, we express the problem as a problem in the geometric dual graph. Then, we find the linear programming dual of this problem. The result is a minimum-cost flow problem in a related planar graph, such that the balance constraints are defined by the costs of the original problem, and the costs are defined by the capacities of the original problem.

As an application for our transformation, we show an $O(n \log^2 n)$ time algorithm for the minimum-cost flow problem in an n-vertex outerplanar graph, which takes advantage of the simple structure of the dual graphs of outerplanar graphs.

3.20 Network sparsification for Steiner problems on planar and bounded-genus graphs

Marcin Pilipczuk (University of Warsaw, PL)

Joint work of Pilipczuk, Marcin; Pilipczuk, Michał; Sankowski, Piotr; van Leeuwen, Erik Jan
Main reference M. Pilipczuk, M. Pilipczuk, P. Sankowski, E. J. van Leeuwen, "Network Sparsification for Steiner Problems on Planar and Bounded-Genus Graphs," arXiv:1306.6593v1 [cs.DS], 2013.
URL http://arxiv.org/abs/1306.6593v1

We propose polynomial-time algorithms that sparsify planar and bounded-genus graphs while preserving optimal solutions to Steiner problems. Our main contribution is a polynomial-time algorithm that, given a graph G embedded on a surface of genus g and a designated face f bounded by a simple cycle of length k, uncovers a set $F \subseteq E(G)$ of size polynomial in g and k that contains an optimal Steiner tree for *any* set of terminals that is a subset of the vertices of f.

We apply this general theorem to prove that:

- given a graph G embedded on a surface of genus g and a terminal set $S \subseteq V(G)$, one can in polynomial time find a set $F \subseteq E(G)$ that contains an optimal Steiner tree T for S and that has size polynomial in g and $|E(T)|$;
- an analogous result holds for the Steiner Forest problem;
- given a planar graph G and a terminal set $S \subseteq V(G)$, one can in polynomial time find a set $F \subseteq E(G)$ that contains an optimal edge multiway cut C separating S (i.e., a cutset that intersects any path with endpoints in different terminals from S) and that has size polynomial in $|C|$.

In the language of parameterized complexity, these results imply the first polynomial kernels for Steiner Tree and Steiner Forest on planar and bounded-genus graphs (parameterized by the size of the tree and forest, respectively) and for Edge Multiway Cut on planar graphs (parameterized by the size of the cutset).

3.21 The k-disjoint paths problem on directed planar graphs

Michał Pilipczuk (University of Bergen, NO)

Joint work of Cygan, Marek; Marx, Dániel; Pilipczuk, Marcin; Pilipczuk, Michał
Main reference M. Cygan, D. Marx, M. Pilipczuk, M. Pilipczuk, "The Planar Directed K-Vertex-Disjoint Paths Problem Is Fixed-Parameter Tractable," in Proc. of the 54th Annual Symp. on Foundations of Computer Science (FOCS'13), pp. 197–206, IEEE CS, 2013; also available as pre-print as arXiv:1304.4207v1 [cs.DM].
URL http://dx.doi.org/10.1109/FOCS.2013.29
URL http://arxiv.org/abs/1304.4207

It was shown in the 90s by Schrijver that the k-Disjoint Paths problem can be solved in $n^{O(k)}$ time on directed planar graphs [1]. In this work we present an FPT algorithm for the problem working in time $f(k) * n^{O(1)}$ for a doubly-exponential function f.

References

1 Alexander Schrijver. *Finding k Disjoint Paths in a Directed Planar Graph*. SIAM Journal on Computing 23(4), 1994

3.22 A near-optimal planarization algorithm

Saket Saurabh (The Institute of Mathematical Sciences – Chennai, IN)

Joint work of Jansen, Bart M. P.; Lokshtanov, Daniel; Saurabh, Saket

The problem of testing whether a graph is planar has been studied for over half a century, and is known to be solvable in $O(n)$ time using a myriad of different approaches and techniques. Robertson and Seymour established the existence of a cubic algorithm for the more general problem of deciding whether an n-vertex graph can be made planar by at most k vertex deletions, for every fixed k. Of the known algorithms for k-Vertex Planarization, the algorithm of Marx and Schlotter (WG 2007, Algorithmica 2012) running in time $2^{k^{O(k^3)}} \cdot n^2$ achieves the best running time dependence on k. The algorithm of Kawarabayashi (FOCS 2009), running in time $f(k)n$ for some $f(k) \in \Omega\left(2^{k^{\Omega(k^3)}}\right)$ that is not stated explicitly, achieves the best dependence on n.

In this paper we present an algorithm for k-Vertex Planarization with running time $2^{O(k \log k)} \cdot n$, significantly improving the running time dependence on k without compromising the linear dependence on n. Our main technical contribution is a novel scheme to reduce the treewidth of the input graph to $O(k)$ in time $2^{O(k \log k)} \cdot n$. It combines new insights into the structure of graphs that become planar after contracting a matching, with a Baker-type subroutine that reduces the number of disjoint paths through planar parts of the graph that are not affected by the sought solution. To solve the reduced instances we formulate a dynamic programming algorithm for Weighted Vertex Planarization on graphs of treewidth w with running time $2^{O(w \log w)} \cdot n$, thereby improving over previous double-exponential algorithms.

While Kawarabayashi's planarization algorithm relies heavily on deep results from the graph minors project, our techniques are elementary and practically self-contained. We expect them to be applicable to related edge-deletion and contraction variants of planarization problems.

3.23 Approximation algorithms for Euler genus, and related problems

Anastasios Sidiropoulos (University of Illinois – Urbana Champaign, US)

Joint work of Chekuri, Chandra; Sidiropoulos, Anastasios

The Euler genus of a graph is a fundamental and well-studied parameter in graph theory and topology. Computing it has been shown to be NP-hard by [Thomassen '89 & '93], and it is known to be fixed-parameter tractable. However, the approximability of the Euler genus is wide open. While the existence of an $O(1)$-approximation is not ruled out, only an $O(\sqrt{(n)})$-approximation [Chen, Kanchi, Kanevsky '97] is known even in bounded degree graphs. In this paper we give a polynomial-time algorithm which on input a bounded-degree graph of Euler genus g, computes a drawing into a surface of Euler genus $poly(g, log(n))$. Combined with the upper bound from [Chen, Kanchi, Kanevsky '97], our result also implies a $O(n^{1/2-\alpha})$-approximation, for some constant $\alpha > 0$. Using our algorithm for approximating the Euler genus as a subroutine, we obtain, in a unified fashion, algorithms with approximation

ratios of the form $poly(\mathrm{OPT}, log(n))$ for several related problems on bounded degree graphs. These include the problems of orientable genus, crossing number, and planar edge and vertex deletion problems. Our algorithm and proof of correctness for the crossing number problem is simpler compared to the long and difficult proof in the recent breakthrough by [Chuzhoy 2011], while essentially obtaining a qualitatively similar result. For planar edge and vertex deletion problems our results are the first to obtain a bound of form $poly(\mathrm{OPT}, log(n))$.

We also highlight some further applications of our results in the design of algorithms for graphs with small genus. Many such algorithms require that a drawing of the graph is given as part of the input. Our results imply that in several interesting cases, we can implement such algorithms even when the drawing is unknown.

3.24 Separators in planar graphs with applications

Christian Wulff-Nilsen (University of Copenhagen, DK)

For planar graphs, we have separators with small size and/or nice structural properties. I present some of these and give applications to shortest path and min cut/max flow problems. Minor-free and shallow minor-free graps also have good separators but we do not know how to compute them in linear time. I will present faster separator algorithms for these graph classes and give applications to shortest paths and maximum matching.

4 Open Problems

4.1 Planarization by vertex deletion

Proposed by Peter Rossmanith (rossmani@cs.rwth-aachen.de)

Given a graph with m edges, can it be made planar by deleting at most $m/6$ vertices?

4.2 All-pair distances on the infinite face

Proposed by Shay Mozes (smozes@idc.ac.il)

Here is a special case of multiple-source shortest paths: Given a planar graph, one can compute all the shortest path trees rooted at vertices on the infinite face f_∞, in total time $O(n \log n)$ (Klein, Cabello-Chambers-Erickson), and there is a matching lower bound. Can one compute all-pair distances between vertices of f_∞ in time $O(n + |f_\infty|^2)$?

4.3 Two-edge-connected planar subgraph

Proposed by Philip Klein (klein@brown.edu)

Given a planar graph G with edge weights, find a subgraph H such that $V(H) = V(G)$, H is 2-edge-connected, and the total weight of H is minimum. Does there exist an efficient PTAS, that is, a PTAS with running time $f(\epsilon)n^c$, where c is an absolute constant (independent of ϵ)?

The problem has linear-time PTAS (Grigni) in the special case when all edge weights are 1, and there is a PTAS with running time $n^{f(\epsilon)}$ for general weights. The problem is APX-hard for general graphs.

A similar problem is two-edge-connected Steiner subgraph where $V(H)$, instead of being equal to $V(G)$, must contain a specified set T of terminal vertices. Is there a PTAS?

4.4 Euclidean multiway cut with unit disks

Proposed by Sergio Cabello (sergio.cabello@fmf.uni-lj.si)

Given k points $s_1 s_2, \ldots s_k$ in the Euclidean plane, and a collection of unit disks (none of which contains any s_i), find a minimum cardinality set of disks that separates every s_i from every s_j. Is there a PTAS?

The problem is NP-hard and has a constant factor approximation. In the special case where $k = 2$, the problem is in P. In the generalization where disks have weights, is there a constant factor approximation? In the variant where the goal is to separate s_1 from every s_i for $i \neq 1$, is the problem FPT?

4.5 Minimum stretch shape shifting

Proposed by Erin Chambers (echambe5@slu.edu)

Given an unweighted triangulated planar graph and two curves ℓ and r, "morph" ℓ into r by a sequence of elementary moves while minimizing the length of the longest intermediate curve. There are two types of elementary moves: either replace one edge of a triangle t by the other two edges of t, or vice versa, or, when the curve goes through a vertex x, insert $(x, y), (y, x)$ into the curve, where $\{x, y\}$ is an edge.

A logarithmic approximation is known by divide-and-conquer using a shortest path, even if the edges have weights. It is not known whether the problem is in NP, nor whether it is NP-hard. In the unweighted triangulation case it is also not known whether the optimal sequence is monotonic, that is, never traverses the same triangle twice. (There is an example where the motion is not purely monotonic, but it is in a graph with appropriate weights.)

4.6 Vertex-capacitated max flow in directed planar graphs

Proposed by Jeff Erickson (jeffe@illinois.edu)

Here is a table showing best known strongly-polynomial running times for some variants of max flow in directed planar graphs:

	source s, sink t	sources s, s', sink t
edge capacities	$O(n \log n)$	$O(n \log n)$
vertex capacities	$O(n \log n)$	$O(n^2 \log n)$

The top-right variant can be addressed by first finding a max st-flow and then finding a max $s't$-flow in the residual graph. The bottom-left variant can be addressed by using a reduction from vertex-capacities to edge-capacities.

For the fourth variant, there is no planarity-exploiting algorithm known!

1. Cannot use the reduction because it leads to violation at one vertex.
2. Cannot use the residual graph because there isn't a residual graph with respect to vertex capacities.

4.7 k-*minimum spanning tree* problem

Proposed by Alina Ene (aene@cs.princeton.edu)

- *input:* undirected edge-weighted graph, integer k
- *output:* subgraph tree T with at least k vertices
- *goal:* minimize the weight of the tree

In general graphs, this is SNP-hard, and there is a 2-approximation algorithm due to Garg, using primal-dual techniques. There exists a PTAS for the Euclidean case. What happens in planar graphs? Is there a PTAS? Is the problem APX-hard? Indications are that the spanner approach will not work.

4.8 Prize-collecting Steiner forest

Proposed by Alina Ene (aene@cs.princeton.edu)

Prize-collecting Steiner forest is APX-hard in series-parallel graphs.

The best approximation we have for bounded treewidth is one for general graphs (the approximation ratio is 2.54). Can we do better? If so, it would yield an improved approximation ratio for planar graphs as well.

4.9 Achieving subexponential time for several parameterized problems in planar graphs

Proposed by Marcin Pilipczuk (malcin@mimuw.edu.pl)

1. directed k-path
2. weighted (undirected) k-path: path on k vertices of minimum weight
3. Exact k-cycle
4. Steiner tree parameterized by number of terminals
5. k-MSR parameterized by k
6. Subgraph isomorphism parameterized by size of subgraph

All but the fourth can be solved in $(1+\varepsilon)^k n^{\varepsilon^{-2}}$ so we don't expect a ETH lower bound.

They should have running times of the form $2^{\sqrt{k}\text{polylog } k} n^c$ but it would be interesting to have $2^{o(k)} n^c$.

- For (undirected) planar disjoint paths, the bound is $f(k) \cdot \text{poly}$ where $f(k)$ is triply exponential in k. Can we get $2^{\text{poly } k} \cdot \text{poly}$?
- Finding an independent set of size $\frac{m}{4} + k$ – is this fixed-parameter tractable?
- Kernels for feedback vertex set, dominating set. Can the dependence of the kernel size on k be reduced, from around $100k$ to, say, $10k$?
- Vertex cover. There is a $2k$-vertex kernel. It is believed that one cannot do better because it might imply a better approximation than 2, which we don't expect. But what about planar graphs?
- k-vertex deletion to get a planar graph, or, more generally, to get a graph that excludes a fixed graph H as a subgraph. Hitting all minors leads to a bound that is doubly exponential in treewidth. Can we get a bound of $2^{\text{poly } tw} \cdot \text{poly(input graph)}$?

4.10 Recognizing map graphs

Proposed by Daniel Lokshtanov (daniello@ii.uib.no)

A map graph derives from a modified notion of planarity in which two (connected) regions of a map are considered adjacent when they share a point of their boundaries (not an edge, as standard planarity requires) (Chen, Grigni, Papadimitriou, STOC'98). How quickly can we recognize them? Thorup has $n^{\geq 120}$. Can $O(n^{10})$ be achieved?

4.11 Planar local TSP

Proposed by Rolf Niedermeier (rolf.niedermeier@tu-berlin.de)

- *input:* edge-weighted planar graph, Hamilton cycle C, given by a permutation π of the vertices, integer $k \geq 0$

- *output:* Does there exist a permutation π' such that $\lambda(\pi, \pi') \leq k$ and such that π' gives a shorter tour?

Here λ is a funtion measuring distance between permutations. There are two versions of λ that are the same up to a factor of 2:

- number of edges in symmetric difference
- number of reversals

For λ =number of swaps, the problem is known to be FPT.

4.12 Subgraph isomorphism in planar graphs with a twist

Proposed by Dániel Marx (dmarx@cs.bme.hu)

Testing whether H is isomorphic to a subgraph of G, where the parameter k is the difference $|E(G)| - |E(H)|$. Is this problem FPT?

For the parameter being zero, the problem is isomorphism, which is solvable in polynomial time.

Can easily achieve n^kpoly (n). If H is also 3-connected, the problem is FPT.

4.13 Mincost flow in planar graphs

Proposed by Jeff Erickson (jeffe@illinois.edu)

- Every vertex has a supply value (could be negative or zero).
- Every arc has a capacity and a cost.

Assign a flow value to every arc such that net flow out of a vertex equals the supply of the vertex. Flow values are nonnegative but no more than capacity.

Goal: minimize cost.

- Without loss of generality, can assume each vertex's supply is zero , which yields the problem *min-cost circulation*.
- Alternatively, can assume capacities are infinite, which yields the *transshipment* problem.

Can you give an algorithm whose running time is strongly polynomial-time and beats $O(n^2 \log n)$?

If maximum cost is $O(1)$ or or max capacity is $O(1)$, can achieve $O(n^{1.5})$ time (Cornelsen and Karrenbauer; Cornelsen, Karrenbauer, Li).

Participants

- MohammadHossein Bateni
Google – New York, US
- Ivona Bezakova
Rochester Institute of Technology, US
- Therese Biedl
University of Waterloo, CA
- Glencora Borradaile
Oregon State University, US
- Sergio Cabello
University of Ljubljana, SI
- Erin Moriarty Wolf Chambers
St. Louis University, US
- Eric Colin de Verdière
ENS – Paris, FR
- Sabine Cornelsen
Universität Konstanz, DE
- Arnaud de Mesmay
ENS – Paris, FR
- Frederic Dorn
SINTEF – Trondheim, NO
- Alina Ene
Princeton University, US and University of Warwick, GB
- Jeff Erickson
University of Illinois – Urbana Champaign, US
- Jittat Fakcharoenphol
Kasetsart Univ. – Bangkok, TH
- Kyle Jordan Fox
University of Illinois – Urbana Champaign, US
- Petr A. Golovach
University of Bergen, NO
- Michelangelo Grigni
Emory University, US
- MohammadTaghi Hajiaghayi
University of Maryland, US
- Marcin Kaminski
University of Warsaw, PL
- Philip N. Klein
Brown University, US
- Yusuke Kobayashi
University of Tokyo, JP
- Nitish Korula
Google – New York, US
- Daniel Lokshtanov
University of Bergen, NO
- Dániel Marx
Hungarian Acad.of Sciences, HU
- Claire Mathieu
Brown University, US
- Tamara Mchedlidze
KIT – Karlsruhe Institute of Technology, DE
- Matthias Mnich
Universität des Saarlandes, DE
- Shay Mozes
Interdisciplinary Center Herzliya, IL
- Matthias Müller-Hannemann
Martin-Luther-Universität Halle-Wittenberg, DE
- Amir Nayyeri
University of Illinois – Urbana Champaign, US
- Rolf Niedermeier
TU Berlin, DE
- Yahav Nussbaum
Tel Aviv University, IL
- Marcin Pilipczuk
University of Warsaw, PL
- Michał Pilipczuk
University of Bergen, NO
- Peter Rossmanith
RWTH Aachen, DE
- Ignaz Rutter
KIT – Karlsruhe Institute of Technology, DE
- Saket Saurabh
The Institute of Mathematical Sciences – Chennai, IN
- Anastasios Sidiropoulos
University of Illinois – Urbana Champaign, US
- Erik Jan van Leeuwen
MPI für Informatik – Saarbrücken, DE
- Oren Weimann
Haifa University, IL
- Christian Wulff-Nilsen
University of Copenhagen, DK

Report from Dagstuhl Seminar 13422

Nominal Computation Theory

Edited by

Mikołaj Bojańczyk[1], Bartek Klin[2], Alexander Kurz[3], and Andrew M. Pitts[4]

1 **University of Warsaw, PL,** `bojan@mimuw.edu.pl`
2 **University of Warsaw, PL,** `klin@mimuw.edu.pl`
3 **University of Leicester, GB,** `kurz@mcs.le.ac.uk`
4 **University of Cambridge, GB,** `andrew.pitts@cl.cam.ac.uk`

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13422 "Nominal Computation Theory". The underlying theme of the seminar was nominal sets (also known as sets with atoms or Fraenkel-Mostowski sets) and they role and applications in three distinct research areas: automata over infinite alphabets, program semantics using nominal sets and nominal calculi of concurrent processes.

Seminar 13–16. October, 2013 – www.dagstuhl.de/13422
1998 ACM Subject Classification F.1.1 Models of computation – automata, Turing machines, computability theory, relations between models, F.3.2 Semantics of programming languages – denotational semantics, operational semantics, process models, F.4.1 Mathematical logic – logic and constraint programming, mechanical theorem proving, set theory, F.1.2 Modes of computation – alternation and nondeterminism, D.3.3 Language constructs and features – control structures, data types and structures
Keywords and phrases nominal sets, Fraenkel-Mostowski sets
Digital Object Identifier 10.4230/DagRep.3.10.58
Edited in cooperation with Joanna Ochremiak

1 Executive Summary

Mikołaj Bojańczyk
Bartek Klin
Alexander Kurz
Andrew M. Pitts

The short Dagstuhl seminar 13422 "Nominal Computation Theory" took place from October 13th to 16th, 2013. The topic of the seminar was the theory of nominal sets and their applications to Computer Science. The seminar arose from a recent exciting and unexpected confluence of the following three distinct research directions.

- The research in automata theory on automata over infinite alphabets with applications to querying XML and databases.
- The research in program semantics on nominal sets, with many applications to the syntax and semantics of programming language constructs that involve binding, or localising names.

Nominal Computation Theory, *Dagstuhl Reports*, Vol. 3, Issue 10, pp. 58–71
Editors: Mikołaj Bojańczyk, Bartek Klin, Alexander Kurz, and Andrew M. Pitts

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

- The research in concurrency on nominal calculi (π-calculus, etc) with applications to the automatic verification of process specifications.

In each of these three topics, an important role is played by name (or atom) symmetries and permutations, albeit for a priori different reasons. In the first case they arise from the way automata use registers to store letters, in the second case they are used to define the notion freshness, in the third case they are needed to minimize automata. In all three cases there is a connection with mathematical model theory, which is aimed at studying classes of mathematical structures definable by logical theories. The permutations allowed on atomic names can be usefully understood as automorphisms of a relational structure on those names. Model-theoretic notions such as homogeneity, algebraic closure and oligomorphic groups turn out very useful in describing those relational structures on atoms that yield meaningful theories of nominal sets.

The aim of the seminar was to profit from the excitement created by the confluence described above and to explore new directions with a new mix of research communities from computer science and mathematical logic. The main topics of interest included: automata and complexity theory in nominal sets, verification of nominal automata, symmetry in domain theory, and nominal programming.

The seminar was attended by 30 participants from 8 countries; 20 of them gave presentations, whose abstracts are included in this document. Four of the presentations (A. Pitts, M. Bojańczyk, B. Klin and N. Tzevelekos) were extended tutorials that presented various points of view on the background topics of the meeting. Other speakers presented the current state of the art in the field, with topic varying from mathematical insights into the nature of nominal sets (A. Blass, D. Petrisan), to applications in automata theory (V. Ciancia, T. Colcombet, S. Lasota, T. Suzuki) and computation theory (A. Dawar, S. Toruńczyk), semantics and domain theory (R. Crole, J. Gabbay, S. Loesch, A. Murawski), process calculi (U. Montanari), Petri nets (R. Lazic), logic programming (J. Cheney) and theorem proving (C. Urban).

2 Table of Contents

3 Overview of Talks

3.1 Set-theoretic aspects of orbit-finiteness

Andreas R. Blass (University of Michigan, US)

I show that a set in a permutation model of set theory is orbit-finite, in the sense of nominal computation theory, if and only if its power set in the permutation model is Dedekind-finite there. I give some equivalent characterizations of Dedekind-finiteness of the power set, and I also discuss some other notions of finiteness in set theory lacking the axiom of choice.

3.2 Orbit finiteness

Mikołaj Bojańczyk (University of Warsaw, PL)

Joint work of Bojańczyk, Mikołaj; Klin, Bartosz; Lasota, Sławomir; Toruńczyk, Szymon

There are two principal topics of this tutorial:

1. Sets with atoms can be based on an arbitrary logical structure. So instead of a countable set of names with equality, the atoms can be the rational numbers with order, or the integers with successor.
2. In the presence of atoms, one can consider a relaxed version of finiteness, called orbit-finiteness. When that logical structure for the atoms is oligomorphic (which is a notion from model theory that is equivalent to being omega-categorical), then the notion of orbit-finiteness is well behaved in that:
   a. there are several equivalent definitions of orbit-finiteness;
   b. orbit-finite sets are closed under finite products and finitely supported subsets.

Furthermore, under natural effectivity assumptions on the atoms (e.g. a decidable first-order theory), orbit-finite sets can be represented by data structures and manipulated by algorithms.

3.3 Nominal logic programming

James Cheney (University of Edinburgh, GB)

Joint work of Cheney, James; Urban, Christian; Momigliano, Alberto
Main reference J. Cheney, C. Urban, "Nominal logic programming," ACM Trans. Program. Lang. Syst. 30, 5, Article 26 (September 2008), 47 pp., 2008.
URL http://dx.doi.org/10.1145/1387673.1387675

Nominal logic programming is an extension to logic programming to incorporate features of nominal logic, such as freshness constraints, names, abstraction, and the new-quantifier. It is suitable for direct translations of many typical formal systems (type systems, functional and concurrency calculi, and languages with generative features such as references). In this talk I present examples of programs in an implemented nominal logic programming language

called alphaProlog, and show how we can search automatically for shallow counterexamples to desired properties of such programs. I also present an overview of the least Herbrand model and least fixed-point semantics of nominal logic programs, and describe the challenges involved in unification modulo equivariance and how to overcome them.

3.4 A decidable class of (nominal) omega-regular languages over an infinite alphabet

Vincenzo Ciancia (CNR – Pisa, IT)

Joint work of Ciancia, Vincenzo; Sammartino, Matteo
Main reference V. Ciancia, M. Sammartino, "A decidable class of (nominal) omega-regular languages over an infinite alphabet," arXiv:1310.3945v1 [cs.FL], 2013.
URL http://arxiv.org/abs/1310.3945v1

We define a class of languages of infinite words over infinite alphabets, and the corresponding automata. The automata used for recognition are a generalisation of deterministic Muller automata to the setting of nominal sets. Remarkably, the obtained languages are determined by their ultimately periodic fragments, as in the classical case. Closure under complement, union and intersection, and decidability of emptiness and equivalence are preserved by the generalisation. This is shown by using finite representations of the (otherwise infinite-state) defined class of automata.

3.5 On the Use of Guards for Logics with Data

Thomas Colcombet (CNRS / University Paris-Diderot, FR)

Joint work of Colcombet, Thomas; Ley, Clemens; Puppis, Gabriele
Main reference T. Colcombet, C. Ley, G. Puppis, "On the Use of Guards for Logics with Data," in Proc. of the 36th Int'l Symp. of Mathematical Foundations of Computer Science (MFCS'11), LNCS, Vol. 6907, pp. 243–255, Springer, 2011.
URL http://dx.doi.org/10.1007/978-3-642-22993-0_24

The notion of orbit finite data monoid (i.e., nominal monoids) was recently introduced by Bojańczyk as an algebraic object for defining recognizable languages of data words. Following Büchi's approach, we introduce the new logic 'rigidly guarded MSO' and show that the data languages definable in this logic are exactly those recognizable by orbit finite data monoids. We also establish, following this time the approach of Schützenberger, McNaughton and Papert, that the first-order variant of this logic defines exactly the languages recognizable by aperiodic orbit finite data monoids. Finally, we give a variant of the logic that captures the larger class of languages recognized by non-deterministic finite memory automata.

3.6 Nominal Lambda Calculus: An Internal Language for FM-Cartesian Closed Categories

Roy L. Crole (University of Leicester, GB)

Joint work of Crole, Roy L.; Nebel, Frank

Theories in equational logic (also known as algebraic theories, or theories for equational reasoning) have been studied for many years. Nominal Equational Logic provides judgements both of expression equality, and of freshness of atoms. Just as Equational Logic can be enriched with function types to yield the lambda-calculus, we have developed Nominal Lambda Calculus (NLC) by enriching NEL with (atom-dependent) function types and abstraction types.

We introduce pure NLC and look in detail at some the the rules of the NLC formal system. We show that NLC is sound by giving a categorical semantics, explaining some of the complexities that arise from a dependent type system. We then consider completeness and illustrate that we cannot define a complete categorical semantics for pure NLC. However, by adding in a novel form of dependently typed atom abstraction, we get a system that is complete. We demonstrate some of the rules for such atom abstraction, and outline how the problem with completeness can be solved in a neat manner.

3.7 Symmetric Circuits and Fixed-Point Logics

Anuj Dawar (University of Cambridge, GB)

Joint work of Anderson, Matthew; Dawar, Anuj

We study queries on graphs (and other relational structures) defined by families of Boolean circuits that are invariant under permutations of the vertices. In particular, we study circuits that are symmetric, that is, circuits whose invariance is explicitly witnessed by automorphisms of the circuit induced by the permutation of their inputs. We show a close connection between queries defined on structures by uniform families of symmetric circuits and definability in fixed-point logics.

3.8 Nominal duality theory for new, forall, and lambda.

Jamie Gabbay (Heriot-Watt University Edinburgh, GB)

Joint work of Gabbay, Jamie; Gabbay, Michael; Petrisan, Daniela; Litak, Tadeusz
Main reference Murdoch J. Gabbay, "Semantics out of context: nominal absolute denotations for first-order logic and computation," arXiv:1305.6291v2 [cs.LO], 2013.
URL http://arxiv.org/abs/1305.6291v2

Nominal algebra lets us axiomatise substitution and quantifiers, and thus the new-quantifier, first-order logic, and the lambda-calculus. Nominal lattice theory lets us characterise binders

as greatest and least upper bounds subject to freshness conditions; this is possible for "forall" and "exists" and also for "lambda".

From this follow soundness, completeness, representation, and topological duality results for algebraic/lattice-theoretic theories in nominal sets and topological spaces. A great deal of structure is revealed by this, which I will outline.

References

1 Murdoch J. Gabbay. *Semantics out of context: nominal absolute denotations for first-order logic and computation.* http://arxiv.org/abs/1305.6291

2 Murdoch J. Gabbay and Michael J. Gabbay. *Representation and duality of the untyped lambda-calculus in nominal lattice and topological semantics, with a proof of topological completeness.* http://arxiv.org/abs/1305.5968

3 Murdoch J. Gabbay, Tadeusz Litak, Daniela Petrisan. *Stone Duality for Nominal Boolean Algebras with NEW.* http://dx.doi.org/10.1007/978-3-642-22944-2_14

3.9 Programming in nominal sets: a survey

Bartek Klin (University of Warsaw, PL)

An important benefit of nominal techniques is a neat syntactic presentation of computations over certain infinite data sets, e.g., α-equivalence classes of λ-terms or state spaces of automata over infinite alphabets. A few programming languages have been designed that aim to offer that presentation to the programmer, hiding the necessary symbolic manipulations in the compiler. I briefly describe and compare:

- FreshML [3] and its later version FreshO'Caml [4]: functional languages aimed for computation of data structures with binding over equality atoms,
- Nλ [1]: an experimental functional language, implemented as an extension of Haskell, for computing orbit-finite data structures over arbitrary atom symmetries, with no direct treatment of binding,
- While programs with atoms [2]: an imperative language with a very weak type system, with a purpose very similar to that of Nλ.

References

1 M. Bojańczyk, L. Braud, B. Klin, S. Lasota. *Towards nominal computation.* Proceedings of POPL 2012.

2 M. Bojańczyk, S. Toruńczyk. *Imperative programming in sets with atoms.* Proceedings of FSTTCS 2013.

3 M. J. Gabbay, A. M. Pitts. *A metalanguage for programming with bound names modulo renaming.* Proceedings of MPC, volume 1837 of LNCS, 2000.

4 M. R. Shinwell. *The fresh approach: functional programming with names and binders.* PhD thesis, University of Cambridge, 2005.

3.10 A machine-independent characterization of timed languages

Sławomir Lasota (University of Warsaw, PL)

Joint work of Lasota, Sławomir; Bojańczyk, Mikołaj
Main reference M. Bojańczyk, S. Lasota, "A Machine-Independent Characterization of Timed Languages," in Proc. of the 39th Int'l Colloquium on Automata, Languages, and Programming (ICALP'12), Part II, LNCS, Vol. 7392, pp. 92–103, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-31585-5_12

We use a variant of sets with atoms (known also as Fraenkel-Mostowski sets, or as generalized nominal sets, cf. [5, 6, 3]) as a framework suitable for stating and proving the following two results on timed automata [2]. As atoms, we suitably choose the structure of reals with the order $<$ and $+1$ predicate. The first result is a machine-independent characterization of languages of deterministic timed automata, in the style of Myhill-Nerode theorem. As a second result we distinguish a subclass of automata with atoms, called by us timed register automata, that extends timed automata and is effectively closed under minimization. The class is obtained naturally by restricting to those sets with atoms that are quantifier-free definable, using the vocabulary of atoms, namely $<$ and $+1$. Timed register automata are also related to timed automata with updates [4]. Thus sets with atoms provide an appropriate framework for minimization of timed automata, cf. [7, 1].

References

1 R. Alur, C. Courcoubetis, N. Halbwachs, D. L. Dill, and H. Wong-Toi. Minimization of timed transition systems. In *CONCUR*, pages 340–354, 1992.
2 R. Alur and D. L. Dill. A theory of timed automata. *Theor. Comput. Sci.*, 126(2):183–235, 1994.
3 M. Bojańczyk, B. Klin, and S. Lasota. Automata with group actions. In *Proc. LICS'11*, pages 355–364, 2011.
4 P. Bouyer, C. Dufourd, E. Fleury, and A. Petit. Updatable timed automata. *Theor. Comput. Sci.*, 321(2-3):291–345, 2004.
5 M. Gabbay. Foundations of nominal techniques: logic and semantics of variables in abstract syntax. *Bulletin of Symbolic Logic*, 17(2):161–229, 2011.
6 M. Gabbay and A. M. Pitts. A new approach to abstract syntax with variable binding. *Formal Asp. Comput.*, 13(3-5):341–363, 2002.
7 M. Yannakakis and D. Lee. An efficient algorithm for minimizing real-time transition systems. *Formal Methods in System Design*, 11(2):113–136, 1997.

3.11 Nominal nets

Ranko Lazic (University of Warwick, GB)

I discussed variants of 'nominal nets', i.e. Petri nets in which tokens are labelled by names, focussing on relative expressiveness and computational complexity of coverability and reachability problems. I also mentioned connections with other formalisms, including logics and automata on words and trees over infinite alphabets. The main message was that too many questions are open, and some of them seem difficult.

3.12 Concurrent Domain Theory with Nominal Sets

Steffen Loesch (University of Cambridge, GB)

Joint work of Lösch, Steffen; Pitts, Andrew M.; Winskel, Glynn

HOPLA by Nygaard and Winskel is a concurrent metalanguage with a fully abstract domain theory. This talk introduces HOPLA and shows how it can be extended with nominal sets, in order to obtain a more expressive metalanguage.

3.13 Named Graphs and HD Automata for Network-Conscious pi-Calculus

Ugo Montanari (University of Pisa, IT)

Joint work of Montanari, Ugo; Sammartino, Matteo
Main reference U. Montanari, M. Sammartino, "Network Conscious Pi-calculus: A Concurrent Semantics," in Proc. of the 28th Conf. on the Mathematical Foundations of Programming Semantics (MFPS'12), ENTCS, Vol. 286, pp. 291–306, Elsevier, 2012.
URL http://dx.doi.org/10.1016/j.entcs.2012.08.019

The semantics of name-passing calculi is often defined employing coalgebraic models over permutation algebras/nominal sets or over presheaf categories. Both these elegant theories lack finiteness properties, hence they are not apt for verification purposes. Coalgebras over named sets, called history-dependent (HD) automata, are better suited for the purpose due to locality of names. The three models are equivalent. Named sets are generalised by the categorical model of families, that is, free coproduct completions, indexed by symmetries, whenever certain conditions are satisfied.

In the talk we survey various notions of HD automata introduced in the last 18 years and we outline the symmetry indexed family construction. Then we observe that traditional process calculi usually abstract away from network details, modeling only communication over shared channels. They, however, seem inadequate to describe new network architectures, such as Software Defined Networks, where programs are allowed to manipulate the infrastructure. Thus we present the Network Conscious pi-calculus (NCPi), a proper extension of the pi-calculus with an explicit notion of network: network links and nodes are both represented as names. Finally, we show that the categorical model of NCPi, equipped with graphs rather than sets of names as resources, actually satisfies the conditions mentioned above, thus admitting an HD automaton construction, suitable for verification.

3.14 Algorithmic games for full ground references

Andrzej Murawski (University of Warwick, GB)

Joint work of Murawski, Andrzej; Tzevelekos, Nikos
Main reference A. S. Murawski, N. Tzevelekos, "Algorithmic Games for Full Ground References," in Proc. of the 39th Int'l Colloquium on Automata, Languages, and Programming (ICALP'12), Part II, LNCS, Vol. 7392, pp. 312–324, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-31585-5_30

We present a full classification of decidable and undecidable cases for contextual equivalence in a finitary ML-like language equipped with full ground storage (both integers and reference names can be stored). The simplest undecidable type is $\mathsf{unit} \to \mathsf{unit} \to \mathsf{unit}$. At the technical level, our results marry game semantics with automata-theoretic techniques developed to handle infinite alphabets. On the automata-theoretic front, we show decidability of the emptiness problem for register pushdown automata extended with fresh-symbol generation.

3.15 On Orbit finiteness and Limits in Nominal Sets

Daniela Petrisan (University of Leicester, GB)

Joint work of Kurz, Alexander; Petrisan, Daniela; Severi, Paula; de Vries, Fer-Jan
Main reference A. Kurz, D. Luan Petrisan, P. Severi, F.-J. de Vries, "Nominal Coalgebraic Data Types with Applications to Lambda Calculus," Logical Methods in Computer Science, 9(4:20); available also as arXiv:1311.1395v2 [cs.LO].
URL http://dx.doi.org/10.2168/LMCS-9(4:20)2013
URL http://arxiv.org/abs/1311.1395v2

In this talk we discuss presentations of limits in nominal sets and its relevance to infinitary data types involving binders. To this end we introduce a notion of bound variable relative to a map on nominal sets and we define a map to be safe when in the fiber above each element there exists an element with a maximal number of bound variables. We also introduce maps with orbit-finite fibers as maps for which the inverse image of each element is a finitely presentable (or orbit-finite) nominal subset. These maps have very nice closure properties, are safe maps, and allow us to represents limits of quotients in Nom as a quotient of a pullback computed in Set. As an application, one can find representatives of infinitary terms up to alpha-equivalence.

3.16 Nominal Sets: Introduction and Survey

Andrew M. Pitts (University of Cambridge, GB)

Main reference A. M. Pitts, "Nominal Sets: Names and Symmetry in Computer Science," Cambridge Tracts in Theoretical Computer Science, Vol. 57, Cambridge University Press, 2013.
URL http://www.cambridge.org/us/academic/subjects/computer-science/programming-languages-and-applied-logic/nominal-sets-names-and-symmetry-computer-science

Names and constructs that bind names are ubiquitous in formal languages in general and programming languages in particular. Nominal sets provide a mathematical theory of

structures involving names, based on some simple, but subtle ideas going back to the symmetric models of set theory with atoms of Fraenkel and Mostowski. The theory was introduced by Gabbay and Pitts in 1999 and has since been developed and applied to programming language semantics, machine-assisted theorem proving and the design of functional and logical metaprogramming languages. This work uses the so-called equality symmetry; more recently Bojańczyk et al have considered more general symmetries and their application to automata theory.

This tutorial introduced the theory of nominal sets, mainly from a category- theoretic rather than set-theoretic perspective; it surveyed some of its applications, concentrating on the case of the equality symmetry; and it highlighted some potential areas for further research.

3.17 Nominal Automata and Regular Expressions

Tomoyuki Suzuki (Academy of Science – Prague, CZ)

Joint work of Kurz, Alexander; Suzuki, Tomoyuki; Tuosto, Emilio
Main reference A. Kurz, T. Suzuki, E. Tuosto, "Nominal Regular Expressions for Languages over Infinite Alphabets," arXiv:1310.7093v1 [cs.FL], 2013.
URL http://arxiv.org/abs/1310.7093v1

We discuss regular expressions to abstractly model and study properties of resource-aware computations. Inspired by nominal techniques – as those popular in process calculi – we extend classical regular expressions with names (to model computational resources) and suitable operators (for allocation, deallocation, scoping of, and freshness conditions on resources). We discuss classes of such nominal regular expressions, show how such expressions have natural interpretations in terms of languages over infinite alphabets, and give Kleene theorems to characterise their formal languages in terms of nominal automata. In the end, we also discuss our ongoing works.

3.18 Turing machines with atoms

Szymon Toruńczyk (University of Warsaw, PL)

Joint work of Toruńczyk, Szymon; Klin, Bartek; Lasota, Sławomir; Bojańczyk, Mikołaj
Main reference M. Bojańczyk, B. Klin, S. Lasota, S. Toruńczyk, "Turing Machines with Atoms," in Proc. of the 28th IEEE/ACM Symp. on Logic in Computer Science (LICS'13), pp. 183–192, IEEE CS; available as pre-print from the author's webpage.
URL http://dx.doi.org/10.1109/LICS.2013.24
URL http://www.mimuw.edu.pl/~bojan/papers/atomturing.pdf

We study Turing machines over sets with atoms, also known as nominal sets. Our main result is that deterministic machines are weaker than nondeterministic ones; in particular, P≠NP in sets with atoms. Our main construction is closely related to the Cai-Fürer-Immerman graphs used in descriptive complexity theory.

3.19 Names, games and automata

Nikos Tzevelekos (Queen Mary University of London, GB)

URL http://www.tzevelekos.org/talks/Dagstuhl_Oct13.pdf

Names constitute a pervasive feature in programming languages. They appear in every computational scenario where entities of specific kinds can be created at will and, moreover, in such a manner that newly created entities are always fresh, i.e. distinct from any other created thus far. For example, references, objects and exceptions in languages like ML or Java can be seen as names. The behaviour of languages which feature names is in general very subtle due to issues of privacy, visibility and flow of names, and the ensuing notion of local state.

This talk is about formal reasoning techniques for programs with names which have emerged in the last years. We will focus on a specific such formalism, called game semantics, which models computation as a formal interaction (a game) between a program and its environment. We moreover see how these models can be given algorithmic representations by means of abstract machines operating on infinite alphabets of names.

3.20 Nominal Isabelle, or, How Not to be Intimidated by the Variable Convention

Christian Urban (King's College London, GB)

Main reference C. Urban, "Nominal Techniques in Isabelle/HOL," J. of Automatic Reasoning, 40(4):327–356, 2008.
URL http://dx.doi.org/10.1007/s10817-008-9097-2

If researchers in programming languages want to formalise and check their work in a theorem prover, they need to deal with binders, renaming of bound variables, capture-avoiding substitution, etc. This is very often a major problem in formal proofs. In informal proofs one often assumes a variable convention and thus side-steps all these problems. In the talk I will show how strong induction principles can be derived that have the variable convention already built-in. However, I will also show that this convention is in general an unsound reasoning principle and requires restrictions in order to be safe. The aim of this work is to provide all proving technology necessary for reasoning conveniently about programming languages. This work has previously been reported in

- C. Urban, Nominal Techniques in Isabelle/HOL. In Journal of Automatic Reasoning, 2008, Vol. 40(4), 327–356.
- C. Urban and C. Kaliszyk, General Bindings and Alpha-Equivalence in Nominal Isabelle. Journal of Logical Methods in Computer Science, Volume 8 (2:14), 2012

Participants

- Andreas R. Blass
University of Michigan, US
- Mikołaj Bojańczyk
University of Warsaw, PL
- James Cheney
University of Edinburgh, GB
- Vincenzo Ciancia
CNR – Pisa, IT
- Thomas Colcombet
CNRS / Univ. Paris-Diderot, FR
- Roy L. Crole
University of Leicester, GB
- Anuj Dawar
University of Cambridge, GB
- Jamie Gabbay
Heriot-Watt University
Edinburgh, GB
- Fabio Gadducci
University of Pisa, IT
- Tomasz Gogacz
University of Wroclaw, PL
- Bartek Klin
University of Warsaw, PL
- Alexander Kurz
University of Leicester, GB
- Sławomir Lasota
University of Warsaw, PL
- Ranko Lazic
University of Warwick, GB
- Steffen Lösch
University of Cambridge, GB
- Justus Matthiesen
University of Cambridge, GB
- Stefan Milius
Univ. Erlangen-Nürnberg, DE
- Ugo Montanari
University of Pisa, IT
- Andrzej Murawski
University of Warwick, GB
- Joanna Ochremiak
University of Warsaw, PL
- Daniela Petrisan
University of Leicester, GB
- Andrew M. Pitts
University of Cambridge, GB
- Luc Segoufin
ENS – Cachan, FR
- Alexandra Silva
Radboud Univ. Nijmegen, NL
- Ian Stark
University of Edinburgh, GB
- Tomoyuki Suzuki
Academy of Science – Prague, CZ
- Szymon Toruńczyk
University of Warsaw, PL
- Emilio Tuosto
University of Leicester, GB
- Nikos Tzevelekos
Queen Mary University of
London, GB
- Christian Urban
King's College London, GB

Report from Dagstuhl Seminar 13431

Real-World Visual Computing

Edited by

Oliver Grau[1], Marcus A. Magnor[2], Olga Sorkine-Hornung[3], and Christian Theobalt[4]

1 **Intel Visual Computing Institute, Saarbrücken, DE,** oliver.grau@intel.com
2 **TU Braunschweig, DE,** magnor@cg.cs.tu-bs.de
3 **ETH Zürich, CH,** hornung@disneyresearch.com
4 **MPI für Informatik – Saarbrücken, DE,** theobalt@mpi-inf.mpg.de

Abstract

Over the last decade, the tremendous increase in computational power of graphics hardware, in conjunction with equally improved rendering algorithms, have led to the situation today where real-time visual realism is computationally attainable on almost any PC, if only the digital models to be rendered were sufficiently detailed and realistic.

With rapidly advancing rendering capabilities, the modeling process has become the limiting factor in realistic computer graphics applications. Following the traditional rendering paradigm, higher visual realism can be attained only by providing more detailed and accurate scene descriptions. However, building realistic digital scene descriptions consisting of 3D geometry and object texture, surface reflectance characteristics and scene illumination, character motion and emotion is a highly labor-intensive, tedious process.

Goal of this seminar is to find new ways to overcome the looming stalemate in realistic rendering caused by traditional, time-consuming modeling. One promising alternative consists of creating digital models from real-world examples if ways can be found how to endow reconstructed models with the flexibility customary in computer graphics. The trend towards model capture from real-world examples is bolstered by new sensor technologies becoming available at mass-market prices, such as Microsoft's Kinect and time-of-flight 2D depth imagers, or Lytro's Light Field camera. Also, the pervasiveness of smart-phones containing camera, GPS and orientation sensors allows for developing new capturing paradigms of real-world events based on a swarm of networked smart-phones. With the advent of these exciting new acquisition technologies, investigating how to best integrate these novel capture modalities into the digital modeling pipeline or how to alter traditional modeling to make optimal use of new capture technologies, has become a top priority in visual computing research.

To address these challenges, interdisciplinary approaches are called for that encompass computer graphics, computer vision, and visual media production. The overall goal of the seminar is to form a lasting, interdisciplinary research community which jointly identifies and addresses the challenges in modeling from the real world and determines which research avenues will be the most promising ones to pursue over the course of the next years.

Seminar 20.–25. October, 2013 – www.dagstuhl.de/13431

1998 ACM Subject Classification I.2.10 Vision and Scene Understanding, I.3.3 Picture/Image Generation, I.3.5 Object Modeling, I.3.7 Three-Dimensional Graphics and Realism, I.4.1 Digitization and Image Capture, I.4.5 Reconstruction, I.4.8 Scene Analysis, I.4.10 Image Representation, I.4.9 Applications, I.3.1 Display devices

Keywords and phrases Image Aquisition, Scene Modeling/Rendering, Image/3D Sensors, Photorealism, Visual Effects, Motion Reconstruction, Animation

Digital Object Identifier 10.4230/DagRep.3.10.72

Edited in cooperation with Felix Klose

1 Executive Summary

Marcus Magnor
Oliver Grau
Olga Sorkine-Hornung
Christian Theobalt

Dagstuhl seminar 13431 "Real-World Visual Computing' took place October 20–25, 2013. 45 researchers from North America, Asia, and Europe discussed the state-of-the-art, contemporary challenges, and promising future research directions in the areas of acquiring, modeling, editing, and rendering of complex natural scenes and events. The seminar was encompassed an introductory and a closing session, 9 scientific presentation sessions, two book organizational sessions as well as one special session on the Uncanny Valley problem. The seminar brought together junior and senior researchers from computer graphics, computer vision, 3D animation and visual special effects, both from academia and industry, to address the challenges in real-world visual computing. Participants included international experts from Kyoto University, Tsinghua University, University of British Columbia, University of Alberta, University of North Carolina, University of Kentucky, Yale University, Technion – Haifa, Filmakademie Baden-Wuerttemberg, Hochschule der Medien Stuttgart, Disney Research Zurich, BBC Research & Development, Intel Visual Computing Institute, Nvidia Corp., Adobe Systems Inc., metaio GmbH as well as many more research institutions and high-tech companies.

Motivating this seminar was the observation that digital models of real-world entities have become an essential part of innumerous computer graphics applications today. With ever-increasing graphics hardware and software capabilities, however, so does the demand for more and more realistically detailed models. Because the traditional, labor-intensive process of digital model creation by hand threatens to stall further progress in computer graphics, conventional manual modeling approaches are giving way to new approaches that aim at capturing complex digital models directly from the real world. The seminar picked up on recent trends in acquisition hardware for real-world events (e.g., Microsoft Kinect, Lytro light field camera, swarm of smartphone sensors, ...) as well as in visual computing applications (e.g., 3D movies, Streetview, digital mock-ups, free-viewpoint systems, ...). It brought together experts from academia and industry working on contemporary challenges in image-based techniques, geometry modeling, computational photography and videography, BRDF acquisition, 3D reconstruction, 3D video, motion and performance capture etc. Collectively we fathomed the full potential of real world-based modeling approaches in computer graphics and visual computing.

Over the past decade, computer graphics has evolved into a mainstream area of computer science. Its economic impact and social pervasion range from professional training simulators to interactive entertainment, from movie production to trauma therapy, from geographic information systems to Google Earth. As a result, expectations on computer graphics performance are rising continuously. In fact, thanks to the progress in graphics hardware as well as rendering algorithms, visual realism is today within easy reach of off-the-shelf PCs, laptops, and even handheld devices. With rapidly advancing rendering capabilities, however, in many application areas of computer graphics the modeling process is becoming the limiting factor. Higher visual realism can be achieved only from more detailed and accurate scene descriptions. So far, however, digitally modeling 3D geometry and object

texture, surface reflectance characteristics and scene illumination, motion and emotion is a labor-intensive, tedious process performed by highly trained animation specialists. The cost of conventionally creating models of sufficient complexity to engage the full potential of modern GPUs increasingly threatens to stall progress in computer graphics.

To overcome this bottleneck, an increasing number of researchers and engineers worldwide is investigating alternative approaches to create realistic digital models directly from real-world objects and scenes: Google and Microsoft already digitize entire cities using panorama video footage, 3D scanners, and GPS; RTT AG in Munich creates highly realistic digital mock-ups for the car industry from CAD data and measured surface reflectance characteristics of car paint; at Disney Research, algorithms are being developed to create stereoscopic movies from monocular input; and BBC R&D has developed various 3D sports visualization methods based on analyzing live-broadcast footage.

In recent years, special effects in movies and computer games have reached a new level of complexity. In their aim to construct convincing virtual environments or even virtual actors, VFX companies are more and more relying on techniques to capture models from the real world. Currently available reconstruction tools, however, are still in their infancy. A lot of time is still spent on manual post-processing and modeling. The research community has responded to this trend by investigating new image- and video-based scene reconstruction approaches that can capture richer and more complex models. An example are performance capture methods that estimate more detailed shape and motion models of dynamics scenes than do commercially available systems. Similar methods for reconstruction of entire sets are also currently investigated, but many algorithmic problems remain to be solved.

The trend towards model capture from real world-examples is additionally bolstered by new sensor technologies becoming available at mass-market prices, such as Microsoft's Kinect, time-of-flight 2D depth imagers, or Lytro's Light Field camera. Also the pervasiveness of smartphones containing a camera, GPS, and orientation sensors allows for developing new capturing paradigms of real-world events based on a swarm of networked handheld devices. With the advent of these exciting novel acquisition technologies, investigating how to best integrate these new capture modalities into the computer graphics modeling pipeline, or how to alter traditional modeling to make optimal use of the new capture approaches, has become a top priority in visual computing research.

Researchers working on all of these problems from different direction came together at the seminar to share their experiences and discuss the scientific challenges. Questions discussed were both theoretical and practical in nature. The seminar participants discussed the contemporary scientific challenges in modeling from the real world and determined which research avenues are likely to be the most promising and interesting ones to pursue over the course of the next years.

Among the questions and issues that have been addressed in the seminar are how to capitalize on new sensors for capture (computational cameras, light field cameras, Time-of-flight sensors, Kinect, omni-visual systems, ...), how to capture different object/scene aspects (geometry, reflectance, texture, material/fabric, illumination, dynamics, ...), how to digitally represent real-world objects/scenes (meshes, voxels, image-based, animation data, ...), how to convincingly & intuitively manipulate real-world models (relighting, motion editing, constrained manipulation, sketch-based, example-based, ...), how to realistically compose/augment and real-time render new scenes (F/X, movie post-production, games, perceptual issues, ...), and how to exploit the immense amount of community image and video data that are captured with handheld devices to build detailed models of the world (buildings, acting/dancing performances, sports events, fish tanks, ...). Also, the challenges

arising from the large data sets of real-world models have been addressed. A special session on perceptional issues in animation (the Uncanny Valley problem) set out to identify the most important factors that are still unrealistic in computer animation. As the single most important area, facial animation was identified and some research directions for improvements were discussed.

The overall goal of the seminar to form a lasting, interdisciplinary research community was impressively underlined by the willingness of many seminar participants to work together on an edited book on the topic of the seminar. The book will be published with CRC Press. Completion of the manuscript is scheduled for August 2014.

2 Table of Contents

3 Overview of Talks

3.1 Automatic Structures for Shape Modeling, Processing and Analysis

Tamy Boubekeur (Télécom Paris Tech, FR)

In this talk, I start by quickly reviewing the recent work of my group in structure-based shape analysis. I present how skeletons, cages and on- surface structures and descriptors can be computed automatically and efficiently for a wide range of applications, including interactive modeling, shape retrieval, surface simplification and animated data processing. In the second part of the talk, I focus on our most recent work on this topic, which relates to medial structures. In particular, I briefly introduce Sphere- Meshes and the Progressive Medial Axis Filtration method.

3.2 Dynamic Facial Processing and Capture: Basic Research to VFX

Darren Cosker (University of Bath, GB)

In this talk I will outline some of the challenges in creating facial performances and using facial models in visual effects. In particular, I will attempt to distinguish between academic challenges and industrial demands, and attempt to highlight some of the shared challenges.

I will also describe some of the work that myself and my group have been performing in the area of 4D facial processing. I will describe how this has led to us stepping back to focus on first solving more 'basic' (or fundamental) computer vision research problems – particularly in the area of optical flow, non-rigid tracking and shadow removal.

3.3 BBC Research & Development and the work of the Immersive and Interactive Content team.

Robert Dawes (BBC Research & Development – London, GB)

BBC R&D is the technical research department of the BBC. It is charged with researching and developing advanced and emerging media technologies for the benefit of the corporation, and wider UK and European media industries, and is also the technical design authority for a number of major technical infrastructure transformation projects for the UK broadcasting industry.

This presentation provides some background information about BBC R&D and its role and will describe the board range of work conducted by the department. It describes in detail the work of the Immersive and Interactive Content team and how it has used Visual Computing to enhance the BBC's output. This work includes virtual studio and sports telestration systems, multi-camera 3D reconstruction and metadata extraction. This work has had to deal with the constraints of broadcasting with particular emphasis on real-time operation on live video.

3.4 Texturing from Images

Jean-Michel Dischler (University of Strasbourg, FR)

Joint work of Vanhoey, Kenneth; Sauvage, Basile; Larue, Frédéric; Dischler, Jean-Michel
Main reference K. Vanhoey, B. Sauvage, F. Larue, J.-M. Dischler, "On-the-Fly Multi-Scale Infinite Texturing from Example," ACM Trans. Graph., 32(6):208:1–208:10, 2013.
URL http://dx.doi.org/10.1145/2508363.2508383

Providing efficient solutions for rendering detailed realistic environments in real-time applications, like games or flight/driving simulators, has always been a major focus in computer graphics. Details can be efficiently rendered using textures. But despite improvements of acquisition technology, graphics hardware, memory capacity and data streaming techniques, which allowed over the recent years for increased scene complexity, creating and rendering efficiently textures still remains a challenging issue. In this talk I'll present three techniques for texturing surfaces, the textures beening constructed from collections of images / photographs. The first method reconstructs noise-based textures from exemplars. It fits parameters based on a metric that computes texture descriptors. The second method manages multi-scale detail transitions by using one input image per represented scale. Undersampling artifacts are avoided by accounting for fine-scale features while colors are transferred between scales. Finally, a robust least-squares based method for fitting 2D parametric color functions on sparse and scattered data is presented. This technique allows us to reconstruct artefact-less surface lightfields over scanned 3D objects using sparse sets of photographs.

References

1 K. Vanhoey, B. Sauvage, F. Larue, J-M Dischler. *On-the-Fly Multi-Scale Infinite Texturing from Example*. ACM Trans. Graph., SIGGRAPH Asia, (2013)
2 K. Vanhoey, B. Sauvage, O. Génevaux, F. Larue, J-M Dischler. *Robust Fitting on Poorly Sampled Data for Surface Light Field Rendering and Image Relighting*. Comput. Graph. Forum 32(6): 101-112 (2013)
3 G. Gilet, J-M Dischler, D. Ghazanfarpour. *Multiple kernels noise for improved procedural texturing*. The Visual Computer 28(6-8): 679-689 (2012)
4 G. Gilet, J-M Dischler. *An Image-Based Approach for Stochastic Volumetric and Procedural Details*. Comput. Graph. Forum 29(4): 1411-1419 (2010)

3.5 Image-based Analysis and Modelling

Peter Eisert (Fraunhofer-Institut – Berlin, DE)

Joint work of Eisert, Peter; Blumenthal, David; Hilsmann, Anna

The presentation describes methods for the estimation of scene properties like geometry, motion, and deformation from multiple camera views. For 3D geometry reconstruction, a method is presented that estimates highly detailed mesh models from a pair of uncalibrated views from standard digital cameras [3, 1]. For better accuracy, lighting changes are explicitly considered in the optimization framework that jointly estimates vertex positions considering smoothness priors via Laplacians. In order to enhance the reconstruction of fine structures like hair, texture gradients are additionally exploited in the smoothness constraints. Comparisons with existing methods show better qualities and more fine details in the surface

reconstruction. The method is for example used in the context of head reconstruction for virtual actor animation in the European project React. For the creation of a full head model, several partial models are fused and semantically labelled using a morphable head model as prior [4]. Another application presented is the modelling and tracking of the lower leg for navigated knee surgery. Based on a morphable leg model adapted to the reconstruction, motion and deformation is tracked over time with the optimization framework providing knowledge about bone movements. Finally, the method is applied for image-based rendering of complex objects like clothes [2] to create image-warps for view interpolation of deformable objects.

References

1 D. Blumenthal-Barby, P. Eisert. *High-Resolution Depth For Binocular Image-Based Modelling*. Computers & Graphics, accepted for publication, 2013.
2 A. Hilsmann, P. Fechteler, P. Eisert. *Pose Space Image-based Rendering*. Computer Graphics Forum (Proc. Eurographics 2013), May 2013.
3 D. C. Schneider, M. Kettern, A. Hilsmann, P. Eisert. *Deformable Image Alignment as a Source of Stereo Correspondences on Portraits*. CVPR'11 Workshop NORDIA, June 2011.
4 D. Schneider, P. Eisert. *Algorithms for Automatic and Robust Registration of 3D Head Scans*. Journal of Virtual Reality and Broadcasting, Oct. 2010.

3.6 New Perspectives on Uncalibrated Photometric Stereo

Paolo Favaro (Universität Bern, CH)

Joint work of Favaro, Paolo; Papadhimitri, Thoma
Main reference T. Papadhimitri, P. Favaro, "A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima," Int'l J. of Computer Vision, 16 pp., October 2013.
URL http://dx.doi.org/10.1007/s11263-013-0665-5

The Lambertian reflectance model is far from being a realistic model of light propagation on real objects. Yet, it remains one of the most widely adopted models for shading in computer vision and has led to very effective algorithms for 3D reconstruction. Surprisingly, this simple model has still more to offer. We illustrate this in the well-known uncalibrated photometric stereo problem, where one aims at reconstructing the normal map and the albedo of an object and the scene illumination given only a collection of images. At the core of uncalibrated photometric stereo is the integrability constraint, which allows to obtain a solution up to 3 free parameters, the so-called generalized bas-relief (GBR) ambiguity. Existing approaches are not able to fix these parameters as they employ heuristics that yield infinite solutions. In contrast, we show that a careful study of the Lambertian model yields two possible exact solutions: one exploiting locations of maximal diffuse brightness [2] and another exploiting the perspective effect in the camera projection model [1]. Both methods are validated on synthetic and real data and achieve state-of-the-art results.

References

1 Thoma Papadhimitri and Paolo Favaro. *A New Perspective on Uncalibrated Photometric Stereo*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013
2 Thoma Papadhimitri and Paolo Favaro. *A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima*. International Journal of Computer Vision, pp 1–16, 2013

3.7 Applications of Rapid Manufacturing Tools to Real-World Visual Computing

Martin Fuchs (Universität Stuttgart, DE)

Joint work of Fuchs, Martin; Hasan, Milos; Kächele, Markus; Lensch, Hendrik P. A.; Matusik, Wojciech; Pfister, Hanspeter; Raskar, Ramesh; Rusinkiewicz, Szymon; Seidel, Hans-Peter

In the last years, rapid manufacturing tools have begun to have an increasing impact on my work on the visual computing cycle. In this talk, I present several works I contributed to for which these tools were crucial in enabling success.

At SIGGRAPH 2008, we introduced passive reflectance field displays, which optically shape incident illumination into view-dependent pictures which can be programmed in up to six dimensions [1]. This requires precise alignment of numerous optical components, which was achieved by 3D printing enclosing structures with a fused deposition modeling process.

Two years later, at SIGGRAPH 2010, 3D printing stackings of selected base materials permitted us to manufacture objects with defined subsurface scattering properties [3], and recently [2], we performed CNC milling and laser cutting to inexpensively create precisely designed faceted mirrors which enabled light field acquisition with a catadioptric setup.

Open problems include simplified access to the internal representations of the manufacturing machines so as to bridge the gap between visual computing and mechanical engineering, and the research into a principled approach to assess the gamut of a fabrication pipeline for visual models.

References

1 Martin Fuchs, Ramesh Raskar, Hans-Peter Seidel and Hendrik P. A. Lensch. *Towards Passive 6D Reflectance Field Displays*. ACM Transactions on Graphics (Proc. SIGGRAPH special issue) 27(3) , 2008

2 Martin Fuchs, Markus Kächele and Szymon Rusinkiewicz. *Design and Fabrication of Faceted Mirror Arrays for Light Field Capture*. VMV 2012: Vision, Modeling & Visualization, 2012. An extended version is to appear at Computer Graphics Forum, DOI: 10.1111/cgf.12201.

3 Milos Hasan, Martin Fuchs, Wojciech Matusik, Hanspeter Pfister, and Szymon Rusinkiewicz. *Physical reproduction of materials with specified subsurface scattering*. ACM Transactions on Graphics (Proc. SIGGRAPH special issue) 29(4), 2010

3.8 Image-Based Approaches for Photo-Realistic Rendering of Clothing

Anna Hilsmann (HU Berlin, DE)

Joint work of Hilsmann, Anna; Fechteler Philipp; Eisert, Peter
Main reference A. Hilsmann, P. Fechteler, P. Eisert, "Pose Space Image Based Rendering," Computer Graphics Forum, 32(2pt3):265–274, 2013.
URL http://dx.doi.org/10.1111/cgf.12046

One principal intention of computer graphics is the achievement of photorealism. Although modeling, animation and simulation tools for rendering of complex objects – e.g. human bodies, faces, or clothes – have been developed in the last decades, achieving photorealism by physically simulating material properties and illumination is still computationally demanding

and time-consuming. Especially clothing is extremely difficult to simulate physically, because cloth deformation and drapery exhibit many degrees of freedom and wrinkles produce complex shading and texture deformations. Yet, these complex details are essential for a realistic appearance of the rendered clothes.

An alternative to synthesis and reconstruction is observation of appearance through a number of images. This is addressed by image-based rendering (IBR) approaches, which synthesize new images by appropriately interpolating and merging a database of prerecorded images. However, the database images typically show various viewpoints of a rigid and stationary object or scene and the synthesis of new images is limited to viewpoint changes. More complex animations are typically not possible.

For clothing that roughly follows the shape of a human body, it is a reasonable assumption that wrinkling depends on the articulated pose of a human body. Under this assumption, a new image-based rendering approach is proposed, which synthesizes new images of such types of clothing from a database of pre-recorded images based on pose information [2]. Image warps, i.e. transformation rules between the images, implicitly extract pose-dependent appearance and shading from the images. These warps are extracted both in the spatial as well as in the photometric domain [1]. For rendering, the images and warps are parametrized and interpolated in pose-space, i.e. the space of body poses, using scattered data interpolation methods that have already been successfully exploited in example-based animation methods. The high-dimensionality of the pose-space is handled by subdividing the space into subspaces, which we assume can be handled independently, assuming that wrinkling of tight-fitting clothes mainly depends on the nearest joints. This reduces the dimensionality of the interpolation domain and thereby allows for a larger variety of possible poses that can be synthesized.

Altogether, the presented approach shifts computational complexity from the rendering to an a-priori training phase. The use of real images and warp-based extraction of deformation and shading allow a photo-realistic visualization and modification of clothes, including fine details, without computationally demanding simulation of the underlying scene and object properties.

References

1 A. Hilsmann and D.C. Schneider and P. Eisert. *Realistic Cloth Augmentation in Single View under Occlusion*. Computers & Graphics 34(5), pp. 567–574, Oct. 2010.

2 A. Hilsmann and P.Fechteler and P. Eisert. *Pose Space Image Based Rendering*. Computer Graphics Forum 32(2) (Proc. of Eurographics 2013), pp. 265–274, May 2013.

3.9 Accurate Binary Image Selection from Inaccurate User Input

Jan Kautz (University College London, GB)

Joint work of Subr, Kartic; Paris, Sylvain; Soler, Cyril; Kautz, Jan

Main reference K. Subr, S. Paris, C. Soler, J. Kautz, "Accurate Binary Image Selection from Inaccurate User Input," Computer Graphics Forum, 32(2pt1):41–50, 2013.

URL http://dx.doi.org/10.1111/cgf.12024

Selections are central to image editing, e.g., they are the starting point of common operations such as copy-pasting and local edits. Creating them by hand is particularly tedious and scribble-based techniques have been introduced to assist the process. By interpolating a few strokes specified by users, these methods generate precise selections. However, most of the

algorithms assume a 100% accurate input, and even small inaccuracies in the scribbles often degrade the selection quality, which imposes an additional burden on users. We propose a selection technique tolerant to input inaccuracies. We use a dense conditional random field (CRF) to robustly infer a selection from possibly inaccurate input. Further, we show that patch-based pixel similarity functions yield more precise selection than simple point-wise metrics. However, efficiently solving a dense CRF is only possible in low-dimensional Euclidean spaces, and the metrics that we use are high-dimensional and often non-Euclidean. We address this challenge by embedding pixels in a low-dimensional Euclidean space with a metric that approximates the desired similarity function. The results show that our approach performs better than previous techniques and that two options are sufficient to cover a variety of images depending on whether the objects are textured.

3.10 A Reconfigurable Camera Add-On for High Dynamic Range, Multispectral, Polarization, and Light-Field Imaging

Oliver Klehm (MPI für Informatik – Saarbrücken, DE)

Joint work of Manakov, Alkhazur; Restrepo, John F.; Klehm, Oliver; Hegedüs, Ramon; Eisemann, Elmar; Seidel, Hans-Peter; Ihrke, Ivo
Main reference A. Manakov, J. F. Restrepo, O. Klehm, R. Hegedüs, E. Eisemann, H.-P. Seidel, I. Ihrke, "A reconfigurable camera add-on for high dynamic range, multispectral, polarization, and light-field imaging," ACM Trans. Graph., 32(4): 47:1–47:14, 2013..
URL http://dx.doi.org/10.1145/2461912.2461937

We propose a non-permanent add-on that enables plenoptic imaging with standard cameras. Our design is based on a physical copying mechanism that multiplies a sensor image into a number of identical copies that still carry the plenoptic information of interest. Via different optical filters, we can then recover the desired information. A minor modification of the design also allows for aperture sub-sampling and, hence, light-field imaging. As the filters in our design are exchangeable, a reconfiguration for different imaging purposes is possible. We show in a prototype setup that high dynamic range, multispectral, polarization, and light-field imaging can be achieved with our design.

3.11 Garment Replacement in Monocular Video Sequences

Felix Klose (TU Braunschweig, DE)

Joint work of Rogge, Lorenz; Klose, Felix; Stengel, Michael; Eisemann, Martin; Magnor, Marcus

In this talk a semi-automatic approach for replacing clothes in a monocular video sequences will be presented. From a monocular video sequence the actors body shape and pose is reconstructed using a parameterized body model. Using this animated body model as a mannequin arbitrary virtual garments can now be simulated. To obtain plausible renderings of this garment an additional scene lighting reconstruction is performed to replicate the original lighting situation. The remaining body shape and pose misalignments are then corrected purely in image space by computing and applying a pixel wise 2D image warp to the rendered image of the simulated garment. The results are then composed back onto the original video sequence.

3.12 Online Scene Reconstruction using Point-Based Data Fusion

Andreas Kolb (Universität Siegen, DE)

Joint work of Keller, Maik; Lefloch, Damien; Lambers, Martin; Izadi, Shahram; Weyrich, Tim; Kolb, Andreas
Main reference M. Keller, D. Lefloch, M. Lambers, S. Izadi, T. Weyrich, A. Kolb, "Real-time 3D Reconstruction in Dynamic Scenes using Point-based Fusion," in Proc. of the 2013 Int'l Conf. on 3D Vision (3DV'13), pp. 1–8, IEEE, 2013.
URL http://dx.doi.org/10.1109/3DV.2013.9

Online range image acquisition extremely evolved in the last view years due to new sensing devices like Time-of-Flight (ToF) camera or the Microsoft Kinect. Even though, this kind of sensors are well suited for accessing individual range images, using this kind of device for online scene acquisition is a challenging task. Processing paradigms that tackle this challenge need to deal with the high level of noise and the high frame rate / bandwidth of these sensors.

This talk discusses an approach based on KinectFusion using solely point-based representations. This approach allows for low-level handling of scene dynamics. Furthermore, open problems are discussed and more general lines of future research are sketched.

3.13 Capture Close Interacting Motions

Yebin Liu (Tsinghua University – Beijing, CN)

Joint work of Liu, Yebin; Wang, Yangang; Theobalt, Christian; Chai, Jinxiang; Dai, Qionghai; Gall, Juergen; Min, Jianyuan; Stoll, Carsten; Zhang, Jianjie; Xu, Feng; Seidel, Hans-Peter
Main reference Y. Liu, J. Gall, C. Stoll, Q. Dai, H.-P. Seidel, C. Theobalt, "Markerless Motion Capture of Multiple Characters Using Multi-view Image Segmentation," IEEE Trans. Pattern Anal. Mach. Intell., 35(11):2720–2735, 2013.
URL http://dx.doi.org/10.1109/TPAMI.2013.47
Main reference Y. Wang, J. Min, J. Zhang, Y. Liu, F. Xu, Q. Dai, J. Chai, "Video-based Hand Manipulation Capture Through Composite Motion Control," ACM Trans. Graph., 32(4):43:1–43:14, 2013.
URL http://dx.doi.org/10.1145/2461912.2462000

The talk presents marker-less based motion capture of close interacting motions. Close interacting motion is common in everyday life and there are important for graphics animation, user interaction, sport analysis, biomechanics and so on. The two typical cases are human body interacting with human body and hand interacting with object. Even for maker motion capture, it is challenging to have these motions, because of the serious occlusion, possible collision and the requirement for capture of subtle contact phenomena. We start the talk from traditional single character or single object motion capture. Given the observed image data, we usually try to search in the pose space, to find out the pose that minimize the difference between the observed images and the rendering result. With a sample in the pose space, we usually first skinning the pose to a 3D mesh and then possibly texture mapping on it to get the appearance. Such way is an analysis by synthesis approach with sampling then rendering and then measurement. In this talk, I will two works showing how analysis by synthesis approach can be used in reconstruction of close interacting motions. 1) Multi-view body labeling using Analysis by synthesis in multi- person motion capture; 2) Analysis by synthesis integrated with Motion Control for hand-object motion capture.

3.14 Multiview HDR systems

Celine Loscos (Universite de Reims, FR)

Joint work of Loscos, Celine; Bonnard, Jennifer; Valette, Gilles; Lucas, Laurent; Remion, Yannick; Ismael, Muhannad; Blache, Ludovic; Prevost, Stephanie; Nocent, Olivier
Main reference L. Lucas, C. Loscos, Y. Remion, (eds.), "3D Video: From Capture to Diffusion," 325 pp., ISBN: 978-1-84821-507-8, Wiley-ISTE, 2013.

Multiview systems are nowadays used in many applications, but they can serve different objectives: stereo acquisition for 3D visualization and multi-viewpoint acquisition for 3D reconstruction. In this presentation, I will go through the details of two different approaches, that are work in progress conducted by a team of researchers of CReSTIC-SIC at the University of Reims on: - 3D HDR video acquisition for auto-stereoscopic visualization, providing HDR pixel values to eight aligned views, - a 30-camera acquisition studio for reconstructing animated models of a moving actor, as well as a hybrid 4D modeling algorithm, associating visual hull with multiscopic approaches to gain precision.

3.15 Towards Perfect Rendering of an Imperfect (Virtual) World

Marcus A. Magnor (TU Braunschweig, DE)

Main reference M. Mustafa, L. Lindemann, M. Magnor, "EEG analysis of implicit human visual perception," in Proc. of the 2012 ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI'12), pp. 513–516, ACM, 2012.
URL http://dx.doi.org/10.1145/2207676.2207746

Noise, outliers, and reconstruction errors will likely be accompanying real world-acquired digital models and scene representations for a long time to come. This talk presents strategies to assess the impact of model and rendering artifacts on visual perception based on electroencephalography (EEG).

The advantages of incorporating knowledge about our human visual system into graphics and visualization algorithms have been apparent for a long time. But while a considerable amount of work has been done on measuring cognitive processing in graphics and visualization through the use of external processes, e.g. surveys and direct observation, much less research has so far been devoted to take advantage of brain measurements. Due to recent hardware advances in the consumer market, primarily driven by applications to control games by thought (e.g. Emotiv's brain-computer interface), electroencephalography (EEG) has become an affordable technique to measure the visual processing in the brain.

It has long been understood that visual processing occurs without conscious perception or attention, and that conscious awareness of a visual stimulus is preceded by complex visual decision-making processes. Directly recording brain activity via EEG enables tapping into these implicit visual processes. A correctly designed EEG experiment is able to robustly and reproducibly disclose novel aspects of our visual perception, undeterred by decision-forming processes that necessarily precede any conscious answer to the questions in a conventional user study. On the other hand, successfully conducting an EEG experiment requires specialized knowledge of signal processing in the brain, in particular the strengths and limitations of the EEG technique. EEG measurements do not supplant traditional, question-based user studies, but offer access to new aspects of visual perception that are otherwise inaccessible.

3.16 3D Shape from Silhouettes in Water for Online Novel-View Synthesis

Shohei Nobuhara (Kyoto University, JP)

This paper is aimed at presenting a new algorithm for full 3D shape reconstruction and online free-viewpoint rendering of objects in water. The key contributions are (1) a new calibration model for the refractive projection, and (2) a new 3D shape reconstruction algorithm based on shape-from- silhouette (SfS) concept. We also propose an online free-viewpoint rendering system as a practical application.

3.17 Photographic Time Machines: Taking Control over the Time of Day in Photos

Sylvain Paris (Adobe Systems Inc. – Cambridge, US)

Joint work of Laffont, Pierre-Yves; Bousseau, Adrien; Drettakis, George; Shih, YiChang; Freeman, William; Durand, Frédo; Paris, Sylvain

Taking a picture in the middle of the day or during sunset makes a huge difference in the produced image. The changes of illumination throughout the day are a critical aspect of outdoor photography. These variations open creative opportunities but also constrain how outdoor photographers work. I will present two approaches to alter the illumination conditions in pictures of outdoor scenes to render them as if they had been taken at a different time of day. The first one relies on internet photo collections and partial 3D reconstructions, and the second on a dataset of time-lapse videos and scene matching. I will dedicate part of my talk to a discussion of the pros and cons of each approach. This is joint work with Pierre- Yves Laffont, Adrien Bousseau, George Drettakis, YiChang Shih, Bill Freeman, and Frédo Durand.

3.18 Videos in Panoramic Contexts

Fabrizio Pece (University College London, GB)

Joint work of Pece, Fabrizio; William, Steptoe; James, Tompkin; Rajvi, Shah; Simon, Julier; Fabian, Wanner; Tim, Weyrich; Sharham, Izadi; Christian, Theobalt; Jan, Kautz; Anthony, Steed
Main reference F. Pece, W. Steptoe, S. Julier, F. Wanner, T. Weyrich, J. Kautz, A. Steed, "PanoInserts: Mobile Spatial Teleconferencing," in Proc. of the 2013 ACM SIGCHI Conf. on Human Factors in Computing Systems (CHI'13) pp. 1319–1328, ACM, 2013.
URL http://dx.doi.org/10.1145/2470654.2466173

Panoramic images and video are now common, with the world quickly being mapped at street level by companies and tourists alike. On a spectrum between 3D virtual environments and 2D images, panoramas lie somewhere in between – a 360 degrees panorama can surround a user, but the scene has only spherical geometry and is effectively flat. For these reasons, panorama are an attractive basis for video- conferencing, as they provide a full 360 degree

view of an environment in a single image, but they are also a convenient context to temporally and spatially relate videos within large collections.

In this talk I have presented two videos-in-panoramic-context systems, together with their evaluations.

Firstly, I have introduced PanoInserts (Pece at al. 2013): a novel tele- conferencing system that uses smartphone cameras to create a surround representation of meeting places. I have showed the results of a user study comparing our system with fully-panoramic video and conventional webcam video conferencing for two spatial reasoning tasks, and I have discussed the representational properties and usability of varying video presentations, exploring how they are perceived and how they influence users when performing spatial reasoning tasks.

I have then presented Vidicontexts (Tompkin et al.): a video-collections+context interface that embeds videos into a panorama. The system features a spatio-temporal index and tools for fast exploration of the space and time of the video collection. I have presented the results of our study on the effect of our video-collection+context system to spatio-temporal localization tasks.

Finally, I have showed various adaptations of VidiContexts to several display devices, discussing the implications of of each of them on videos-in-panoramic- context interfaces.

References

1 J. Tompkin, F. Pece, R. Shah, S. Izadi, J. Kautz and C. Theobalt. *Video Collections in Panoramic Contexts*. Symposium on User Interface Software and Technology (UIST 2013) St Andrews, UK, Oct. 8-11, 2013.

2 F. Pece, W. Steptoe, S. Julier, F. Wanner,T. Weyrich, J. Kautz, and A. Steed. *PanoInserts: Mobile Spatial Teleconferencing*. ACM CHI Human Factors in Computing Systems (CHI 2013) Paris, France, April 22-May 2 2013.

3.19 Recent work on material models from captured data.

Holly E. Rushmeier (Yale University, US)

Joint work of Rushmeier, Holly E.; Dorsey, Julie; Xue, Su; Lockerman, Yitzchak; Wu, Hongzhi
Main reference Y. D. Lockerman, S. Xue, J. Dorsey, H. E. Rushmeier, "Creating Texture Exemplars from Unconstrained Images," Report no. TR1483, Yale University, 2013.
URL http://graphics.cs.yale.edu/site/tr1483
Main reference H. Wu, J. Dorsey, H. E. Rushmeier, "Inverse Bi-Scale Material Design," ACM Trans. Graph., 32(6):163:1–163:10, 2013.
URL http://dx.doi.org/10.1145/2508363.2508394

We discuss the extraction of uniform textures from natural images. First, we present a method of clustering textures in an image using diffusion distances. We then modify this approach using simple user input for the scale and sample location of the desired texture to form an interactive technique for extracing the texture.

We also describe bi-scale reflectance modeling. First we show how the large scale BRDF can be computed from small scale BRDF and geometry by formulating this as the multiplication of large matrices which are then simplified using a randomization technique. Inverting this process is possible by using a large library of geometries and BRDF that are represented by small, approximate matrices.

3.20 Accessing and Interacting with the Real World via the 3D-Web with XML3D

Philipp Slusallek (DFKI – Saarbrücken, DE)

Joint work of Klein, Felix; Rubinstein, Dmitri; Sons, Kristian; Einabadi, Farshad; Herhut, Stephan; Slusallek, Philipp
Main reference F. Klein, D. Rubinstein, K. Sons, F. Einabadi, S. Herhut, P. Slusallek, "Declarative AR and Image Processing on the Web with Xflow," in Proc. of the 18th Int'l Conf. on Web 3D Technology (Web3D'13), pp. 157–165, ACM, 2013.
URL http://dx.doi.org/10.1145/2466533.2466544
URL https://graphics.cg.uni-saarland.de/2013/declarative-ar-and-image-processing-on-the-web-with-xflow/

A key element for Real World Visual Computing is the ability of non-experts to get access to what we come up with in our research. In the past this has been more than problematic mainly because of the large efforts of developing advanced interactive Visual Computing applications as well as the problems of deploying them across different devices, hardware architectures, and operating systems. The availability of the low-level WebGL interface in the latest browsers has not really changed this situation much.

With XML3D we provide a high-level interface to interactive 3D Graphics and Visual Computing. XML3D is an extension to HTML-5 that runs in all modern browsers via an implementation in JavaScript that efficiently simulates native support by the browser. We use HTML to describe the 3D scene, reusing existing Web technologies (like CSS, DOM events, JS, etc.) wherever possible. This makes it possible for millions of experienced Web developers to easily make use of and integrate interactive 3D graphics into their Web applications.

But XML3D goes significantly further than traditional scene graphs: Its novel generic data handling concept allows highly flexible and GPU-friendly operations while data can be asynchronously be retrieved via simple URLs from anywhere in the Internet (just like HTML images). Semantic annotations allow for intelligent processing (even AI) within the 3D scenes and linking the objects back to their origin, such that distributed Web services can easily be created. Real-time synchronization of 3D scenes among many users is currently being added to XML3D.

A declarative flow-graph approach integrates efficient animation, image processing, and AR tracking operators flexibly and directly into XML3D. This allows for portable and interactive AR application directly in any mobile device. Parts of the flow-graph can be efficiently and portably be mapped to the available hardware (e.g. via vertex or compute shaders). Our new material model "shade.js" provides portable material descriptions that allow XML3D scenes to be rendered equally well using rasterization (forward or deferred) and ray tracing.

In my presentation, I give an overview of the XML3D technology and highlight some of its core capabilities and applications that have already been realized. During the workshop I would like to explore opportunities how XML3D can be used more widely also in the research community – both as a tool for the research itself as well as for delivering the research results directly and interactively to users.

References

1 Felix Klein, Dmitri Rubinstein, Kristian Sons, Farshad Einabadi, Stephan Herhut, and Philipp Slusallek, *Declarative AR and Image Processing on the Web with Xflow*, Proceedings of the 18th International Conference on Web 3D Technology of Web3D'13 , pp. 157–165, June 2013

2 Jozef Dobos, Kristian Sons, Dmitri Rubinstein, Philipp Slusallek, and Anthony Steed, *XML3DRepo: A REST API for Version Controlled 3D Assets on the Web*, Proceedings of the 18th International Conference on 3D Web Technology of Web3D'13 , pp. 47–55, June 2013
3 Jacek Jankowski, Sandy Ressler, Kristian Sons, Yvonne Jung, Johannes Behr, and Philipp Slusallek, *Declarative Integration of Interactive 3D Graphics into the World-Wide Web: Principles, Current Approaches, and Research Agenda*, Proceedings of the 18th International Conference on 3D Web Technology of Web3D'13 , pp. 39–45, June 2013
4 Kristian Sons, Christian Schlinkmann, Felix Klein, Dmitri Rubinstein, and Philipp Slusallek, *xml3d.js: Architecture of a Polyfill Implementation of XML3D*, 6th Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS), 2013
5 Felix Klein, Kristian Sons, Dmitri Rubinstein, Sergiy Byelozyorov, Stefan John, and Philipp Slusallek, *Xflow – Declarative Data Processing for the Web*, Proceedings of the 17th International Conference on Web 3D Technology , pp. 37–45, August 2012
6 Kristian Sons, Felix Klein, Dmitri Rubinstein, Sergiy Byelozyorov, and Philipp Slusallek, *XML3D – Interactive 3D Graphics for the Web*, Proceedings of the 15th International Conference on 3D Web Technology, pp. 175–184, July 2010

3.21 Challenges of high spatio-angular-temporal resolution image data

Alexander Sorkine-Hornung (Disney Research – Zürich, CH)

Joint work of Sorkine-Hornung, Alexander; Sorkine-Hornung, Olga; Kim, Changil; Zimmer, Henning; Pritch, Yael;Perazzi, Federico; Yücer, Kaan; Wang, Oliver; Gross, Markus
Main reference C. Kim, H. Zimmer, Y. Pritch, A. Sorkine-Hornung, M. Gross, "Scene Reconstruction from High Spatio-Angular Resolution Light Fields," ACM Trans. Graph., 32(4):73:1–73:12, 2013.
URL http://dx.doi.org/10.1145/2461912.2461926

Sensors capture the world at ever higher spatial, temporal, and angular resolution. For example, todays professional cinema cameras record 120fps at 4k and beyond. For the industry, one of the most obvious resulting problems is data management. However, an additional challenge arises, since many of our existing algorithms simply have not been designed to handle such densely sampled input. Besides the issue of increased computational complexity, in the past we have carefully tuned our existing methods to address issues of sparse sampling. For instance, standard strategies in image-based reconstruction and optical flow involve global regularization to handle ambiguous, incomplete, or noisy measurements. The corresponding benchmarks for stereo and flow rarely feature image resolutions over 1MP. In this presentation it is argued that for very densely sampled input, such as light fields or high frame rate video, existing strategies do not necessarily get the best out of the data, and that this type of data poses novel interesting research challenges beyond simply improving speed and scalability of existing techniques. As an example, a method for 3D scene reconstruction from "gigaray" light fields is proposed, which breaks with a number of established practices in image-based reconstruction in order to achieve improved reconstruction quality and speed.

3.22 Visualizing Point Sets

Ayellet Tal (Technion – Haifa, IL)

Joint work of Tal, Ayellet; Katz, Sagi; Basri, Ronen

If you try to draw a large point cloud, you my find that it is difficult to understand what the cloud represents. In this talk I will discuss the reduction of points, such that the visual comprehension of the set is improved. Two simple operators are utilized, the first computes the visible points from a given viewpoint and the second is its dual. Both operators work directly on the point cloud, skipping surface reconstruction.

References

1 Sagi Katz and Ayellet Tal and Ronen Basri. *Direct Visibility of Point Sets*. ACM Transactions on Graphics, Volume 26, Issue 3, 2007, 24/1-24/11
2 Sagi Katz and Ayellet Tal. *Improving the Visual Comprehension of Point Sets*. CVPR, 2013

3.23 Finite Element Based Tracking of Deforming Surfaces

Stefanie Wuhrer (Universität des Saarlandes, DE)

Joint work of Wuhrer, Stefanie; Lang, Jochen; Tekieh, Motahareh; Shu, Chang
Main reference S. Wuhrer, J. Lang, M. Tekieh, C. Shu, "Finite Element Based Tracking of Deforming Surfaces," arXiv:1306.4478v1 [cs.CV], 2013.
URL http://arxiv.org/abs/1306.4478v1

We present an approach to robustly track the geometry of an object that deforms over time from a set of input point clouds captured from a single viewpoint. The deformations we consider are caused by applying forces to known locations on the object's surface. Our method combines the use of prior information on the geometry of the object modeled by a smooth template and the use of a linear finite element method to predict the deformation. This allows the accurate reconstruction of both the observed and the unobserved sides of the object. We present tracking results for noisy low-quality point clouds acquired by a stereo camera and a Kinect sensor, and simulations with point clouds corrupted by different error terms, and show that our method is also applicable to large non-linear deformations.

Participants

- Philippe Bekaert
Hasselt Univ.. – Diepenbeek, BE
- Tamy Boubekeur
Télécom Paris Tech, FR
- Edmond Boyer
INRIA Rhône-Alpes, FR
- Gabriel Brostow
University College London, GB
- Darren Cosker
University of Bath, GB
- Carsten Dachsbacher
KIT – Karlsruhe Institute of Technology, DE
- Robert Dawes
BBC Research & Development – London, GB
- Jean-Michel Dischler
University of Strasbourg, FR
- Bernd Eberhardt
Hochschule der Medien – Stuttgart, DE
- Peter Eisert
Fraunhofer-Institut – Berlin, DE
- Paolo Favaro
Universität Bern, CH
- Dieter W. Fellner
TU Darmstadt, DE
- Jan-Michael Frahm
University of North Carolina – Chapel Hill, US
- Martin Fuchs
Universität Stuttgart, DE
- Bastian Goldlücke
Universität Heidelberg, DE
- Oliver Grau
Intel Visual Computing Inst. – Saarbrücken, DE
- Volker Helzle
Filmakademie Baden-Württemberg – Ludwigsburg, DE
- Anna Hilsmann
HU Berlin, DE
- Adrian Hilton
University of Surrey, GB
- Martin Jagersand
University of Alberta, CA
- Jan Kautz
University College London, GB
- Oliver Klehm
MPI für Informatik – Saarbrücken, DE
- Felix Klose
TU Braunschweig, DE
- Andreas Kolb
Universität Siegen, DE
- Hendrik P. A. Lensch
Universität Tübingen, DE
- Christian Lipski
metaio GmbH – München, DE
- Yebin Liu
Tsinghua Univ. – Beijing, CN
- Céline Loscos
Universite de Reims, FR
- Marcus A. Magnor
TU Braunschweig, DE
- Shohei Nobuhara
Kyoto University, JP
- Sylvain Paris
Adobe Systems Inc. – Cambridge, US
- Fabrizio Pece
University College London, GB
- Kari Pulli
NVIDIA Corporation – Santa Clara, US
- Bodo Rosenhahn
Leibniz Univ. Hannover, DE
- Holly E. Rushmeier
Yale University, US
- Hans-Peter Seidel
MPI für Informatik – Saarbrücken, DE
- Alla Sheffer
University of British Columbia – Vancouver, CA
- Philipp Slusallek
DFKI – Saarbrücken, DE
- Olga Sorkine-Hornung
ETH Zürich, CH
- Alexander Sorkine-Hornung
Disney Research – Zürich, CH
- Ayellet Tal
Technion – Haifa, IL
- Christian Theobalt
MPI für Informatik – Saarbrücken, DE
- Stefanie Wuhrer
Universität des Saarlandes, DE
- Ruigang Yang
University of Kentucky, US
- Remo Ziegler
LiberoVision AG – Zürich, CH

Report from Dagstuhl Seminar 13441

Evaluation Methodologies in Information Retrieval

Edited by

Maristella Agosti[1], Norbert Fuhr[2], Elaine Toms[3], and Pertti Vakkari[4]

1 **University of Padova, IT,** agosti@dei.unipd.it
2 **Universität Duisburg-Essen, DE,** norbert.fuhr@uni-due.de
3 **Sheffield University, GB,** e.toms@sheffield.ac.uk
4 **University of Tampere, FI,** pertti.vakkari@uta.fi

Abstract

This report documents the program and the outcome of Dagstuhl Seminar 13441 "Evaluation Methodologies in Information Retrieval", which brought together 42 participants from 11 countries. The seminar was motivated by the fact that today's information retrieval (IR) applications can hardly be evaluated based on the classic test collection paradigm, thus there is a need for new evaluation approaches. The event started with five introductory talks on evaluation frameworks, user modeling for evaluation, evaluation criteria, measures, evaluation methodology, and new trends in IR evaluation. The seminar participants then formed working groups addressing specific aspects of IR evaluation, such as reliability and validity, task-based IR, learning as search outcome, searching for fun, IR and social media, graph search, domain-specific IR, interaction measures and models, and searcher-aware information access systems.

Seminar 27. October – 01. November, 2013 – www.dagstuhl.de/13441
1998 ACM Subject Classification H.3.4 Systems and Software: Performance evaluation (efficiency and effectiveness), H.5.2 User Interfaces: Benchmarking, Evaluation/methodology
Keywords and phrases evaluation, testbeds, user studies, measures
Digital Object Identifier 10.4230/DagRep.3.10.92
Edited in cooperation with Vu Tran

1 Executive Summary

Maristella Agosti
Norbert Fuhr
Elaine Toms
Pertti Vakkari

Evaluation of information retrieval (IR) systems has a long tradition. However, the test-collection based evaluation paradigm is of limited value for assessing today's IR applications, since it fails to address major aspects of the IR process. Thus there is a need for new evaluation approaches, which was the focus of this seminar.

Before the event, each participant was asked to identify one to five crucial issues in IR evaluation methodology. Pertti Vakkari presented a summary of this homework, pointing out that there are five major themes deemed relevant by the participants: 1) Evaluation frameworks, 2) Whole session evaluation and evaluation over sessions, 3) Evaluation criteria: from relevance to utility, 4) User modeling, and 5) Methodology and metrics.

Evaluation Methodologies in Information Retrieval, *Dagstuhl Reports*, Vol. 3, Issue 10, pp. 92–126
Editors: Maristella Agosti, Norbert Fuhr, Elaine Toms, and Pertti Vakkari

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Based on the evaluation model proposed in Saracevic & Covi [1], the seminar started with four introductory talks covering major areas of IR evaluation: Nick Belkin gave a survey over "Framework(s) for Evaluation (of whole-session) IR", addressing the system components to be evaluated and the context to be considered. In his presentation "Modeling User Behavior for Information Retrieval Evaluation", Charlie Clarke described efforts for improving system-oriented evaluation through explicit models of user behavior. Kal Järvelin talked about "Criteria in User-oriented Information Retrieval Evaluation", characterizing them as different types of experimental variables and distinguishing between output- and (task-)outcome related criteria. "Evaluation Measures in Information Retrieval" by Norbert Fuhr outlined the steps necessary for defining a new metric and the underlying assumptions, calling for empiric foundation and theoretic soundness. Diane Kelly presented problematic issues related to "Methodology in IR Evaluation", such as the relationship between observation variables and criteria, the design of questionnaires, the difference between explanatory and predictive research and the appropriateness of statistical methods when dealing with big data. The round of introductory talks was concluded with Maristella Agosti's presentation "Future in Information Retrieval Evaluation", where she summarized challenges identified in three recent workshops in this area.

For the rest of the week, the participants then formed working groups described in the following.

"From Searching to Learning" focused on the learning as search outcome and the need for systems supporting this process. Learning may occur at two different levels, namely the content level and the search competence level. There is a need for understanding of the learning process, its relationship to the searcher's work task, the role of the system, and the development of appropriate evaluation methods. Approaches may address different aspects of the problem, such as the system, the interaction, the content, the user and the process. For evaluation, the framework from Ingwersen and Jarvelin [2] suggests criteria and measures at the levels of information retrieval, information seeking, the work task and the social-organizational and culture level.

"Social Media" allow users to create and share content, with a strong focus on personal connections. While web search engines are still the primary starting point for many information seeking activities, information access activities are shifting to more personalized services taking into account social data. This trend leads to new IR-related research issues, such as e.g. utility, privacy, the influence of diverse cultural backgrounds, data quality, authority, content ownership, and social recommendations. Traditional assumptions about information seeking will have to be revised, especially since social media may play a role in a broad range of information spaces, ranging form everyday life and popular culture to professional environments like journalism and research literature.

"Graph Search and Beyond" starts from the observation that an increasing amount of information on the Web is structured in terms of entities and relationships, thus forming a graph, which, in turn allows for answering more complex information needs. For handling these, search engines should support incremental structured query input and dynamic structured result set exploration, Thus, in contrast to the classical search engine result page, graph search calls for an incremental query exploration page, where entries represent the answers themselves (in the form of entities, relationships and sub-graphs). The new possibilities of querying and result presentation call for the development of adequate evaluation methods

"Reliability and Validity" is considered as the most central issue in IR evaluation, especially in the current situation where there is increasing discussion in the research community about

reproducibility and generalizability of experimental results. Thus, this working group decided to start the preparation of a book on best practices in IR evaluation, which will cover the following aspects: Basic definitions and concepts, reliability and validity in experimentation, reporting out experiments, failure analysis, definition of new measures and methods, guidelines for reviewing experimental papers.

"Domain Specific Information Retrieval" in specific domains like e.g. in cultural heritage, patents and medical collections is not only characterized through the specifics of the content, but also through the typical context(s) in which this information is accessed and used, which requires specific functionalities that go beyond the simple search interaction. Also, context often plays an important role, and thus should be considered by the information system. However, there is a lack of appropriate evaluation methods for considering contexts and new functions.

"Task-Based IR" typically refers to research focusing on the task or goal motivating a person to invoke an IR system, thus calling for systems being able to recognize the nature of the task and to support the accompanying search process. As task types, we can distinguish between motivating tasks, seeking tasks, and search tasks. Task-based IR approaches should be able to model people as well as the process, and to distinguish between the (task-related) outcome and the (system) output.

"Searching for Fun" refers to the interaction with an information system without a specific search objective, like e.g. online window shopping, watching pictures or movies, or reading online. This type of activity requires different evaluation criteria, e.g. with regard to stopping behavior, dwell time and novelty. Also, it is important to distinguish between system criteria and user criteria, where the latter may be subdivided into process criteria and outcome criteria. A major problem in this area is the design of user studies, especially since the starting points (e.g. casual or leisure needs) are difficult to create under experimental conditions. A number of further issues was also identified.

The working group "The Significance of Search, Support for Complex Tasks, and Searcher-aware Information Access Systems" addressed three loosely related challenges. The first topic addresses the definition of IR in the light of the dramatic changes during the last two decades, and the limited impact of our research. The second topic is the development of tools supporting more complex tasks, and their evaluation. Finally, information systems should become more informed about the searcher and the progress in user's task.

"Interaction, Measures and Models" discussed the need for a common framework for user interaction models and associated evaluation measures, especially as a means for achieving a higher degree of reliability in interactive IR experiments. This would allow for evaluating the effect of the interaction and the interface on performance. A possible solution could consist of three components, namely an interaction model, a gain model and a cost model.

Finally, many of the attendees were planning to continue to collaborate on the topics addressed during the seminar since the fruitful discussions were a useful base for future cooperation.

References

1 Tefko Saracevic, Lisa Covi (2000). Challenges for digital library evaluation. In D. H. Kraft (Ed.), Knowledge Innovations: Celebrating Our Heritage, Designing Our Future. *Proceedings of the 63rd Annual Meeting of the American Society for Information Science. Washington, D.C.: American Society for Information Science.* pp. 341–350.

2 Peter Ingwersen, Kalervo Järvelin (2005). The Turn: Integration of Information Seeking and Retrieval. In *Context. Dortrecht, NL: Springer.* ISBN 1-4020-3850-X

2 Table of Contents

3 Overview of Talks

3.1 A Summary of Homework

Pertti Vakkari (University of Tampere, FI)

The major themes in the issues of IR evaluation methodology are presented. They include 1) Evaluation frameworks, 2) Whole session evaluation and evaluation over sessions, 3) Evaluation criteria: from relevance to utility, 4) User modeling, and 5) Methodology and metrics.

3.2 Framework(s) for Evaluation (of whole – session) IR

Nicholas J. Belkin (Rutgers University – New Brunswick, US)

Main reference T. Saracevic,L. Covi, "Challenges for digital library evaluation," in D. H. Kraft (Ed.), Knowledge Innovations: Celebrating Our Heritage, Designing Our Future – Proc. of the 63rd Annual Meeting of the American Society for Information Science, pp. 341–350, American Society for Information Science, 2000.

This presentation uses the structure proposed by Saracevic and Covi (ti00ti) to discuss the constructs and contexts that could specify framework(s) for evaluation of interactive information retrieval (IR). These constructs and contexts are considered from the point of view of the following overall goal for IR systems in general: The goal of (IR) systems is to support people in resolution of the tasks or goals that led them to engage in information seeking in an IR system, through effective interaction with information objects. I propose the following.

An IR system consists of:

- An information resource
- Methods for organizing and representing IOs
- People who have "information problems"
- Methods for representing information problems
- Methods for retrieving and presenting IOs in response to information problems
- Methods for supporting interaction of the people and the other components of the IR system

Most of these elements of the IR system should be evaluated, although some to a greater extent than others. Evaluating information resources is a crucial issue for Web search engines. Evaluating methods for organizing and representing IOs is a classic IR issue. Evaluating people is clearly not our problem, but understanding them is. Methods for evaluating representation of information problems has been ignored, but is increasingly realized as being important. Methods for retrieving and presenting IOs is a classic IR issue, although ideas of presentation have been rather limited.

Saracevic and Covi suggest that systems can be evaluated at Social, Institutional, Individual, Interface, Engineering, Processing and Content levels. It seems likely that the Social level is probably not relevant to IR evaluation; the Institutional level is peripheral; the Individual level is crucial; the Interface level is crucial; the Engineering level is also central;

the Processing level is crucial, but focuses on two aspects: do algorithms work as intended, and do algorithms do what is intended; Content level needs to be evaluated.

If we construe an information seeking episode as a sequence of different kinds of interactions of the information seeker with information objects, with the various IR techniques (i.e. Methods as above) adapting to support the different interactions, then evaluation of their support can be tailored to the goals of each of the different kinds of interaction.

3.3 Modeling User Behavior for Information Retrieval Evaluation

Charles Clarke (University of Waterloo, CA)

Information retrieval systems may be evaluated through user oriented studies or system-oriented tests. User-oriented studies are based on actual user behavior, including laboratory experiments, A/B testing, and the analysis of interaction logs. Unfortunately, these studies can be expensive, requiring substantial time, money, and data. System-oriented tests, often called batch-style or "Cranfield-style" tests, provide a lost-cost and repeatable alternative. Unfortunately, these tests may be criticized for lacking a clear connection with actual user behavior and preferences, and for reporting results in meaningless units.

This presentation describes various efforts to improve system-oriented testing through the addition of explicit models of user behavior. As a specific example, we examine time-biased gain (TBG). TBG provides a unifying framework for information retrieval evaluation, generalizing many traditional effectiveness measures while accommodating aspects of user behavior not captured by these measures. By using time as a basis for calibration against actual user data, TBG can reflect aspects of the search process that directly impact user experience, including document length, near-duplicate documents, and summaries. Unlike traditional measures, which must be arbitrarily normalized for averaging purposes, TBG is reported in meaningful units, such as the total number of relevant documents seen by the user. TBG also provides a method for incorporating user variance into system-oriented tests. The modeling of user variance is critical to understanding the impact of effectiveness differences on the actual user experience. If the variance of a difference is high, the effect on user experience will be low. By incorporating per-query variance, TBG allows for the measurement of the effect size of differences, which allows researchers to understand the extent to which predicted performance improvements matter to real users. The development of TBG is joint work with Mark Smucker appearing in SIGIR 2012, CIKM 2012 and HCIR 2012.

In addition, the presentation provides an overview of a SIGIR 2013 workshop on Modeling User Behavior for Information Retrieval Evaluation (MUBE 2013). The workshop brought together researchers interested in improving Cranfield-style evaluation of information retrieval through the modeling of user behavior. After two invited talks and ten short paper presentations, the workshop participants brainstormed research questions of interest and formed breakout groups to explore these questions in greater depth. This presentation summarizes some of the important questions raised by the workshop and briefly outlines some resulting research directions for the improvement of information retrieval evaluation. The organization of the workshop was a joint effort with Luanne Freund, Mark Smucker, and Emine Yilmaz.

3.4 Criteria in User-oriented Information Retrieval Evaluation

Kalervo Järvelin (University of Tampere, FI)

In the presentation, criteria are the dimensions of evaluation. The presentation discusses research designs where the criteria are used as dependent, independent and controlled variables.

The presentation first discussed nested evaluation frameworks for Information Retrieval (IR) – from a specific IR context to increasingly contexts including the information seeking context, the work task context, and the socio-organizational context. Exemplary evaluation criteria were given for each. It was stressed that if the broader contexts are neglected, there is the risk of sub-optimization in IR system development. It was also pointed out that evaluation and theory development in IR go hand in hand because, on the one hand, evaluation requires a model (a theory) of the system being evaluated and some goal to be achieved and, on the other hand, theory grows in instrumental research (like IR) through evaluation. Evaluation requires experimental designs where the evaluation criteria are used as dependent, independent and controlled variables. This allows evaluation / theory development where the effects of some independent variables are tested on one or more dependent variables.

In a nested evaluation framework, the dependent variables (of a narrow framework) may indirectly affect some dependent variable (of a broader framework) that remains outside the evaluation design. The presentation discussed some experimental evaluation designs. It was underlined, that typically in information retrieval research, especially in test-collection based evaluation studies, the evaluation design is specified such that the dependent variable is the search engine's ranking effectiveness (measured through some metric), and the independent variables consist of document representation and topic methods, and of matching methods for comparing the former. However, the ultimate dependent variable is effective information interaction, and it is often believed that the latter is positively correlated with the former. Alternative experimental designs seek to identify context, searcher and system criteria affecting information access and ultimately effective information interaction. The controlled variables may contain some context variables, searcher variables and/or system variables. The independent variables may belong to the same categories. Ingwersen and Järvelin (2005, Chapter 7) discuss these categories of variables.

Technology alone is insufficient in explaining the effectiveness of interactive IR. In order to develop the technology sensibly we need to understand how technology together with users-in-context produces the desired outcomes in information access and the ultimate benefits (Järvelin 2013). Failing to take context and searchers into account in many study designs may be one reason for the views that IR is not a very theoretical field but rather pragmatic. However, it is exactly more theory that is required to manage the complexity of interactions in IR. Experiments for theory building may be based on test collections and simulation, real-user experiments in test collections, or operational systems evaluations. A study design used in Baskaya & al. (2013) was discussed.

In general, user studies are useful for IR systems development when (1) they inform design or (2) guide design. The former may be based on deliberately incorporating systems variables in the study designs. The latter may be based on identifying user or interaction variables that contribute to the dependent variables AND that may be affected by (future) systems variables. However, user studies may also be useful when there is no instrumental

(system design) interest. This happens when they focus research on fruitful areas, or help us understand information interaction – in order to later support it.

The ideas presented above are discussed at more depth in the contributions cited below.

References

1 Ingwersen, P. & Järvelin, K. (2005). The Turn: Integration of Information Seeking and Retrieval. In *Context. Dortrecht, NL: Springer.* ISBN 1-4020-3850-X

2 Baskaya, F. & Keskustalo, H. & Järvelin, K. (2013). Modeling Behavioral Factors in Interactive Information Retrieval. In: *He, Q. & al. (Eds.). Proceedings of the 22nd International Conference on Information and Knowledge Management (CIKM 2013)*, Burlingame, CA, Oct 27 – Nov 1, 2013. New York, NY: ACM Press, pp. 2297–2302. doi>10.1145/2505515.2505660

3 Järvelin, K. (2007). An Analysis of Two Approaches in Information Retrieval: From Frameworks to Study Designs. *Journal of the American Society for Information Science and Technology (JASIST)* 58(7): 971–986.

4 Järvelin, K. (2013). User-Oriented Evaluation in IR. In: Agosti, M. & al. (eds.), *PROMISE Winter School 2012, Heidelberg: Springer, Lecture Notes in Computer Science vol. 7757*, pp. 86–91.

5 Järvelin, K. (2011). Evaluation. In: Ruthven, I. & Kelly, D. (eds.), I*nteractive Information Seeking, Behaviour and Retrieval. London, UK: Facet Publishing*, pp. 113–138. ISBN: 978-1-85604-707-4

6 Järvelin, K. (2009). Explaining User Performance in Information Retrieval: Challenges to IR evaluation. In: *L. Azzopardi & al. (Eds.), Proceedings of the 2nd International Conference on the Theory of Information Retrieval. 2009, Heidelberg: Springer, Lecture Notes in Computer Science vol. 5766*, pp. 289–296.

3.5 Evaluation Measures in Information Retrieval

Norbert Fuhr (Universität Duisburg-Essen, DE)

There is a wide variety of IR measures, but many of them are defined in an ad-hoc way, Basically, the definition of a new metric should consist of the following steps:

- Starting from the chosen criterion, assume a specific user behavior (e.g. stopping after a certain number of relevant documents)
- Define preferences (e.g. the smaller the number of documents seen, the better).
- Define the basic metric obeying the preferences (e.g. precision).
- Furthermore, one can assume a user population, and compute a weighted average of the metric values according to this population (e.g., for average precision, it is assumed that at each relevant document, the same number of users stops).
- Finally, for getting a single result for a set of queries or sessions, an aggregation method has to be chosen (e.g. arithmetic mean).

Many of the current metrics suffer from a number of weaknesses

- The underlying assumptions are not made explicit and/or lack lack empiric foundation (e.g. for mean average precision, Robertson 2008 reconstructed the underlying assumptions; moreover, the assumption of a uniform distribution of users over the possible stopping points seems unrealistic).

- They have theoretic flaws (e.g., reciprocal rank can hardly be seen as interval-scaled, which, however, is a prerequisite for computing mean values).

Based on these observations, one can formulate a number of requirements for the development of new metrics. They should

- allow for more complex user behavior (beyond going through a linear list),
- be able to consider more complex benefits, (like e.g. dependency between documents, or a user searching for fun would like to be entertained all the time),
- have a proper empiric foundation (e.g. with respect to the stopping behavior of a user population),
- be theoretically sound by complying to the fundamentals of measurement theory as well as to basic axioms.

References

1 Luca Busin, Stefano Mizzaro: Axiometrics: An Axiomatic Approach to Information Retrieval Effectiveness Metrics. *ICTIR 2013:8*

2 Stephen Robertson, A new interpretation of average precision, *Proc. SIGIR 2008*, pp. 689–690

3 Warren S. Sarle: Measurement Theory: Frequently asked questions. *Disseminations of the International Statistical Applications Institute, 4th edition*, 1995, Wichita: ACG Press, pp. 61–66. *Revised March 18, 1996.*

3.6 Methodology in IR Evaluation

Diane Kelly (University of North Carolina – Chapel Hill, US)

This talk presents several potentially problematic issues related to evaluation methodology in IR evaluation. The distinction between methodology and methods is made, and questions regarding typical measurement practices, including the convenient practices of associating available and easily obtainable signals (e.g., dwell time) with a number of different constructs (e.g., usefulness, relevance, engagement) without clearly or formally developing a measurement model, and the ad-hoc development of questionnaire items to assess user experience, are raised. A standard psychometric theory is presented, along with a set of steps in which one might engage to create a valid and reliable measure. The talk then examines the differences between explanatory and predictive research and issues related to sample size, power analysis and effect size. A recent study, which questions the appropriateness of some statistical methods to the analysis of big data, is reviewed. The talk closes with some questions to guide workshop discussions about methodology in IR evaluation.

3.7 Future in Information Retrieval Evaluation

Maristella Agosti (University of Padova, IT)

The talk has presented some new challenges in Information Retrieval Evaluation that have been identified thanks to the CULTURA project (http://www.cultura-strep.eu/), the SIGIR 2013 workshop on Exploration, navigation and retrieval of information in cultural heritage (ENRICH 2013, http://www.cultura-strep.eu/events/enrich-2013), and the PROMISE Retreat on Prospects and Opportunities for Information Access Evaluation (Brainstorming workshop held on May 30–31, 2012, Padua, Italy). In fact relevant aspects of the CULTURA project and environment together with ENRICH 2013 and the PROMISE Retreat Report give examples of evaluation challenges that need to be addressed in the next future.

References

1 S. Lawless, M. Agosti, P. Clough, O. Conlan. Exploration, navigation and retrieval of information in cultural heritage: ENRICH 2013. *Proc. of the 36th ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2013)*. ACM, New York, USA, page 1136

2 PROMISE Retreat Report on Prospects and Opportunities for Information Access Evaluation. *Brainstorming workshop*, May 30–31, 2012, Padua, Italy

4 Working Groups

4.1 From Searching to Learning

Luanne Freund (University of British Columbia – Vancouver, CA), Jacek Gwizdka (University of Texas – Austin, US), Preben Hansen (Stockholm University, SE), Jiyin He (CWI – Amsterdam, NL), Noriko Kando (National Institute of Informatics – Tokyo, JP), Soo Young Rieh (University of Michigan – Ann Arbor, US)

Search systems to date have been viewed more as tools for the retrieval of content to satisfy information needs, than as environments in which humans interact with information content in order to learn. However, as full-text, information rich search systems become the norm, there is growing recognition of the importance of learning as a search outcome and of the need to provide support for it (Allan et. al., 2012). This is particularly true for environments in which learning is an acknowledged priority, such as collaborative, workplace, and academic search, but learning may also be an important general outcome of search that is not well-served by the drive for ever-greater customization and efficiency. In order to design systems that support learning, we need to investigate when and how learning occurs and develop reliable methods and measures to assess learning through search.

4.1.1 Concepts

Search provides an opportunity for learning on multiple levels, which should be distinguished in order to develop appropriate assessment tools. The primary level relates to learning

about the content being searched, which may include acquiring subject knowledge and/or an understanding of the searcher's problem space in relation to the content. At a secondary level, the searcher may also learn about the search system and develop search skills and competencies. Searching may also provide opportunities to learn about oneself and about society through the lens of the content searched. While only a few studies exist that focus on search as learning (Jansen, Booth, & Smith, 2009; Wilson & Wilson, 2013), there is a substantial literature on relevant concepts and frameworks of learning (e.g. Bloom et al, 1956; Kaptelenin & Nardi, 2006; Kintsch, 1998). Across these frameworks, learning is characterized in diverse ways, including learning as knowledge acquisition, learning as sense-making, learning as interpreting, and learning as synthesizing (Säljö 1979; Smith, 2013). Given the breadth of approaches to learning, researchers seeking to assess search as learning need to be explicit about the theoretical framework employed.

4.1.2 Issues

Some of the key issues related to the evaluation of search as learning are:

- How does learning occur through search?
- How does the learning process fit into the searcher's broader Work Task?
- Which system functionalities, components and features of search systems influence learning outcomes?
- What signals are indicators of learning?
- What are appropriate methods and measures?
- Can methods be imported from other fields (learning science, education, cognitive science)?
- Can these incorporated into a methodology to understand the learning process?

4.1.3 Approaches

When approaching the area of search as learning, we may take different viewpoints and look at the problem from different dimensions. Possible approaches could be:

System – How do information access systems including IR systems and tools facilitate learning? When building IR systems, functionalities that support learning should be considered. *Interaction* – How can we design systems that support subject learning? At this dimension, aspects from interactive IR, HCI and Interaction design could be used. How can a system assess the knowledge state of the user? Interaction with content: *Search trails* – predefined trails through content to optimize the searcher learning experience. *Information/resources* – For example, we may prioritize novel content and how to manipulate the quality, quantity of results. The informativeness of the information may be considered. *User* – Several aspects of searchers' knowledge and status are related: how to use the IR system; search strategies and tactics; domain knowledge; task knowledge; socio-cultural background; reading and comprehension ability; ability to conceptualize and integrate; ability to information use *Process* – the learning process need to be acknowledged and taken into account. For example: *Session track research* – that focuses on how systems information should be displayed over the course of a session, depending on the user learning. (e.g.Bates: Berry-picking model). Furthermore, search as learning is just one component of the whole learning process and thus contextual aspects need to be considered.

4.1.4 Measures

Ingwersen and Järvelin provide a conceptual framework for interactive IR evaluation (2005). Based on the framework, learning can occur across the process of search and as an outcome

of the search at all four levels. Corresponding measures at each level need to be defined. Table 1 identifies some indicators of learning that could be considered.

Information retrieval Level – Criteria and measures Patterns of query formulation and reformulation, query length, term and terminological variety; number of documents viewed, saved, and downloaded; number of documents assessed and time spent on assessing; pace of interaction, informativeness measures.

Information seeking Level – Criteria and measures Diversity of information seeking strategies; depth, breath, and richness of searchers' understanding of the subject area; searchers' knowledge level and confidence; comprehension test scores (T/F, summary writing), interaction metrics (annotation, notes, writing); searchers' cognitive load, mental workload, affect (happiness, frustration, engagement).

Work Task Level – Criteria and Measures Amount, quality, diversity of the outputs of the searcher's work task, e.g., work report, essay, and decisions made.

Social-organizational and culture Level – Criteria and Measures Success of the organization or social unit, job satisfaction, job promotion, evidence of lifelong learning

4.1.5 Methods

There are a wide range of data collection methods from diverse academic disciplines that could be applied in order to evaluate learning through searching. For example, transaction logs, eye-tracking, think-aloud, observation, self-reports, and interviews could be used. Some methods could be domain- and content-specific, while others are more generic. Methods developed from learning sciences (LS) could provide a useful toolbox. These include pre- and post- tests using instruments such as multiple-choice tests, domain-term lists, concept mapping, essays, comprehension tests, and text understanding measures (sentence verification tasks (SVT), Salmeron et al. 2010).

4.1.6 Conclusion

This summary report aims at providing an initial outline of frameworks and approaches to the evaluation of learning in the context of search in order to better understand how learning takes place and to inform the design of interactive information systems to better support people learning.

References

1 Allan, J., Croft, B., Moffat, A., and Sanderson, M. (2012). Frontiers, challenges, and opportunities for information retrieval: *Report from SWIRL 2012 the second strategic workshop on information retrieval in Lorne. SIGIR Forum 46:1* (May 2012), pp. 2–32, ACM Press, New York, NY, USA. http://doi.acm.org/10.1145/2215676.2215678

2 Bloom, B. S., Englehard, E., Furst, W., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. New York: McKay.

3 Jansen, B. J., Booth, D., Smith, B. (2009). Using the taxonomy of cognitive learning to model online searching. *Information Processing and Management, 45,* pp. 643–663.

4 Kaptelinin, V. and Nardi, B. (2006). Acting with Technology: Activity Theory and Interaction Design. *Cambridge: MIT Press.*

5 Kintsch, W. (1998) Comprehension: A paradigm for cognition. *New York: Cambridge University Press.*

6 Säljö, R. (1979). Learning in the learner's perspective. I. Some common-sense conceptions, *Reports from the Institute of Education, University of Gothenburg, 76.*

7 Salmeron, L., Gil, L., Braten, I., & Stromso, H. (2010). Comprehension effects of signalling relationships between documents in search engines. *Computers in Human Behavior, 26(3)*, pp. 419–426. doi: 10.1016/j.chb.2009.11.013.
8 Smith, M. K. (2003). Learning theory, in Encyclopedia of Informal Education. http://infed.org/mobi/learning-theory-models-product-and-process/. Retrieved:October 31, 2013.
9 Wilson, M. J. and Wilson, M. L. (2013). A comparison of techniques for measuring sense-making and learning within participant-generated summaries, *Journal of the American Society for Information Science and Technology, 64(2)*, pp .291–306.

4.2 Social Media

Omar Alonso (Microsoft Research – Mountain View, US), Ann Blandford (University College London, GB), Maarten de Rijke (University of Amsterdam, NL), Norbert Fuhr (Universität Duisburg–Essen, DE), Peter Mutschke (GESIS – Köln, DE), Doug Oard (University of Maryland – College Park, US), Max L. Wilson (University of Nottingham, GB)

4.2.1 Introduction

Social media refers to the interaction among people who share different types of information in a particular Internet service. When researchers and practitioners invoke social media, it is usually in the context of social networks like Facebook or Twitter.

All these services have a strong focus on personal connections (e.g., friends, followers) and on user-generated content that is shaped at least in part by those social connections. The use of people to create and enhance content is not new, Wikipedia being but one less "social" example. Having said that, social media has a strong focus on personalization: you are the query.

While search engines like Google and Bing receive millions of queries per day, information dissemination and consumption is also a prominent feature of services with a focus on social characteristics. This phenomenon is changing the landscape of how users access and share information. With users multi-tasking between different information services to get what they are looking for, there is an increased interest to incorporate, to some extent, social data into well-established services. For example, Bing introduced the annotations of Web links with social connections from Facebook, and Google implemented a similar feature using Google+.

Given the huge adoption of social media, what are the implications for the IR community? How do we evaluate the contribution of social data for the next generation of search engines? We need to further investigate user needs, user intent, and the utility of this new source of content and behavioral evidence.

4.2.2 Task categorization

Thinking about tasks and components

We propose the following levels of contexts for characterizing user tasks. We also include components in parentheses.

- IR context: TREC-like, Reputation management (effect prediction), Be the query (contextual search)
- Seeking context: Social utilities, Social load balancing (dynamic routing), Ideation (prognostication detection)
- Socio-organizational & cultural context: A task ecosystem, Buzz exploration (causal reasoning), Event monitoring (interestingness ranking), Groupalization (community detection), Tweet to Powerpoint (contextualization)

Facets

We can further break down the tasks and components into facets to get a different perspective:
- Directionality: Encountering, Monitoring, Influencing, Joining the conversation
- Object: Information, People
- Actor: Human, Machine [on behalf of some human(s)]

4.2.3 Issues

Some of the key issues dealing with social data are:
- Utility: how useful is this data and how can it be used?
- Privacy: how do we explain how the data will be used?
- Differing cultural expectations on privacy
- Controversial content (e.g., adult, racism, etc.) and unsanctioned content (modeling censorship)
- Data quality in an adversarial environment: buying followers (e.g., by celebrities)
- Inferred content (e.g., implicit geo-tagging)
- Informal use of language
- Data cleaning and provenance
- Estimating interestingness (societal, personal, transience)
- Authority detection, personal resolution and the filter bubble
- Content ownership, evolution and curation
- Influence of social recommendations on information seeking behavior
- Feedback, network and virality effects of social media on knowledge dissemination and community building

4.2.4 Opportunities

New assumptions

We need to challenge the traditional assumptions on how users interact with a search engine or IR system. Questioning the established beliefs is essential to understand the potential of social media for the next generation of IR systems. We suggest the following "new" assumptions:
- Information seeking is only one part of the story
- There may be no explicit query; the user and his/her online and offline presence are the query
- The content being searched is neither stable nor bounded
- Emergence can be as important as intent
- Item-based evaluation may concern aboutness ("what"), framing ("how") as well as reception

- SERP evaluation criteria may include different types of diversity (content, perspective, speed, etc)

Interacting information spaces

Social media does not exist in isolation; people use social media to react to content they find in other media, and other media react in some ways to the activity on social media. Some of these interactions include:

- Social media and the research literature
- Social media and journalism
- Social media and popular culture
- Social media and real life (hybrids of the online and offline worlds)
- Social signals for other IR tasks
- Social media as one lens in an Internet-scale social science "observatory"

Value proposition

Similar to a patient that needs to go to the hospital, we can summarize the main points as follows:

- Know when to go (when to use social media)
- Understand what they say (aggregate and summarize what the "crowd" is saying)
- Learn what they can't tell you (which kind of expertise/knowledge social media can produce)
- Construct strength from adversity (re-construct a story, extract different perspectives)
- Inform their decision making (summarize findings, perceived utility)

4.2.5 Building Bridges

Social media touches much of what we have discussed here at Dagstuhl. Here are some relationships to other discussion groups that formed:

- Task-based retrieval – grounds what we are doing.
- Search as learning – from each other
- Our cultural heritage – is what social media is constructing
- Reliability and validity – are what make our research relevant
- Modeling users – not a single user but lots of users.

4.2.6 Conclusions

This summary report suggest some ways of thinking about social data in the context of IR, and the potential advantage of doing so. We have presented a characterization of tasks and components, identified issues with this type of content, and shed some light on opportunities moving forward.

References

1 O. Alonso, C. Marshall & M. Najork. Are some tweets more interesting than others? #HardQuestion. HCIR, 2013.
2 R. Baeza-Yates. Searching the Future. *SIGIR MF/IR Workshop*, 2005.
3 J. Cook, K. Kenthapadi & N. Mishra. Group chats on Twitter. *WWW*, 225-236, 2013.
4 L. Dantonio, S. Makri, & A. Blandford. Coming across academic social media content serendipitously. *ASIST*, 49(1): 1-10, 2012.
5 M. Efron. Information search and retrieval in microblogs. *JASIST 62(6)*: 996-1008, 2011.

6 S. Erdelez. Investigation of information encountering in the controlled research environment. *IPM 40(6)*, 1013-1025, 2004.
7 J. Hurlock & M. Wilson. Searching Twitter: Separating the Tweet from the Chaff. *ICWSM*, 2011.
8 G. Mishne & M. de Rijke. A study of blog search. *ECIR*, 289-301, 2006
9 P. Mutschke & M. Thamm. Linking social networking sites to scholarly information portals by ScholarLib. *Web Science 2012*, 315-320, 2012.
10 S. Petrovic, M. Osborne & V. Lavrenko. Streaming first story detection with application to Twitter. *HLT-NAACL*, 181-189, 2010.

4.3 Graph Search and Beyond

Omar Alonso (Microsoft Research – Mountain View, US), Jaap Kamps (University of Amsterdam, NL)

4.3.1 Motivation

Information on the Web is increasingly structured in terms of entities and relations from large knowledge resources, geo-temporal references and social network structure, resulting in a massive multidimensional graph. This graph essentially unifies both the searcher and the information resources that played a fundamentally different role in traditional IR, and offers major new ways to access relevant information. You are the query.

Graph search affects both query formulation as well as result exploration and discovery. On the one hand, it allows for incrementally expressing complex information needs that triangulate information about multiple entities or entity types, relations between those entities, with various filters on geo-temporal constraints or the sources of information used (or ignored), and taking into account the rich profile and context information of the searcher (and his/her peers, and peers of peers, etc). On the other hand, it allows for more powerful ways to explore the results from various aspects and viewpoints, by slicing and dicing the information using the graph structure, and using the same structure for explaining why results are retrieved or recommended, and by whom.

This new graph based approach introduces great opportunities, but also great challenges, both technical ranging from data quality and data integration to user interface design, as well as ethical challenges in terms of privacy; transparency, bias and control; and avoiding the so-called filter bubbles. The best examples at the time of writing are Facebook Graph Search and related efforts at Bing, Google and other commercial search engines. Similar approaches can be applied to other highly structured data, just to give an example, the hansards or parliamentary proceedings are fully public data with a clear graph structure linking every speech to the respective speaker, their role in parliament and their political party.

4.3.2 Issues

We view the notion of "graph search" as searching information from your personal point of view (you are the query), over a highly structured and curated information space. This goes beyond the traditional two-term queries and ten blue links results that users are familiar, requiring a highly interactive session covering both query formulation and result exploration.

Two Step Interaction

Incremental Structured Query Input: Creating a graph query requires incremental construction of a complex query using a variety of building blocks. Current search engines treat this as a form of query suggestion or query completion, which offers tailored suggestions trying to promote longer queries that cover multiple entity types and relations and various filters. Suggestions and entity types may be based on the user's own activity. This goes beyond prevailing autocompletion techniques, with previews and surrogates from traditional result pages or SERPs (Search Engine Results Page) moving to a more dynamic query suggestion.

Dynamic Structured Result Set Exploration: Results are highly personalized: they are unique for the searcher at a given point in time. The result set is highly structured: rather than just showing the top-10 results from an almost infinite list, a faceted exploration based on your interests is needed. The structure is dynamically derived from the graph structure and the user's point of view, rather than a rigid facet and facet value hierarchy.

When to Use Graph Search?

Rather than a universal solution, the graph search is particularly useful for specific types of information needs and queries. This is also depending on the character of the data available. E.g., Facebook Graph Search emphasizes the social network structure, friends and other persons, locations and location-tagged objects. Social network data is abundantly available (although getting access presents a major barrier) but also notoriously skewed. Rather the searcher personal point of view, it can also be used to show results from the viewpoint of any person in the network. There are many interesting sets of data – both historically or modern – that capture both the persons and related information: think of parliamentary data in public government, or intranet data in organizations.

Query Classification

Graph search also requires a new query classification, beyond the traditional division into navigational, informational, and transactional queries. Is there a new way to characterize queries in this new model? Does the notion of information need change? It is the ultimate form of personalization, with the searcher not only responsible for the query but also determining the (slice of) the data being considered. What shifts in control and transparency are needed to accomplish this?

Graph Search Evaluation

This also presents a range of new evaluation problems. How to evaluate the overall process, given its personalized and interactive nature? How to evaluate the first stage as essentially a form of query autocomplete? And how to evaluate the second stage as to explore and exploit the result set?

4.3.3 Methods

Graph search requires a highly interactive session covering both query formulation and result exploration.

Query Exploration

There is a radical shift towards the control of the searcher, necessitating new tools that help a searcher construct the appropriate graph search query, and actively suggest refinements or filters to better articulate their needs, or explore further aspects. This leads to a far more dynamic interaction than with traditional result lists, or modern hit lists showing summaries of a static set of results.

This suggests a new form of "query autocomplete" that invites and allows users to issue longer queries constructed based on entities, relationships, and templates. In constrast to SERP, we define IQEP as the Incremental Query Exploration Page. IQEP allows the user to explore more the result set as part of the input query. We can think of IQEP as an interactive mechanism that promotes relevant results selected by the user from the traditional SERP to the input box. Figure 1 shows IQEP as a bi-directional channel that moves results from the search list to the input box or viceversa.

There are a range of suitable evaluation methods. The obvious way is by direct evaluation of query suggestion, query recommendations (are they any good?). There is also a range of criteria useful for behavioral observation for in the wild testing: users should issue longer queries, multiple filters, dwell-time, active engagement, structured-query templates. There are query segments where this type of querying is expected to be most useful: torso and tail queries; exploratory scenarios. Traditional head or navigational queries seem less interesting, although these could be part of a more complex underlying information need.

This goes beyond Broder's taxonomy: queries are all navigational, informational, and transactional but they are entity-focused. Queries may aim to return a single or a small set (not unlike traditional Boolean querying over structured data), or there is a need for data analytics on the whole set of results.

Result Exploration

There is a radical shift towards the control of the searcher – small changes in the query can lead to radically different result sets – necessitating active exploration of slices of the data to explore further aspects.

This suggest a new form of search results unique for every user. Similarly to the query exploration mechanism, this interaction encourages users to explore over entities, relationships, and filters. Unlike traditional facetted search options, the result space is highly dynamic, and requires adaptive exploration options tailored to the context and searcher, at every stage of the process.

This is a radical departure from the traditional "ten blue links". The IQEP moves from links to answers, and from answers to suggesting (expressions of) needs. This is an complete shift from the traditional dichotomy between query (the searcher's responsibility) and results (the system's responsibility). Traditional search results have moved to a hit list of result summaries (still a fix set of results, but the shown summaries are tailored to the searcher and her query). These summaries in terms of entities are now answers rather than links to answers. Now these results, or previews of them, are moving into the search box, in the form of structured query suggestions with some sort of preview indicating of the consequences on the result set (often in terms of numbers of results, or entity previews).

There are many options for the evaluation of components: (adaptive) captioning, (adaptive) filters, graph query templates. E.g., captioning should describe (relative to the entity), explain (relative to the user), and be contrastive (relative to the IQEP). There are standard experimental evaluation methods from HCI and UI/UX design. With a running

service, evaluation in the wild is very suitable. There are various implicit and explicit criteria: users should explore the result set, usage of multiple filters, dwell-time, active engagement, structured-query templates. Torso and tail queries, and exploratory scenarios are the most suitable query segments.

4.3.4 Conclusions

Graph search gives amazing power, and unleashes the potential of semantically annotated information with many entities, and relations between entities. It brings the control back to the searcher. Graph-based search systems also have the potential to solve part of the old IR problem of conceptual search.

In terms of IR research and required evaluation methods, as discussed in the sections above, there are some open problems. What we need is to work on sharable research data, that exemplifies most of the characteristics we want to study. There is no need to be on Facebook or Twitter, or hand over your personal data. Similar small data sets and systems are available (e.g., so.cl, NYT, Parliamentary data, etc.) It will be hard to share a realistic subset of social network data (unless there are enough volunteers?) but we could work on a simulated set. What would be a concrete task to study on this data? Instead of implementing all features, it is would be useful to select a few components like query suggestion box, filters as facets, and captions to show the potential.

Search engine user interfaces has been very stable in the last 15 years. The input box and the 10 blue links are the still the most optimal way to show search results. Can we do better in terms of user experience? This would be give users a lot of flexibility and options. However, remains to be seen if users would adopt such dynamic interface.

At a high level, graph search seems limited to familiar entity types (e.g. Facebook entities) and templates. How far can this scale? Will this work on truly open domains? Finally, there are a number of ethical issues such as privacy, transparency, bias and control, and filter bubbles.

4.4 Reliability and Validity – A Guide To Best Practices in IR Evaluation

Nicola Ferro (University of Padova, IT), Hideo Joho (University of Tsukuba, JP), Diane Kelly (University of North Carolina – Chapel Hill, US), Dirk Lewandowski (HAW – Hamburg, DE), Christina Lioma (University of Copenhagen, DK), Heather O'Brien (University of British Columbia – Vancouver, CA), Martin Potthast (Bauhaus–Universität Weimar, DE), C.J. Keith van Rijsbergen (University of Cambridge, GB), Paul Thomas (CSIRO – Canberra, AU), Vu Tran (Universität Duisburg–Essen, DE), Arjen P. de Vries (CWI – Amsterdam, NL)

4.4.1 Motivation

Experimental evaluation is one of the backbones of the information retrieval field since its inception. Over the years, it provided both qualitative and quantitative evidence as to which methods, algorithms, and techniques are more effective. Moreover, due to its early and

systematic adoption of strong evaluation methodologies, the IR community is often regarded as "leading" in this respect by computer science people but there still many open questions.

Indeed, carrying out thorough experiments is a challenging activity where many "traps" are hidden. For example, there is increasing discussion in the research community about reproducibility and generalizability of our experiments, as it may be difficult to re-use research, methods, measures, data and results.

Moreover, people in IR come from different backgrounds and there is a need to consolidate ideas/expertise from different fields and to establish a common ground around some key concepts (reliability, validity, ...) as well as an understanding of their trade-offs and design decisions.

Finally, a better support for students is needed in order to avoid them to learn best practices in a very fragmented and sometimes inconsistent way, not to say the risk of adopting approaches which have been discarded with the passing of time due to lack of robustness.

Therefore, there is an overall need for a reasoned guide to best practices for IR evaluation which will turn around the two key concepts: reliability is the extent to which a [measure/-method] produces similar results under stated conditions for a stated period of time [inspired by ISO 9126]; and, validity is the extent to which a [measure/method] accurately reflects the phenomena it is intended to reflect.

4.4.2 Goals and Scope

The proposed best practices have the following goals:

- to produce research results with confidence: for communicating with the research and stakeholder community; for assessing their impact, longevity, and generalizability;
- to gain an appreciation of the "trade-offs" and limitations inherent in our studies;
- to encourage good practices for novices and experts;
- to enable/promote repeatability/reproducibility of the experiments.

The proposed best practices have the following scope:

- to understand whether an IR/IIR experiment is valid and reliable, including design, carrying out, analysis, and results presentation;
- to better communicate results
- to make the context explicit: types of methods for IR/IIR evaluation and beyond; experiments as our way to evaluate (lab, insitu, crowdsourcing, log analysis, ...); the kind of context itself (IR context, Seeking context, ...).

The target audience of the proposed best practices are:

- graduate students;
- PhD students;
- researchers (and reviewers).

4.4.3 Structure of the Best Practices

The proposed best practices will be structured as follows:

- Pillar Definitions and Concepts: starting from the definitions of reliability and validity provided above, we will explore and detail them for different methods and measures as well as provide example of factors that demonstrate and/or enhance reliability/validity of measures and methods;

- Reliability and Validity in Experimentation: we will discuss how to ensure reliability and validity in carry out actual experiments, by covering hypotheses, sampling, methods, environments, data analysis, measurement and procedures, as well as pointing at other issues such as ethical/legal issues and privacy and intellectual property rights.
- Reporting Out Experiments: we will discuss how to present experimental results, their limitation, to acknowledge alternative interpretations of the results, to report data analysis as well as verifiable/falsifiable outcomes and we will deal also with archiving and infrastructures for experimental data management (data curation).
- Failure analysis: this is deemed one of the most important activities to actually understand how and why a system behaves differently than expected and why it fails to achieve the desired performances. Unfortunately, this is an extremely demanding activity in terms of time and effort needed to carry it out.
- Definition of new measures/methods: we will cover the steps and the process needed for establishing and motivating a new measure and its trade-offs as well as the checks to ensure its reliability and validity.
- For reviewers: we will look at the previously introduced concept and best practices from the angle of reviewers in order to support them in effectively and fairly reviewing papers reporting experimental data.

References

1 M. H. Birnbaum. Psychological Experiments on the Internet. *Academic Press*, USA, 2000.
2 D. K. Harman and C. Buckley. Overview of the Reliable Information Access Workshop. *Information Retrieval, 12(6)*, pp. 615-641, 2009.
3 D. K. Harman. Information Retrieval Evaluation. *Morgan & Claypool Publishers*, USA, 2011.
4 D. K. Harman and E. M. Voorhees. TREC. Experiment and Evaluation in Information Retrieval. *MIT Press*, USA, 2005.
5 K. Hornbaek: Some Whys and Hows of Experiments in Human-Computer Interaction. *Foundations and Trends in Human-Computer Interaction (FnTHCI), 5(4)*, pp. 299-373, 2013.
6 ISO/IEC 25010:2011. Systems and software engineering – *Systems and software Quality Requirements and Evaluation (SQuaRE)* – System and software quality models.
7 K. Järvelin: An analysis of two approaches in information retrieval: From frameworks to study designs. *JASIST 58(7)*, pp. 971-986, 2007.
8 D. Kelly. Methods for evaluating interactive information retrieval systems with users. *Foundations and Trends in Information Retrieval (FnTIR)*, 3(1-2), pp. 1-224, 2009.
9 F. W. Lancaster. Information Retrieval Systems: Characteristics, Testing and Evaluation. *John Wiley & Sons Inc; 2nd edition*, 1979.
10 H. L. O'Brien, E. G. Toms: The development and evaluation of a survey to measure user engagement. *JASIST 61(1)*, pp. 50-69, 2010.
11 M. Sanderson. Test Collection Based Evaluation of Information Retrieval Systems. *Foundations and Trends in Information Retrieval (FnTIR)*, 4(4), pp. 247-375, 2010.
12 K. Spärck Jones. Information Retrieval Experiment. *Butterworths*, UK, 1981.
13 P. E. Spector. Summated rating scale construction: An introduction. *Sage Publication*, USA, 1992
14 J. Zobel. Writing for Computer Science. *Springer*, Germany, 2004.

4.5 Domain Specific Information Retrieval

Maristella Agosti (University of Padova, IT), Floriana Esposito (University of Bari, IT), Ragnar Nordlie (Oslo University College, NO), Vivien Petras (HU Berlin, DE), Christa Womser–Hacker (Universität Hildesheim, DE)

The working group domain-specific information retrieval met for one day. After defining domain-specific information retrieval and information systems, the focus of the group was directed at discussing information retrieval and evaluation issues in the domains that were relevant for their use cases (cultural heritage, patent retrieval).

4.5.1 Definition Domain Specific Information System

Domain specific information systems collect, store, preserve, organize, search and display domain specific objects or their (metadata) representations in a digital environment. Good examples of domain of interest are: cultural heritage, patents, and medical collections.

4.5.2 Motivation

For the domains of interest, there may be a challenge to manage collections of documents that the user wants to interact with not only through a query function. This means that some specific features of the domain need to be taken into account when envisaging a system that has to manage the document collection. This also has consequences for information retrieval evaluation.

4.5.3 Summary of Challenges

For the users of some kinds of domain specific information systems to start searching the system – i.e. starting the interaction with an information system through a query – may not always be the optimal mode of access. Domain specific information systems have responded by providing exploratory interaction functionalities like curated digital exhibitions, featured objects, or user-provided stories to present alternative starting options to the user. Other than studying whether users "liked" these features, IR evaluation has not progressed towards a formalized mode of evaluation that would allow comparing the "usefulness" of these features in different applications with respect to the goals of the system (or the user). These efforts have not succeeded in recommendation system improvements based on evaluation. In a domain specific information system, a significant facility will be to provide context for the represented objects, either through links to other objects in the managed collection or through associated text that can be user-provided or producer-provided. Evaluating the quality of this context is a challenge which may be met in different ways: through establishing some measure of semantic similarity between the object and the context, or measuring some degree of user satisfaction and relevance judgment of the context provided. What is needed can only be decided by looking at the outcome of the interactive process with a strong priority on the experts' involvement. For example Patent people are very conservative and want to continue their work e.g. with Boolean systems. If we stay with Boolean systems no progress will be possible. We – as IR scientists – have to convince them that there are new innovative approaches. Automatically judging the quality of retrieval functions based on observable user behavior could allow for making retrieval evaluation faster, cheaper, and more user centered.

However, the relationship between observable user behavior and retrieval quality is not yet completely investigated.

4.5.4 The No-search / Exploratory Access Evaluation Problem in Cultural Heritage

Cultural heritage user types can possibly be divided into two groups: humanities scholars utilizing cultural heritage information systems for their research and information "tourists" utilizing cultural heritage information systems to get informed or be entertained about or by cultural artefacts. IR evaluation has traditionally been focused on users searching (i.e. more or less targeted querying) a predetermined document pool. For information tourists user types in cultural heritage information systems, search – i.e. starting the interaction with an information system through a query – is probably not the optimal mode of access, because they (a) don't know what content the system provides and (b) often do not have a specific information need in mind that can be translated into a query. Cultural heritage information systems have responded by providing exploratory interaction functionalities like curated digital exhibitions, featured objects, or user-provided stories to present alternative starting options to the user. Other than studying whether users "liked" these features, IR evaluation has not progressed towards a formalized mode of evaluation that would allow comparing the "usefulness" of these features in different applications with respect to the goals of the system (or the user). These efforts have not succeeded in recommendation system improvements based on evaluation. One challenge for IR evaluation in cultural heritage information systems is therefore to develop evaluation scenarios that do not have the conventional query-output (maybe iteration thereof) process in mind, but alternative exploratory options (which might lead to a retrieval-based outcome nevertheless). This would consequently also require the development of new assessment approaches and the creation or adaptation of appropriate measures.

4.5.5 Contextualization and Evaluation of Context in Cultural Heritage

The information retrieval systems for cultural heritage are embedded in a rich context. Documents no longer exist on their own; they are connected to other documents, they are associated with users and they can be mapped onto a variety of ontologies. Retrieval tasks are interactive and are solidly embedded in a user's social and historical context. New challenges in information retrieval will not come from smarter algorithms that better exploit existing information sources, but from new retrieval algorithms that can intelligently use and combine new sources of contextual metadata. Machine learning methods (multirelational learning) could be used: - to automatically create the markup or metadata for existing unstructured documents - to create, merge, update, and maintain ontologies. In a cultural heritage system, a significant facility will be to provide context for the represented objects, either through links to other objects in the database or through associated text, user-provided or producer-provided. Evaluating the quality of this context is a challenge which may be met in different ways: through establishing some measure of semantic similarity between the object and the context, or measuring some degree of user satisfaction and relevance judgment of the context provided. The system's ability to limit or extend the amount of context, encourage the pursuit of context etc should also be evaluated.

4.5.6 Innovative Applications vs. Traditional Values in Patent Retrieval

We do not have fully automatic systems that could perform patent retrieval successfully, yet. Patent retrieval is per se an interactive task sharing human and system intelligence. Since we have more than 10 million patent applications p.a., the need for a solution is very important. On the other hand, we have many system approaches delivering solutions for specific sub-tasks. Here, the need for evaluation steps in. How can proposed solutions be evaluated? What is needed can only be decided by looking at the outcome of the interactive process with a strong priority on the experts' involvement. Patent searchers are very conservative and want to continue their work, e.g. with Boolean systems. If we stay with Boolean systems, no progress will be possible. We – as IR scientists – have to convince them that there are new innovative approaches. How? By showing them that they do better work in shorter time in the interactive scenario. However, control should stay with the experts.

4.5.7 Automatic Observation of User Behavior

Automatically judging the quality of retrieval functions based on observable user behavior could allow for making retrieval evaluation faster, cheaper, and more user centered. However, the relationship between observable user behavior and retrieval quality is not yet completely investigated. A paper studying this relationship for a search engine operating on the arXiv.org e-print archive has shown that none of the eight well known absolute usage metrics (e.g., number of clicks, frequency of query reformulations, abandonment) reliably reflect retrieval quality for the considered sample. Learning techniques have been applied in information retrieval (IR) applications generally for information extraction, relevance feedback, information filtering, text classification and text clustering. Recently, online learning models have been proposed for interactive IR with the aim of providing results of maximum utility to the user. The interaction between human and system takes the following form. The user issues a command (e.g. query) and receives a result in response (ranking). The user then interacts with the results (clicks), thereby providing implicit feedback about the user's utility function. Using online learning models (for example, coactive learning algorithms), the feedback can be inferred from observable user behavior from clicks during-search. In each iteration, a user, drawn from an unknown but fixed distribution, presents a context (e.g., query) to the system and receives a ranking in response. The user is represented by a utility function that determines the actions (e.g. clicks) and therefore the feedback to the learning algorithm. The same utility function also determines the value of the presented ranking. The goal is to learn a ranking function that has high social utility, which is the expected utility over the user distribution.

4.5.8 Support of Exploration and Content Enriching Tools

A domain specific information system has to support the exploration of the managed collection, so it needs to support traditional search-based exploration, but also has to move beyond it. It can support a range of innovative normalization and natural language processing technologies that allow entities and relationships to be extracted from the collection and visualized using a range of specially designed visualizations. A domain specific information system has also to provide for entity oriented search, and allows users to crosswalk from one tool to another, ensuring that their exploration of the collection is flexibly supported. The system has also to provide a comprehensive set of bookmarking, and annotating tools that make it a powerful aid to both extensive and intensive work on content collections.

References

1 M. Agosti, M. Manfioletti, N. Orio, C. Ponchia. Evaluating the Deployment of a Collection of Images in the CULTURA Environment. In: *Proc. of the Int. Conf. on Theory and Practice of Digital Libraries (TPDL 2013)*, Valletta, Malta. LNCS Vol. 8092, Springer, Berlin, 2013, pp. 180–191

2 F. Esposito, C. d'Amato, N. Fanizzi. Fuzzy Clustering for Semantic Knowledge Bases. *Fundamenta Informaticae*, Vol. 99 2/2010, pp. 187–205

3 N. Fanizzi, C. d'Amato, F. Esposito. Inductive Classification of Semantically Annotated Resources Through Reduced Coulomb Energy Networks. *International Journal On Semantic Web And Information Systems*, 2009, Vol. 5(4), pp. 19–38

4 N. Fanizzi, C. d'Amato, F. Esposito. Metric-Based Stochastic Conceptual Clustering For Ontologies. *Information Systems*, 2009, Vol. 34(8), pp. 792–806

5 N. Ferro, R. Berendsen, A. Hanbury, M. Lupu, V. Petras, M. de Rijke, G. Silvello. Context Evaluation. In: *PROMISE Retreat Report. Prospects and Opportunities for Information Access Evaluation*. Brainstorming workshop held on May 30–31, 2012, Padua, Italy

6 J. Jürgens, P. Hansen, C. Womser-Hacker. Going beyond CLEF-IP: The "Reality" for Patent Searchers? In: Catarci, T.; Forner, P.; Hiemstra, D.; Peñas, A.; Santucci, G. (Eds.), *Information Access Evaluation. Multilinguality, Multimodality, and Visual Analytics Third International Conference of the CLEF Initiative – CLEF 2012*, Rome, Italy, September 2012, pp. 30–35

7 S. Lawless, M. Agosti, P. Clough, O. Conlan. Exploration, navigation and retrieval of information in cultural heritage: ENRICH 2013. In: *Proc. of the 36th ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2013)*. ACM, New York, USA, p. 1136

8 V. Petras, J. Stiller, M. Gäde. Building for Success (?) - Evaluating Digital Libraries in the Cultural Heritage Domain. In: Cool, C., NG, K. B (Eds.), *Recent Developments in the Design, Construction and Evaluation of Digital Libraries: Case Studies*, ICI Global, 2013

9 M. Sweetnam, M. O. Siochru, M. Agosti, M. Manfioletti, N. Orio, C. Ponchia. Stereotype or Spectrum: Designing for a User Continuum. In: *Proc. of the Workshop of the 36th ACM SIGIR Conference on Exploration, navigation and retrieval of information in cultural heritage (ENRICH 2013)*, Dublin, 2013, pp. 23–30

4.6 Task–Based Information Retrieval

Nick Belkin (Rutgers University – New Brunswick, US), Kalervo Järvelin (University of Tampere, FI), Evangelos Kanoulas (Google Switzerland – Zürich, CH), Birger Larsen (Aalborg University Copenhagen, DK), Thomas Mandl (Universität Hildesheim, DE), Elaine Toms (Sheffield University, GB), Pertti Vakkari (University of Tampere, FI)

4.6.1 Core concepts & definitions

What is meant by task-based analysis of IR?

There are several possible answers to this question. One answer is to focus on the task or goal that motivates a person to engage in information seeking in an IR system. Examples of motivating tasks or goals are work tasks, hobbies, everyday life tasks and leisure time interests. From this point of view, task-based analysis of IR means understanding and responding to

motivating tasks and goals, and designing IR systems which can support accomplishment of a variety of such tasks and goals. This answer requires that an IR system be able to recognize different tasks. Another possible answer is understanding the nature of the task or tasks of specific groups, and the design of IR systems which are tailored to support those people as they are engaged in those specific types of tasks. This answer requires the design of many different IR systems.

Task types

We use the following categorization of task for IR purposes:

- Motivating tasks, sometimes called "work tasks", accomplishment of which have led the person to engage in an IR system. These may lead to more than one information seeking or information search session.
- Seeking tasks. These involve deciding where, or with whom to engage in order to obtain information which will be useful in accomplishing the motivating task. This can be from a variety of sources and systems, may include several sessions over time, but may also involve only other people or a combination of both people and systems.
- Search tasks. These are the tasks which a person engages in during an information seeking session, trying to accomplish their intentions while using an IR system. These will involve one or more sequences of behaviors over a search session. Examples of such tasks are formulating a query, learning about a domain, comparing search results, judging usefulness or relevance of search results, etc.

4.6.2 How is task-based IR different from traditional IR, and what does this mean for evaluation of IR system performance?

Traditional IR evaluates performance according the system's response to a single query. However, motivating tasks typically generate several information-seeking intentions, leading to multiple sessions and multiple search tasks within sessions. Task-based IR studies IR as a process, with sequences of behaviors associated with different search tasks during the course of a search session, or over multiple sessions. Evaluation of IR system performance, from the point of view of task-based IR, must be based on some description of what it would take to accomplish the "work task" (i.e. the motivating task context), and of how accomplishment of that task could be measured. Then the IR system's support techniques, and measures for their evaluation, must be justified according to hypotheses about how those techniques will support accomplishment of the motivating task, and of the search tasks, and to how the measures reflect such accomplishment.

4.6.3 Steps toward task-based evaluation of IR system performance

We suggest that the following issues need to be addressed in a coordinated research program, to provide the basis for being able to perform evaluation from the task-based IR point of view

Persons

Minimally, we need to learn more about motivations for engaging in information seeking behavior, about the effects of people's knowledge of task, topic and system, of stage in task accomplishment, and of individual and cultural differences on behaviors and intentions in

search sessions. We need to expand upon or perhaps integrate the different classifications of task types that have been proposed by various researchers.

Process of task-based IR

We need a better understanding of the intentions of search tasks, and especially of sequences of search tasks during search sessions of different types. A start toward this goal would be a series of studies, both lab and live, which both observe search behaviors, and elicit search intentions related to the different behaviors.

Outputs and outcomes

Outputs are the products delivered by the system, outcomes are the benefits for the user produced by the system, e.g. task accomplishment. Outputs as well as outcomes are highly task-dependent and may require different measures for evaluating system performance.

4.6.4 Possible actions

There are three overall issues which task-based IR needs to address in order to move toward appropriate evaluation. The first is mapping the territory of task-based IR and identifying thereby areas where knowledge is shallow or nonexistent (white patches). One should also check the borders (or outline) of the map. The second is developing a research program that seeks to systematically analyze the connections of motivating task features, search task features, search behavior features, IR system features, output features and outcome features. And the third is to develop study designs which include variables from larger tasks, consequent search tasks including search processes, outputs and outcomes, and system features. Two kinds of designs are needed: field studies and experimental designs. Due to the complexity of the phenomena investigated, field studies can be used to reveal mechanisms connecting larger tasks with search tasks and consequent search processes and outcomes. These results should provide information for designing evaluation experiments and systems. Experimental designs should identify system features that can most strongly be expected to have a connection to the motivating task features and then examine that connection.

4.7 Searching for fun

Ragnar Nordlie (Oslo University College, NO), Ann Blandford (University College London, GB), Floriana Esposito (University of Bari, IT), Douglas W. Oard (University of Maryland – College Park, US), Vivien Petras (HU Berlin, DE), Max L. Wilson (University of Nottingham, GB), Christa Womser–Hacker (Universität Hildesheim, DE)

4.7.1 Definition

"Searching for fun" might have the double sense of "searching for something that is fun" or "having fun while searching". Our discussion was concerned with the second sense: the activity of interacting with an information system without having a specific search objective in mind. For short, it may be called FII: fun information interaction. This may involve activities such as: online window shopping with nothing to buy, reading online, like reading

fiction or the news, watching funny videos or finding funny pictures. It may perhaps also include examples of pursuing more traditional information needs, as in situations where not finding a result is no great concern. Even a traditional search process may waste time, pique interest, be fun.

4.7.2 Motivation

Many information systems are constructed, at least partly, with the objective of inducing the user to interact with the system without a predefined purpose, and to retain this user in interaction; to encourage unexpected discovery, to encourage a certain kind of learning, to support a certain kind of shopping behavior, to expose the user to advertising, or for a number of other reasons. System users frequently engage, also for a number of reasons, in this kind of non-intentional interaction. Evaluation of this kind of system or this kind of system activity calls for different evaluation criteria and measures from those employed for goal-directed information retrieval.

4.7.3 Prior work

Related, but maybe not identical, issues have been discussed at recent workshops, particularly:

- 'Entertain Me' workshop on supporting complex search tasks at SIGIR 2011; followed up by a "contextual suggestion" track at TREC 2012 and 2013
- Searching4Fun workshop at ECIR2012

Despite these efforts, we feel that the evaluation challenges presented by non-outcome-focused interaction have not been exhaustively discussed. This report is a first effort to address these challenges. Discussions How does FII differ from traditional IR evaluation? It changes our assumptions about searching (and browsing, and whatever other activities involve finding things). This changes our criteria, and thus our interpretations of measures. Instances of this include:

- Stopping behavior: Stopping may mean running out of things to find; finding a good result, may be reason to continue (not to stop, as in IR in general);
- Time spent: More time can be good;
- Novelty: Novelty and Repetition might be equally important.

What criteria may be applied to the evaluation of FII? Evaluation of FII may be considered from a system or a user perspective. The system's motivation may be expressed in a simplified manner as "get them in – keep them in – convert entry into experience (learning, shopping, entertainment...)". The user motivation may be to be entertained, to spend time, to make unexpected discoveries...For the user evaluation a number of criteria may identify one or more of these experiences; we discussed, among others: engagement, flow, cognitive load, stimulation, currentness, social engagement, novelty, sensemaking, meaningmaking (contextualization), outcome state, state change, user empowerment. We provisionally concluded that these criteria can be comprised in a process criterion: engagement, and an outcome criterion: state change. A good engagement level for the user involves, for instance, avoiding bad disengagement, avoiding over-engagement. State change may imply changes such as bored to not bored, stressed to not stressed, sad to happy, or changes via a transforming state, such as stressed to relaxed via horrified and surprised. It is difficult to design study conditions for systematic measurements of these criteria. Casual or leisure needs are by nature intrinsic and hard to create under experimental conditions. How can we make participants bored or stressed or sad so that they will naturally entertain themselves,

try to relax or be amused? How can we measure engagement when, for instance, a measure like time spent may be either positive or negative depending on the circumstances? Rather than define measures, we identify some open questions regarding FII IR, which may influence the choice of measures, such as

- Is there more to FII than just the distinction between visceral and conscious needs (Taylor, 1962)?
- How does FII relate to things like serendipity?
- Are there gaming measures that are relevant?
- Can we have FII within Serious and Project leisure (Stebbins, 2009)?
- Can we optimize systems for FII behavior?
- Can we detect certain state-change targets (bored to not-bored, stressed to relaxed, etc)?
- How do different demographics differ?
- In what way is the journey more important than the objects found?

What are the current challenges for IR evaluation of FII? We developed the following (incomplete and unordered) list of challenges for FII evaluation:

- Actually studying fun information interaction in action
- Discovering fruitful scenarios/contexts
- Identifying successful FII strategies (If there are strategies for this?)
- Correlating system interactions with study findings
- Determining measures for Fun Information Interaction
- Designing simulated user interaction models that relate to FII
- Create systems that increase engagement
- Identify ways systems can support FII

4.7.4 Actions

Some time has passed since the Searching4Fun workshop, which in itself did not focus primarily on evaluation methods. There might be scope for a new workshop, for instance at IIiX 2014 in Regensburg.

References

1 Ilaria Bordino, Yelena Mejova & Mounia Lalmas (2013). Penguins in Sweaters, or Serendipitous Entity Search on User-Generated Content, CIKM.

2 Mihaly Csikszentmihalyi (1990). Flow: The Psychology of Optimal Experience, Harper Collins.

3 Heather O'Brien (2010). The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. Interacting with Computers, 22(5), pp. 344–352. *Psyclopedia, Effort Recovery Model.* http://www.psych-it.com.au/Psychlopedia/article.asp?id=356

4 Kateriina Saarinen & Pertti Vakkari (2013). A sign of a good book: readers's methods of accessing fiction in the public library. *Journal of Documentation*, 69(5):736–754.

5 Robert Stebbins (2009). Leisure and Consumption: Common Ground, Separate World, Palgrave Macmillan.

6 Robert Taylor (1962). The Process of Asking Questions. *American Documentation*, 13(3):391–396.

4.8 The Significance of Search, Support for Complex Tasks, and Searcher–aware Information Access Systems

Jaap Kamps (University of Amsterdam, NL)

This abstract documents three loosely related challenges. The first challenge is the role and significance of the field in general. There are massive challenges in the way the information available is changing in quantity and in character, and in the ways we create, publish, share, and use information in the always-online world. This urges us to keep 'reinventing search' and redefine the field of information retrieval and its key research problems and research methods. How do these changes affect the core questions we address in the field of IR and what sort of evidence do we need for addressing these questions? How can we factor the larger scope and context into IR evaluation? It is interesting to consider a publication like Salton's "Developments in automatic text retrieval" published in Science in 1991. Salton (1991) is from before the Web happened and discusses all the basic IR aspects: retrieval models, indexing structures, but also hypertext, knowledge resources and semantic search. Articles like Salton (1991) still look surprisingly modern! This raises two question that are perhaps not unrelated: First, why hasn't our research field changed in a dramatic way to suit the revolutionary changes in the information environment. Second, why isn't our field making a larger impact outside our field (Salton published 2 Science articles in 1991) given the dramatic increased role and importance of "search" nowadays.

The second challenge is to work on information access tools that support complex tasks. That is, to build and evaluate information access tools that actively supports a searcher to articulate a whole search task, and to interactively explore the results of every stage of the process. In the prolonged search session, how should we evaluate the overall effectiveness as well as the success at various stages? How can evaluation reflect the different goals of each stage? There is a striking difference in how we ask a person for information, giving context and articulating what we want and why, and how we communicate with current search engines. Current search technology requires us to slice-and-dice our problem into several queries and sub-queries, and laboriously combine the answers post hoc to solve our tasks. Combining different sources requires opening multiple windows or tabs, and cutting-and-pasting information between them. Current search engines may have reached a local optimum for answering micro information needs with lighting speed. Supporting the overall task opens up new ways to significantly advance our information access tools, by develop tools that are adapted to our overall tasks rather than have searchers adapt their search tactics to the "things that work."

The third challenge is to make information access systems more informed about the searcher. Can we make a retrieval system aware of the searcher's stage in the information seeking process, tailor the results to each stage, and guide the searcher through the overall process? How to evaluate the utility of this (accuracy of the prediction, usefulness of the support, etc)? Can we equate evaluation with observing preferred information interaction patterns? A search session for a non- trivial search task consists of stages with different sub-goals (e.g., problem identification) and specific search tactics (e.g., reading introductory texts, familiarizing with terminology). Making a system aware of a searcher's information seeking stage has the potential to significantly improve the search experience. Searchers are stimulated to actively engage with the material, to get a grasp on the information need and articulate effective queries, to critically evaluate retrieved results, and to construct a

comprehensive answer. This may be of particularly great help for those searchers having poor information or media literacy. This is of obvious importance in many situations: e.g., education, medical information, and search for topics "that matter". Some special domains, such as patent search and evidence based practices in medicine, have clearly prescribed a particular information seeking process in great detail. Here building a systems to support (and enforce) this process is of obvious value.

4.9 Interaction, Measures and Models

Gianmaria Silvello (University of Padova, IT), Leif Azzopardi (University of Glasgow, GB), Charlie Clarke (University of Waterloo, CA), Matthias Hagen (Bauhaus–Universität Weimar, DE), Robert Villa (University of Sheffield, GB)

A common framework for user interaction models and a common framework in which to place evaluation measures (i.e., the units of measurement) should be consistent but does not yet exist. Current measures are not comparable as the units used are not clearly defined in terms of real-world outcomes, and vary between measures. Since most measures encode some form of user behaviour as an underlying user interaction model, having measures that use the same unit of measure would enable comparisons between different user interaction models across different systems. As well as making it possible to compare between measures themselves (opposed to viewing them independently in different units).

4.9.1 Motivation

The main goal is to enable assessment of the performances of the system as a whole or specific components in particular. For that we need a repeatable way to say that a system is better than another on a gain base (utility, usefulness, happiness, ...). Ideally, the effect of user attributes that are not salient to the evaluation itself should be minimized (e.g. "what the user had for breakfast"). The measures should be comparable; that is, defined using the same units (i.e. gain, cost, or gain/cost). We would also like to be able to determine the effects of the interface and interaction on the actual performance.

4.9.2 Proposed solution

Integrate the interaction with an IR interface into the measures, e.g. in a TREC- style evaluation, individual IR systems may submit conventional ranked lists. Systems can then be evaluated based on different models of user interfaces or interactions. To extend TREC-style evaluations to accommodate more realistic interfaces, individual systems might submit responses to a variety of user actions, which would then be evaluated across more complex and detailed interfaces and interaction models.

One possible solution would be to decompose measures into components: Interaction model (I) (traditionally: when the user stops) Gain model (G) (traditionally: number of viewed relevant docs) Cost model (C) (traditionally: number of viewed docs with unit costs) An evaluation measure could then be parameterized by the components as M(I,G,C).

An interaction model might be characterized by a sequence of states and for each state some specific interaction with the system taken; potentially depending on the intent and task

of the user (e.g., a recall oriented task). The interface of the system could be encoded in the cost model. The gain would model the documents returned to the user (e.g., degree of relevance)

For example, we can deconstruct DCG into the three main components outlined above: the interaction model is "defined" by the discounting function, the gain model is how we sum up the weights of viewed (relevant) documents and the cost model is the number of viewed documents (with a fixed cost for each document). This means that we can fix the gain and the cost models while changing the interaction model still being able to compare measurements.

We could define an idealized interaction between the user and the system (including its interface). Idealized in this case would mean the optimal behavior where users are able to make decisions towards the best possible gains. System comparison based on such idealized interactions seem to be much more reasonable and comparable than based on arbitrary and possibly sub optimal decisions. This approach would also enable us to drop from the models a number of parameters that are difficult to estimate, such as click and query reformulation probabilities.

References

1 Allan, J., Croft, B., Moffat, A., and Sanderson, M. (2012). Frontiers, challenges, and opportunities for information retrieval: *Report from SWIRL 2012 the second strategic workshop on information retrieval in Lorne. SIGIR Forum 46:1* (May 2012), pp. 2–32, ACM Press, New York, NY, USA . http://doi.acm.org/10.1145/2215676.2215678

2 Azzopardi, L. (2009). Usage based effectiveness measures: monitoring application performance in information retrieval, *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM'09)*, pp. 631–640, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=1646034

3 Azzopardi, L. (2011). The economics in interactive information retrieval. *Proceedings of the 34th international ACM SIGIR International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'11)*, pp. 15–24, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2009916.2009923

4 Azzopardi, L., Kelly, D., and Brennan, K. (2013). How query cost affects search behavior. *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*, pp. 23–32, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2484049

5 Azzopardi, L., Järvelin, K., Kamps, J. and Smucker, M. D. (2011). Report on the SIGIR 2010 workshop on the simulation of interaction, *SIGIR Forum 44:2 (December 2010)*, pp. 35-47, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=1924475.1924484

6 Baskaya, F., Keskustalo, H., and Järvelin, K. (2012). Time drives interaction: Simulating sessions in diverse searching environments. *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'12)*, pp. 105–114, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2348301

7 Belkin, N., Dumais, S., Kando, N. and Sanderson, M. (2012). N*II Shonan meeting on Whole-Session Evaluation of Interactive Information Retrieval Systems, Held October 2012*. To appear. http://www.nii.ac.jp/shonan/seminar020/

8 Carterette, B. (2011). System effectiveness, user models, and user utility: a conceptual framework for investigation, *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'11)*, pp. 903–912, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2009916.2010037

9 Carterette, B., Kanoulas, E. and Yilmaz, E. (2012). Incorporating variability in user behavior into systems based evaluation. *Proceedings of the 21st ACM international Conference*

on Information and Knowledge Management (CIKM'12), pp. 135–144, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2396761.2396782

10 Sakai, T. and Dou, Z. (2013). Summaries, ranked retrieval and sessions: a unified framework for information access evaluation, *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*, pp. 473–482, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2484028.2484031

11 Smucker, M. and Clarke, C. (2012). Time-based calibration of effectiveness measures, *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'12)*, pp. 95–104, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2348283.2348300

12 Smucker, M. and Clarke, C. (2012). Modeling User Variance in Time-Biased Gain, *Proceedings of the Sixth Symposium on Human-Computer Interaction and Information Retrieval (HCIR'12)*, Article 3, 10 pages, ACM Press, New York, NY, USA. http://dl.acm.org/citation.cfm?id=2391224.2391227

Participants

- Maristella Agosti
University of Padova, IT
- Omar Alonso
Microsoft Research – Mountain View, US
- Leif Azzopardi
University of Glasgow, GB
- Nicholas J. Belkin
Rutgers University – New Brunswick, US
- Ann Blandford
University College London, GB
- Charles Clarke
University of Waterloo, CA
- Maarten de Rijke
University of Amsterdam, NL
- Arjen P. de Vries
CWI – Amsterdam, NL
- Floriana Esposito
University of Bari, IT
- Nicola Ferro
University of Padova, IT
- Luanne Freund
University of British Columbia – Vancouver, CA
- Norbert Fuhr
Universität Duisburg-Essen, DE
- Jacek Gwizdka
University of Texas – Austin, US
- Matthias Hagen
Bauhaus-Universität Weimar, DE
- Preben Hansen
Stockholm University, SE
- Jiyin He
CWI – Amsterdam, NL
- Kalervo Järvelin
University of Tampere, FI
- Hideo Joho
University of Tsukuba, JP
- Jaap Kamps
University of Amsterdam, NL
- Noriko Kando
National Institute of Informatics – Tokyo, JP
- Evangelos Kanoulas
Google Switzerland – Zürich, CH
- Diane Kelly
University of North Carolina – Chapel Hill, US
- Birger Larsen
Aalborg Univ. Copenhagen, DK
- Dirk Lewandowski
HAW – Hamburg, DE
- Christina Lioma
University of Copenhagen, DK
- Thomas Mandl
Universität Hildesheim, DE
- Peter Mutschke
GESIS – Köln, DE
- Ragnar Nordlie
Oslo University College, NO
- Heather O'Brien
University of British Columbia – Vancouver, CA
- Doug Oard
University of Maryland – College Park, US
- Vivien Petras
HU Berlin, DE
- Martin Potthast
Bauhaus-Universität Weimar, DE
- Soo Young Rieh
University of Michigan – Ann Arbor, US
- Gianmaria Silvello
University of Padova, IT
- Paul Thomas
CSIRO – Canberra, AU
- Elaine Toms
Sheffield University, GB
- Vu Tran
Universität Duisburg-Essen, DE
- Pertti Vakkari
University of Tampere, FI
- C .J. Keith van Rijsbergen
University of Cambridge, GB
- Robert Villa
University of Sheffield, GB
- Max L. Wilson
University of Nottingham, GB
- Christa Womser-Hacker
Universität Hildesheim, DE