DAGSTUHL
REPORTS

Volume 3, Issue 1, January 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
June, 2013

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.3.1.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

Report from Dagstuhl Seminar 13021

Symbolic Methods in Testing

Edited by

Thierry Jéron[1], Margus Veanes[2], and Burkhart Wolff[3]

1 **INRIA – Rennes – Bretagne Atlantique, FR,** `Thierry.Jeron@inria.fr`
2 **Microsoft Research – Redmond, US,** `margus@microsoft.com`
3 **Université Paris Sud, FR,** `wolff@lri.fr`

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13021 "Symbolic Methods in Testing". The aim of the seminar was to bring together leading researchers of this field; the seminary ended up with 38 participants from 10 countries: France, The Netherlands, The Unites States, Germany, Switzerland, United Kingdom, Brazil, Norway, Estonia and Italy. Through a series of presentations, discussions, and working group meetings, the seminar attempted to get a coherent picture of the field, which transcends the borders of applications and disciplines, of existing approaches and problems in formal testing. The seminar brought together, on the one hand, researchers from the different camps and various tools. The main outcome of the seminar is the exchange of information between different groups and the discussion of new trends (parallelization, cloud-computing).

Seminar 06.–11. January, 2013 – www.dagstuhl.de/13021
1998 ACM Subject Classification D.2.5 Testing and Debugging (Symbolic execution, Testing tools (e.g., data generators, coverage testing), Tracing), D.2.4 Software/Program Verification (Formal Methods, Assertion Checkers, Class invariants, Programming by Contract, Validation), B.2.3 Reliability, Testing, and Fault-Tolerance (Test generation)
Keywords and phrases Automated Deduction, White-box testing, Black-box Testing Fuzz-Testing Unit-Testing Theorem prover-based Testing
Digital Object Identifier 10.4230/DagRep.3.1.1

1 Executive Summary

Thierry Jéron
Margus Veanes
Burkhart Wolff

Recent breakthroughs in deductive techniques such as satisfiability modulo theories (SMT), abstract interpretation, model-checking, and interactive theorem proving, have paved the way for new and practically effective techniques in the area of software testing and analysis. It is common to these techniques that statespaces, model-elements, program-fragments or automata are represented symbolically making systems amenable to analysis that have formerly been out of reach. Several research communities apply similar techniques to attack the classical problem of state space explosion by using symbolic representation and symbolic execution: parametrized unit testing, fuzz testing, model-based testing, theoremprover based test case generation techniques, and real-time system testing. Moreover, several areas where symbolic methods are used in testing, are often considered more closely related to verification

Symbolic Methods in Testing, *Dagstuhl Reports*, Vol. 3, Issue 1, pp. 1–29
Editors: Thierry Jéron, Margus Veanes, and Burkhart Wolff

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

and end up in conferences specialized on those topics rather than at testing conferences. There is little synergy between the different communities although many of them use similar underlying symbolic techniques.

In the following areas, symbolic analysis techniques have recently had significant impact, both industrially as well as in academia. The following areas capture some topics of interest for the proposed seminar, assuming focus on the use of symbolic techniques in each area: Unit Testing, Symbolic Automata Theory in Testing, Model Based Testing, Fuzz Testing, Security Testing, Real-time System Testing, Theorem-Prover-based Test-Case Generation, Hybrid System Testing, and Mutation Testing.

2 Table of Contents

3 Overview of Talks

3.1 Symbolic methods for efficient mutation testing

Sébastien Bardin (CEA – Gif sur Yvette, FR)

Joint work of Bardin, Sébastien; Cheynier, François

Automatic white-box test data generation techniques have recently made huge progress, especially through the Dynamic Symbolic Execution (DSE) paradigm. However, these methods are limited to rather basic coverage criteria, typically instruction or decision coverage. On the other hand, mutation is a powerful testing criterion, but it is poorly supported by current automatic testing tools. We present in this talk some ongoing work aiming both at efficiently automating mutation testing and at leveraging DSE to sophisticated coverage criteria. We show that (a subset of) weak mutations can be reduced to predicate reachability, allowing to reuse all the standard machinery developed for software verification in the framework of weak mutations. Especially, we focus on the following issues: automatic test generation – ATG – in order to achieve high mutation score (using DSE and smart instrumentation), automatic mutant-equivalence checking (using static analysis and theorem proving) and efficient computation of the mutation score (again through instrumentation). Especially, for ATG, while a direct instrumentation yields an exponential blowup of the search space, we present a tight instrumentation with only a linear growth of the search space.

3.2 Towards symbolic and timed testing with JTorX

Axel Belinfante (University of Twente, NL)

We describe our model-based testing tool JTorX and the current plans for extensions to improve the support of symbolic and timed testing. JTorX [1, 3] is a tool for online (on-the-fly) test derivation and execution, based on the ioco-theory and its extensions uioco (underspecification), tioco (time) and sioco (data). Tests are derived from models given in Graphml, Aldebaran (.aut), GraphViz, mCRL2, STS-as-XML (.sax) or as a network of timed automata. Alternatively, JTorX interfaces to all models accessible via the LTSmin or CADP tool sets. JTorX accesses models on-the-fly (on demand), which allows it to deal with infinite models, as long as they are finitely branching. It can be used in 2 modes: manual (interactive), or automatic (random). In either mode, an optional test purpose (model to guide test derivation) can be specified. It has built-in adapter support to connect to a model used as system-under-test, and to connect to implementations that interact using labels of the model, over standard input/output, or via TCP. The latter adapter also comes in a time-aware variant. JTorX can be used (in automatic mode) from the command line, or via a GUI. During a test run, the GUI offers visualization of the testing progress in the model, and in a message sequence chart. In addition, the GUI contains a built-in interactive simulator, and a checker (by Lars Frantzen) that checks whether two non-symbolic models are (u)ioco-related. JTorX is used for teaching and for case studies; recent ones include testing of a software bus at Neopost [5], of an X-ray detector at PANalytical, and of railway interlocking

software. For most of above-mentioned formalisms, the model is accessed through a uniform model-access interface, via separate, formalism-specific, exploration components (for a few automaton-based formalisms, support is built-in in JTorX). The model-access interface gives access, not to an LTS, QTS or STS, but to a PTS: a parameterized transition system. A PTS has two kinds of transitions: those labeled with a parameterized label, and those labeled with an instantiation. A parameterized label is like an LTS label, but with parameters (like an open term) and a constraint over the parameters (and possibly, over parameters that have been introduced in an earlier transition on the path from the initial state, and that have not yet been instantiated). An instantiation is a list of parameter-value bindings. The interface allows JTorX to request(1) the (id of the) initial state, and (2) the outgoing transitions of a given a state id, where for each transition a parameterized label and the destination state id are given. Moreover, it allows JTorX to propose an instantiation, to feed back to the model, the concrete parameter values that were used during an interaction with the system under test. When the instantiation succeeds, i.e. when the PTS contains a corresponding transition, the (id of the) destination is given to JTorX. Semantic manipulation of parameterized labels is not done in JTorX, but in the formalism-specific model explorers. When JTorX needs concrete values for the parameters in a parameterized label that it wants to use as stimulus, it obtains those from a formalism-specific instantiator (which can either be part of the formalism-specific model explorer, or a separate tool component). JTorX uses this interface to access models via Lars Frantzen's STSimulator [4], and Henrik Bohnenkamp's timed automata explorer [2]. (An extension to STSimulator has been proposed and implemented in a JTorX variant: lazy-on-the-fly, by David Farago). For timed testing, each label contains a parameter (or, when instantiated, a value) that represents the time interval (or timestamp) at which the corresponding interaction with the SUT has to take (or has taken) place. The time-aware adapter variant takes care of applying stimuli in time, and of time-stamping observations.

References

1 Axel Belinfante. *JTorX: A Tool for On-Line Model-Driven Test Derivation and Execution*. In: Proceedings of TACAS 2010. LNCS, vol. 6015, pp. 266-270. Springer (2010)
2 Henrik Bohnenkamp, Axel Belinfante. *Timed testing with TorX*. In: FM 2005, Newcastle, UK. LNCS, vol. 3582, pp. 173-188. Springer (2005)
3 *JTorX Web page*: http://fmt.ewi.utwente.nl/tools/jtorx/
4 *STSimulator project page*: https://stsimulator.dev.java.net/
5 Sijtema, M. and Stoelinga, M.I.A. and Belinfante, A.F.E. and Marinelli, L. *Experiences with Formal Engineering: Model-Based Specification, Implementation and Testing of a Software Bus at Neopost*. In: Proceedings of FMICS 2011. LNCS, vol. 6959, pp. 117-133. Springer (2011)

3.3 An update on Z3

Nikolaj Bjørner (Microsoft Research – Redmond, US)

The overview talk provides a high-level summary of current directions for the SMT solver Z3 from Microsoft Research and newer applications. These include efficient engines for solving non-linear polynomial arithmetic, checking satisfiability of Horn clauses for symbolic model

checking of software. I will also summarize the main (newer) API features of Z3 and illustrate some of the ways they can be used by applications. Z3 was recently made available as shared source on z3.codeplex.com.

3.4 Counterexamples for Isabelle: Ground and Beyond

Jasmin Christian Blanchette (TU München, DE)

Joint work of Berghofer, Stefan; Bulwahn, Lukas; Nipkow, Tobias;
Main reference J.C. Blanchette, L. Bulwahn, T. Nipkow, "Automatic proof and disproof in Isabelle/HOL," in FroCoS 2011, pp. 12–27, LNAI 6989, Springer, 2011.
URL http://dx.doi.org/10.1007/978-3-642-24364-6_2

Users of the Isabelle/HOL proof assistant can rely on two counterexample generators to test their conjectures: Nitpick and Quickcheck. Nitpick grounds problems to SAT, but some of its optimizations are symbolic. Quickcheck combines three ground strategies and one symbolic strategy under one roof: random testing, bounded exhaustive testing, mode-based predicate compilation, and narrowing. For future work, we would like to expand the supported HOL fragment of both tools through more symbolic methods.

3.5 Model-based Conformance Testing of Security Properties

Achim D. Brucker (SAP Research – Karlsruhe, DE)

Joint work of Brucker, Achim D.; Brügger, Lukas; Wolff, Burkhart
Main reference A.D. Brucker, L. Brügger, P. Kearney, B. Wolff, "Verified Firewall Policy Transformations for Test-Case Generation," in Third Int'l Conf. on Software Testing, Verification, and Validation (ICST), pp. 345–354, IEEE CS, 2010.
URL http://www.brucker.ch/projects/hol-testgen

Modern systems need to comply with large and complex security policies (e. g., based on company rules, privacy laws, or regulations such as HIPAA or SOX) that need to be enforced at runtime. These policies are often expressed in declarative languages such as XACML (in case of high-level access control policies) or iptables (in case of firewall policies) and enforced by highly-efficient (and, thus, difficult to implement) enforcement engines. The correct configuration of such systems is highly error-prone, mainly due to their complexity. However, for the overall security both the correct implementation of the enforcement engines as well as their correct configuration is crucial. We are addressing these issues by presenting an approach for the model-based conformance testing of security policies. It is using HOL-TestGen [5, 4], a mode-based testing tool based on an interactive theorem proving environment. In more detail, we present a model-based testing approach encompassing the complete testing process using modular specifications of security policies (e.g., access control policies [2] or firewall policies [1, 3]). The generated test cases can be used for both testing the correctness of the security infrastructure as well as the conformance of its configuration to a high-level security policy. A particular emphasis is put on a partial solution to the scalability problem inherent in model-based policy testing.

References

1 Achim D. Brucker, Lukas Brügger, Paul Kearney, and Burkhart Wolff. *Verified firewall policy transformations for test-case generation*. In Third International Conference on Software Testing, Verification, and Validation (ICST), pages 345–354. IEEE Computer Society, 2010.
2 Achim D. Brucker, Lukas Brügger, Paul Kearney, and Burkhart Wolff. *An approach to modular and testable security models of real-world health-care applications*. In ACM SACMAT, pages 133–142. ACM Press, 2011.
3 Achim D. Brucker, Lukas Brügger, and Burkhart Wolff. *Model-based firewall conformance testing*. In Kenji Suzuki and Teruo Higashino, editors, Testcom/FATES 2008, number 5047 in LNCS, pages 103–118. Springer, 2008.
4 Achim D. Brucker and Burkhart Wolff. *HOL-TestGen: An interactive test-case generation framework*. In Marsha Chechik and Martin Wirsing, editors, Fundamental Approaches to Software Engineering (FASE), number 5503 in LNCS, pages 417–420. Springer, 2009.
5 Achim D. Brucker and Burkhart Wolff. *On theorem prover-based testing*. Formal Aspects of Computing, 2012.

3.6 Symbolic Execution for Evolving Software

Cristian Cadar (Imperial College London, GB)

Joint work of Cadar, Cristian; Marinescu, Paul; Collingbourne, Peter; Kelly, Paul

One of the distinguishing characteristics of software systems is that they evolve: new patches are committed to software repositories and new versions are released to users on a continuous basis. Unfortunately, many of these changes bring unexpected bugs that break the stability of the system or affect its security. In this talk, I describe two techniques based on symbolic execution for testing and verifying evolving software: a technique for automatic and comprehensive testing of code patches, which combines symbolic execution with several novel heuristics based on static and dynamic program analysis; and a technique that combines symbolic execution and crosschecking to test and verify the correctness of optimizations such as SIMD and GPGPU optimizations.

3.7 Collaborative Verification and Testing with Explicit Assumptions

Maria Christakis (ETH Zürich, CH)

Joint work of Christakis, Maria; Müller, Peter; Wüstholz, Valentin;
Main reference M. Christakis, P. Müller, V. Wüstholz, "Collaborative Verification and Testing with Explicit Assumptions," in FM 2012: Formal Methods, pp. 132–146, LNCS, 7436, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-32759-9_13

Most mainstream static program checkers make a number of compromises to increase automation, improve performance, and reduce both the number of false alarms and the annotation overhead for the programmer. These compromises include not checking certain program properties, and making implicit, unsound assumptions. As a result, static checkers that make such compromises cannot provide definite guarantees about program correctness,

thus rendering unclear which properties remain to be tested. To address this problem, we have proposed a tool architecture that: 1) makes the compromises of static checkers explicit and their verification results precise with a simple language extension that facilitates collaborative verification, i.e., the integration of multiple, complementary static checkers, and 2) reinforces static checking with automated, specification-based testing to make up for any soundness limitations of the upstream checkers. In this context, we have also discussed an approach for testing object invariants.

3.8 A Certified Constraint Solver over Finite Domains

Catherine Dubois (ENSIIE – Evry, FR)

Joint work of Carlier, Matthieu;Dubois, Catherine; Gotlieb, Arnaud;
Main reference M. Carlier, C. Dubois, A. Gotlieb, "A Certified Constraint Solver over Finite Domains," in Formal Methods – 18th International Symposium (FM 2012) , Vol. 7436 of LNCS: 116–131, 2012, Paris, France.
URL http://dx.doi.org/10.1007/978-3-642-32759-9_12

Constraint solvers are often used within verification or testing tools. These tools are complex, implement clever heuristics and thus may contain bugs. When these tools are used for critical software, a skeptical regard on the implementation of constraint solvers especially when the result is that a constraint problem has no solution, i.e., unsatisfiability. We review some state-of-the art solutions allowing more confidence. We present a Coq formalisation of a constraint filtering algorithm and a simple labeling procedure, focing on arc-consistency and bound-consistency. The proof of their soundness and completeness has been completed using Coq. As a result, a formally certified constraint solver written in OCaml (the first one as far as we know) has been automatically extracted from the Coq development. The solver, yet not as efficient as specialized existing (unsafe) implementations, can be used to formally certify that a constraint system is unsatisfiable.

3.9 Diagnosis Modulo Theories for Hybrid Systems

Juhan Ernits (Tallinn University of Technology, EE)

Joint work of Ernits, Juhan; Dearden, Richard

Diagnosis of hybrid systems involves tracking the state of the system on the basis of observations and control input to distinguish nominal behaviour from faulty. If faulty behaviour is encountered the approach proceeds to identify the causes of faults. Consistency-based diagnosis (CBD) is one of possible approaches designed to achieve such goals. In CBD the nominal and faulty behaviour of the system are modelled in terms of constraints present in each state and transitions between those states. To date the models used are mostly discrete, so system variables must be discretised before diagnosis can be performed. We introduce a new approach that uses a satisfiability modulo theories (SMT) solver in the implementation of the conflict directed A* algorithm – the core of the consistency-based diagnosis procedure. It is thus possible to use constraints from the theories supported by the SMT solver in the model, which is novel in diagnosis and provides a new application for

SMT solvers. The application has become practical only recently due to the integration of efficient non linear arithmetic theory decision procedures into SMT solvers.

3.10 Theorem-Prover Based Test Generation for Circus

Abderrahmane Feliachi (Université Paris Sud, FR)

Joint work of Feliachi, Abderrahmane; Gaudel, Marie-Claude; Wolff, Burkhart

HOL-TestGen [2] is a theorem-prover based environment for specification and test generation. Starting from a data-oriented (HOL) specification of a system under test, HOL-TestGen automatically derives test cases and test data for this system. Built on top of the Isabelle/HOL theorem prover, it allow for combining test generation tactics with symbolic computations and proof methods in a sound formal way. Since real complex systems combines complex data and behavioral aspects. We introduce, in the basis of Isabelle/HOL, a formal environment [1] for specifying and verifying complex systems. Specifications are written in Circus [3], a combination of Z and CSP, with a well defined and unified semantics. The semantics embedding of the Circus language in HOL is the basis of our environment. It makes it possible to reason on Circus specifications using HOL standard rules and proof methods. This environment is combined with HOL-TestGen, to provide a test generation environment covering complex data and behavioral aspects. Different symbolic representations and computations are used to define and generate tests from Circus specifications. This includes the embedding of symbolic variables, constraints over them and the resolution of these constraints for test data generation. A concrete application of this testing environment on a real system is presented in this talk.

References

1 Abderrahmane Feliachi, Marie-Claude Gaudel, and Burkhart Wolff. *Isabelle/Circus: A process specification and verification environment*. In Rajeev Joshi, Peter Müller, and Andreas Podelski, editors, Verified Software: Theories, Tools, Experiments, volume 7152 of Lecture Notes in Computer Science, pages 243–260. Springer Berlin / Heidelberg, 2012.

2 Achim Brucker and Burkhart Wolff. *On theorem prover-based testing*. Formal Aspects of Computing, pages 1–39, 2012. 10.1007/s00165-012-0222-y.

3 Jim Woodcock and Ana Cavalcanti. *The semantics of Circus*. In Proceedings of the 2nd International Conference of B and Z Users on Formal Specification and Development in Z and B, (ZB '02), pages 184–203, London, UK, UK, 2002. Springer-Verlag.

3.11 Off-Line Test Case Generation for Timed Symbolic Model-Based Conformance Testing

Christophe Gaston (CEA – Gif sur Yvette, FR)

Joint work of Bannour, Boutheina; Escobedo, Jose Pablo; Gaston, Christophe; Le Gall, Pascale
Main reference B. Bannour, J. P. Escobedo, C. Gaston, P. Le Gall, "Off-Line Test Case Generation for Timed Symbolic Model-Based Conformance Testing," in Proc. of the 24th Int'l Conf. on Testing Software and Systems, ICTSS 2012, pp. 119–135, LNCS, Volume 7641, Springer.

Model-based conformance testing of reactive systems consists in taking benefit from the model for mechanizing both test data generation and verdicts computation. On-line test case generation allows one to apply on- the-fly analysis techniques to generate the next inputs to be sent and to decide whether or not observed outputs meet intended behaviors. On the other hand, in off-line approaches, test suites are pre-computed from the model and stored under a format that can be later performed on test benches. In this seminar, we presented an off-line approach in two phases: (1) for the test generation part, a test suite is a predefined timed sequence of input data; (2) For the verdict production part, a post treatment is performed based on an analysis of the timed sequence of output data produced by the system under test with respect to the model. Our models are Timed Input Output Symbolic Transition Systems. Therefore, our off-line algorithm involves symbolic execution and constraint solving techniques.

References

1 Boutheina Bannour, Jose Pablo Escobedo, Christophe Gaston and Pascale Le Gall. *Off-Line Test Case Generation for Timed Symbolic Model-Based Conformance Testing*. In Proceedings of the 24th IFIP WG 6.1 International Conference on Testing Software and Systems, ICTSS 2012. Springer, LNCS, Volume 7641, pages 119–135. Aalborg, Denmark, November 19–21, 2012.

3.12 Distributed and Asynchronous Model Based Testing

Robert M. Hierons (Brunel University, GB)

Some systems interact with their environment at several physically distributed interfaces, called ports, and when testing such a system it is normal to place a local tester at each port. If the local testers cannot interact with one another during testing and there is no global clock, then each local tester observes only the sequence of inputs and outputs at its interfaces (a local trace). This can make it impossible to reconstruct the global trace that occurred. Similarly, we might not directly observe the input and output of the system under test (SUT) if there is an asynchronous communications channel between the tester and the SUT: the observation of output produced by the SUT is delayed, as is the SUT receiving input from the tester. Both situations lead to some loss of information regarding the sequence of events that the SUT performed and so they change the nature of testing and require us to define new implementation relations [6, 1]. They also affect certain standard testing problems. For example, in distributed testing it is undecidable whether there is a test case that is guaranteed to take a model to a particular state or to distinguish two given states

and this result holds even if we are testing from a deterministic finite state machine [3]. It is also undecidable whether there is a test case that is capable of distinguishing two states [2] and the Oracle problem is NP-hard [5]. There are currently fewer results for asynchronous testing but some problems, such as checking conformance, are known to be undecidable [1]. However, some test generation problems are decidable when there are FIFO channels and we are testing from a finite state model that is not output-divergent (there are no states from which one can take an infinite sequence of transitions without receiving an input) [4].

References

1 R. M. Hierons. *Implementation relations for testing through asynchronous channels*. The Computer Journal, to appear.
2 R. M. Hierons. *Verifying and comparing finite state machines for systems that have distributed interfaces*. IEEE Transactions on Computers, to appear.
3 Robert M. Hierons. *Reaching and distinguishing states of distributed systems*. SIAM Journal on Computing, 39(8):3480–3500, 2010.
4 Robert M. Hierons. *The complexity of asynchronous model based testing*. Theoretical Computer Science, 451:70–82, 2012.
5 Robert M. Hierons. *Oracles for distributed testing*. IEEE Transactions on Software Engineering, 38(3):629–641, 2012.
6 Robert M. Hierons, Mercedes G. Merayo, and Manuel Núñez. *Implementation relations and test generation for systems with distributed interfaces*. Distributed Computing, 25(1):35–62, 2012.

3.13 Model based conformance testing with ioco/tioco and Symbolic techniques

Thierry Jéron (INRIA Bretagne Atlantique – Rennes, FR)

Joint work of Jéron, Thierry; Jard, Claude; Jeannet, Bertrand; Rusu, Vlad; Zinovieva, Elena; Bertrand, Nathalie; Stainer, Amélie; Krichen, Moez; Chédor, Sébastien; Morvan, Christophe

This talk reviews some of the works of our group related to automatic test generation for reactive systems in the ioco/tioco testing framework. Starting with the finite state model case, we review the main ingredients of the ioco testing theory. We also generalize ioco to the io-refinement/abstraction pre-order relation. As the io-abstraction relation preserves the soundness of test cases, we show that it is a key for test generation when undecidability questions arise, due to infinity. We recall the principles of off-line test generation using test purposes, in the finite state case. The generation process is based on suspension, determinization and co-reachability analysis. We then extend test generation to infinite state systems, trying to mimic the finite state case. In particular, we focus on symbolic and approximation methods used to overcome problems due to infinite state as well as undecidability problems. First for models with data (IOSTS), limited to the deterministic case (determinization is an issue for this model), we show how abstract interpretation can be used for approximate co-reachability analysis. The consequences for test generation are examined: preservation of soundness and exhaustiveness, but loss of control to target. We then consider the case of timed automata (TAIO) in the general case (with internal actions and non-determinism, invariants for urgency). The principles of approximate determinization using games and symbolic represenations by regions or zones, which produces an io-abstraction

of a TAIO, is sketched. On the other side, co-reachability is exact, thanks to the abstract symbolic representation by zones and regions. Overall the test generation process allows to preserve the soundness of test cases, but may lead to a loss of exhaustiveness, stricness and precision only when determinization is not exact. recursion or time (in the tioco extension for real-time systems). Finally, the last (unpresented) part deals with recursive models in the form of recursive tiles systems (RTS), a sort of graph grammars. Determinization of RTSs being an issue, off-line test generation is restricted to a determinizable class, weighted RTSs, while on-line test generation can be applied on the full model. On the other hand, co-reachability is exact in the general case and can be decided on the RTS. Overall, the properties of soundness, exhaustiveness, strictness and precision of test cases are preserved in both cases.

References

1 C. Jard, T. Jéron. *TGV: theory, principles and algorithms, A tool for the automatic synthesis of conformance test cases for non-deterministic reactive systems.* Software Tools for Technology Transfer (STTT), 6, October 2004.

2 B. Jeannet, T. Jéron, V. Rusu, E. Zinovieva. Symbolic *Test Selection based on Approximate Analysis.* In 11th Int. Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'05), LNCS, Volume 3440, pages 349–364, Edinburgh (Scottland), April 2005.

3 N. Bertrand, T. Jéron, A. Stainer, M. Krichen. *Off-line Test Selection with Test Purposes for Non-Deterministic Timed Automata.* In 17th International Conference on Tools and Algorithms for the Construction And Analysis of Systems (TACAS'11), LNCS, Volume 6605, pages 96–111, Saarbrücken, Germany, April 2011.

4 S. Chédor, T. Jéron, C. Morvan. *Test generation from recursive tiles systems.* In 6th International Conference on Tests & Proofs (TAP'12), LNCS, Volume 7305, pages 99–114, Prague, May 2012.

3.14 Using State Infection Conditions to Detect Equivalent Mutants and Speed up Mutation Analysis

René Just (University of Washington – Seattle, US)

Joint work of Just, René; Ernst, Michael D.; Fraser, Gordon

Main reference R. Just, M.D. Ernst, G. Fraser, "Using State Infection Conditions to Detect Equivalent Mutants and Speed up Mutation Analysis," arXiv:1303.2784v1 [cs.SE].

URL http://arxiv.org/abs/1303.2784

Mutation analysis evaluates test suites and testing techniques by measuring how well they detect seeded defects (mutants). Even though well established in research, mutation analysis is rarely used in practice due to scalability problems — there are multiple mutations per code statement leading to a large number of mutants, and hence executions of the test suite. In addition, the use of mutation to improve test suites is futile for mutants that are equivalent, which means that there exists no test case that distinguishes them from the original program. This paper introduces two optimizations based on state infection conditions, i.e., conditions that determine for a test execution whether the same execution on a mutant would lead to a different state. First, redundant test execution can be avoided by monitoring state infection conditions, leading to an overall performance improvement. Second, state infection

conditions can aid in identifying equivalent mutants, thus guiding efforts to improve test suites.

References

1 R. Just, G. M. Kapfhammer, and F. Schweiggert. *Using non-redundant mutation operators and test suite prioritization to achieve efficient and scalable mutation analysis*. In Proc. of the 23rd ISSRE. IEEE Computer Society, 2012.
2 R. Just, F. Schweiggert, and G. M. Kapfhammer. *MAJOR: An efficient and extensible tool for mutation analysis in a Java compiler*. In Proc. of the 26th ASE, IEEE Computer Society, 2011.

3.15 A Symbolic Approach to Model-based Online Testing

Marko Kaeaeramees (Tallinn University of Technology, EE)

Joint work of Kääramees, Marko;
Main reference M. Kääramees, "A Symbolic Approach to Model-based Online Testing," PhD thesis, 2012. Tallinn University of Technology, Department of Computer Science, Estonia.
URL http://digi.lib.ttu.ee/i/?806

Non-deterministic control structures and data components provide a powerful means of abstraction for high level modelling of complex systems, at the expense of making automated test generation more challenging. Online model- based testing where test inputs are computed from the model and outputs fed back to the tester at the time of testing provides an approach where testing non-deterministic systems is possible. One of the restrictions to more widespread use of online model-based testing is the relatively high omputational overhead at runtime. We introduce an approach that addresses the computational overhead issue of online testing by pre-computation of test strategies based on the model and a formally specified test purpose. The proposed method allows the model of the IUT to be formalised as an Extended Finite State Machine over different first- order background theories. Both reachability and coverage oriented test purposes can be expressed using constraints attributed to edges of the model, called traps. We show how a testing strategy can be represented symbolically by a set of constraints and generated from the model and test purpose offline using symbolic backwards reachability analysis. The strategy can be used in online testing for efficient test input generation that guides the IUT towards fulfilment of the test purpose. The method is supported by the latest achievements of Satisfiability Module Theories (SMT) solver technology.

References

1 Marko Kääramees. *A Symbolic Approach to Model-based Online Testing*. PhD thesis, 2012. Tallinn University of Technology, Department of Computer Science, Estonia.
2 Danel Ahman and Marko Kääramees. *Constraint-based heuristic on-line test generation from non-deterministic I/O EFSMs*. In Alexander K. Petrenko and Holger Schlingloff, editors, MBT, volume 80 of EPTCS, pages 115–129, 2012.
3 Marko Kääramees, Jüri Vain, and Kullo Raiend. *Synthesis of on-line planning tester for non-deterministic EFSM models*. In Leonardo Bottaci and Gordon Fraser, editors, Testing – Practice and Research Techniques, volume 6303 of LNCS, pages 147- -154. Springer Berlin / Heidelberg, 2010.

4 Jüri Vain, Marko Kääramees, and Maili Markvardt. *Dependability and Computer Engineering : Concepts for Software-Intensive Systems*, chapter Online testing of nondeterministic systems with reactive planning tester, pages 113–150. IGI Global, Hershey, PA, 2011.

3.16 Critical Systems Development Methodology using Formal Techniques

Dominique Méry (LORIA – Nancy, FR)

Joint work of Méry, Dominique; Singh, Neeraj, Kumar;
Main reference D. Méry, N.K. Singh, "Critical systems development methodology using formal techniques," in Proc. of the 3rd Symp. on Information and Communication Technology (SoICT'12), pp. 3–12, ACM, 2012.
URL http://hal.inria.fr/hal-00747305/
URL http://dx.doi.org/10.1145/2350716.2350720

Formal methods have emerged as an alternative approach to ensuring the quality and correctness of the high confidence critical systems, overcoming limitations of the traditional validation techniques such as simulation and testing. We present a methodology for developing critical systems from requirement analysis to automatic code generation with standard safety assessment approach. This methodology combines the refinement approach with various tools including verification tool, model checker tool, real-time animator and finally, produces the source code into many languages using automatic code generation tools. This approach is intended to contribute to further the use of formal techniques for developing critical systems with high integrity and to verify complex properties, which help to discover potential problems. Assessment of the proposed methodology is given through developing a standard case study: the cardiac pacemaker. Further work remain to start for improving the development cycle by integrating testing phase for validating elements of the medical domain on the resulting system, namely the pacemaker.

References

1 D. Méry, N. K. Singh. *Functional Behavior of a Cardiac Pacing System*. International Journal of Discrete Event Control Systems (IJDECS), 2010.
2 D. Méry, N. K. Singh. *A generic framework: from modeling to code* In Innovations in Systems and Software Engineering (ISSE), pp. 1–9, 2011.
3 D. Méry, N. K. Singh. *Real-Time Animation for Formal Specification*. In Complex Systems Design & Management 2010, M. Aiguier, F. Bretaudeau, D. Krob (ed.), Springer, pp. 49–60. Paris, France, 2010.
4 D. Méry, N. K. Singh. *Trustable Formal Specification for Software Certification*. In 4th International Symposium On Leveraging Applications of Formal Methods – ISOLA 2010, T. Margaria, B. Ste (ed.), LNCS, 6416, Springer, pp. 312–326. Heraklion, Crete, Greece, 2010.
5 D. Méry, N. K. Sinhg. *Critical systems development methodology using formal techniques*.In 3rd International Symposium on Information and Communication Technology – SoICT 2012, ACM, pp. 3–12. Ha Long, Viet Nam, 2012.
6 D. Méry, N. K. Singh. *Formalization of Heart Models Based on the Conduction of Electrical Impulses and Cellular Automata* In Foundations of Health Informatics Engineering and Systems, Z. Liu, A. Wassyng (ed.), LNCS 7151, Springer Berlin Heidelberg, pp. 140–159. 2012.

3.17 Testing Real-time Systems under Uncertainty

Brian Nielsen (Aalborg University, DK)

Joint work of Nielsen, Brian; David, Alexandre; Larsen, Kim Guldstrand; Li, Shuhao; Mikucionis, Marius
Main reference A. David, K.G. Larsen, S. Li, M. Mikucionis, B. Nielsen , "Testing Real-Time Systems under Uncertainty," in Proc. of 9th Int'l Symp. on Formal Methods for Components and Objects (FMCO'10), pp. 352–371, LNCS, Vol. 6957, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-25271-6_19

Model-based testing is a promising technique for improving the quality of testing by automatically generating an efficient set of provably valid test cases from a system model. Testing embedded real-time systems is challenging because it must deal with timing, concurrency, processing and computation of complex mixed discrete and continuous signals, and limited observation and control. Whilst several techniques and tools have been proposed, few deals systematically with models capturing the indeterminacy resulting from concurrency, timing and limited observability and controllability. This paper proposes a number of model-based test generation principles and techniques that aim at efficient testing of timed systems under uncertainty.

3.18 Identifying suspicious values in programs with floating-point numbers

Michel Rueher (Université de Nice, FR)

Joint work of Rueher, Michel; Ponsini, Olivier; Michel, Claude
Main reference O. Ponsini, C. Michel, M. Rueher, "Refining abstract interpretation based value analysis with constraint programming techniques," in Proc. of CP 2012, LNCS, Vol. 7514, pp. 593–607, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-33558-7_43

Programs with floating-point computations control complex and critical physical systems in various domains such as transportation, nuclear energy, or medicine. Floating-point computations are an additional source of errors and famous computer bugs are due to errors in floating-point computations. Value analysis is often used to check the absence of run-time errors, such as invalid integer or floating-point operations, as well as simple user assertions. Value analysis can also help with estimating the accuracy of floating-point computations with respect to the same sequence of operations in an idealized semantics of real numbers. Existing automatic tools are mainly based on abstract interpretation techniques. For instance, a state-of-the-art static analyser like FLUCTUAT computes an over-approximation of the domains of the variables for a C program considered with semantics on real numbers. It also computes an over- approximation of the error due to floating-point operations at each program point. However, these over-approximations may be very coarse even for usual programming constructs and expressions. As a consequence, numerous false alarms may be generated. We introduce here a hybrid technique -called RAiCP- for value analysis of floating- point programs that combines abstract interpretation and constraint programming techniques. We show that constraint solvers over floating-point and real numbers can significantly refine the over-approximations computed by abstract interpretation. Experiments show that RAiCP is substantially more precise than FLUCTUAT, especially on C programs that are difficult to handle with abstract interpretation techniques. This is mainly due to the refutation

capabilities of filtering algorithms over the real numbers and the floating-point numbers used in RAiCP. Reducing these two over-approximations is a critical issue for program testing since it shrinks the set of suspicious values that have to be tested. RAiCP could also eliminate 13 false alarms generated by FLUCTUAT on a set of 57 standard benchmarks proposed by D'Silva et al to evaluate CDFL, a program analysis tool that embeds an abstract domain in the conflict driven clause- learning algorithm of a SAT solver. Moreover, RAiCP is on average at least 5 times faster than CDFL on this set of benchmarks.

References

1 Olivier Ponsini, Claude Michel and Michel Rueher. *Refining abstract interpretation based value analysis with constraint programming techniques*. Proc of CP 2012 , Springer Verlag, LNCS 7514, pp. 593-607, 2012

3.19 Pex4Fun: Serious Gaming powered by Symbolic Execution

Nikolai Tillman (Microsoft Research – Redmond, US)

Pex4Fun (http://www.pex4fun.com/) is a web-based serious gaming environment for learning and teaching programming. With Pex4Fun, a student edits code in any browser, supported by auto-completion. Pex4Fun then executes the code and analyzes it in the cloud. The analysis is based on the automated test generator Pex which uses symbolic execution and the SMT solver Z3 to generate high coverage test suites and find counterexamples to assertions. In particular, Pex4Fun finds interesting and unexpected input values that help students understand what their code is actually doing. The real fun starts with Coding Duels where students write code to implement a teacher's specification. Pex4Fun finds discrepancies in behavior between the student's code and the specification. The student wins when Pex4Fun cannot find any behavioral differences. This tutorial will discuss the technologies that lie beneath the Pex4Fun platform, show how to use Pex4Fun in teaching and learning, explore existing course materials and illustrate how to create new puzzles.

3.20 Symbolic Automata

Margus Veanes (Microsoft Research – Redmond, US)

The symbolic automata toolkit lifts classical automata analysis to work modulo rich alphabet theories. It uses the power of state-of-the-art constraint solvers for automata analysis that is both expressive and efficient, even for automata over large finite alphabets. The toolkit supports analysis of finite symbolic automata and transducers over strings. It also handles transducers with registers. Constraint solving is used when composing and minimizing automata, and a much deeper and powerful integration is also obtained by internalizing automata as theories.

3.21 Using Interpolation for Test-Case Generation for Security Protocols

Luca Vigano (Università di Verona, IT)

Joint work of Dalle Vedove, Giacomo; Rocchetto, Marco; Vigano, Luca; Volpe, Marco

Interpolation has been successfully applied in formal methods for model checking and test-case generation for sequential programs. Security protocols, however, exhibit such idiosyncrasies that make them unsuitable to the direct application of such methods. In this paper, we address this problem and present an interpolation-based method for test-case generation for security protocols. Our method starts from a formal protocol specification and combines Craig interpolation, symbolic execution and the standard Dolev-Yao intruder model to search for goals (i.e., test cases representing possible attacks on the protocol). Interpolants are generated as a response to search failure in order to prune possible useless traces and speed up the exploration. We illustrate our method by means of two concrete examples.

3.22 Paths to property violation: a structural approach for analyzing counter-examples

Hélène Waeselynck (LAAS – Toulouse, FR)

Joint work of Bochot, Thomas; Virelizier, Pierre; Waeselynck, Hélène; Wiels, Virginie;
Main reference T. Bochot, P. Virelizier, H. Waeselynck, V. Wiels, "Paths to Property Violation: A Structural Approach for Analyzing Counter-Examples," in Proc. of IEEE 12th Int'l Symp. onHigh-Assurance Systems Engineering (HASE'10), pp. 74–83, IEEE, 2010.
URL http://dx.doi.org/10.1109/HASE.2010.15

At Airbus, flight control software is developed using SCADE formal models, from which 90% of the code can be generated. Having a formal design leaves open the possibility of introducing model checking techniques. But, from our analysis of cases extracted from real software, a key issue concerns the exploitation of counterexamples showing property violation. To address this issue, we propose an automated structural analysis that identifies paths of the model that are activated by a counterexample over time. This analysis allows us to extract minimal information to explain the observed violation. It is also used to guide the model checker toward the search for different counterexamples, exhibiting new path activation patterns. The approach is closely related to path-based analysis techniques developed for testing purposes.

3.23 An Introduction to Model-based Testing with Isabelle/HOL-TestGen

Burkhart Wolff (Université Paris Sud, FR)

Main reference A.D. Brucker, B. Wolff, "On Theorem Prover-based Testing," Accepted the 07-08-2011. In Formal Aspects of Computing.
URL http://dx.doi.org/10.1007/s00165-012-0222-y

Techniques for the automated generation of test-cases – be it from specifications in form of pre- and postconditions, from transition systems or from annotated programs – suffer from state-space explosion similarly to model-checking techniques. One possible anwer to the challenge is to use symbolic representations of models, their normal forms, symbolic states, the resulting test-cases (partitions of test-data) and constraint-solving techniques for test-data selection. HOL-TestGen is a model-based test-generation environment based on the Theorem-proving based approach. In this talk, we will give an introduction into architecture, accommodation scenarios for a wide range of testing problems ranging from unit, sequence, reactive testing scenarios, and present an overview over major case-studies done with the system. This introduction was accompanied by a tutorial available from the material website.

References

1 Achim D. Brucker, Burkhart Wolff: *On Theorem Prover-based Testing.*. In: Formal Aspects of Computing (FAOC). DOI: 10.1007/s00165-012-0222-y. Springer, 2012.
2 Achim D. Brucker, L. Brügger and Burkhart Wolff: *Model-Based Firewall Conformance Testing.* ICTSS 08, LNCS 5047, pp. 103-118, http://dx.doi.org/10.1007/978-3-540-68524-1_9, Springer, 2008.
3 Achim D. Brucker, Abderrahmane Feliachi, Yakoub Nemouchi, and Burkhart Wolff. *Test Program Generation for a Microprocessor: A Case-Study*. In TAP 2013: Tests And Proofs. To appear in LNCS, Springer-Verlag, 2013.
4 Lukas Brügger. *A Framework for Modelling and Testing of Security Policies*. ETH Dissertation No. 20513. ETH Zurich,2012.
5 Achim D. Brucker, Lukas Brügger, Paul Kearney, and Burkhart Wolff. *An Approach to Modular and Testable Security Models of Real-world Health-care Applications*. In ACM symposium on access control models and technologies (SACMAT)., pages 133-142, ACM Press, 2011.
6 Achim D. Brucker, Lukas Brügger, Paul Kearney, and Burkhart Wolff. *Verified Firewall Policy Transformations for Test-Case Generation*. In Third International Conference on Software Testing, Verification, and Validation (ICST), pages 345-354, IEEE Computer Society , 2010.
7 Matthias P. Krieger. *Test Generation and Animation Based on Object-Oriented Specifications*. University Paris-Sud XI,2011.
8 Achim D. Brucker, Matthias P. Krieger, Delphine Longuet, and Burkhart Wolff. *A Specification-based Test Case Generation Method for UML/OCL*. In MoDELS Workshops. LNCS 6627, pages 334-348, Springer-Verlag , 2010.
9 Achim D. Brucker and Burkhart Wolff. *HOL-TestGen: An Interactive Test-case Generation Framework*. In Fundamental Approaches to Software Engineering (FASE09). LNCS 5503, pages 417-420, Springer-Verlag , 2009.
10 Achim D. Brucker, Lukas Brügger, and Burkhart Wolff. *Verifying Test-Hypotheses: An Experiment in Test and Proof*. Fourth Workshop on Model Based Testing (MBT 2008). In ENTCS, 220 (1), pages 15-27, 2008.

3.24 Dijkstra's Verdict Considered Harmful

Burkhart Wolff (Université Paris Sud, FR)

Dijkstra's Verdict "Testing can only show the absence of errors, not the presence" has been very influential and even more misleading in the scientifico-political debate between model-checking, testing and deductive verification communities. Having seen too many Phds using Dijkstra's Verdict in the introduction to motivate their model-checking and proof-based verification tool in a hopelessly exaggerating way, denying the importance of testing and experimentation, I recall that no approach can guarantee the absence of bugs. In contrast, the danger of uncritical application of verification tools (based on complex memory models, implicit methodological assumptions, etc) is very real. In particular, deductive methods can in practice NOT guarantee the absence of errors in programs, as far as they are written in real programming languages and run on real machines. Strictly speaking, it can not even be safely argued that deductive methods are BETTER than testing methods, since the underlying assumptions are incomparable.

3.25 Online Verification of Value-Passing Choreographies through Property-Oriented Passive Testing

Fatiha Zaïdi (Université Paris Sud, FR)

Joint work of Huu-Nghia, Nguyen; Pascal, Poizat; Fatiha, Zaïdi
Main reference Huu Nghia Nguyen, P. Poizat, F. Zaïdi, "Online Verification of Value-Passing Choreographies through Property-Oriented Passive Testing," in Proc. of the 14th IEEE Int'l High Assurance Systems Engineering Symposium (HASE 12), pp. 106–113, IEEE, 2012.
URL http://dx.doi.org/10.1109/HASE.2012.15

Choreography supports the specification, with a global perspective, of the interactions between roles played by peers in a collaboration. Choreography conformance testing aims at verifying whether a set of distributed peers collaborates wrt. choreography. Such collaborations are usually achieved through information exchange, thus taking data into account during the testing process is necessary. We address this issue by using a non-intrusive passive testing approach based on functional properties. A property can express a critical (positive or negative) behaviour to be tested on an isolated peer (locally) or on a set of peers (globally). We support online verification of these kind of properties against local running traces of each peer in a distributed system where no global clock is needed. Our framework is fully tool supported.

3.26 Symbolic Model-Based Testing of Real-Time Systems using SYMBOLRT

Wilkerson de Lucena Andrade (Federal University of Campina Grande, BR)

The state space explosion problem is one of the challenges to be faced by test case generation techniques, particularly when data values need to be enumerated. This problem gets even worse for Real-Time Systems that also have time requirements. The usual solution consists in enumerating data values (restricted to finite domains) while treating time symbolically, thus leading to the classical state explosion problem. We propose a symbolic model for conformance testing of real-time systems software named TIOSTS that addresses both data and time symbolically [1, 3]. Moreover, a test case generation process was defined to generate tests through TIOSTS models based on a combination of symbolic execution and constraint solving for the data part and symbolic analysis for timed aspects. All the process is supported by the SYMBOLRT tool [2]. SYMBOLRT is a model-based test generation tool developed to automatically generate test cases from symbolic models in the context of real-time systems. SYMBOLRT handles both data and time requirements symbolically, avoiding the state explosion problem during the test case generation.

References

1 W. L. Andrade, P. D. L. Machado, T. Jéron, H. Marchand. *Abstracting time and data for conformance testing of real-time systems*. In 7th Workshop on Advances in Model Based Testing (A-MOST 2011) / 2011 IEEE Fourth International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Pages 9-17, IEEE Computer Society, March 2011.

2 W. L. Andrade, D. R. Almeida, J. B. Cândido, P. D. L. Machado. *SYMBOLRT: A Tool for Symbolic Model-Based Test Case Generation for Real-Time Systems*. In 19th Tools Session of the 3rd Brazilian Conference on Software: Theory and Practice (CBSoft 2012), Pages 31-37, Natal, Brazil, September 2012.

3 W. L. Andrade, P. D. L. Machado. *Generating Test Cases for Real-Time Systems Based on Symbolic Models*. In IEEE Transactions on Software Engineering, to appear.

3.27 High-performance Analysis and Symbolic Online Test Generation

Jaco van de Pol (University of Twente, NL)

This overview talk consists of three parts: (1) A description of the LTSmin toolset for high-performance analysis, based on symbolic and multi-core model checking. The basis for the multi-core algorithms is a concurrent hashtable, with a very nice measured speedup on multi-core machines. Selecting one particular algorithm, I explain a parallel NDFS algorithm for LTL model checking. The initial variant had a small error, which was really hard to find. This work could be connected to research in symbolic testing in two directions: (a) a challenge for testing: could the error above be found by (symbolic) testing? (b) LTSmin could be used for high-performance test algorithms, e.g. determinisation, strategy synthesis, etc. (2) I shortly review my previous work in testing. First, we implemented simulated time

in TTCN-3 and applied it to model based testing of Railway Interlocking systems. Next, we proposed a symbolic online test generation algorithm, along the following lines:
- compute a data abstraction of the system (e.g. by homomorphism)
- generate an abstract test case towards some test purpose goal
- infer data parameters by constraint solving
- follow the test path for a while; re-plan when you cannot continue.

(3) Finally, I believe that the following themes are both challenging and promising as future work:
- developing robust black-box coverage criteria for symbolic testing
- optimal test case generation as strategy synthesis from a test game

As a closing remark: An important methodological issue concerning all research on testing is: What is a common yardstick to evaluate progress in testing research?

References

1 S.C.C. Blom, N. Ioustinova, J.C. van de Pol, A. Rennoch, and N. Sidorova, *Simulated Time for Testing Railway Interlockings with TTCN-3.* Proc. Formal Approaches to Software Testing (FATES'05, Edinburgh), LNCS 3997, Springer, pp. 1-15, 2006.

2 S.C.C. Blom, T. Deiß, N. Ioustinova, A. Kontio, J.C. van de Pol, A. Rennoch, and N. Sidorova, *Simulated Time for Host-Based Testing with TTCN-3.* Software Testing, Verification and Reliability, 18(1):29-49, 2008.

3 J.R. Calamé, N. Ioustinova, J.C. van de Pol, and N. Sidorova, *Data Abstraction and Constraint Solving for Conformance Testing.* Proc. Asia-Pacific Software Engineering Conference (APSEC'05, Taipei, Taiwan), IEEE, pp. 541-548, 2005.

4 Stefan Blom, Jaco van de Pol, and Michael Weber, *LTSmin: Distributed and Symbolic Reachability.* Proc. Computer Aided Verifciation (CAV'10, Edinburgh), LNCS 6174, Springer, pp. 354-359, 2010.

5 Alfons Laarman, Jaco van de Pol, and Michael Weber, *Multi-Core LTSmin: Marrying Modularity and Scalability.* Proc. Nasa Formal Methods (NFM'11, Pasadena, USA), LNCS 6617, Springer, pp. 506-511, 2011.

6 Alfons Laarman, Jaco van de Pol and Michael Weber, *Boosting Multi-Core Reachability Performance with Shared Hash Tables.* Proc. Formal Methods in Computer Aided Design (FMCAD'10, Lugano, Switzerland), IEEE, pp. 247-256, 2010.

7 Alfons Laarman, Rom Langerak, Jaco van de Pol, Michael Weber, and Anton Wijs, *Multi-Core Nested Depth-First Search.* Proc. Automated Technology for Verification and Analysis (ATVA'11, Tapei, Taiwan), LNCS 6996, Springer, pp. 321-335, 2011.

8 Tom van Dijk, Alfons Laarman and Jaco van de Pol, *Multi-core and/or Symbolic Model Checking*, Proc. Automated Verification of Critical Systems (AVOCS'12, Bamberg, Germany), ECEASST 53, 2012.

3.28 A Conformance Testing Relation for Symbolic Timed Automata

Sabrina von Styp (RWTH Aachen, DE)

Main reference Sabrina von Styp, Henrik Bohnenkamp, and Julien Schmaltz. A Conformance Testing Relation for Symbolic Timed Automata. In Proc. FORMATS 2010. pages 243-255. Volume 6246 of LNCS. Springer-Verlag, 2010.

We introduce Symbolic Timed Automata, an amalgamation of symbolic transition systems and timed automata, which allows to express nondeterministic data- dependent control

flow with inputs, outputs and real-time behaviour. In particular, input data can influence the timing behaviour. We define two semantics for STA, a concrete one as timed labelled transition systems and another one on a symbolic level. We show that the symbolic semantics is complete and correct w.r.t. the concrete one. Finally, we introduce symbolic conformance relation stioco, which is an extension of the well-known ioco conformance relation. Relation stioco is defined using FO-logic on a purely symbolic level. We show that stioco corresponds on the concrete semantic level to Krichen and Tripakis' implementation relation tioco for timed labelled transition systems.

4 Working Groups

4.1 Working Group Report: Towards a Competition in Model-based testing

Burkhart Wolff (Université Paris-Sud, FR)

The relative success of Tool- or Modeling competitions à la VSTTE or SMT- Competitions raises the question if the MBT community should develop a similar institution for its field. Competitions tend to reward otherwise neglected tool development, give indications on the technical state-of-the-art, and produce a strong feedback for choices in the foundational theories. We discuss certain scenarios and present a strategy that seems most appropriate to our field, which has the problem that *modeling* is a key issue and therefore a plethora of different modeling-languages make a simple comparison of solutions difficult. First, we identified a number of scenarios:

1. The "Archive Scenario" (Woodcock's FM Archive) ...
2. The "Grand Challenge Scenario" (The production cell scenario, MONDEX case study, etc.)
3. Open Format Competition à la VSTTE (open format, co-loc, Jury, fixed time)
4. Fixed Format Competition à la SMT and Casc (fixed format, co-location with a Conf.)

In principle, these scenarios sketch already a progression, and a strategy towards a competition. The latter should pave the path to a Modeling Competition a la VSTTE, where informal specifications of small problems were given, and hard time constraints for solutions were required by participating teams. Evaluation of solutions and tools by were performed by a Jury. The challenge for really setting this up consists essentially finding in a notable jury that sets up a call-for-paper, a prestigious prize and a reasonable conference to co-locate with ...

4.2 Working Group Report: Proof and Test

Cathérine Dubois (Université Évry, FR), Burkhart Wolff (Université Paris-Sud, FR)

Test and proofs seem to have complementary properties: while (formal) proofs are based on formal manipulation / computation, prove properties over models "once and for all", relate

models by logically complete arguments, testing is traditionally seen as experimentation with the goal to finding bugs, which can relate models with concrete systems, but are inherently incomplete.

The complementary characteristic is sometimes expressed as: *Beware of bugs in the above code; I have only proved it correct, not tried it.* Given that modern deductive verification tools make a lot of assumptions over memory model, machine-model, and methodology, the risk that a formally proven program is still buggily executed on a concrete machine is very real (see the nice example on Maria Christakis Slides presented here). The *validation* of models and system assumptions is therefore a necessary complement to deductive verification.

Over the tombstones of this old debate, there is in research communities the growing consensus that there is a lot of common ground between these two and that they complement each other. Testing can help Proof by:

- MBT for model validation,
- counter example finding and checking,
- using testing techniques on executable models,
- testing for verifying assumptions for proof tools,
- ...

Proof can contribute to testing :

- use prover for test-case generation,
- proof to pruning test objectives,
- basic technology: constraint solving, SMT's,
- mature interactive formal development environments (such as Isabelle),
- ...

Quality in software development (and, by the way, high-level certifications such as Common Criteria) require a clever combination of static analysis, proving, testing, model checking, constraint solving, reviewing ...

References

1 B. Wolff. Using Theorem Provers for Testing. https://www.lri.fr/~wolff/talks/UsingTheoremProversforTesting.pdf Invited Talk at the Seminary of the Digiteo Foundation http://www.digiteo.fr/-seminaires-digiteo-, Paris, 20 march 2013.

4.3 Working Group Report: Testing and the Cloud

Wolfgang Grieskamp (Google, Seattle, US), Burkhart Wolff (Université Paris-Sud, FR)

Cloud computing is a major trend in computing, both as a technological development as well as a scientific endeavor. With respect to testing, this trend generates to directions or research:

1. *Testing the Cloud*: How to adapt testing techniques to the new hybrid, highly distributed, highly virtualized infra-structures that become part of our world-wide IT infra-structure?
2. *Testing with the Cloud*: How can the massive computing power inherent in these new infra-structures be effectively used for old and new testing techniques?

With respect to the former question, we agreed that old problems are coming back big: the problems of scalability of techniques, the problem of isolation of components under test from the environment, the security and privacy problems as well as the problems of test configuration and deployment. However, some new aspects also pop up: the monitoring makes

runtime verification ("live" testing) easier and opens new ways on the virtualization level. The dynamics makes clusters of discrete systems increasingly continuous, which increases the need for symbolic methods. On the other hand, cloud computing makes the billing, concretely: the cost of testing, more explicit. In the cloud, cost a factor of availability, and production typically needs 99% availability. Here testing has an advantage over production: test departments can get away with perhaps 75% availability (which can be substantially cheaper).

Wrt. the latter question, we observe the phenomenon that more than two orders of magnitude or more of affordable computing power becomes available to testing — which can have an impact to methodology (less brain-power, more brute force ?). The other concern is, that the new infra-structures will not necessarily be *accessible* to everyone, it is a new era of company-owned computing centers, possibly a post-PC era. Further interest attracted the question, what type of testing techniques are, generally speaking, "cloudifiable". This is the case if they follow the map-reduce pattern, i.e.

- inexpensive pre-computation splitting the problem in a fixed number of independent sub-problems
- sub-problems executed on different machines, where communication via remote procedure calls is possible occurring rare in practice, and
- results can be collected easily and a potential start over to re-computation is possible.

This is the case for online model-based testing (symbolic or not), model-checking (some approaches already exist), concolic execution, and load testing.

4.4 Working Group Report: Machine Learning and Testing

Margus Veanes (Microsoft Research, Redmond, USA), Juhan Ernits (Tallinn University of Technology, Estonia)

Testing of complex systems, such as the behavior of a *quad copter* that is a real-time system controlled by hand gestures recognized through Kinect, can be very challenging. The 20 second youtube video clip[1] illustrates some aspects of what may go wrong and clearly raises questions as to how to best test such systems. There are several reasons for why testing of robotics and augmented reality applications is difficult in general and raises the difficulty bar compared to traditional approaches to testing:

- How do we define the test oracle for the system?
- How do we generate tests for the system?
- How do we measure coverage, i.e., when have we tested enough?

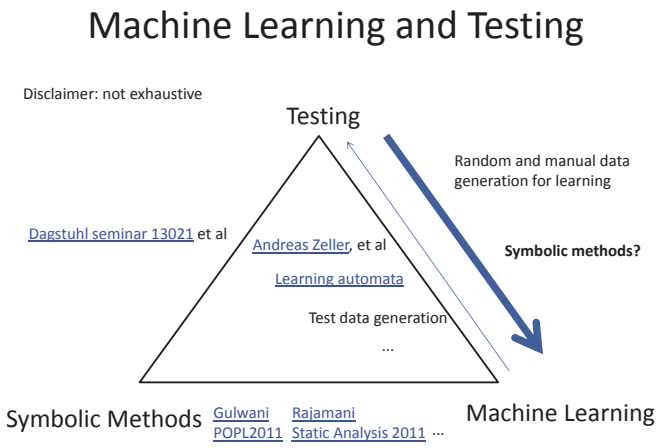
One observation we can make about the concrete scenario in the video is that, at some point, the intended gesture that was supposed to lower the copter was misinterpreted, and instead, the copter flew up. We know that the Kinect sensor uses *decision trees*[2] to detect body parts and to recognize gestures. Hypothetically, the flaw that caused the wrong behavior was in the logic of the decison tree. The decision tree has been constructed after *machine learning*

[1] https://www.youtube.com/watch?v=OQnRA6wZM-A
[2] http://dx.doi.org/10.1109/FPL.2012.6339226

of intended gestures for the given task. The question that is central to the theme of this seminar is:

- Could symbolic testing techniques be applied to decision tree classifiers in order to improve their accuracy?

This question is very interdisciplinary. The relation between the three fields of research: *testing*, *symbolic methods*, and *machine learning* is quite interesting and is illustrated in the following figure:



Abtractly, a machine learning algorithm produces a classifier, e.g., in form of a decision tree, that can be characterized as a function $F : X_1 \times \cdots \times X_n \to Y$ where elements of X_i are inputs to the classifier (e.g. 2D coordicates). The output value is the value depending on the inputs, typically Y is finite and is for example the shape of the detected skeleton. There are at least two different uses of classifiers in the context of testing.

- Classifier itself is the system under test (SUT).
- Classifiers are used as test data generators for the SUT (e.g. the SUT is the copter in the above scenario).

There are several open research questions related to the above uses.

- How could a white box approach be used, e.g., can a decision tree be treated as a white box?
- How could approximate distance based methods be used for data generation?
- Can one make minimal modification to fail classifiers?[3]

There are many more open questions. For eaxmple, could mutation based testing be used to improve classifiers? One can also apply other symbolic approaches such as: predicate extraction for explaining class-boundary conditons, and precondition and postcondition checking, or invariant checking such as: "In a similar domain every *even* input is classified as class-1, is it true for this classifier?".

[3] See for example CV Dazzle: http://cvdazzle.com/.

5 Programme

Programme for Monday 7th of January:

09:00–09:30: Margus Veanes: Welcome Talk of the Organizers
09:30–10:30: Nikolaj Bjørner: An update on Z3: Features and Applications
10:30–11:00:Coffee Break
11:00–12:00:Elevator Pitch Sessions: Position Statements, Perspectives, New Ideas, Value Propositions ...
12:00–14:00:Lunch
14:00–15:30 :Logic in Testing
– L.Viganó:Using Interpolation for Test-Case Generation for Security Protocols
– B. Wolff: An Introduction to Model-based Testing with Isabelle/HOL-TestGen
– J. v.d. Pol: Formal Methods and Tools: Testing U Twente
15:30–16:00:Coffee Break 16:00–17:30:Symbolic IOCO
– T. Jéron: Model based conformance testing with iocotioco and Symbolic techniques
– C. Gaston: Symbolic execution for Off-line Test Case Generation For Timed Model-Based Testing

Programme for Tuesday 8th of January:

09:00–10:30:Symbolic Testing involving Time
– A. Belinfante: Towards symbolic and timed testing with JTorX
– S. v. Styp: A Conformance Testing Relation for Symbolic Timed Automata
– W. Andrade: Symbolic Model-Based Testing of Real-Time Systems using SYMBOLRT
10:30–11:00:Coffee Break
11:00–12:00: N. Tillmann: Pex4Fun: Serious Gaming powered by Symbolic Execution
12:00–14:00: Lunch
14:00–15:30: Combined Process & Data Testing
– F. Zaidi: Online Verification of Value-Passing Choreographies through Property-Oriented Passive Testing
– A. Feliachi: Theorem-Prover Based Test Generation for Circus
– Discussion
15:30–16:00:Coffee Break
16:00–17:30: Counter-Example Generation
– J. Blanchette: Counterexamples for Isabelle: Ground and Beyond
– H. Waeselynck: Paths to property violation: a structural approach for analyzing counter-examples
– Discussion

Programme for Wednesday 9th of January:

9:00–10:30: Alternative Approaches
– M. Christakis: Collaborative Verification and Testing with Explicit Assumptions
– J. Ernits: Diagnosis Modulo Theories
– M. Kaeaeramees: A Symbolic Approach to Model-based Online Testing
Coffee
11:00–12:00: A. Brucker, L. Brügger: Tutorial: Model-based Testing of Security-critical Systems
12:00: Dagstuhl Foto

12:00–14:00 Lunch
14:00–14:30: D. Molnar: SAGE
15:00–21:00 Excursion to Saarburg Winery

Programme for Thursday 10th of January:

09:00–10:00: Mutation Testing
– S. Bardin: Symbolic methods for efficient mutation testing
– R. Just: Using State Infection Conditions to Detect Equivalent Mutants and Speed up Mutation Analysis
10:00–10:30 2nd elevator pitch sessions
Coffee
11:00–12:00: Group session : 4 Groups
– Should there be Tool Competition or an Archive on MBT
– Proof and Test
– Machine Learning and Testing
– Testing in and of the Cloud
14:30–15:30:
– D. Mery: Critical Systems Development Methodology
– B. Nielsen: Testing real-time systems under uncertainty
Coffee
16:00–17:30
– C. Cadar: KATCH: High-Coverage Testing of Software Patches
– M. Veanes: Symbolic automata

Programme for Friday 11th of January:

09:00–10:00
– C. Dubois: A Certified Constraint Solver over Finite Domains.
– M. Rueher: Combining AI & CP to identify suspicious values
Coffee Break
10:30–12:00: Summary of groups (1h) 10mn/groups
– B. Wolff: Should there be Tool Competition or an Archive on MBT
– C. Dubois: Proof and Test
– J. Ernits: Machine Learning and Testing
– W. Grieskamp: Testing in and of the Cloud
Discussion

Participants

- Sébastien Bardin
CEA – Gif sur Yvette, FR
- Axel Belinfante
University of Twente, NL
- Nikolaj Bjorner
Microsoft Res. – Redmond, US
- Jasmin Christian Blanchette
TU München, DE
- Achim D. Brucker
SAP Research – Karlsruhe, DE
- Lukas A. Brügger
ETH Zürich, CH
- Cristian Cadar
Imperial College London, GB
- Maria Christakis
ETH Zürich, CH
- Sylvain Conchon
Université Paris Sud, FR
- Wilkerson de Lucena Andrade
Federal University of Campina Grande, BR
- Catherine Dubois
ENSIIE – Evry, FR
- Juhan Ernits
Tallinn Univ. of Technology, EE
- Abderrahmane Feliachi
Université Paris Sud, FR
- Christophe Gaston
CEA – Gif sur Yvette, FR
- Arnaud Gotlieb
Simula Reseach Laboratory – Lysaker, NO
- Wolfgang Grieskamp
Google – Sammamish, US
- Robert M. Hierons
Brunel University, GB
- Thierry Jéron
INRIA Rennes – Bretagne Atlantique, FR
- René Just
University of Washington – Seattle, US
- Marko Kääramees
Tallinn Univ. of Technology, EE
- Pascale Le Gall
Ecole Centrale – Paris, FR
- Martin Leucker
Universität Lübeck, DE
- Delphine Longuet
Université Paris Sud, FR
- Dominique Méry
LORIA – Nancy, FR
- David Molnar
Microsoft Res. – Redmond, US
- Brian Nielsen
Aalborg University, DK
- Grgur Petric Maretic
ETH Zürich, CH
- Frank Rogin
Biotronik SE&Co.KG – Berlin, DE
- Michel Rueher
Université de Nice, FR
- Nikolai Tillmann
Microsoft Res. - Redmond, US
- Jan Tretmans
Embedded Systems Institute – Eindhoven, NL
- Jaco van de Pol
University of Twente, NL
- Margus Veanes
Microsoft Res. – Redmond, US
- Luca Vigano
Università di Verona, IT
- Sabrina von Styp
RWTH Aachen, DE
- Hélène Waeselynck
LAAS – Toulouse, FR
- Burkhart Wolff
Université Paris Sud, FR
- Fatiha Zaïdi
Université Paris Sud, FR

Report from Dagstuhl Seminar 13022

Engineering Resilient Systems: Models, Methods and Tools

Edited by

Maritta Heisel[1], Mohamed Kaaniche[2], Alexander Romanovsky[3], and Elena Troubitsyna[4]

1 **Universität Duisburg-Essen, DE,** maritta.heisel@uni-duisburg-essen.de
2 **LAAS – Toulouse, FR,** mohamed.kaaniche@laas.fr
3 **Newcastle University, GB,** alexander.romanovsky@newcastle.ac.uk
4 **Abo Akademi University – Turku, FI,** Elena.Troubitsyna@abo.fi

Abstract

Software-intensive systems are becoming widely used in such critical infrastructures as railway, air- and road traffic, power management, health care and banking. In spite of drastically increased complexity and need to operate in unpredictable volatile environment, high dependability remains a must for such systems. Resilience – the ability to deliver services that can be justifiably trusted despite changes – is an evolution of the dependability concept. It adds several new dimensions to dependability concepts including adaptability to evolving requirements and proactive error prevention. To address these challenges we need novel models, methods and tools that enable explicit modeling of resilience aspects and reasoning about them. The Dagstuhl Seminar 13022 "Engineering Resilient Systems: Models, Methods and Tools" discussed the most promising techniques for achieving resilience both at the system design stage and at runtime. It brought together researchers from dependability, formal methods, fault tolerance and software engineering communities that promoted vivid cross-disciplinary discussions.

Seminar 7.–11. January, 2013 – www.dagstuhl.de/13022
1998 ACM Subject Classification B.8.1 Reliability, Testing, and Fault-Tolerance, D.2.1 Requirements/Specifications, D.2.2 Design Tools and Techniques, D.2.11 Software Architectures, D.4.5 Reliability, F.3.1 Specifying and Verifying and Reasoning about Programs
Keywords and phrases Resilience, modelling, verification, evaluation, fault tolerance, evolution
Digital Object Identifier 10.4230/DagRep.3.1.30

1 Executive Summary

Maritta Heisel
Mohamed Kaaniche
Alexander Romanovsky
Elena Troubitsyna

The Dagstuhl Seminar 13022 – Engineering Resilient Systems: Models, Methods and Tools has brought together prominent researchers from different fields to discuss the problems of engineering resilient systems. The seminar was run in a highly interactive manner. The discussions were centered around the following topics:

- defining resilience
- resilience in modelling languages for requirement analysis and system design
- resilience in implementation languages and frameworks
- verifying resilience using testing, model checking and static analysis
- assessing resilience using probabilist models
- resilience mechanisms at architectural and implementation level

The concept of resilience has been introduced to capture the move towards a greater adaptability and flexibility. However, the notion of resilience is still a subject of debates. The seminar has discussed various proposed definitions and converged to defining resilience as *dependability in presence of changes*.

Over the last decades a remarkable progress has been achieved in engineering of highly dependable systems, i.e., the systems that can be justifiably trusted to provide critical services to a society. However, novel computing paradigms pose new scientific and technological challenges to the dependability field. To deliver critical services in a dependable way, the systems should smoothly adapt to changes. At the seminar, we had a dedicated session discussing the nature of changes. Among the proposed categories were

- evolving user requirements
- changing operating environment
- unforeseen failure modes
- scalability challenge

Resilience is strongly linked with the entire life-cycle of a system. Engineering of resilient systems should empower the systems with capabilities to cope with changes in a predictable way, cater for evolution and ensure robust behavior in spite of faults. These require novel techniques that explicitly address resilience through the entire system development cycle. Our seminar has explored challenges in formal modelling and verification of resilient systems. At the seminar we discussed suitable models for resilience, resilience-explicit development methods and verification techniques enabling both quantitative and qualitative resilience evaluation.

Modelling is the primarily vehicle driving development of resilient systems. However, system modelling area is still highly fragmented. The most acute problems are caused by

- the gap between the requirements and models and
- heterogeneity of models used to represent different aspects of system behaviour

Indeed, over the last few years the problem of poor flow-down of system requirements to software requirements has started to receive a proper attention. The vast majority of development relate the severe design problems with the flawed requirements and misunderstandings about what the software should do. Requirements tend to focus on describing nominal behaviour while omitting or poorly describing off-nominal conditions, safety constrains and fault tolerance mechanisms.

During the seminar we have brainstormed the examples of requirements that would be specific to resilient systems and tried to linked them with the modelling techniques.

While developing resilient systems the designers use dedicated models to reason about different (often antagonistic) aspects of system behaviour. Hence, the design space is inherently heterogeneous. On the one hand, specialised models provide the designers with expressive and powerful techniques to analyse various aspects of system behaviour. On the other hand, it becomes hard to obtain a holistic view on the system characteristics and analyse trade-offs between several potentially conflicting goals, define the mechanisms for

adapting to volatile operating conditions and devise appropriate mechanisms for proactive fault tolerance.

We have discussed the advances in formal modelling of resilient systems and in particular proactive fault tolerance and adaptive fault tolerance mechanisms at various frameworks. We have reviewed the advances achieved in the area of formal modelling of resilient systems and brain-stormed the techniques leveraging an integration of various models to facilitate emergence of integrated modelling approaches.

Essentially, any design flow can be seen as a set of well-defined abstraction levels. The design flow should allow the designer to optimize design decision at each level and move freely between abstraction layers. At our seminar we discussed the principles of mapping abstract models onto architectural models and design implementation. We addressed the problem of achieving architectural plasticity and brain-stormed architectural patterns supporting adaptation as well as mechanisms guaranteeing adequate predictable system reaction on changes. A significant attention has also been paid to the methods and tools for resilience assessment.

Open Problems

Engineering resilient systems is a young research area. The participants of the seminar have agreed that often it is hard to distinguish a traditional dependability research from the resilience research. We have converged to the view that the system ability to scale, cope with changes and evolve emphasizes the resilience aspect.

It was also noted that the area of resilience engineering lacks a comprehensive reference guide that would allow the designers of resilient systems understand how various proposed methods and tools can facilitate design of resilient systems. The participants of the seminar has decided to work on such a book.

2 Table of Contents

3 Overview of Talks

3.1 Testing and monitoring of dynamic systems: a governance-based framework

Antonia Bertolino (ISTI-CNR, IT)

For nowadays dynamic systems, resilience to failures needs to be planned in advance by collaborative agreements among involved stakeholders, ruling how components and services must be designed, documented, deployed, assessed. In our view resilience of dynamic systems should be supported by a governance framework establishing policies to be followed for off-line and on-line validation. By monitoring we can timely detect deviations of behaviour from functional and non-functional requirements. However, if the monitored system is obtained by the dynamic composition of independently developed services, who is the actor to whom monitor should report? and what action should (could) be taken? Moreover, whereas monitoring can only passively detect problems after they have occurred, a proactive approach of triggering selected behaviours might help anticipate possible future problems. We called such approach on-line testing. The idea of continuing to test a system after deployment and during real execution is appealing, but its implementation poses complex challenges: how to prevent or mitigate side effects? how can a tester simulate real or realistic service requests? In this presentation I will overview ongoing work along such directions by my group within the ongoing Choreos European Project. In particular, we are currently implementing a governance-based framework for on-line testing and monitoring of service choreographies. The presentation will outline some preliminary approaches with proposed tools and policies, but will mostly focus on open challenges for steering discussion and potential collaborations.

An annotated list of some related papers:

References

1. Antonia Bertolino, Guglielmo De Angelis, Sampo Kellomaki, Andrea Polini: Enhancing Service Federation Trustworthiness through Online Testing. IEEE Computer 45(1):66–72 (2012): *gives an overview of the idea beyond collaborative on-line testing and how this can be implemented within a service federation.*
2. Guglielmo De Angelis, Antonia Bertolino, Andrea Polini: Validation and Verification Policies for Governance of Service Choreographies. WEBIST 2012:58–70: *introduces some policies that could be considered for governing the testing of services.*
3. Antonia Bertolino and Andrea Polini. 2009. SOA Test Governance: Enabling Service Integration Testing across Organization and Technology Borders. In Proc. IEEE International Conference on Software Testing, Verification, and Validation Workshops (ICSTW '09). IEEE Computer Society, Washington, DC, USA, 277–286 : *A keynote talk introducing the motivation and vision for test governance framework.*
4. Francesco De Angelis, Andrea Polini, Guglielmo De Angelis: A Counter-Example Testing Approach for Orchestrated Services. ICST 2010:373–382: *the solution implementing within Choreos on-line testing approach*
5. Antonia Bertolino, Antonello Calabro, Francesca Lonetti, Antinisca Di Marco, Antonino Sabetta: Towards a Model-Driven Infrastructure for Runtime Monitoring. SERENE 2011:130–144: *outlines the GLIMPSE monitoring framework*

3.2 Model-based dependability and performance assessment in evolving contexts: the CONNECT experience

Felicita De Giandomenico (ISTI-CNR, IT)

The European FP7 Future and Emerging Technology Project CONNECT aimed at enabling seamless and dependable interoperability among networked systems in spite of technology diversity and evolution. The ambitious goal of the project was to have eternally functioning distributed systems within a dynamically evolving open-world context. This is pursued through the on-the-fly synthesis of the CONNECTors through which heterogeneous networked systems can communicate in dependable and secure way. Indeed, effective interoperability requires ensuring that such on-the-fly CONNECTed systems provide the required nonfunctional properties and continue to do so even in presence of evolution, thus calling for enhanced and adaptive assessment frameworks.

In the context of the CONNECT project, approaches to both off-line and runtime analysis have been investigated to analyze and ensure the synthesis of CONNECTors with required dependability and performance levels. In particular, an assessment framework has been proposed which combines continuous on-line assessment of non-functional properties through a lightweight flexible monitoring infrastructure with stochastic model-based analysis. The goal is to assess complex dependability and performance metrics through accurate analysis that adapts to the evolving context. Although not novel in its basic principles, this off-line and run-time integrated framework is proposed as a general, automated approach to fulfill the dependability and performance assessment needs in dynamic and evolving contexts.

3.3 A Model-based Assessment Framework to Analyse the Impact of Interdependencies in Power Systems

Felicita De Giandomenico (ISTI-CNR, IT)

Critical Infrastructures (CI) are complex and highly interdependent systems, networks and assets that provide essential services in our daily life. Given the increasing dependence upon such critical infrastructures, research and investments in identifying their vulnerabilities and devising survivability enhancements are recognized paramount by many countries. Understanding and analyzing interdependencies and interoperabilities between different critical infrastructures and between the several heterogeneous subsystems each infrastructure is composed of, are among the most challenging aspects faced today by designers, developers and operators in these critical sectors. Assessing the impact of interdependencies on the ability of the system to provide resilient and secure services is of primarily importance; following this analysis, steps can be taken to mitigate vulnerabilities revealed in critical assets. This presentation focuses on Electric Power Systems (EPS), one of the prominent representatives of CI systems, and overviews a model-based assessment framework for EPS, which explicitly accounts for interdependencies between the two infrastructures composing EPS: the Electric Infrastructure (EI) and the information infrastructure (II). The major achievements in this research line by the dependability group in Pisa along several years are shown.

3.4 Assessing Self-Organising Systems Resilience using DREF

Giovanna Di Marzo Serugendo (University of Geneva, CH)

One of the central features of self-organizing (SO) systems is their natural resilience to changes and faults, due to their ability to adapt their behavior. Assessing this resilience is generally done with experiments and simulations. Robustness and adaptation to some changes is obtained through specific self-organizing mechanisms, which have their limits and do not help overcoming any type of faults or change. For instance, digital pheromone in ant-based systems help overcome the appearance of obstacles in their environment or the disappearance of food, but is of limited help in case of faults (malicious or not) in the agent behavior (e.g. not properly following the pheromone). Therefore, in the process of development of a SO system, a developer will often want to achieve better resilience by adding, removing or modifying the system's behavior, then testing the system to see whether and how it has improved. Due to the complex behaviors of SO systems, however, it is not easy to quantify how a new version of a SO system compares to the one it replaces. Informal methods of comparison are generally effective only for relatively simple and small-scale systems. As SO systems are often used to model large complex behaviors, a structured, systematic and repeatable way to compare the properties of different versions of a system is necessary.

In this article we show how the evolution process taking place during the development of a SO system can benefit from a quantification of the satisfaction of properties by different versions of the system. To this end, we will enrich the classical "trial and error" development process (varying parameters and performing simulations) with a formal framework for the quantification of resilience called DREF (Dependability and Resilience Engineering Framework). The main goal of this article is to show how a precise, quantitative definition of resilience measures helps the developer in the choice of a particular version of a system.

3.5 Analytical Architecture Fault Models

Peter H. Feiler (CMU – Pittsburgh, US)

This presentation discusses how challenges in safety-critical software-intensive systems are addressed by the concept of an analyzable architecture fault model – expressed in SAE AADL and its revised Error Model Annex standard. The revised Error Model Annex includes a multi-set based type system and an error propagation type ontology. The presentation illustrates its use in avionics systems to address safety and resilience concerns through analysis early in the development life cycle. Examples show fault impact analysis, compositional specification and analysis of tolerance to failures, discuss the interaction between operational and failure modes, and conclude with an illustration of resilience to the intricacies of timing behavior in safety-critical systems.

References

1 P. Feiler, "Analytical Architecture Faults Models", Keynote presentation at the 3rd International Analytic Virtual Integration of Cyber-Physical Systems Workshop held in conjunction with RTSS 2012.
2 P. Feiler, SAE AS5506/3 Revised Error Model Annex Standard, Draft Oct 2012.
3 P. Feiler, D. Gluch, "Model-Based Engineering with AADL: An Introduction to the SAE Architecture Analysis & Design Language", Addison Wesley, Sept 2012.
4 P. Feiler, L. Wrage, J. Hansson, System Architecture Virtual Integration: A Case Study Embedded Real-time Software and Systems Conference (ERTS2010), May 2010.
5 P. Feiler, A. Rugina, Dependability Modeling with the Architecture Analysis & Design Language (AADL), Software Engineering Institute (SEI) Technical Report CMU/SEI-2007-TN-043, July 2007. http://www.sei.cmu.edu/library/abstracts/reports/07tn043.cfm

3.6 A model checking approach in the engineering of resilient systems

Stefania Gnesi (CNR – Pisa, IT)

Software Product Line Engineering (SPLE) is a paradigm for developing a diversity of software products and software-intensive systems based on the underlying architecture of an organisation's product platform. In the context of Software Product Lines (SPLs) the introduction of variability in the software development cycle has been extensively studied [5, 7]. At all abstraction levels, a product line description is composed of a constant part and a variable part. Variability among products is made explicit by variation points, i.e., places in design artifacts where a specific decision is reduced to several features but the feature to be chosen for a particular product is left open (like optional, mandatory, or alternative features). Variety from a single product platform is achieved by identifying such variability points. Variability management is the key aspect differentiating SPLE from conventional software engineering. Modelling variability in product families has been studied extensively in the literature on SPLs, especially that concerning feature modeling [6]. Formal methods have been developed to show that a product belongs to a family or to derive instead a product from a family. Deontic-style logics [1, 8] have become popular to formalize descriptional and behavioural aspects of computer systems, mainly because they provide a natural way to formalise concepts like violation, obligation, permission and prohibition. Intuitively, these concepts permit one to distinguish correct (normative) states and actions from non-compliant ones. Hence, deontic logic is a natural candidate for expressing the conformance of members of a family of products with respect to variabilities. A number of models, logics and associated tools for the qualitative analysis of variability aspects and their use to deal with adaptability and evolvability of systems have recently been proposed. In these lectures, we will focus on the approach presented in [2, 3, 4], where the introduction of the action-based branching-time temporal logic MHML allows expressing constraints over the products of a family as well as constraints over their behaviour in a single logical framework. Based on model-checking techniques for MHML, a modelling and verification framework will be presented that can automatically generate all the family's valid products, visualise the family/products behaviour and efficiently model check properties expressed in MHML over products and families alike. The use of the above methods, techniques and tools will be applied to a scenario derived from a family of resilient systems.

References

1 Åqvist, L.: Deontic logic. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, 2nd edition, vol. 8, pp. 147–264. Kluwer (2002)
2 Asirelli, P., ter Beek, M.H., Fantechi, A., Gnesi, S.: A logical framework to deal with variability. In: Mery, D., Merz, S. (eds.) IFM 2010. Lecture Notes in Computer Science, vol. 6396, pp. 43–58. Springer (2010)
3 Asirelli, P., ter Beek, M.H., Fantechi, A., Gnesi, S.: A verification environment for families of services. In: Bruni, R., Dingel, J. (eds.) FMOODS/FORTE 2011. Lecture Notes in Computer Science, vol. 6722, pp. 44–58. Springer (2011)
4 Asirelli, P., ter Beek, M.H., Gnesi, S., Fantechi, A.: Formal description of variability in product families. In: SPLC 2011, pp. 130–139. IEEE (2011)
5 Clements, P.C., Northrop, L.: Software Product Lines-Practices and Patterns. Addison-Wesley (2002)
6 Kang, K., Choen, S., Hess, J., Novak, W., Peterson, S.: Feature Oriented Domain Analysis (FODA) Feasibility Study. Technical Report SEI-90-TR-21, Carnegie Mellon University (1990)
7 Pohl, K., Bockle, G., van der Linden, F.: Software Product Line Engineering-Foundations, Principles, and Techniques. Springer (2005)
8 Meyer, J.-J.Ch., Wieringa, R.J. (eds.): Deontic Logic in Computer Science – Normative System Specification. John Wiley & Sons (1994)

3.7 Adaptability Metrics for QoS-driven Adaptable Systems

Vincenzo Grassi (Università di Roma "Tor Vergata", IT)

Joint work of Grassi, Vincenzo; Mirandola Raffaela

One of the major current research trends in Software Engineering is the focus on the development of new methodologies and techniques to deal efficiently with the design of more resilient systems that are able to evolve and adapt to rapid changes of their requirements or their execution environment. We present some ideas about the definition of metrics able to quantify and evaluate the adaptability of a software system at the architectural level. In particular, we focus on metrics measuring the system adaptability with respect to its ability to fulfill Quality of Service requirements. These metrics could be used by the software architect to drive design decisions concerning QoS-oriented adaptability features. They could also be helpful during runtime – when human intervention is not possible – to drive self-adaptation actions based on the impact they can have on the architecture adaptability.

References

1 A. DeLoach, V. A. Kolesnikov "Using Design Metrics for Predicting System Flexibility" FASE 2006, LNCS 3922, pp. 184–198, 2006
2 E. Kaddoum, C. Raibulet, J-P. Georg, G. Picard, M-P. Gleizes "Criteria for the evaluation of self-* systems" ACM SEAMS (2010), pp. 29–38
3 D. Perez-Palacin, R. Mirandola, J. Merseguer, "Software Architecture Adaptability Metrics for QoS-based Self-Adaptation" ACM QoSA (2011)
4 C. Raibulet, Masciadri, "Evaluation of dynamic adaptivity through metrics: an achievable target?" IEEE WICSA/ECSA (2009), pp. 341–344

5 P. Reinecke, K. Wolter, A. van Moorsel.,"Evaluating the adaptivity of computing systems" Performance Evaluation 67 (2010), pp. 676–693
6 H. Schmeck, C. Muller-Schloer, E. Cakar, M. Mnif, U. Richter "Adaptivity and self-organization in organic computing systems" ACM Trans. on Autonomous and Adaptive Systems, 5(3)10, Sept. 2010
7 N. Subramanian, L. Chung "Metrics for software adaptability" Software Quality Management (2001) pp. 95–108

3.8 Pattern and Component-based Development of Dependable Systems

Denis Hatebur (Universität Duisburg-Essen, DE)

The presentation discusses a process for pattern and component-based development of dependable systems. Dependability is an important aspect of resilient systems. The process is based on an description of the environment and covers analysis, design, implementation and testing. It is supported by an extended UML tool that checks the consistency of different artifacts. The process is extended by patterns for dependability requirements (confidentiality, integrity, availability, and reliability) that are part of a pattern system used to identify missing and conflicting requirements. It also covers a structured development of the architecture for dependable systems that fulfils the requirements.

3.9 Resilience – The ReSIST perspective

Mohamed Kaaniche (LAAS – Toulouse, FR)

This talk is aimed at presenting the definition of Resilience that has been proposed in the context of The ReSIST Network of Excellence funded by the European Commission. This definition is used as a starting point for initiating interactions and discussions about: 1) the difference between resilience and other similar concepts such as dependability, trustworthiness, survivability, etc., 2) whether this definition needs to be extended, and 3) the main new challenges that need to be addressed for the development of resilient systems compared to traditional existing approaches used the development of fault tolerant and dependable computing systems.

3.10 Developing Mode-Rich Satellite Software by Refinement in Event-B

Linas Laibinis (Abo Akademi University – Turku, FI)

One of the guarantees that the designers of on-board satellite systems need to provide, so as to ensure their dependability, is that the mode transition scheme is implemented correctly, i.e. that the states of system components are consistent with the global system mode. There is still, however, a lack of scalable approaches to developing and verifying systems with complex mode transitions. This paper presents an approach to formal development of mode-rich systems by refinement in Event-B. We formalise the concepts of modes and mode transitions as well as deriving specification and refinement patterns which support correct-by-construction system development. The proposed approach is validated by a formal development of the Attitude and Orbit Control System (AOCS) undertaken within the ICT DEPLOY project. The experience gained in the course of developing such a complex industrial system as AOCS shows that Event-B refinement provides the engineers with a scalable formal technique. Moreover, the case study has demonstrated that Event-B can facilitate formal development of mode-rich systems and, in particular, proof-based verification of their mode consistency.

3.11 Assessing the resilience of medical device user interfaces with verification tools

Paolo Masci (Queen Mary University of London, GB)

Medical device regulators such as the US Food and Drug Administration (FDA) aim to make sure that medical devices are reasonably safe before entering the market. To expedite the approval process and make it more uniform and rigorous, regulators are considering the development of reference models that encapsulate safety requirements against which software incorporated in to medical devices must be verified. Safety, insofar as it relates to interactive systems and its regulation, is generally a neglected topic, particularly in the context of medical systems. An example is presented that illustrates how the interactive behaviour of a commercial Patient Controlled Analgesia (PCA) infusion pump can be verified against a reference model. Infusion pumps are medical devices used in healthcare to deliver drugs to patients, and PCA pumps are particular infusion pump devices that are often used to provide pain relief to patients on demand. The reference model encapsulates the Generic PCA safety requirements provided by the FDA, and the verification is performed using a refinement approach.

3.12 A "SERENE" overview of the SOTA on Engineering Resilient Systems

Henry Muccini (Univ. degli Studi di L'Aquila, IT)

This short talk will introduce a preliminary analysis and classification of papers on software resilience presented in the past SERENE workshops. It wants to serve the purpose to create a starting point for realizing a survey on methods, approaches, and tools for engineering resilient systems.

3.13 From observations to models of resilient systems

Andras Pataricza (Budapest Univ. of Technology & Economics, HU)

Modeling, analysis, design of resilience related phenomena all need empirical substantiation due to the complexity of the typical target system and the underlying mechanisms and phenomena, as well. While model based design paradigms are well-proven and highly efficient in addressing complex problems, the faithfulness of the model is a crucial factor deciding in the very first moment all the quality of the results generated by the subsequent analysis/-synthesis steps. Accordingly, the extraction of a proper model from initial observations is the indispensable precondition for a well-substantiated approach in any resilience related engineering problem. Analyzing empirical resilience related data (like operation logs, or data sequences gained in fault injection campaigns or benchmarking experiments) is extremely difficult in the terms of statistics despite of decades long research efforts.

Analysis of resilience related data is clearly a big data problem over long sequences of many dimensional data reaching frequently the order of magnitude of several gigasamples of tens of thousands of signals. Moreover, as systems are dependable enough, fault manifestations are typically only rare outliers in a huge amount of samples corresponding to the correct behavior. As a consequence most popular algorithms widely used in other fields of statistics are of little use, as they simply suppress outliers and the hard class of rare event processing has to be used instead.

The problem of model creation is in itself a multi-phase process. The first, exploratory phase serves on the identification and rough characterization of the phenomena observed including the analysis of the individual signals, detection of the outliers and the relation between different signals. The derived characteristics help in the later phases of resilience analysis the estimation of principal factors leading to the individual failure modes, control clustering identifying typical operation domains and failure modes, etc.

Typically, in this initial phase of analysis, there is no or little statistical knowledge on the observed data to be analyzed. Exploratory data analysis (EDA) is an effective visual analysis approach to extract the main characteristics without the necessity of using not well founded and unnecessarily restrictive mathematical assumptions. The outcome of EDA is on the one hand a qualitative phenomenological model to be used in the formal analysis, and guiding heuristics for the detailed, already algorithmic analysis.

3.14 Engineering an open-source platform for mission planning of autonomous quadrotors

Patrizio Pelliccione (Univ. degli Studi di L'Aquila, IT)

Several projects exist to specify environmental monitoring missions. However, existent projects specify missions by means of programming languages which are too distant from the knowledge and terminology of the kind of users typically involved in such tasks. In this paper we propose an open-source platform, which enables the specification of monitoring missions that will be performed by fleets of autonomous quadrotors. The specification is performed at a high-level of abstraction and permits to graphically specify missions in the ground station. The mission will be then automatically decomposed by the platform in instructions to be performed by each quadrotor in order to fulfill the common goal. The platform enables users with limited IT knowledge, but domain experts in environmental missions to plan missions easily. A reconfiguration engine is specifically designed to autonomously react to faults and external events in order to accomplish the designed mission. Moreover, under some limitations explained in the paper, the reconfiguration engine permits to change the mission at run-time.

3.15 The shape of resilience

Matteo Risoldi (University of Luxembourg, LU)

There exist several different definitions of resilience, based on different scientific domains, intended purposes of the definition, requirements, and other scientific and cultural biases. We think however that, in most cases, these different definitions simply identify different ways that resilience manifests itself, rather than actually different "types" of resilience. In this talk, we present DREF, a formal framework for dependability and resilience, that allows quantification and visualization of many concepts related to resilience. In addition to being used for the quantitative assessment of resilience, DREF can be usefully employed for visualizing and identifying the fundamental ways that resilient behaviour manifests itself, through the variation of resilience-related parameters. We give a few examples of how different definitions of resilience may be represented in DREF, and give a (non-exhaustive) list of common traits of resilient behaviour.

3.16 Concurrency & Resilience – challenges in modern IT systems

Thomas Santen (European Microsoft Innovation Center – Aachen, DE)

Performance is one of the key requirements on modern IT systems from a user's perspective. This has major implications on the way that those systems are designed and implemented.

Many safe technologies that avoid certain types of faults by design such as type safe languages often cannot be used because they would impact the performance of the resulting system in an unacceptable way. True concurrency as induced by many-core systems is another source of complexity. To handle this complexity, techniques like formal code verification and model-based approaches can be helpful. Looking at the future of service-based systems resilience in providing services despite failures in data centers, communication, or other effects on the service infrastructure, is an increasingly relevant concern, and more technology to cope with those effects is needed that should take into account the other constraints of industrial software production like the key requirement of performance.

3.17 Reengineering of systems supposed to be resilient

Rolf Schumacher (Ingenieur-Büro Rolf Schumacher – Buchholz, DE)

With the event of clever virtual (as opposed to physical) attacks to public cyber physical systems or the ever increasing frequency of technological changes there will be a demanding need to improve the architecture of existing systems regarding yet unconsidered resilient properties. However, many of today's average software systems, including cyber physical systems, lack a reliable and complete set of maintainable requirements and architecture from where to start improvements in a controlled manner. As there are many thinkable approaches for re-engineering existing software systems in place and many of them failed or turned out to be overly expensive this contribution presents a process that worked in practice. The presented process has been applied to improve software for dependability. The process proceeds to be applied to similar average software systems, proving it to be general enough to be applied to a variety of existing software systems to be improved.

After a description of some challenging dependability aspects, a use-case driven top-down approach is presented for a behavioral model. It limits the significance of static models in favor of the more robust service oriented view. It leads to a set of requirements and a manageable architecture fulfilling its desired goals for the re-engineered system, e.g. dependability assessment. It is observed that one key success factor has been the renouncement of completeness of system states in favor of completeness of component functions usage. Having this experience in place we can concentrate on resilience requirements improving existing systems.

3.18 Resilience in Cyber-Physical Systems

Janos Sztipanovits (Vanderbilt University, US)

DARPA Adaptive Vehicle Make (AVM) Program is a major DARPA program a decade after MoBIES:

- End-to-end model- and component-based design and integrated manufacturing of a next generation amphibious infantry vehicle – a complex, real-life cyber-physical sys-

tem. From infrastructure to manufactured vehicle prototype in five yeas (2010-2014). Engineering/economic goals:

- Decrease development time by 80% in defense systems (brings productivity consistent with other industries)
- Enable the adoption of fabless design and foundry concept in CPS
- "Democratize" design by open source tool chain, crowed- sourced model library and prize-based design challenges

Scientific challenge: achieve AVM goals by pushing the limits of "correct-by-construction" design

- "Separation of concerns" principle need to be re-examined META pursues multi-physics, multi-abstraction and integrated cyber-physical design flows: modeling cross-domain interactions is in focus
- Multi-modeling makes model integration a fundamental challenge in the META design automation tool chain. META extensively uses model integration and includes CYPhy – a model integration language.

Participants

- Antonia Bertolino
CNR – Pisa, IT
- Felicita Di Giandomenico
CNR – Pisa, IT
- Giovanna Di Marzo Serugendo
University of Geneva, CH
- Peter H. Feiler
CMU – Pittsburgh, US
- Stefania Gnesi
CNR – Pisa, IT
- Vincenzo Grassi
Università di Roma "Tor Vergata", IT
- Denis Hatebur
Universität Duisburg-Essen, DE
- Maritta Heisel
Universität Duisburg-Essen, DE
- Mohamed Kaaniche
LAAS – Toulouse, FR
- Linas Laibinis
Abo Akademi University – Turku, FI
- Paolo Masci
Queen Mary University of London, GB
- Henry Muccini
Univ. degli Studi di L'Aquila, IT
- Andras Pataricza
Budapest Univ. of Technology & Economics, HU
- Patrizio Pelliccione
Univ. degli Studi di L'Aquila, IT
- Matteo Risoldi
University of Luxembourg, LU
- Alexander Romanovsky
Newcastle University, GB
- Thomas Santen
European Microsoft Innovation Center – Aachen, DE
- Rolf Schumacher
Ingenieur-Büro Rolf Schumacher – Buchholz, DE
- Janos Sztipanovits
Vanderbilt University, US
- Anton Tarasyuk
Abo Akademi University, FI
- Elena Troubitsyna
Abo Akademi University, FI
- Marco Vieira
University of Coimbra, PT

Report from Dagstuhl Seminar 13031

Computational Counting

Edited by

Peter Bürgisser[1], Leslie Ann Goldberg[2], Mark Jerrum[3], and Pascal Koiran[4]

1 **Universität Paderborn, DE,** pbuerg@upb.de
2 **University of Liverpool, GB,** L.A.Goldberg@liverpool.ac.uk
3 **Queen Mary University of London, GB,** m.jerrum@qmul.ac.uk
4 **ENS – Lyon, FR,** Pascal.Koiran@ens-lyon.fr

Abstract

Dagstuhl Seminar 13031 "Computational Counting" was held from 13th to 18th January 2013, at Schloss Dagstuhl – Leibniz Center for Informatics. A total of 43 researchers from all over the world, with interests and expertise in different aspects of computational counting, actively participated in the meeting.

Seminar 13.–18. January, 2013 – www.dagstuhl.de/13031
1998 ACM Subject Classification F.2 Analysis of Algorithms and Problem Complexity
Keywords and phrases Computational complexity, counting problems, graph polynomials, holographic algorithms, statistical physics, constraint satisfaction.
Digital Object Identifier 10.4230/DagRep.3.1.47
Edited in cooperation with Kitty Meeks

1 Executive Summary

Peter Bürgisser
Leslie Ann Goldberg
Mark Jerrum
Pascal Koiran

Introduction

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

Specific topics covered by the meeting include

- Techniques for exact counting, including moderately exponential algorithms for intractable problems, fixed parameter tractability, and holographic algorithms and reductions;

- techniques for approximate counting including Markov Chain Monte Carlo (MCMC);
- computational complexity of counting, including complexity in algebraic models; and
- applications, for example to models in statistical physics, and to constraint satisfaction.

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

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

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

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

Participation and Programme

The seminar brought together 43 researchers from Canada, China, Europe, India, Israel, Japan and the United States with interests and expertise in different aspects of computational counting. Among them there was a good mix of senior participants, postdoctoral researchers and PhD students. Altogether, there were 32 talks over the week.

If the spread of talks at the meeting is a reliable guide, the most active topics in the field at the moment are: algorithms and complexity in algebraic models, the complexity of Counting CSPs (Constraint Satisfaction Problems), and holographic algorithms and the holant framework. Other topics covered included: graph polynomials, MCMC (Markov Chain Monte Carlo) algorithms, parameterised complexity, phase transitions/decay of correlation and its relation to computational complexity, streaming algorithms, and exponential-time

exact algorithms. In addition to the technical presentations listed in the online programme, there were tutorial-style talks on topics featured in the Seminar. On Monday afternoon, Tyson Williams introduced the audience to holant problems and holographic transformations, and on Tuesday, Thore Husfeldt provided a similar service for newcomers to ETH and #ETH (the "Exponential Time Hypothesis" and its counting analogue).

One of the main aims of the seminar was to bring together researchers from different, but related fields, covering all aspects of computational counting with the goal of fostering the exchange of knowledge and to stimulate new research. This goal was fully achieved according to our opinion and the participant's feedback. The programme was as usual a compromise between allowing sufficient time for participants to present their work, while also providing unstructured periods for informal discussions. New contacts and maybe even friendships were made.

Snow and an early sunset did not the prevent the traditional Wednesday "hike" from taking place, though they did curtail it somewhat. The scenery was enhanced by the recent snowfall.

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

2 Table of Contents

3 Overview of Talks

3.1 The complexity of the noncommutative determinant

Markus Bläser (Universität des Saarlandes, DE)

We consider the complexity of computing the determinant over arbitrary finite-dimensional algebras. We first consider the case that A is fixed. We obtain the following dichotomy: If $A/\mathrm{rad}A$ is noncommutative, then computing the determinant over A is hard. "Hard" here means #P-hard over fields of characteristic 0 and ModP_p-hard over fields of characteristic $p > 0$. If $A/\mathrm{rad}A$ is commutative and the underlying field is perfect, then we can compute the determinant over A in polynomial time.

We also consider the case when A is part of the input. Here the hardness is closely related to the nilpotency index of the commutator ideal of A. The commutator ideal of A is the ideal generated by all elements of the form $xy - yx$ with $x, y \in A$. We prove that if the nilpotency index of the commutator ideal is linear in n, where $n \times n$ is the format of the given matrix, then computing the determinant is hard. On the other hand, we show the following upper bound: Assume that there is an algebra $B \subseteq A$ with $B = A/\mathrm{rad}(A)$. (If the underlying field is perfect, then this is always true.) The center $Z(A)$ of A is the set of all elements that commute with all other elements. It is a commutative subalgebra. We call an ideal J a complete ideal of noncommuting elements if $B + Z(A) + J = A$. If there is such a J with nilpotency index $o(n/\log n)$, then we can compute the determinant in subexponential time. Therefore, the determinant cannot be hard in this case, assuming the counting version of the exponential time hypothesis.

Our results answer several open questions posed by Chien et al.

3.2 Fast and Slow Mixing in the Ferromagnetic Potts Model

Magnus Bordewich (University of Durham, GB)

Joint work of Bordewich, Magnus; Greenhill, Catherine; Patel, Viresh

The Potts model is a statistical physics model of magnetism closely related to the Tutte polynomial of a graph in combinatorics. One element of interest is the Glauber dynamics of the model – a Markov chain process on a vertex colourings of the underlying graph. Each step of the Markov chain involves recolouring a single vertex. The state of the chain converges to a stationary distribution which is a weighted distribution on all colourings of the graph. This convergence can either happen in polynomial time in the number of vertices of the graph (rapid mixing), or it can take an exponential number of steps (torpid mixing). In this talk we explore what properties of the graph and the parameters of the model lead to fast or slow mixing.

3.3 On the average number of roots of a real sparse polynomial

Irénée Briquel (Université d'Orléans, FR)

The average number of real zeros of a random real polynomial is a well-studied problem. For instance, when the coefficients are identically tossed following a normal distribution, the number of zeros is asymptotically logarithmic in the degree. We here consider random sparse polynomials: polynomials for which a (small) number of nonzero coefficients are tossed. A criterion from Descartes tells us that the number of positive zeros of a sparse polynomial cannot be greater than the number of nonzero coefficients, no matter what the degree is. But no better result is known in the average. Yet, we suspect that the number could be much smaller in the average. We will discuss the tools we could use to estimate this number of real zeros and what implications it could have in complexity theory.

3.4 Geometric Complexity Theory and Counting

Peter Bürgisser (Universität Paderborn, DE)

Joint work of Bürgisser, Peter; Ikenmeyer, Christian
Main reference P. Bürgisser, C. Ikenmeyer, "Explicit Lower Bounds via Geometric Complexity Theory," in Proc. of STOC 2013, ACM, 2013.
URL http://arXiv.org/abs/1210.8368v2

We prove the lower bound $3/2m^2 - 2$ on the border rank of m x m matrix multiplication by exhibiting explicit representation theoretic (occurence) obstructions in the sense of Mulmuley and Sohoni's geometric complexity theory (GCT) program. While this bound is weaker than the one recently obtained by Landsberg and Ottaviani, these are the first significant lower bounds obtained within the GCT program. Behind the proof is the new combinatorial concept of obstruction designs H, which encode highest weight vectors f_H in $\mathrm{Sym}^d \otimes 3C^n$ and provide new insights into Kronecker coefficients. Deciding whether f_H equals the zero polynomial is not easy: we show that for a simple family of obstruction designs H_n, this is equivalent to the Alon-Tarsi Conjecture on Latin squares. The Alon-Tarsi Conjecture is also relevant for the permanent versus determinant problem. Kumar showed that this conjecture implies restrictions on the possible candidates for occurences obstructions. This excludes an asymptotic approach to the problem.

3.5 The complexity of counting CSP with complex weights

Xi Chen (Columbia University, US)

We prove a complexity dichotomy theorem for all Counting Constraint Satisfaction Problems (#CSP) with complex weights. We give three conditions for tractability. Let F be any finite set of complex functions, then we show that the #CSP defined by F is solvable in polynomial

time if all three conditions are satisfied; and is #P-hard otherwise. Our dichotomy generalizes a long series of important results on counting problems.

In this talk, we will focus on some of the most interesting ingredients of our algorithm, including a new polynomial-time operation over relations that share a common Mal'tsev polymorphism. We will then describe the framework of utilizing this operation to solve any #CSP efficiently when all three tractability conditions are satisfied.

3.6 Weighted counting of k-matchings is #W[1]-hard

Radu Curticapean (Universität des Saarlandes, DE)

Joint work of Bläser, Markus; Curticapean, Radu
Main reference M. Bläser, R. Curticapean, "Weighted Counting of k-Matchings Is #W[1]-Hard," in Proc. of the 7th Int'l Symp. on Parameterized and Exact Computation (IPEC'12), LNCS, Vol. 7535, pp. 171–181, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-33293-7_17

In the seminal paper for parameterized counting complexity (Flum, Grohe 2002), the following problem is conjectured to be #W[1]-hard: Given a bipartite graph G and a number k, which is considered as a parameter, count the number of matchings of size k in G.

We prove hardness for a natural weighted generalization of this problem: Let $G = (V, E, w)$ be an edge-weighted graph and define the weight of a matching as the product of weights of all edges in the matching. We show that exact evaluation of the sum over all such weighted matchings of size k is #W[1]-hard for bipartite graphs G.

As an intermediate step in our reduction, we also prove #W[1]-hardness of the problem of counting k-partial cycle covers, which are vertex-disjoint unions of cycles including k edges in total.

3.7 Enumeration complexity of query problems: constant delay and quantifier elimination methods

Arnaud Durand (University Paris-Diderot, FR)

In this talk, we present some results on the enumeration complexity for query problems and CSP. We first survey the main complexity measures defined in the context of enumeration and comment on some connexions between enumeration and counting. We then focus on the notion of constant delay enumeration. This class contains all problems for which solutions can be enumerated with a polynomial (linear in our case) precomputation and a constant delay between two consecutive solutions. This notion was introduced in the context of query answering and, in this setting, "constant" means "depending on the formula size". We show that several natural classes of queries can be enumerated with constant delay. Surprisingly, in all cases, methods based on quantifier elimination are developed to obtain these upper bounds.

3.8 Generating random regular graphs and digraphs

Martin Dyer (University of Leeds, GB)

We consider the problem of sampling a regular graph uniformly at random, in time polynomial in the size of the graph. We review various approaches to this problem, which work in different ranges for the degree of the graph as a function of its size. In particular, we consider Markov chain methods for generating such a graph. Jerrum and Sinclair gave a general solution, but we consider more natural chains, which have applications to the design of peer-to-peer networks. These methods use "switches" and "flips", which are simple modifications of the graph. Finally, we consider the extension of these methods to regular digraphs.

3.9 The Potts/Tutte connection with an external field

Jo Ellis-Monaghan (Saint Michael's College – Colchester, US)

Joint work of Ellis-Monaghan, Jo; Moffatt, Iain
Main reference J. Ellis-Monaghan, I. Moffatt, "The Tutte-Potts connection in the presence of an external magnetic field," Advances in Applied Mathematics, Vol. 47, Issue 4, October 2011, pp. 772–782, 2011.
URL http://dx.doi.org/10.1016/j.aam.2011.02.004

The Potts model in statistical mechanics is a rapidly emerging and increasingly applicable predictive model for complex systems in which very simple interactions at the microscale level determine the macroscale properties of the system. This model plays an important role in the theory of phase transitions and critical phenomena in physics, and has applications as widely varied as tumor migration, foam behaviors, and social demographics.

Here we define the V-polynomial, which lifts the classical relationship between the Tutte polynomial of graph theory and the zero field Potts model to encompass external magnetic fields. The classical relationship between the Tutte polynomial of graph theory and the zero field Potts model has resulted in valuable interactions between the disciplines. Unfortunately, it does not include the external magnetic fields that appear in most Potts model applications. Thus the current work unifies an important segment of Potts model theory and brings previously successful combinatorial machinery, including complexity results, to bear on a wider range of statistical mechanics models.

3.10 On fixed-polynomial size circuit lower bounds for uniform polynomials in the sense of Valiant

Hervé Fournier (University Paris-Diderot, FR)

We consider the problem of fixed-polynomial lower bounds on the size of arithmetic circuits computing uniform families of polynomials. Our first result is that for all k, there exist polynomials with coefficients in MA having no arithmetic circuits of size $O(n^k)$ over the

complex field (i.e. allowing any complex constant). We also investigate links between fixed-polynomial size circuit bounds in the Boolean and arithmetic settings. In particular, NP without n^k size circuits or NP=MA imply lower bounds on the circuit size of uniform polynomials in VNP over the complex field.

3.11 On the connection between interval size functions and path counting

Andreas Goebel (University of Liverpool, GB)

Joint work of Evangelos Bampas, Aris Pagourtzis, Aris Tentes

We investigate the complexity of hard counting problems that belong to the class #P but have easy decision version; several well-known problems such as #Perfect Matchings, #DNFSat share this property. We focus on classes of such problems which emerged through two disparate approaches: one taken by Hemaspaandra et al. who defined classes of functions that count the size of intervals of ordered strings, and one followed by Kiayias et al. who defined the class TotP, consisting of functions that count the total number of paths of NP computations. We provide inclusion and separation relations between TotP and interval size counting classes, by means of new classes that we define in this work. Our results imply that many known #P-complete problems with easy decision are contained in the classes defined in Hemaspaandra et al.—but are unlikely to be complete for these classes under certain types of reductions. We also define a new class of interval size functions which strictly contains FP and is strictly contained in TotP under reasonable complexity-theoretic assumptions. We show that this new class contains some hard counting problems.

3.12 Approximating the partition function of planar two-state spin systems

Leslie Ann Goldberg (University of Liverpool, GB)

Joint work of Goldberg, Leslie Ann; Jerrum, Mark; McQuillan, Colin
Main reference L.A. Goldberg, M. Jerrum, C. McQuillan, "Approximating the partition function of planar two-state spin systems," arXiv:1208.4987v2 [cs.CC].
URL http://arxiv.org/abs/1208.4987v2

We consider the problem of approximating the partition function of the hard-core model on planar graphs of degree at most 4. We show that when the activity lambda is sufficiently large, there is no fully polynomial randomised approximation scheme for evaluating the partition function unless NP=RP. The result extends to a nearby region of the parameter space in a more general two-state spin system with three parameters. We also give a polynomial-time randomised approximation scheme for the logarithm of the partition function.

3.13 Factoring bivariate lacunary polynomials without heights

Bruno Grenet (ENS-Lyon, FR)

Joint work of Arkadev Chattophyay, Pascal Koiran, Natacha Portier, Yann Strozecki

The lacunary, or supersparse, representation of a multivariate polynomial P is a list of pairs (c, e) where c is a coefficient of P and e is a vector of exponent. Each pair defines a term of the polynomial, and P equals the sum of all these terms. The factorization of lacunary polynomials has been investigated in a series of papers by Cucker, Koiran and Smale (J. Symb. Comput., 1999), Lenstra (Number Theory in Progress, 1999), and Kaltofen and Koiran (ISSAC 2005 & 2006). In this paper, we are interested in more elementary proofs for some of these results. We focus on Kaltofen and Koiran's results dealing with linear factors of bivariate lacunary polynomials. In particular, we give a polynomial-time algorithm to find linear factors of bivariate polynomials that is not based on heights of algebraic numbers. This simplification allows us to give a similar result for some fields of positive characteristic. Our main technical result is an upper bound on the valuation of polynomials of the form $P(X, 1 + X)$ where P is a bivariate lacunary polynomial, and can be viewed as a generalization of a result of Hajós.

3.14 A Complete Dichotomy Rises from the Capture of Vanishing Signatures

Heng Guo (University of Wisconsin – Madison, US)

Joint work of Cai, Jin-Yi; Williams, Tyson
Main reference J.-Y. Cai, H. Guo, T. Williams, "A Complete Dichotomy Rises from the Capture of Vanishing Signatures," arXiv:1204.6445v1 [cs.CC].
URL http://arxiv.org/abs/1204.6445v1

We prove a complexity dichotomy theorem for Holant problems over an arbitrary set of complex-valued symmetric constraint functions F on Boolean variables. This extends and unifies all previous dichotomies for Holant problems on symmetric constraint functions taking values in a field of characteristic zero. We define and characterize all symmetric vanishing signatures. They turned out to be essential to the complete classification of Holant problems. The dichotomy theorem has an explicit tractability criterion. The Holant problem defined by a set of constraint functions F is solvable in polynomial time if it satisfies this tractability criterion, and is #P-hard otherwise. The tractability criterion can be intuitively stated as follows: the set F is tractable if (1) every function in F has arity at most 2, or (2) F is transformable to an affine type, or (3) F is transformable to a product type, or (4) F is vanishing, combined with the right type of binary functions, or (5) F belongs to a special category of vanishing type Fibonacci gates. The proof of this theorem utilizes many previous dichotomy theorems on Holant problems and Boolean #CSP.

3.15 The Parity of Directed Hamiltonian Cycles

Thore Husfeldt (IT University of Copenhagen, DK)

We present a deterministic algorithm that given any directed graph on n vertices computes the parity of its number of Hamiltonian cycles in $O(1.618^n)$ time and polynomial space.

3.16 Approximate Counting of Matchings in Uniform Hypergraphs

Marek Karpinski (Universität Bonn, DE)

Joint work of Karpinski, Marek; Rucinski, Andrzej; Szymanska, Edyta
Main reference M. Karpinski, A. Rucinski, E. Szymanska, "Approximate Counting of Matchings in Sparse Uniform Hypergraphs," arXiv:1204.5335v1 [cs.DS].
URL http://arxiv.org/abs/1204.5335v1

We give a fully polynomial *randomized approximation scheme (FPRAS)* for counting matchings in k-uniform hypergraphs without *wide* edges, or equivalently hypergraphs whose intersection graphs are claw-free. The method generalizes to the weighted monomer-polymer (trimer, tetramer, pentamer, etc.) problems for the corresponding class of hypergraphs.

The method adopts and generalizes the canonical path method of Jerrum and Sinclair to those restricted hypergraph classes. We prove also that the problem of counting matchings in k-uniform hypergraphs without the above restriction is approximation hard for all $k \geq 6$. It leaves very interesting open problems on approximate counting matchings in arbitrary k-uniform hypergraphs for $k = 3, 4$ and 5, and also some connected issues on approximate counting independent sets in some classes of intersection graphs.

References

1 O.J. Heilmann and E.H. Lieb. *Theory of Monomer-Dimer Systems*. Commun. Math. Physics 25, pp. 190–232, 1972
2 M. Jerrum and A. Sinclair. *Approximating the Permanent*. SIAM J. Comput. 18, pp. 1149–1178, 1989
3 M. Karpinski, A. Rucinski and E. Szymanska. *Computational Complexity of the Perfect Matching Problem in Hypergraphs with Subcritical Density*. Int. J. Found. Comput. Sci. 21, pp. 905–924, 2010
4 M. Karpinski, A. Rucinski and E. Szymanska. *Approximate Counting of Matchings in Sparse Uniform Hypergraphs*. Proc. 13th SIAM ANALCO (2013), pp. 71–78
5 A. Sly. *Computational Transition at the Uniqueness Threshold*. Proc. 51st IEEE FOCS (2010), pp. 287–296

3.17 Enumerating monomials

Meena Mahajan (The Institute of Mathematical Sciences – Chennai, IN)

Joint work of de Rugy, Nicolas; Saurabh, Nitin; Sreenivasaiah, Karteek; Strozecki, Yann

In this talk, I discuss some ongoing work on the following problems:

Given an arithmetic circuit C computing a polynomial p,

1. deterministically enumerate (without repetition) all monomials of p,
2. and given also a term m, compute the coefficient of m in p.

We look for restrictions that allow us to achieve these tasks with preprocessing time / delay / computation time polynomial in the size of C.

Joint work with Nicolas de Rugy, Nitin Saurabh, Karteek Sreenivasaiah and Yann Strozecki. Funded by an IFCPAR (CEFIPRA) project.

3.18 A computational framework for the study of partition functions and graph polynomials

Johann Makowsky (Technion – Haifa, IL)

Partition functions and graph polynomials have found many applications in combinatorics, physics, biology and even the mathematics of finance. Studying their complexity poses some problems. To capture the complexity of their combinatorial nature, the Turing model of computation and Valiant's notion of counting complexity classes seem most natural. To capture the algebraic and numeric nature of partition functions as real or complex valued functions, the Blum-Shub-Smale (BSS) model of computation seems more natural. As a result many papers use a naive hybrid approach in discussing their complexity or restrict their considerations to sub-fields of $\mathbb{C}$ which can be coded in a way to allow dealing with Turing computability.

In this paper we propose a unified natural framework for the study of computability and complexity of partition functions and graph polynomials and show how classical results can be cast in this framework.

3.19 Lower bounds for restricted arithmetic computations

Guillaume Malod (University Paris-Diderot, FR)

Joint work of Dvir, Zeev; Malod, Guillaume; Perifel, Sylvain; Yehudayoff, Amir
Main reference Z. Dvir, G. Malod, S. Perifel, A. Yehudayoff, "Separating multilinear branching programs and formulas," in Proc. of the 44th Symp. on Theory of Computing Conference (STOC'12), pp. 615–624, ACM, 2012.
URL http://dx.doi.org/10.1145/2213977.2214034

Valiant's theory contains classes VPe, VPws, VP, VNP which we can use to classify the complexity of polynomials such as the permanent or the determinant. As in other areas of computational complexity, no class separation is known. I will present several lower bound results: Nisan's beautiful lower bounds for non-commutative computations, Raz's multilinear

techniques, and an application separating formulas from branching programs, from a joint work with Dvir, Perifel and Yehudayoff.

3.20 Approximating Holant problems

Colin McQuillan (University of Liverpool, GB)

Main reference C. McQuillan, "Approximating Holant problems by winding," arXiv:1301.2880v1 [cs.CC].
URL http://arxiv.org/abs/1301.2880v1

I will discuss the complexity of approximately evaluating Holant problems. I will present a hardness result for Holant problems with edge weights, and an FPRAS for Holant problems with even, odd, and not-all-equal constraints [1].

References

1 Colin McQuillan, *Approximating Holant problems by winding*, CoRR (2013), abs/1301.2880.

3.21 Structural Tractability of Counting of Solutions to Conjunctive Queries

Stefan Mengel (Universität Paderborn, DE)

Joint work of Durand, Arnaud; Mengel, Stefan;
Main reference A. Durand, S. Mengel, "Structural Tractability of Counting of Solutions to Conjunctive Queries," accepted for the 16th Int'l Conf. on Database Theory (ICDT2013).
URL http://arxiv.org/abs/1303.2059

Conjunctive queries (CQs) are a fundamental class of logical queries. The corresponding decision problem consist of evaluating an existential conjunctive first-order formula over a finite structure. This is equivalent to answering Select-Project-Join queries in database theory and has several equivalent definitions, in particular, in terms of constraint satisfaction or homomorphism problems. Generally, answering CQs is NP-hard, but there has been huge progress in identifying so-called 'islands of tractability', i.e. subclasses of queries which can be evaluated in polynomial time.

I will survey recent joint results with Arnaud Durand on the complexity of counting solutions to CQs and discuss how the situation differs from #CSP.

3.22 The complexity of approximating conservative counting CSP

David Richerby (University of Liverpool, GB)

Joint work of Chen, Xi; Dyer, Martin; Goldberg, Leslie Ann; Jerrum, Mark; Lu, Pinyan, McQuillan, Colin; Richerby, David
Main reference X. Chen, M. Dyer, L.A. Goldberg, M. Jerrum, P. Lu, C. McQuillan, D. Richerby, "The complexity of approximating conservative counting CSPs," arXiv:1208.1783v2 [cs.CC].
URL http://arxiv.org/abs/1208.1783v2

We discuss the complexity of approximately solving the weighted counting constraint satisfaction problem (#CSP) in the conservative case, where all unary weight functions are assumed

to be available. We define the notions of weak log-modularity and weak log-supermodularity. For this presentation, we restrict to constraint languages with at-most binary functions. If such a constraint language F is weakly log-modular, then #CSP(F) can be solved exactly in polynomial time; otherwise, if it is weakly log-supermodular, then #CSP(F) is equivalent to the problem #BIS of counting independent sets in bipartite graphs; otherwise, #CSP(F) is NP-hard to approximate.

3.23 Effective De Rham Cohomology

Peter Scheiblechner (Hochschule Luzern, CH)

Main reference P. Scheiblechner, "Effective de Rham cohomology: the hypersurface case," in Proc. of the 37th Int'l Symp. on Symbolic and Algebraic Computation (ISSAC'12), pp. 305–310, ACM, 2012.
URL http://dx.doi.org/10.1145/2442829.2442873

A long standing open problem in computational algebraic geometry is to find an algorithm which computes the topological Betti numbers of a semialgebraic set in single exponential time. There has been recent progress on the corresponding problem over the complex numbers. A fundamental theorem of Grothendieck states that the Betti numbers of a smooth complex variety can be computed via its algebraic de Rham cohomology, which is defined in terms of algebraic differential forms on the variety. In this talk, we discuss single exponential degree bounds on these differential forms and their importance for the algorithmic computation of Betti numbers.

References

1 P. Scheiblechner. Effective de Rham cohomology – the Hypersurface Case. Proc. ISSAC 2012. arXiv:1112.2489v1.
2 P. Scheiblechner. Effective de Rham cohomology – the General Case. submitted, 2012. arXiv:1203.5706v1.

3.24 Inapproximability of the Partition Function for the Spin Models (Antiferromagnetic Ising, Hard-Core Models, and more)

Daniel Stefankovic (University of Rochester, US)

Recent inapproximability results of Sly (2010), together with an approximation algorithm presented by Weitz (2006) establish a picture for the computational complexity of approximating the partition function of the hard-core model. Let $L_c(T_D)$ denote the critical activity for the hard-model on the infinite D-regular tree. Weitz presented an FPTAS for the partition function when $L < L_c(T_D)$ for graphs with constant maximum degree D. In contrast, Sly showed that for all $D >= 3$, there exists $\epsilon > 0$ such that (unless RP=NP) there is no FPRAS for approximating the partition function on graphs of maximum degree D for activities L satisfying $L_c(T_D) < L < L_c(T_D) + \epsilon$.

We prove the complementary result that for the antiferromagnetic Ising model without external field that, unless RP=NP, for all $D >= 3$, there is no FPRAS for approximating the partition function on graphs of maximum degree D when the inverse temperature lies

in the non-uniqueness regime of the infinite tree T_D. Our results extend to a region of the parameter space for general 2-spin models. Our proof works by relating certain second moment calculations for random D-regular bipartite graphs to the tree recursions used to establish the critical points on the infinite tree. We will also report on the progress for multi-spin systems.

3.25 Counting Arbitrary Subgraphs in Data Streams

He Sun (MPI für Informatik – Saarbrücken, DE)

Joint work of Kane, Daniel; Mehlhorn Kurt; Sauerwald Thomas
Main reference D. Kane, K. Mehlhorn, T. Sauerwald, H. Sun, "Counting Arbitrary Subgraphs in Data Streams," in Pro. of the 39th Int'l Colloquium on Automata, Languages and Programming (ICALP'12), LNCS, Vol. 7392, pp. 598–609, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-31585-5_53

We study the subgraph counting problem in data streams. We provide the first non-trivial estimator for approximately counting the number of occurrences of an *arbitrary* subgraph H of constant size in a (large) graph G. Our estimator works in the turnstile model, i.e., it can handle both edge-insertions and edge-deletions and is applicable in a distributed setting. Prior to this work, estimators were known for only a few non-regular graphs in the case of edge-insertions, leaving the problem of counting general subgraphs in the turnstile model wide open. Further we demonstrate the applicability of our estimator by analyzing its concentration for several graphs H and the case where G is a power law graph.

3.26 What are the Chances that P=NP?

Leslie Valiant (Harvard University, US)

Main reference L.G. Valiant, "Some observations on holographic algorithms," in Proc. 9th Latin American Theoretical Informatics Symp. (LATIN'10), LNCS, Vol. 6034, pp. 577–590, Springer, 2010.
URL http://dx.doi.org/10.1007/978-3-642-12200-2_50

We argue that while there has been substantial recent progress in the theory of holographic algorithms, these results have not removed the possibility that holographic methods might be sufficient to yield polynomial time algorithms for all of #P. They do, however, exclude some possibilities. Known negative results dichotomize into those that are representation dependent, such as the Cai-Lu collapse theorem, and those that are not. In pursuit of the latter, we discuss the notion of an "elementary holographic transformation to matchgrids," which is sufficient to express many of the known polynomial time holographic algorithms. We discuss a lower bound argument that shows that #SAT, the Boolean satisfiability counting problem, cannot be solved by such an elementary transformation. The constraints on elementarity can be evaded by using interpolation from many individual elementary transformations, or by using exponential size fields. We give examples of such elementarity evasive algorithms for various parity problems related to graph coloring, connected independent sets and forests.

3.27 Analysis of message-passing iterative decoders via zeta functions

Pascal Vontobel (HP Labs – Palo Alto, US)

Joint work of Pfister, Henry; Vontobel, Pascal
Main reference to be submitted to IEEE Trans. Inf. Theory

Graph-based codes and message-passing iterative decoders have become increasingly popular in the last fifteen years. It is fair to say that these codes and decoding algorithms (and ideas related to them) have thoroughly changed much of modern communications. Before this backdrop, a good understanding of these types of communication techniques is obviously highly desirable, especially the understanding of iterative decoding of finite-length codes.

We will focus on a particular message-passing iterative decoder, the so-called sum-product algorithm (SPA) decoder. As the name SPA suggests, this algorithm is counting certain objects, and, indeed, when the underlying factor graph has no cycles then it is clear what the SPA is counting. However, when the underlying factor graph has cycles then the situation is much fuzzier.

In this talk we use graph zeta functions for analyzing the SPA decoder. This approach allows us to connect and count computation tree pseudo-codewords and graph-cover pseudo-codewords, two central objects in the SPA analysis.

3.28 The Complexity of Planar Boolean #CSP with Complex Weights

Tyson Williams (University of Wisconsin – Madison, US)

Joint work of Williams, Tyson; Guo, Heng
Main reference H. Guo, T. Williams, "The Complexity of Planar Boolean #CSP with Complex Weights," arXiv:1212.2284v1 [cs.CC].
URL http://arxiv.org/abs/1212.2284v1

We prove a complexity dichotomy theorem for symmetric complex-weighted Boolean #CSP when the constraint graph of the input must be planar. The problems that are #P-hard over general graphs but tractable over planar graphs are precisely those with a holographic reduction to matchgates. This generalizes a theorem of Cai, Lu, and Xia for the case of real weights. We also obtain a dichotomy theorem for a symmetric arity 4 signature with complex weights in the planar Holant framework, which we use in the proof of our #CSP dichotomy. In particular, we reduce the problem of evaluating the Tutte polynomial of a planar graph at the point $(3, 3)$ to counting the number of Eulerian orientations over planar 4-regular graphs to show the latter is #P-hard. This strengthens a theorem by Huang and Lu to the planar setting. Our proof techniques combine new ideas with refinements and extensions of existing techniques. These include planar pairings, the recursive unary construction, the anti-gadget technique, and pinning in the Hadamard basis.

3.29 Dichotomy for Holant* Problems with Domain Size 3

Mingji Xia (MPI für Informatik – Saarbrücken, DE)

Joint work of Cai, Jin-Yi; Lu, Pinyan
Main reference J.-Y. Cai, P. Lu, M. Xia, "Dichotomy for Holant* Problems with a Function on Domain Size 3," arXiv:1207.2354v1 [cs.CC].
URL http://arxiv.org/abs/1207.2354v1

This talk is about the complexity of Holant* problems defined by one symmetric ternary function in variables from the domain of size 3. Holographic reductions are used for both algorithms and #P-hardness proofs. Three kinds of tractable problems are introduced one by one following the corresponding cases of Boolean domain Holant* problems. Two segments of #P-hardness proof are shown in details, because the second one is closely related with the discovery of the third kinds of tractable problems. One is realizing a binary function of rank 2, and the other is realizing a binary function of rank 2 such that its eigenvector corresponding to eigenvalue 0 is not isotropic.

3.30 Approximation Classification of Complex-Weighted Counting CSPs

Tomoyuki Yamakami (University of Fukui, JP)

We deal with counting constraint satisfaction problems, which are particularly composed of complex-valued constraints over Boolean variables. We show a complete classification theorem on the computational complexity of approximately solving those problems when auxiliary unary constraints are freely available besides input constraints. To clarify the roles of extra free unary constraints, we then present an alternative proof of the classification theorem in which unary constraints are required only at the very end of its argument. This talk is based on recent papers [1, 2].

References

1 Tomoyuki Yamakami. Approximate counting for complex-weighted Boolean constraint satisfaction problems. Information and Computation, 219 (2012) 17–38.
2 Tomoyuki Yamakami. Constant unary constraints and symmetric real-weighted counting CSPs. In: Proc. of the 23rd International Symposium on Algorithms and Computation (ISAAC 2012), Lecture Notes in Computer Science, Springer-Verlag, vol.7676 pp. 237–246, 2012.

3.31 Approximate Counting via Correlation Decay on Planar Graphs

Chihao Zhang (Shanghai Jiao Tong Univ., CN)

Joint work of Yin, Yitong; Zhang, Chihao
Main reference Y. Yin, C. Zhang, "Approximate Counting via Correlation Decay on Planar Graphs," arXiv:1207.3564v1 [cs.DS].
URL http://arxiv.org/abs/1207.3564v1

We show for a broad class of counting problems, correlation decay (strong spatial mixing) implies FPTAS on planar graphs. The framework for the counting problems considered by us is the Holant problems with arbitrary constant-size domain and symmetric constraint functions. We define a notion of regularity on the constraint functions, which covers a wide range of natural and important counting problems, including all multi-state spin systems, counting graph homomorphisms, counting weighted matchings or perfect matchings, the subgraphs world problem transformed from the ferromagnetic Ising model, and all counting CSPs and Holant problems with symmetric constraint functions of constant arity.

The core of our algorithm is a fixed-parameter tractable algorithm which computes the exact values of the Holant problems with regular constraint functions on graphs of bounded treewidth. By utilizing the locally tree-like property of apex-minor-free families of graphs, the parameterized exact algorithm implies an FPTAS for the Holant problem on these graph families whenever the Gibbs measure defined by the problem exhibits strong spatial mixing. We further extend the recursive coupling technique to Holant problems and establish strong spatial mixing for the ferromagnetic Potts model and the subgraphs world problem. As consequences, we have new deterministic approximation algorithms on planar graphs and all apex- minor-free graphs for several counting problems.

Participants

- Markus Bläser
Universität des Saarlandes, DE
- Magnus Bordewich
University of Durham, GB
- Irénée Briquel
Université d'Orléans, FR
- Peter Bürgisser
Universität Paderborn, DE
- Andrei A. Bulatov
Simon Fraser University – Burnaby, CA
- Xi Chen
Columbia University, US
- Radu Curticapean
Universität des Saarlandes, DE
- Holger Dell
University of Wisconsin – Madison, US
- Arnaud Durand
University Paris-Diderot, FR
- Martin Dyer
University of Leeds, GB
- Jo Ellis-Monaghan
Saint Michael's College – Colchester, US
- Hervé Fournier
University Paris-Diderot, FR
- Andreas Goebel
University of Liverpool, GB
- Leslie Ann Goldberg
University of Liverpool, GB
- Bruno Grenet
ENS – Lyon, FR
- Heng Guo
University of Wisconsin – Madison, US
- Miki Hermann
Ecole Polytechnique – Palaiseau, FR
- Thore Husfeldt
IT Univ. of Copenhagen, DK
- Mark Jerrum
Queen Mary University of London, GB
- Marek Karpinski
Universität Bonn, DE
- Pascal Koiran
ENS – Lyon, FR
- Meena Mahajan
The Institute of Mathematical Sciences – Chennai, IN
- Johann A. Makowsky
Technion – Haifa, IL
- Guillaume Malod
University Paris-Diderot, FR
- Colin McQuillan
University of Liverpool, GB
- Kitty Meeks
Queen Mary University of London, GB
- Stefan Mengel
Universität Paderborn, DE
- Mike S. Paterson
University of Warwick, GB
- Natacha Portier
ENS – Lyon, FR
- David Richerby
University of Liverpool, GB
- Peter Scheiblechner
Hochschule Luzern, CH
- Uwe Schöning
Universität Ulm, DE
- Daniel Stefankovic
University of Rochester, US
- Yann Strozecki
University of Versailles, FR
- He Sun
MPI für Informatik – Saarbrücken, DE
- Sebastien Tavenas
ENS – Lyon, FR
- Leslie Valiant
Harvard University, US
- Pascal Vontobel
HP Labs – Palo Alto, US
- William Whistler
University of Durham, GB
- Tyson Williams
University of Wisconsin – Madison, US
- Mingji Xia
MPI für Informatik – Saarbrücken, DE
- Tomoyuki Yamakami
University of Fukui, JP

Report from Dagstuhl Seminar 13041

Civilian Crisis Response Models

Edited by

Bernhard Katzy[1] and Ulrike Lechner[2]

1 Leiden University, NL, bkatzy@liacs.nl
2 Universität der Bundeswehr München, DE, Ulrike.Lechner@unibw.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13041 "Civilian Crisis Response Models".

The vulnerability of modern societies to the threats of man made and natural disaster increases and scale and number of disasters are expected to rise. The earthquakes of Haiti with its subsequent Cholera epidemics, the natural disasters in Pakistan as well as the ongoing situation in Japan illustrate the need for effective and efficient crisis and disaster response organizations as well as humanitarian aid organizations in developing and first world countries. Disaster preparedness is a key to effectiveness and efficiency in case of crisis or disaster – but we observe that natural and human disasters are too often beyond what is being planned for.

There is a need for new and better approaches in disaster and crises response and humanitarian aid. There is a need for well designed systems as well as for models, methods, instruments and tools for analysis and decision making. This Dagstuhl Seminar is motivated by the fact that computer science is an enabler for the changes and should contribute to the body of scientific knowledge and instruments and tools alike.

The Seminar discussed approaches to Crisis Response from a variety of disciplines. In a workshop like setting with talks, panels and discussions, seminar participants worked on a common understanding of crisis and crisis response, characteristics of crisis situations and crisis response and research topics on crisis management. The participants developed on a research agenda for Networked Civilian Crisis Response Models.

Seminar 20.–25. January, 2013 – www.dagstuhl.de/13041
1998 ACM Subject Classification K.4 Computers and Society, K.4.m Miscellaneous, J.7 Computer in Other Systems: Command and Control
Keywords and phrases Crisis Response, Humanitarian Aid
Digital Object Identifier 10.4230/DagRep.3.1.67

1 Executive Summary

Ulrike Lechner
Bernhard Katzy

The vulnerability of modern societies to the threats of man made and natural disaster increases and scale and number of disasters are expected to rise. The earthquakes of Haiti with its subsequent Cholera epidemics, the natural disasters in Pakistan as well as the ongoing situation in Japan illustrate the need for effective and efficient crisis and disaster response organizations as well as humanitarian aid organizations in developing and first world countries. Disaster preparedness is a key to effectiveness and efficiency in case of crisis or disaster – but we observe that natural and human disasters are too often beyond what is being planned for.

Civilian Crisis Response Models, *Dagstuhl Reports*, Vol. 3, Issue 1, pp. 67–93
Editors: Bernhard Katzy and Ulrike Lechner
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

There is a need for new and better approaches in disaster and crises response and humanitarian aid. Think of IT-systems and how well designed systems can help or think of what science can contribute in terms of models, methods, instruments and tools for analysis and decision making. This Dagstuhl Seminar is motivated by the fact that computer science is an enabler for the changes and should contribute to the body of scientific knowledge and instruments and tools alike. This seminar on crisis response Models aims to make a contribution to the systematic development of a body of scientific knowledge for crisis and disaster response and Humanitarian Aid organizations. We invite researchers and practitioners in the field of humanitarian aid and crisis and disaster response as well as researchers in computer science and related disciplines to this Dagstuhl Seminar on Civilian Crisis Response Models. We address with this seminar on crisis response models questions concerning the design of systems in crisis and disaster response and humanitarian aid. Currently, there is a window of opportunity for redesigning the crisis response system as the proliferation of mobile phones, smart phones and social software facilitate novel services and new Command and Control (C2) systems allows for new designs. Many examples demonstrate the increasing use of social media in emergencies: For human and man-made disasters websites and Internet services are created to support the inflicted population as well as the aid organizations. A popular and successful example is Ushahidi, a NGO developed platform in response to civil war in Kenya 2008 mapping incidents of violence. In the ongoing crisis in Japan, Twitter and Facebook messages were compiled to provide guidance of what kind of help is needed. Web services are used widespread to locate missing persons. "Google Crisis" provides its set of services to be deployed in case via the Google website.

These systems, many of which have been created ad-hoc by volunteers illustrate the feasibility of better information systems in crisis response management. In many cases, they turned out to be efficient, precise and easy to operate. From these services, evaluation towards a permanent information system is needed. These novel systems illustrate the need for good governance and the need to analyze and reconsider the whole disaster response system with its information flows. What is the impact of the use of such systems in case of a disaster on communication, logistics, the behavior of the population and the aid organizations? Again, scientific methods eventually might be useful to build new systems and develop new processes and strategies.

With this Dagstuhl Seminar on Civilian Crisis Response Models we go beyond the design of technology and aims at contributing to the scientific body of knowledge of crisis and disaster response and Humanitarian aid. Disaster preparedness is the area in the field of crisis and disaster management that requires well developed, evidence-based quantitative models and theories to feed and guide the simulations, optimizations, serious games, analytical methods, architectures and process models, creative techniques or case studies. Disaster preparedness requires its body of scientific knowledge to be used for exploring disaster preparedness, for building IT-systems, for assessing humanitarian aid and disaster response organizations and for guiding the necessary changes in the crisis response system to adopt it to new threats and new scenarios. Methods and models are crucial for making better decisions in tight financial situations.

The seminar addressed the needs and solution options in a systematic way. The report documents on the presentations, the panel discussions and various workshop and working group sessions and their results.

The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 317382, NITIMesr.

2 Table of Contents

3 Overview of Talks

3.1 Sorting enabling technologies for risk analysis and crisis management

Ivo Häring (Fraunhofer Ernst-Mach-Institut – Efringen-Kirchen, DE)

The talk presented an approach to define crisis response research in terms of relevant categories (aspects, dimensions, properties) for such research, including e.g. types of threats (hazard sources, crisis triggering events), crisis management steps which are supported, hardware technologies employed, disciplines involved or software technologies used. If existing crisis research is sorted into such a categorization scheme, foci of on-going research efforts can be identified as well as missing combinations. Such definition by examples and attribute ranges also avoids exclusions and is open for a truly interdisciplinary (trans-disciplinary) approach.

Besides this top level approach, examples of emerging application tools were given that implement risk management or analysis steps. All tools are currently developed within EU FP7 research projects and at least parts of their functionalities are relevant for crisis management. In BESECURE we apply risk management to structure an urban security enhancement process. It gives access to best practice methods for enhancing security, urban attractiveness and performance. In a similar way in cases of crisis methods can be selected in a systematic way and employed for achieving user-defined objectives. In VITRUV we show how urban planning for countering terrorism can be supported by empirical (data-base driven) and quantitative analysis. Taking measures for susceptibility, vulnerability and risk into account the software supports urban planning at plan and detail level. This allows countering crisis by preventive urban design with focus on non-physical measures, in particular rearrangement of geometries and urban space design. In ENCOUNTER risk analysis of explosive improvised (home-made) devices in urban environments is conducted taking into account organizational and physical counter measures including neutralization and removal options. This provides scenario assessment input for crisis management. In D-BOX we contribute to a toolbox for improving humanitarian demining worldwide. In particular quantitative hazard and damage analysis is applied to this domain. Also interactive databases are provided on hazards sources, neutralization, removal and personal protective equipment. Mine fields are at the core of many (long- term) crisis-like developments in developing countries.

For defining a roadmap for crisis research a step-wise approach is proposed. The steps are conducted in a very exemplary and incomplete way. First, we give three examples for step-wise informed schemes for risk analysis schemes, risk management and crisis management schemes are given, respectively. Second a list of technological and societal enablers is given. Third it is shown how this enablers are expected to influence the respective risk analysis steps, with respect to short-term, medium term and long term effects, respectively. It is indicated how in a similar way the effects of the enablers on risk management and crisis management, e.g. using the residence circle, can be conducted. Finally we provide a summarizing schematic.

3.2 Towards a framework for the conceptualization of command post exercises — an action research approach with staffs of disaster response organizations

Erich Heumüller (Universität der Bundeswehr – München, DE)

Joint work of Erich Heumüller, Sebastian Richter, and Ulrike Lechner

The goal of our research is the support of command post exercises. We develop a framework to support conceptualization of command post exercises emphasizing exercise goal-achievement and exercise evaluation. The presentation at the Dagstuhl Seminar Civilian Crisis Response Models contains the framework with a classification of exercises, a framework for goal-oriented and evaluation-driven exercise conceptualization. The focus lies on the conceptual model of staffs that is the basis for a systematic exercise evaluation. This model is based on scholarly literature on teams and leadership. It analyzes staff processes and has as constructs Resource Management, Decision, Responder, Information Management, Task Coordination and Commander's Intent. The method follows an Canonical Action Research approach with an empirical basis of four command post exercises.

3.3 Usage of Social Media in Crisis Situations

Nicole Krämer (Universität Duisburg-Essen, DE)

Joint work of Nicole C. Krämer, German Neubaum, Leonie Rösner, Astrid Rosenthal-von der Pütten, and Jennifer Klatt

Nowadays, every major crisis is accompanied by massive activity in social media. Against this background the presentation discussed whether and by which means social media (such as e.g. blogs or social networking sites) can support governmental efforts to avoid or reduce the impact of major hazards on the public. From a media psychological as well as social psychological perspective it has to be analyzed which motives the public has for using social media during and after catastrophic risks and what effects can be expected from using social media for crisis communication. With regard to the latter, it is important to – on the one hand – understand how the public's discussion in social media might change their attitudes towards the crisis and whether the dynamics lead to beneficial or unfortunate effects. On the other hand, it has to be analyzed whether governmental institutions can benefit from using social media platforms for addressing the public (given that recent data show that especially younger audiences and target groups tend to turn to the Internet instead of the TV in order receive timely and authentic information). Here, empirical data have to show a) how governmental institutions can present themselves in a trustworthy, believable way within social media platforms and b) which steps have to be taken before a major hazard in order to be able to communicate e.g. within a social networking site when a crisis starts. Additionally, it can be tested whether governmental institutions might be able to use the information that is discussed within social media applications in order to support immediate rescuing (e.g. when information on potentially threatened victims are posted as it was the case during Katrina) or to monitor the public´s fears, attitudes and intended behaviors.

In the presentation, media psychological and social psychological theories were presented which can help to understand the processes. Also, a multimethod study on the usage of a social networking site during and after the Love Parade stampede in Duisburg 2010 (Neubaum et al., 2012) was presented and discussed. Results showed that social media usage fulfilled various needs and functions such as information seeking as well as emotional regulation.

References

1 Neubaum, G., Rösner, L., Presting, P., Muraa, G., von der Pütten, A.M., and Krämer, N.C. (2012). *The Role of Social Media Usage related to the Stampede at the Love Parade 2010.* Paper at the Conference of the International Communication Association 2012, Phoenix, AZ, USA.

3.4 Geoinformation and disaster management

Wolfgang Reinhardt (Universität der Bundeswehr – München, DE)

In the presentation first an overview on relevant national and international programs as well as on activities of the authors group is given. After that the "phases-approach" to disaster management is discussed and some of the used phases models are outlined. After that the usage of Geographic Information (GI) for the prevention and preparedness phases is introduced by means of examples. It is emphasized that especially Geo Web Services play an important role here which is illustrated by examples from the Alpine Space Project "TranSafe-Alp". The main message of the presentation is to demonstrate that GI plays an important role — among other factors — for all phases of disaster management.

3.5 End User Perspective

Heiko Werner (Bundesanstalt Technisches Hilfswerk, DE)

Disaster Management in Germany is a federal business. Across the world, the structure of THW is unique: As a Federal agency, THW belongs to the department of the Federal Ministry of the Interior. However, only one percent of the staff works full-time for the authority. 99 percent of the THW- members work on a voluntary basis for THW. Nationwide more than 80,000 volunteers commit themselves during their leisure time in 668 local sections in order to provide professional help to people in distress. THW flexibly adapts its structures to changing threat situations. Modern equipment and well-trained specialists are the basis of its high efficiency. Security Research is a key element in enabling THW to prepare for future challenges.

3.6 Navigation Support with Landmarks: A Design Case Study

Volker Wulf (Universität Siegen, DE)

The presentation comprises the Design Oriented Research method and various design cases on the design of IT for fire fighters as well as empirical work on the use of social media in the so-called Arab Spring.

References

1 L. Ramirez, T. Dyrks, J. Gerwinski, M. Betz, M. Scholz, V. Wulf (2011). *Landmarke: and ad hoc deployable ubicomp infrastructure to support indoor navigation of firefighters.* Journal Pers. Ubiquit Comput, 2011.

2 V. Wulf, K. Misaki, M. Atam, D. Randall, M. Rohde (2013). *On the Ground in Sidi Bouzid: Investigating Social Media Use during the Tunisian Revolution* CSCW'13, February 23–27, 2013, San Antonio, Texas, USA., ACM, 2013.

4 Plenary

4.1 Civilian Crisis Response Models – Panel Discussion

Ozgur Dedehayir (Tampere University of Technology, FI)

Joint work of Ozgur Dedehayir, Simon French, Bernhard Katzy, Dietmar Kühne, Nils Weidmann

The Setting

Dagstuhl seminars are aimed to bring together leading minds in connected fields of research to generate new, potentially ground breaking pathways in scientific endeavor. The seminar on "Civilian Crisis Response Models" similarly brought together experts from different yet related fields of investigation pertaining to crisis management. Commensurate with the ideology and objectives of Dagstuhl meetings, a panel discussion was organized to establish a consensus on how the group could work together in a trans-disciplinary project. Four panelists with prior experience in such projects were invited to take center stage, including Professor Simon French from the University of Warwick, Professor Bernard Katzy from Leiden University, Mr. Dietmar Kühne from the German Army / Landeskommando Bayern, and Dr. Nils Weidmann from the University of Konstanz. The panel discussion was held on Tuesday 22nd January (from 16:00 until 18:00).

Prior to the panel discussion, the seminar participants had witnessed presentations from several colleagues with respect to their own research agendas, and also took part in workshops designed to align the thinking of participants towards a roadmap of collaborative research. However, these prior events had demonstrated that while seminar participants had come from connected research realms, there were differences on how they approached the 'crisis' phenomena. Different theoretical models were used as lenses through which the phenomena was viewed, and a variety of methodologies were employed in systematically studying the phenomena at hand. Nevertheless, the potential of reaching synergy was also evident and the panel discussion was aimed to unite the participants' thinking towards a path of synergy and future collaboration.

The Questions

Three overarching questions were posed to the panel discussants. The first question centered on what participants could learn from prior experience in trans-disciplinary projects. More specifically, the panelists were asked to comment on their own experiences concerning the pitfalls that should be avoided as well as the issues that should be emphasized or underlined as vital to successive collaboration. Secondly, the panelists were invited to comment on matters concerning the contribution of the project to crisis management science. Specifically, the panelists were asked to comment on the novel outcomes that they could foresee future research attaining, possible research gaps that could form the foci of such research, and the unique methodologies that could aid researchers reach these outcomes. And thirdly, the panelists were asked to remark on the next steps that could be taken for a roadmap of collaborative research.

Perspectives on Civilian Crisis Response Models:

1. The social dimension of "tacit knowledge" of a new research field
   The panel firstly proposed that opportunities to participate in multidisciplinary projects should not be relinquished. Furthermore, they underlined that successful projects generally have a "glue person" that binds the connected yet separate research streams, and that the participants of these projects are recommended to respect the contribution of other fields of investigation on common phenomena. Hence, according to the panel, although difficult to generalize, the starting point rests on identifying the phenomena that need to be analyzed. Moreover, for such an interdisciplinary group, establishing the common grounds or elements is pertinent.
2. Shared research questions
   The justification for the collaborative research agenda seemed to be clear to some participants, considering that the EU commission has recently awarded funding to the NITIM project which aims to study crisis management networks and processes, with an emphasis on key performance indicators such as faster recovery and reduced damage, in the face of a higher number of disasters. Nevertheless, it became apparent from the dialogue of the panelists that the unifying research themes or questions remained masked at this point of the seminar. A set of common expectations and coherent group ideas would pave the way towards more concrete research goals. At this point some of the research themes may include the role of entrepreneurs in crisis management, leadership and the performance of organizations, and the role of technology in new ways of reacting to crises, however, these would require agreement among all partners. The lack of clarity as to the seminar objectives was also expressed by some audience members, who enquired of the central ideas that bound them together. Furthermore, it was also proposed that different research projects and teams were already active in a similar manner across Europe and USA, and therefore the identification of the raison d'être of this seminar was brought forward.
3. Typologies of Crisis and relation to existing work
   The possibility of developing a typology of crises to help the seminar focus its research agenda and to help participants see the types of crisis for which knowledge exists and others where there is no or little knowledge, was subsequently considered. Despite the difficulty of generating such a typology (for example identifying the dimensions that could be used remained ambiguous) and the fact that several attempts had already been made, an all-encompassing framework of types was deemed to be an important contribution to the field of research. Concerning research outcomes, the panelists proposed that the most important point is relevance for society. They also suggested that the novelty of

the research would naturally depend on the outlets that would be targeted; although the panelists agreed that unifying outlets that disseminate project findings would likely to be a futile task. It was recognized by the audience that a number of journals already existed in this field, potentially making novel contribution difficult. However, the role of technological development and its influence on crisis management, e.g. making redundant the role of the government as manager of the crisis, and paving the way for new governance structures, which were highlighted as under-researched themes in these journals.

4. "Technology Push" for social innovation in crisis management
   From the dialogue between the panel and the audience a unifying research theme began to emerge in the form of 'technological factors', i.e. technology enhanced communication at various levels that involve different actors, which can act as tools that help citizens in times of crisis but also pose risks in crisis issues. For instance, the use of social media can help citizens understand what is happening in crises situations. At the same time researchers can acquire data from social media to understand how people actually behave during crises. Nevertheless, the need to narrow the focus of this research theme, towards specific research questions also emerged as a necessary step. An additional though related research theme was proposed to be the leadership or governance matters in crisis situations, especially when influenced by technological factors. Some specific questions that were posed under this theme included: what are the types of leader that emerge in crisis situations? And: how can successful entrepreneurs be trained to be leaders in such situations? With respect to the methodological considerations, panelists were divided over the usability of several crisis cases to derive concrete questions for investigation.

Overall, the seminar group made significant progress through the panel discussion concerning unifying research themes, means of making scientific contribution, and employable methodologies. In the following days, these topics would be honed to arrive at more tangible research questions and suitable theoretical methodologies.

Panel Discussion Notes

- Relevant for society? Identifying the phenomena- How do you go about solving this? Difficult to give general recommendation. Where do I want to get my work published? Up front need to establish the common elements.
- Standardization of interfaces, Software Structure
- Justification for bringing together this project, There is money.Agree with Nils that we have a relevant question. Go from "technology push" to "demand pull", hence, search for a unifying need, More disasters, engage more people, recover is faster, damage is lower (key performance indicators), Unifying question based on these, Joint question that we all believe in. Can we standardize the outlets for our work? not possible. But creating impact is about just being visible, Ability to find common points in discussion, Fragmentation is a problem
- What makes successful multi-disciplinary research. "success is going from one failure to another without losing enthusiasm"- Respect the other disciplines, understand where their theories are coming from and vice versa, Have some people to act as glue from different disciplines, Don't design multi-disciplinary research, you just jump on when they come, time, money, assessment? these are critical attributes for any project
- Need the explicit idea behind this, need to focus on expectations, more coherent group and ideas needed, good outcomes will come out but how?

- More combinations between technological and social
- What is the question we deal with? ("how do we react to crises?"), the role of entrepreneurs is important, leadership and performance of organizations, the role of technology in new ways of reacting to crises
- Do we need a typology of crises? lots of types out there, presence of amateurism which acts as a hindrance, the politicians are perhaps the biggest amateurs in times of crises, be very clear on the types of crises we are looking at, come up with a topography (nobody has been able to come up with)
- Start with a typology – this helps us to see for which types of crises there is knowledge and others where there is none or little
- Crisis could be seen as not a bad thing but an opportunity to change an existing system, not aware of an existing topography, could be based on personal harm
- Database exist? of scenarios?
- In the area of terrorism there are some, especially in USA, department of security, this database has been used to extract frequency of attacks, for example, or scenario generation, skeptical about real content but still used (you can say "it has happened") – get an idea of trends
- In the area of political sciences there exist macro and micro databases (e.g. in a given conflict what events took place), we need to define key concepts – cant see how things will work without key definitions (e.g. crises)
- Not sure if definition is helpful, puzzled with what brings us together – no need for yet another group to tackle this crisis issue, what is it that binds us together? many journals exist already, but one way of creating "impact" would be the factor of technological development, and the government not being the crisis of manager anymore (given technological developments) – hence, new governance structures
- Social systems are coupled more so than before
- Would be beneficial to focus on technological factors that can be tools that help us but also pose risks in crisis issues, need to narrow down the focus
- Leadership is changing, technological drivers which is enabling so called "leadership" or governance in crisis situations, what are the type of leaders that are emerging in crises? can we train successful entrepreneurs are possible leaders in such situations?
- What are the catch words, headline, short abstract of project? if we talk about research roadmap, we need the above
- A concrete question can be formulated by investigating an interesting theme across different crisis cases
- Not able to make conclusions from study of the different cases, agree with Prof. Bakker about the technological factors that help us understand crises management, but leadership issues are not viable, need to focus: on a unifying question and common elements for the group at the outset, conceptual definition (perhaps typology), find topical niche (leadership change in technological change and the role of entrepreneurs)
- Either to learn about what is happening out there (in crisis management), or then come up with new strategies to deal with the issues
- Technological development is the unifying theme, considering that we are in an interdisciplinary group, this is perhaps more relevant than a particular question
- Technology enhanced communication at various levels, new actors are coming into the picture, use of social media can help understand what is happening in crises situations (people actually tell us what they need), technology enhanced communication at various levels that involve different actors

5 Workshops on Research Topics for Crisis Response

The first workshop series was about identifying research topics and discussing networked crisis response from various perspectives. The structure of topics follows the structure of the NITIM project Networked Crisis Management with five topics

1. Crisis Network Management and Governance
2. Communication and collaboration infrastructures for crisis management
3. Coordination and collaboration in heterogeneous actor networks
4. Humanitarian, crisis response logistics
5. Engagement models for entrepreneurs and volunteers

The workshop participants got questions to guide the discussion:

- Phenomenology: What are the phenomena being studied?
- Theories: What are the models and theories used?
- Integration: What are the interfaces and relations to the other topics?
- Education: What do we need to teach?
- What is being done? What to do to be prepared? What to rethink and reconceptualize?
- What are the 3–5 most important topics for a researcher?

5.1 Workshop Humanitarian Crisis Response Logistics

Ulrike Lechner (Universität der Bundeswehr – München, DE)

Logistics is Crisis Management on a day-to-day basis. What distinguishes a crisis from day-to-day business? The workshop draws is motivation from a discussion around business continuity, automatization of computing centers, cases contributed by the workshop participants. Cases that the participants mentioned were Edeka and its role in the crisis response after the Elbe-Flut disaster, and various examples for exercises and logistics in disasters at county level.

The involvement of volunteers and response forces was one topic that was discussed. What happens if forces are in a conflict between helping their families or their own business on the one side and doing their duty in a disaster response organization?

A second topic discussed was public private partnerships as well as engagement models for volunteers. Given the scarcity of financial resources on the one side and the increasing expectations by media and general public as well as the increasing vulnerabilities on the other hand, different models of crisis response logistics need to be developed. Private – public partnership are one way. Think of a major city in which the urban population has stocks of food and water for, say two or three days. Would it possible to engage retailers to provide water and food as the capacities of the disaster response organizations would hardly suffice in such a case? Also, the equipment of the disaster response organizations and the material that they would be able to obtain from, say, logistics organizations might not be compatible and up to date. Such new models however have implications for business continuity management in organizations as well as the information systems supporting this. The workshop participants discuss protocols, legal regulations as well as sensitivity of information shared in such a case.

5.2 Workshop Overview Research

Nico Kaptein (COT – Den Haag, NL)

1. We have recognized that seminar participants come from different backgrounds and disciplines and use different 'languages' and definitions in the field of crisis management and civilian response. However, it is not our aim to harmonize these definitions. We accept the existence of different conceptual frameworks and have agreed to share the ones we prefer and work with – to enable common understanding and exchange of ideas and results. The example we work with (and have amended) at COT is the British Standard for Crisis Management, as published by the British Cabinet Office.
2. We do want to invest in research that identifies the most relevant dimensions in crisis management and civilian response. What characteristics are relevant and make a difference in the way a crisis should be handle and response is best organized. How will crisis and disasters develop in the future? What characteristics can be used to guide strategy? Can we deduct early warnings for potential crisis situations? What existing theoretical frameworks can we then meaningfully use and how to these need development and elaboration in the future?
3. For specific areas we need literature research. We do not aim for a full all-encompassing literature review – we do aim to have an up to date overview for specific areas, for instance on transparency and sharing incidents or lessons learned, with peer organizations as well as with the general public. Another example is on dynamic models in logistics. Another example could be cyber-related crisis management. Both theoretical frameworks and case studies may be relevant.
4. Especially mechanisms to identify and learn from what went wrong may need attention. At least in some areas political pressures to come with positive evaluations seem to block insight and progress.

5.3 Workshop Crisis management and governance

Edwin Bakker (Universiteit Leiden, Campus Den Haag, NL)

The workshop first looked into the concepts of crisis network management and governance, agreeing on the idea that there is no single definition, but also agreeing on the notion that any actors are involved with a wide range of responsibilities, capacities, and/or needs. Regarding today's context of crisis management, the workshop agreed on the impact and importance of rapid changes in technologies and societal changes.

Discussing trends in crisis network management and governance there was a common understanding of the limitations of formal, hierarchical and fixed management structures. These types of organizations might do well in dealing with incidents. When it comes to crises, in particular those with unexpected or new effects/elements there is a need for flexibility requiring new or ad hoc governance structures with different types of actors who have no formal relationship with each other. This notion goes back to the general notion of networked societies and changes in (technological) ways of communication, information sharing and risk analysis.

Regarding the level and willingness of cooperation between a wide variety of actors during times of crisis it was mentioned that a high degree of altruism and the need of out-of-the-box-thinking provides opportunities for innovations in terms of establishing new relationships, networks or even operational standards. "Never waste a good crisis" was mentioned several times, indicating that crises allow for changes for improvement of ways in which we can prevent crises or respond to them.

Regarding theoretical and methodological approaches, the workshop discussed the need to develop new ideas, strategies, structures or models with regard to crisis management, by way of exercises and simulations. The group also stressed the need to learn from cases studies and sharing best practices. Relevant academic sources that were mentioned ranged from decision making theory and network dynamics to e-sociology and big data studies.

5.4 Workshop Engagement Models for Entrepreneurs and Volunteers

Kateryna Bondar (Universität der Bundeswehr – München, DE)

The workshop discussed the following topics:

1. Engaged scholarship — how do you engage people to do things? Fundamental shift of how you deal with people
2. Shift in the institutional mind that there are people there outside who can help in crisis response
3. Engaging people involved in the crisis with the technology
4. It cannot be a specific technology, as the hype of some technology can be gone with time
5. We need to better understand what impact Twitter, Facebook are doing to people on the ground
6. Twitter is not free to the commercial organizations
7. Network effect comes into place
8. We should be clear about the type of the crisis and each type of crisis should have a different response
9. Can we rely on the data coming from people on the ground? How representable and reliable is it?
10. In Wikipedia there is 1% of people who is writing information and 9% who is checking, so the check takes place
11. We need to design what crisis is: 1. Uncertainty of information, 2. Dynamics.
12. Agent-based modeling: we want to understand and predict to a certain extend how agents are moving and acting, hetero-generating the population
13. The technology is very helpful for the scientists to learn about crisis and then bring the information back to people
14. Questions: uncertainty, how do you learn something from the information you get
15. Using the wisdom of the crowd
16. Decentralization of decision-makers
17. In Facebook you cannot access the information if you are not part of the network
18. The problem with Twitter is that the sentences are so short and the linguist cannot extract the information from these sentences
19. What characteristics should the network have in order to promote engagement?

20. What is happening is that the network's behavior is changing
21. How do you make sense of the knowledge you gain? One approach is to go into depth understanding of one case, another approach is to build a single model that could explain the events happening — generalizability
22. Different methodologies: design methodology, descriptive methodology
23. What are the characteristics of the social system that would affect crisis response? And this could have a predictive power by knowing the characteristics of the system
24. How do you establish credibility?
25. Legal issue of information release
26. Paradigm shift from control of information to free access
27. From which sources/organizations will information be credible and trusted?
28. How do credibility and trust move along the networks?
29. Empowerment of individuals through technology
30. How do you coordinate individuals: the balance between autonomy and control?
31. Looking at recovery from entrepreneurial point of view: for example, how do you rebuild the houses?
32. Impact evaluation: e. g., would the Arab spring have happened without social media?
33. Relevance of European research for the US journals
34. Where to publish?
35. Position ourselves in different journals
36. Journal of Computer Supported Corporate Work has a special issue on crisis management, Journal of Risk and Uncertainty, Journal of Disaster Research, Journal Risk Research, Journal of Business Continuity
37. ISCRAM topics: use of mobile phone applications for rescuing, voice recognition; use of social media before, during and after the crisis → citizens' response, crisis mapping
38. How do markets respond to the change of technology?
39. Nature creates opportunities within one crisis
40. Doing experiments: field research, living labs; sense-making; post hoc analysis; multi-method approach

Interests of the workshop participants:

1. Social systems, questions of uncertainty, reliability, patterns in the data
2. Dynamics of crisis situations
3. Analyzing social media processes
4. Integration of governmental organizations into the usage of social media
5. Understanding practices dealing with disasters, what role the IT systems are playing
6. Action research
7. Actor-network theory

Underlining points:

1. Dynamics
2. Trust and credibility
3. Impact

Theories:
1. Use and ratification approach, communication of information availability
2. Mass and personal communication assumptions by Fock
3. Systems theory

Methodology:
1. Descriptive methodology
2. Design-oriented
3. Engaged scholarship (action research)

6 Plenary – A Research Roadmap for Crisis Response

6.1 A Research Roadmap for Crisis Response

Bernhard Katzy, Ulrike Lechner, and Christina Weber

The Setting

Dagstuhl Seminars aim to advance the field and explore new topics. In a workshop setting the participants of the Seminar Civilian Crisis Response Models discussed future topics of Crisis Response. In this discussion eight core future research topics have been identified and discussed. This documentation includes topics and notes from the discussions on these topics.

The Research Topics

- Overview Research
- Social Media
- IS Application
- Coordination
- Dynamic Models
- System Resilience
- Risk Communication
- Network Strategies

Subsequently, each topic is presented with the ideas grouped within a topic and the results of the group discussion.

Topic Overview Research

The ideas:
- Research Overview
- Typology of Crises
- Research Matrix-Crises
- Phases Competencies
- The genesis of crises management system: Actors, Connections, Bottlenecks.

The results from the group discussion:
- Literature review

- typology of crises depending on focused dimensions
- Compilation of COT heuristics
- UN definitions and more already existent typologies for comparison
- exemplification on data/statistics.

Topic Social Media

The ideas:

- Using Social Media to communicate and engage (as opposed to broadcast or for situation awareness)
- What makes networks trustable in hierarchies
- Social Media can enhance performance of networks of organizations in dealing with emergency sites?

The results from the group discussion:

- Theories & Models
- Customer needs, Existing "things" & their combination (e.g. applications, capabilities)
- Communities (actors), Information domains
- Future technologies (e.g. Web 4.0, post-Facebook)
- Social impact
- International dimensions
- Resources required & their management

Information System Application

The ideas:

- IT-Systems, developed with end-users for different scenarios like no electrical energy
- Medical aid etc for different time steps
- How can we use/build technology to advance
- Optimize communication between actors (civilians, institutions)
- Interoperability of applications
- Standard operating procedure (SOP) based Information systems
- Managing information reliability.

The results from the group discussion:

- Customer needs
- Existing "things" & their combination (e.g. applications, capabilities),
- Communities (actors)
- Information domains
- Future technologies (e.g. Web 4.0, post-Facebook)
- Social impact
- International dimensions
- Resources required & their management

Coordination

The ideas:

- Best Practices and bullshit sessions from the field

- What crisis management capabilities do we need when physical and virtual worlds get further intertwined?
- Integrated assessment and operations planning
- Portfolio – / program (cross team) coordination & autonomy
- Handling changing uncertainties through a crisis
- How can we improve the coordination skills or possibilities of decision-makers in a dynamic environment
- Crisis management

The results from the group discussion:

- Volunteers see themselves as a part of a club, but also want to be part of a government (we are the big ones)
- There are only less people in the professional staff who do the operational business
- People accept THW as a governmental organization
- Command system is characterized by "Auftragstaktik" this is a German tradition
- How to find the right people? → internal and external
- Platoon leaders know their team and are responsible for recruitment
- The issue is the find people who are willing to spend amount of time for the leader training ... normally our leaders are also leaders in a company
- We get the information where the respective experts are (bottom up)
- My problem is to find the experts (even scientist) outside the THW
- We are monitoring the field (research, technology, ...) through a lens of technical need, then we bring both sides together ... it is very time consuming
- There is a lack of underwater equipment
- THW has mostly contact to business continuity staff
- Recruitment is normally done by somebody knows somebody
- Performance, number of members depends very strong on the respective leader ... lifecycle of a Ortsverband is about 15 years ("lifetime" of a leader)
- The field of coordination in the THW (professional, volunteers but also force coordination and resource management (equipment)) is relevant and interesting
- What can we learn about management of technical innovation, management of who-is-who?
- Even in international disasters you meet the same people. There is a personal network and know each other.
- Also of interest:
  - Management of the volunteer rescue teams.
  - There is research from the EU, mgmt. of resources in civil protection
  - INKA-project — volunteers in civil protection
  - How is the coordination and management in higher levels?
  - Visibility is important in the coordination of the NGOs.

Dynamic Models

The ideas:

- Loosely coupled coordination
- Understanding Complexity
- Leadership in networks and dynamic crises
- Leadership in Dynamic Crises situations with engaged civilians

System Dynamics

The ideas:
- usability
- acceptance of C2 structures
- system of systems
- resource management
- the evolution of crises management socio-technical systems → innovations in subsystems that facilitate and necessitate changes in other stages

The results from the group discussion:
- Resilient systems
- How can we have systems reconfigure themselves
- Resilient response (when damage occurs, it is about restricting the damage)
- Dynamic system modeling
- Understand the granularity of analysis
- Develop reconfiguration strategies
- Human influence on resilience
- Develop quantities for assessing resilience quality
- Applying engineering models for realistic system modeling
- Operator response strategies, SOP for resilience systems operation and citizen behavior
- Resilience by design
- Validation, field testing, simulation

Risk Communication

The ideas:
- Interdisciplinary
- Study on risk communication based on GIS and simulation methods (technical, social aspects)
- Visibility
- Development of simulation and analysis tools
- new social media applications

The results from the group discussion:
- IS-Application should be end-user driven development
- Transformations of processes and procedures in tasks and workflows
- Standard formats & interfaces (meta-models)
- Scalability (green and white IT)
- Integration and evaluation of heterogeneous data sources eg social media
- Creation of cooperation network toolbox -> structuring according to processes
- Online-offline-efficiency of data transfer and volumes

Network Strategies

The ideas:
- Structuration
- Network evolution pattern
- Network capabilities
- loosely coupled coordination

The results from the group discussion:

- Studies on different Levels of coordination: leadership vs implementation level
- Who is who in organizations — meetings for coordination, building of personal networks between organizations, hubs of knowledge, integration of social media, webplatforms and cell phones
- Comparison of networks in crises management in different regions of the world: highly industrialized vs. least developed regions of the world, actors involved, media used, high or low government involvement
- Analysis of network cooperation in different phases, development of indicators of network performance dependent on size/central or multicentered networks

7 Workshops Towards a Research Roadmap for Crisis Response

The second workshop series was about deepening the understanding of future research topics that emerged from a group discussion process. The topics of the workshops are

1. Resilient Systems
2. Risk Assessment and Communication
3. Considerations for Research (projects) on Social Media used for Crisis Response, Disaster Management, and Civil Protection

The summaries of the workshops are provided below.

7.1 Workshop Research Topic Resilient Systems

Ivo Häring

Research on resilient systems extends classical system analysis, modeling and simulation approaches. Of interest are system architectures and behaviors that are capable of dealing with partial, major or complete system damage in such a way that the intended overall system performance is recovered very fast. Examples for systems include infrastructure grids, critical infrastructure, communication networks, organizational structures, societal structures and their respective technical support. Examples for partial system damage are power grid interruptions due to terror attacks, or effects of earthquakes on urban infrastructure. Hence research on resilient systems will contribute to avoid and mitigate crisis effects as well as to improve crisis prevention and shorten recovery times.

The concept of resilience goes beyond reliability by redundancy or reliability by design. It allows for dynamic reconfiguration, resulting in systems that may differ from the original systems. Further it asks for the analysis of a large system trajectory/evolution space and measures for assessing the system reconfiguration options and performance. The time scales of coupled systems under analysis may be very different, also typically very different disciplines are involved, asking for flexible and scalable system analysis, modeling and simulation.

Resilient systems are allowed to fail partially, thus allowing for optimized overall performance taking system operation as well as recovery costs into account. Along with the technical research questions of defining, analyzing, modeling, simulating, testing and validating resilient systems societal needs and requirements are expected to drive and determine research options.

For instance the level of reconfiguration or the partial loss of system performance must be subject to public involvement, discussion and assessment.

7.2 Workshop Considerations for Research (projects) on Social Media used for Crisis Response, Disaster Management, and Civil Protection

Jens Schwarter

- Theories & Models
- Customer needs
- Existing "things" & their combination (e.g. applications, capabilities)
- Communities (actors)
- Information domains
- Future technologies (e.g. Web 4.0, post-Facebook)
- Social impact
- International dimensions
- Resources required & their management

The workshop participants identified the relation of Theory and Practice of Customers and Practitioners for research in Crisis Response as follows:

a) Research driven b) Customer

a) Theory b) Practitioners

7.3 Requirements for Information Systems in Crisis, Disaster Response and Humanitarian Aid

Bernd Hellingrath

A number of different requirements have to be met by Information Systems in crisis, disaster response and humanitarian aid. First of all there is the primary necessity to adapt to the needs of the users, being in a stressful situation to cope with where these systems have to give a support. Principles of user centered design have therefore to be followed in the design of such systems. The highly diverging situations in which the systems have to be applicable demand a flexible architecture allowing an adaptation to the specific needs and also towards the scale of the crisis situation. This leads to a modular design making it able to configure the applied solutions out of a number of different building blocks. On the functional side, the information systems should allow secure and reliable communication capabilities among the different actors. Information from different sources should be integrated and shared among the users of the system. A joint resource management has to be a basis for the coordination of the different activities being supported by the crisis management system.

7.4 Workshop Research Topic Risk Assessment and Communication

Ivo Häring

The ubiquitous presence of communication means, the increase of technical options for citizen participation and new emerging policies for citizen involvement and engagement ask for a rethinking of risk assessment and communication before, during and after crisis events. For example in case of potential landslide areas, the assessment and communication of risks might be critical because of undesired economic effects of for reasonable counter measures, for avoiding undesired economic effects, etc. Similar arguments hold true for other potential risks of crisis like earth-quakes, flooding, power grid loss, terror events. However, how should the ever increasing analysis options and assessment tools along with their refined visualization options be shared with the public? How should the public and responders become part of a shared risk assessment? Research questions include:

- Public and scientific credibility of means of risk assessment, e.g. for earth quakes, land slides, terror events, flooding, storm, social deprivation, etc.
- Further improvement of analysis options for different types of crisis events
- All risk/all hazard approach, measures for risk comparison
- Big data options for risk analysis, e.g. data-driven risk assessment
- Selection of analysis results relevant for crisis management cycle: type of analysis, resolution of analysis, etc.
- Legal and economic issues, e.g. privacy concerns, self-fulfilling prophecy
- Public perception, involvement and feedback
- Foundations of risk analysis supported or based on social media data
- Real time risk analysis and management supported or based on responder data, sensor data, citizen data
- Risk analysis for recovery phase
- Use of geo data and spatial information service infrastructure
- Mobile applications
- User involvement during design of solutions

7.5 Workshop on Coordination and Collaboration in Crisis Management

Francesc Miralles (Ramon Llul University, ES)

Introduction

The main goal of this workshop was to identify those research strands that can be of interest in the field of crisis management from the perspective of coordination and collaboration. In this workshop, the discussion focus was to propose a first group of components that can help to delineate the scope of the relevant research strands. The main point to start this discussion is that crisis management, in general, and coordination and collaboration issues, in particular, have been affected by new drivers that require new perspectives on this research field.

These workshop outcomes are structured in three blocks. Firstly, a description of the main phenomena to be addressed in a research agenda for coordination and collaboration efforts in disaster situations. Secondly, the set of education areas that have to be updated to be able to face the research challenges that these phenomena pose. Finally, which are the tools and instruments that should be faced to progress in this research endeavors.

Setting the Stage. What is a crisis? And what is crisis management?

The participants in this workshop agreed on a definition of crisis that can be synthesized by the following definition by the National Consortium (START Programme, University of Maryland, http://www.start.umd.edu/start/). A crisis is a "serious disruption of the functioning of a community or a society causing widespread human, material, economic or environmental losses which exceed the ability of the affected community or society to cope using its own resources"

A crisis can be a natural disaster, like a flood, an earthquake or a hurricane; or it can be man-made, like a technology failure, or a terrorism act. Crisis management is the discipline whose aim is to study how a community foresees threats to its citizens, and reacts to those threats effectively. Crisis management governance is moving from a hierarchical perspective to a more networked structured. In this vein, coordination and collaboration research has to provide new light on the research challenges that this move produces.

Key drivers of change The workshop participants agreed on the fact that Information, Information Technology, and Information Systems are in the root of effective Crisis Management approaches and are primary enablers of coordination and collaboration structures. In this vein, two key drivers were identified by the participants to delimitate the changes to crisis management from the perspective of coordination and collaboration: Technology Evolution and Networked Society.

Technology Evolution An important pillar on the evolution of structures for coordination and collaboration in disaster situations is the study of the role of ICT, in a wide sense, as enabler of strategic proposals for Networks of organizations and linkage of crisis management and communication. The Internet, social media, sensor networks, and multiple communications networks, provide multiple sources of data that foster a change in the decision making process and in building and valuing collaborative structures in disaster situations. Consequently, new challenges are posed in the research agenda of coordination and collaboration structures among crisis management agents.

Networked society Furthermore, although Web 2.0 and social media platforms allow for the participation of the public to provide field information and to widen the communication channels, new challenges have appeared. Social Networks and Web 2.0 technologies are becoming more present in today's society and there is an increasing need to analyze their effects on public participation in disaster settings. Moreover, ICTs have offered new infrastructures to make it easier to enhance collaboration among small and medium size companies. Recently, network research has been extended to address multi-stakeholder environments crossing the domains of public and private sector as well as social networks and networks of social production. These extensions seem particularly relevant to study crisis management phenomena.

Why do we talk about Coordination and Collaboration?

Following the trends of the key drivers mentioned above, crisis management can be faced by creating dynamic ad-hoc networked organizations. Consequently, crisis management is shifting from hierarchical organized structures to agile organizations tailored to the situation.

Coordination Coordination is "managing interdependencies between activities, enabling all resources to work together harmoniously in achieving a common goal" [2]. The agility gained in a networked organization comes at the cost of an increased coordination effort, especially when the coordination is performed by people. Coordination efforts in disaster situations have been challenged by the plethora of new agents that can have a relevant role in this kind of situations. Technology evolution allows for a reduction of efforts to propose new coordination structures but increases the complexity of coordination settings.

Collaboration Collaboration refers to groups of stakeholders who work together to face crisis situations. Collaborative governance includes broader participation and better balance among diverse interests of engaged agents. Main benefits of collaboration settings in crisis management are trust-building among participants, improved deliberation, and beneficial conflict resolution [4].

Technology evolution and a networked society pose opportunities for collaboration in disaster situations. However, new research is needed to understand the balance between these new opportunities and their efficiency.

Phases in a Crisis Situation As a starting point, participants agreed on the basis that coordination and collaboration endeavors are different for the different phases or stages of a crisis situation. In this vein, a four-phase schema was adopted: Preparation, emergency, recovery, and mitigation. Although participants realized that this is not the only schema for crisis stage, they agreed on the fact that this framework is good enough to analyze how coordination and collaboration can affect the evolution of crisis management. Additionally, this four phase's framework has been widely used in previous studies on crisis management [1]. Using this four-phase framework, crisis management activity can be considered an ongoing process and all phases should be analyzed as part of an iterative and continuous cycle.

Outcomes – Identifying relevant phenomena

In a kind of brainstorming activity, participants proposed the most important issues for each one of the four phases of a crisis situation. The goal was to identify which aspects have been affected the most by the key drivers under a perspective of coordination and collaboration.

Preparation

- Analysis for preparedness (vulnerability; risk analysis, hazard analysis, damage analysis, susceptibility analysis)
- Identification of relevant actors (reliability)
- Emergency plans
- Risk analysis and communication
- Scenario identification
- Standards for recovery (rebuilding) in a short / long term emergency – raising awareness

Emergency

- Reliability, Responsibility, Communication
- Awareness, Situation awareness
- Sense-making, Decision making
- Dealing with uncertainty
- Real time risk awareness/sense
- Simulation/prediction of damages / cascading effects
- Societal and governmental issues
- Legal issues

- Inter organizational collaboration
- Flexibility to involve additional actors
- Adaptability of processes and structure

Recovery
- Dynamic reconfiguration (resilience behavior)
- Identification most effective actions (operations)
- Self help of people, crowd sourcing, "Internet sociology"
- Citizenship empowerment
- Visualization of progress (communication)
- Allocation of resources
- Measurement / criteria / standardization of "normal stage" -> opportunity for new standardization
- Entrepreneurial action
- Debriefing

Mitigation
- Analysis of the crisis (learnt lessons)
- Adaption of emergency plans etc.
- Plan for new infrastructures
- Adaption of governmental, organizational, structure, legal issues
- Implementation of the lessons learnt
- Network structure
- Scenario update
- Risk mitigation

Outcomes – Summary of the findings to identify relevant phenomena

This activity allowed the participants to propose the following relevant research areas.
- Identification of relevant actors and management of these actors. In a networked society the number of potential actors that can participate in each phase of crisis management grows exponentially, current technological features can make viable the participation of many actors. Citizens, NGO, civil protection agencies, governmental units, and so on can participate in any disaster. Coordination & collaboration of the different agents is a challenge to assure effectiveness.
- Citizenship involvement. Social Media has been the enabler of citizenship involvement in crisis management. Many examples illustrate how the role of citizenship can be important in dealing with the consequences of a crisis situation (a citation is needed here). In this area, citizenship involvement is related to the improvement of chances for citizenship participation. This is a special area that refines the previous one on identification of relevant actors.
- Analysis and Simulations for crisis management. This is a wide area of research. It includes the collection of data during a disaster, all the algorithms that can help in the analysis of the data, the synthesis of the results to provide information to decision makers, and finally the presentation of the outcomes in a suitable way to be part of command and control centers.
- Risk communication is considered a central component in the handling of disaster management. Although this centrality applies to all phases of a crisis, dealing with risk communication requires a specific approach in each one of the phases. Harnessing risk communication for specific issues in a disaster situation can sensitize the citizenship to

specific problems, create legitimacy for the actions of government agencies, and thus enhance the strategic capabilities of crisis managers.
- Resilience Reconfiguration. Resilience has been defined as "the capacity of a system, community, or society potentially exposed to hazards to adapt, by resisting or changing, in order to reach and maintain an acceptable level of functioning and structure. This is determined by the degree to which the social system is capable of organizing itself to increase its capacity for learning from past disasters for better future protection and to improve risk reduction measures." ([3] p. 21). Participants agree to propose that resilience reconfiguration should be a research area where coordination and collaboration can play a specific role.
- Simulating exercises and training.

Outcomes – Education

To identify the research challenges, the participants identified the most relevant education subjects needed to be able to face the aforementioned phenomena. In this vein, the following education subjects were proposed to be included in the crisis management field:
- Comparative studies, terrorism, crisis
- Security studies, safety
- Cultural issues (media, communication studies)
- Transnational studies
- Social resilience
- Data analysis (interactive research methods)
- Crisis models

Outcomes – What is missing in the field? What are possible further topics?

The following research subjects were proposed to progress in these research areas:
- Overarching theoretical approach (kind of ...)
- e-Sociology
- Effects of extreme conditions on models and decision making, Governance
- Hazard & risk approaches (an overall perspective)
- Policy oriented approach applied sciences – assistances for disaster management (processes, guidelines, rules, technologies, ...)
- Tool development

References

1 James, K. (2011). The organizational science of disasterterrorism prevention and response: Theory – building toward the future of the field. Journal of Organizational Behavior, 32(7), 1013–1032.
2 Malone, T. W., and Crowston, K. (1990). What is coordination theory and how can it help design cooperative work systems?. In Proceedings of the 1990 ACM conference on Computer-supported cooperative work (pp. 357–370). ACM.
3 National Science and Technology Council. (2005, June). Grand challenges for disaster reduction: A report of the Subcommittee on Disaster Reduction. Washington, D. C.: National Science and Technology Council, Executive Office of the President, Washington, D.C. Retrieved from http://www.sdr.gov/docs/SDRGrandChallengesforDisasterReduction.pdf
4 Robertson, P. J., and Choi, T. (2012). Deliberation, consensus, and stakeholder satisfaction: A simulation of collaborative governance. Public Management Review, 14(1), 83–103.

Participants

- Edwin Bakker
Universiteit Leiden, Campus Den Haag, NL
- Oliver Block
Landeskommando Bayern, DE
- Kateryna Bondar
Universität der Bundeswehr München, DE
- Matthias Brechmann
Unternehmensberatung H & D GmbH München, DE
- Ozgur Dedehayir
Tampere University of Technology, FI
- Simon French
University of Warwick, GB
- Hanno Friedrich
TU Darmstadt, DE
- Ivo Häring
Fraunhofer Ernst-Mach-Institut, Efringen-Kirchen, DE
- Bernd Hellingrath
Universität Münster, DE
- Erich Heumüller
Universität der Bundeswehr München, DE
- Nico Kaptein
COT – Den Haag, NL
- Bernhard Katzy
Leiden University, NL
- Nicole Krämer
Universität Duisburg-Essen, DE
- Erik Kropat
Universität der Bundeswehr München, DE
- Dietmar Kühne
Landeskommando Bayern, DE
- Ulrike Lechner
Universität der Bundeswehr München, DE
- Francesc Miralles
Ramon Llul University, ES
- Stefan Pickl
Universität der Bundeswehr München, DE
- Wolfgang Reinhardt
Universität der Bundeswehr München, DE
- Jens Schwarter
Bundesministerium für Verteidigung, DE
- Gideon Shimshon
Universiteit Leiden, Campus Den Haag, NL
- Christina Weber
Strascheg Center for Entrepreneurship München, DE
- Nils B. Weidmann
Peace Research Institute – Oslo, NO
- Heiko Werner
Bundesanstalt Technisches Hilfswerk, DE
- Volker Wulf
Universität Siegen, DE

Report from Dagstuhl Seminar 13042

Epidemic Algorithms and Processes: From Theory to Applications

Edited by

Benjamin Doerr[1], Robert Elsässer[2], and Pierre Fraigniaud[3]

1 **MPI für Informatik – Saarbrücken, DE**
2 **Universität Salzburg, AT,** robert.elsaesser@sbg.ac.at
3 **University Paris-Diderot, FR,** pierre.fraigniaud@liafa.univ-paris-diderot.fr

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13042 "Epidemic Algorithms and Processes: From Theory to Applications", which took place from January 20 to 25, 2013 at Schloss Dagstuhl – Leibniz Center for Informatics. Several research topics were covered by the seminar participants, including scientists working in Theoretical Computer Science, as well as researchers from the more practical area of Computer Systems. Most of the participants presented their recent results on the topic of the seminar, as well as some challenging new directions and open problems. The presentations contained a description of the main research area for a wide audience. During the seminar, ample time was reserved for informal discussions between participants working on different topics. In our executive summary, we describe the main field of the seminar, as well as our goals in general. Then, we present the abstracts of the presentations given during the seminar.

Seminar 20.–25. January, 2013 – www.dagstuhl.de/13042
1998 ACM Subject Classification C.2.1 Network Architecture and Design – Network communications, C.2.4 Distributed Systems – Distributed applications, F.2.2 Nonnumerical Algorithms and Problems – Computations on discrete structures / Geometrical problems and computations, G.2.2 Graph Theory – Graph algorithms
Keywords and phrases Message dissemination, Epidemic spreading, Dynamic spreading processes
Digital Object Identifier 10.4230/DagRep.3.1.94
Edited in cooperation with Adrian Ogierman

1 Executive Summary

Benjamin Doerr
Robert Elsässer
Pierre Fraigniaud

The Dagstuhl seminar 13042 "Epidemic Algorithms and Processes: From Theory to Applications" took place from January 20 to 25, 2013, and the main goal of the seminar was to fertilize interaction between theory and applications in this emerging research area. Especially in the algorithmic community several fundamentally new ideas have been developed in recent years. At our Dagstuhl seminar, we explored them further, by mixing various ideas coming from experts working on different fields. Theoretical computer scientists presented their

results and methods, in order to disseminate them to a wider community. Researchers from application areas presented their current findings and new challenging research directions, in order to influence (theoretical) research toward real-world applications. The interaction between the seminar participants led to ample discussions and further research collaborations between different domains.

Epidemic algorithms provide a powerful paradigm for distributed computing. Some of the most interesting application areas are the efficient dissemination of updates in replicated data-bases, as well as data dissemination in peer-to-peer systems or wireless sensor networks. By contacting random neighbors in parallel, and making them join forces, an epidemic like progress can be achieved. Furthermore, epidemic processes inherently possess a high level of simplicity and robustness, and therefore the corresponding algorithms can easily deal with the dynamically changing structure of the networks mentioned before.

Theoretical Computer Science makes these useful observations precise and provides certain performance guarantees. One of the well-known algorithms is the so called *randomized rumor spreading*, which disseminates a piece of information in a network to all nodes in a number of communication rounds. In the corresponding communication model, in each round every informed node (i.e, a node which possesses the message) passes/retrieves the information to/from a randomly chosen neighbor. Since 2008, epidemic algorithms received an increased attention by the theory community, leading to a series of new developments such as the development of new analysis techniques for e.g. the bit-complexity of random phone call algorithms, flooding protocols for dynamic graphs, or relating the performance of an epidemic algorithm to the conductance of the network. On the other side, new algorithm design principles have been introduced, which allow the nodes to remember (and avoid) a certain number of previously contacted neighbors, or the use of intentionally dependent randomized decisions. The first modification resulted in an exponential improvement in the number of message transmissions, and lead to the remarkable result that in social networks information can be spread in sublogarithmic time. The second idea gave rise to a number of high-quality papers ranging from, e.g., a theoretical analysis of the amount of randomness needed to the design of the first epidemic rumor spreading algorithm having a safe termination criterion.

One of the main goals of the seminar was to intensify the collaboration between theory and application fields on epidemic algorithms and processes. We mainly concentrated on two major applications. The first one focuses on the construction and maintenance of peer-to-peer networks in a highly dynamic scenario. Since the epidemic algorithms described above are scalable, robust against edge or node failures, and only require a small amount of message transmissions, they can successfully deal with the challenges imposed in a peer-to-peer environment.

The second focus was on the generation of personalized connections in social networks by using epidemic algorithms. Personalization is applied to fundamental processes such as dissemination, search, and navigation, in order to improve the benefits of social networking. The generated views give rise to certain clusters within the network, and the gossip algorithm for communicating profiles and broadcasting messages distinguishes then between intra-cluster and inter-cluster connections.

2 Table of Contents

3 Overview of Talks

3.1 From Caesar to Twitter: On the Elites of Social Networks

Chen Avin (Ben Gurion University – Beer Sheva, IL)

Joint work of Avin, Chen; Lotker, Zvi; Pignolet, Yvonne-Anne; Turkel, Itzik
Main reference C. Avin, Z. Lotker, Y.-A. Pignolet, I.Turkel, "From Caesar to Twitter: An Axiomatic Approach to Elites of Social Networks," arXiv:1111.3374v3 [cs.SI].
URL http://arxiv.org/abs/1111.3374v3

In many societies there is an *elite*, a relatively small group of powerful individuals that is well connected and highly influential. Since the ancient days of Julius Caesar's senate to the recent days of celebrities on Twitter, the size of the elite is a result of conflicting social forces competing to increase or decrease it. In this paper we formulate these forces as axioms and study their equilibrium and other properties of elites in social networks and complex systems.

Our findings indicate that elite properties such as a *size* of $\Theta(\sqrt{m})$ (where m is the number of edges in the network), disproportionate *influence*, *stability* and *density* are universal and should join an increasing list of common phenomenon that complex systems share such: "small world", power law degree distributions, high clustering, etc. As an approximation for the elite we study the subgraph formed by the highest degree nodes, also known as the rich-club. We analyze the structural properties of the k-rich-club of nine existing complex networks and three theoretical models systematically, where the k-rich-club is the subgraph induced by the k nodes with the highest degree in the network. In all real-life networks we observe similar elite properties for rich-clubs consisting of around $\sqrt{m}$ nodes, however, none of the theoretical models we analyzed captures all the elite properties, and thus they should be either adjusted or extended to address these findings.

3.2 Tight Bounds for Connected Dominating Set Packings, Distributed Construction, and Applications

Keren Censor-Hillel (MIT – Cambridge, US)

Joint work of Censor-Hillel, Keren; Ghaffari, Mohsen; Kuhn, Fabian

We study the connected dominating set (CDS) packing and the closely related connected domatic partition (CDS-Partition) problems. The CDS-Partition problem asks for partitioning the nodes of a graph into as many vertex-disjoint connected dominated sets as possible. CDS-Packing is the fractional relaxation of CDS-Partition, where nodes can be in several CDSs and we aim to maximize the ratio between the total number of CDSs and the maximum number of CDSs that contain any node. The size of any CDS-Packing or CDS-Partition is upper bounded by the vertex connectivity k of the graph. Our main result is an efficient distributed algorithm that constructs a CDS-Packing of size $\Omega(k/\log(n))$. We show that there are graphs on which no better CDS-Packing exists. For the CDS-Partition problem, we describe a fully-random algorithm that does not require any communication between the nodes and which constructs a CDS-Partition of size $\Omega(\sqrt{(}k)/\log(n))$, and we give an efficient distributed algorithm that constructs a CDS-Partition of size $\Omega(k/\log^2(n))$ if $k = \Omega(\sqrt{(}n))$.

As a prime application of our results, we obtain efficient distributed algorithms to construct high-throughput communication backbones for store-and-forward broadcast algorithms. In particular, the CDS-Packing algorithm allows to construct a communication backbone for store-and-forward algorithms with throughput $Omega(k/\log(n))$, which we prove to be optimal. This result also implies that the network coding advantage for simultaneously broadcasting messages – the ratio between the throughput achievable with and without network coding – is a tight $\Theta(\log n)$. As a by-product, we also get an efficient distributed algorithm for approximating the vertex connectivity of a graph, and we identify almost optimal connections between the vertex-connectivity of a graph and its resilience to remain connected in the presence of random node failures.

3.3 Fast Fault Tolerant Rumor Spreading with Minimum Message Complexity

Carola Doerr (MPI für Informatik – Saarbrücken, DE)

Joint work of Doerr, Benjamin; Doerr, Carola; Moran, Shay; Moran, Shlomo
Main reference B. Doerr, S. Moran, S. Moran, C. Winzen, "Fast Fault Tolerant Rumor Spreading with Minimum Message Complexity," arXiv:1209.6158v3 [cs.DS].
URL http://arxiv.org/abs/1209.6158v3

An approach for dissemination in complete networks different from randomized rumor spreading that naturally leads to a simple termination criterion is the *workload splitting* algorithm proposed by Gasieniec and Pelc (Parallel Computing 22:903–912, 1996).

It is based on the following simple idea: Each node, when called, receives a list of other nodes that he has to ensure being called. For the initially informed node, this list consists of all other nodes. A node with such a non-empty workload list calls one node of the list, sends the rumor to it, and also moves half of the own workload list to the new node. This easily obtains the optimal dissemination time of $\lceil \log_2 n \rceil$.

The problem with this approach is that it is not very robust against node failures. The approach taken by Gasieniec and Pelc is that a node repeats calling nodes from its list until successful, and then again splits the remaining workload into equal pieces. This may, though, lead to a dissemination time of $\lceil \log_2(n-k) \rceil + k$ when k nodes do not cooperate.

We first show that this estimate is far to pessimistic when we can assume that the failed nodes are distributed randomly. We then obtain a logarithmic dissemination time even for a constant fraction of failed nodes.

This can easily be used to obtain a good protocol also against adversarial node failure: The initial node simply starts the protocol of Gasieniec and Pelc, however, with randomly permuted node identifiers. The price for this approach is that the random permutation also has to be communicated to all nodes, increasing the maximum message size to linear. We then show how to overcome this short-coming by reducing the number of available permutations from $n!$ to polynomial, thus allowing to transmit the permutation with a logarithmic number of bits.

3.4 Epidemic Algorithms: A "Systems" Perspective

Pascal Felber (Université de Neuchâtel, CH)

Joint work of Felber, Pascal; Kermarrec, Anne-Marie; Leonini, Lorenzo; Rivière, Etienne ; Voulgaris, Spyros
Main reference P. Felber, A.-M. Kermarrec, L. Leonini, E. Rivière, S. Voulgaris, "Pulp: An adaptive gossip-based dissemination protocol for multi-source message streams," Peer-to-Peer Networking and Applications, Vol. 5, Issue 1, pp. 74–91, 2012.
URL http://dx.doi.org/10.1007/s12083-011-0110-x

Epidemic protocols are attractive from a theoretical perspective because of their algorithmic simplicity (typically easy to model, prove correct, study). They are also attractive from a "systems" perspective because they are robust and scalable. They just work in real systems! This talk gives an overview of some classical epidemic algorithms and presents a pragmatic design for a push-pull dissemination protocol that performs efficiently in large-scale networks.

3.5 WhatsUp: a P2P instant news item recommender

Davide Frey (INRIA Bretagne Atlantique – Rennes, FR)

WhatsUp [1] is a collaborative-filtering system for disseminating news items in a large-scale dynamic setting with no central authority. It relies on three layered gossip protocols. The lowest two layers build on overlay network that clusters users with similar opinions on the news items they receive. The top-level protocol uses this overlay to disseminate news items to interested users.

The lowest layer in the architecture is a random peer sampling protocol [2]. It gossips information about other nodes in the network and provides each node with a continuously changing random sample of the network. The middle layer also gossips information about other network nodes, but it aims at clustering nodes that have similar interests. These two protocols provide the basis for BEEP, the top-layer protocol in WhatsUp's architecture. BEEP is a heterogeneous epidemic dissemination protocol. Unlike the other two protocols, it exchanges news items. In doing so, it (1) biases the orientation of its targets towards those with similar interests, and (2) amplifies dissemination based on the level of interest in each news item.

WhatsUp outperforms various alternatives in terms of accurate and complete delivery of relevant news items while preserving the fundamental advantages of standard gossip: namely simplicity of deployment and robustness. Nonetheless, the fact that it is based on the combination of three protocols makes it difficult to analyze. The interaction among the protocols is in fact bidirectional. The top-layer dissemination protocol depends on the two overlay protocols. Yet, its dissemination choices also influence the overlay by determining what items will reach which users. An interesting direction for future research will consist in building models that will describe its behavior and provide insight in tuning its operation.

References

1 A. Boutet, D. Frey, R. Guerraoui, A. Jégou, A.-M. Kermarrec (2013). WhatsUp Decentralized Instant News Recommender. IPDPS 2013. Retrieved from http://hal.inria.fr/hal-00769291

2 M. Jelasity, S. Voulgaris, R. Guerraoui, A.-M. Kermarrec, M. van Steen. 2007. Gossip-based peer sampling. ACM Trans. Comput. Syst. 25, 3, Article 8 (August 2007).

3.6 Gossiping Efficiently in an Asynchronous World

Chryssis Georgiou (University of Cyprus, CY)

Joint work of Georgiou, Chryssis; Gilbert, Seth; Guerraoui, Rachid; Kowalski, Dariusz
Main reference C. Georgiou, S. Gilbert, R. Guerraoui, D.R. Kowalski, "Asynchronous Gossip," Journal of the ACM, Vol. 60, Issue 2, 2013.
URL http://www.cs.ucy.ac.cy/~chryssis/pubs.html

We consider the complexity of epidemic rumor spreading in an asynchronous, message-passing fault-prone distributed system. The underlining philosophy of this work is "to design for the best and hope for the best". That is, the objective is to design rumor spreading algorithms that do not rely on any information on the network's message delay or the relative processes' speeds, but be efficient when such timing bounds indeed hold.

Under this setting, we show that an adaptive adversary can significantly hamper the spreading of a rumor, while an oblivious adversary cannot. Under the latter adversarial type we devise efficient asynchronous epidemic gossip algorithms, and show that when applied as a building block, they can improve the complexity of asynchronous randomized consensus.

3.7 Tight Bounds for Rumor Spreading with Vertex Expansion

George Giakkoupis (INRIA – Rennes, FR)

Main reference G. Giakkoupis, "Tight Bounds for Rumor Spreading with Vertex Expansion," arXiv:1302.6243v1 [cs.DM].
URL http://arxiv.org/abs/1302.6243v1

We establish a bound for the classic Push-Pull rumor spreading protocol on arbitrary graphs, in terms of the vertex expansion of the graph. We show that $\mathcal{O}(\log^2(n)/\alpha)$ rounds suffice with high probability to spread a rumor from a single node to all n nodes, in any graph with vertex expansion at least alpha. This bound matches a known lower bound, and settles the question on the relationship between rumor spreading and vertex expansion asked by Chierichetti, Lattanzi, and Panconesi (SODA 2010). Further, some of the arguments used in the proof may be of independent interest, as they give new insights, for example, on how to choose a small set of nodes in which to plant the rumor initially, to guarantee fast rumor spreading.

3.8 How To Gossip (Multiple Messages)

Bernard Haeupler (MIT – Cambridge, US)

Main reference B. Haeupler, "Simple, Fast and Deterministic Gossip and Rumor Spreading," in Proc. of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'13), pp. 705–716, SIAM, 2013.
URL http://knowledgecenter.siam.org/0236-000055/

Gossip algorithms have recently gained attention as a powerful approach for achieving robust and message efficient multicast communication. This talk presents several ideas to improve the efficiency and applicability of these algorithms.

In particular, we provide the first gossip protocol whose efficiency does not rely on expansion properties of the network but which instead performs well on any topology. We also give a novel analysis that shows that a wide variety of natural gossip processes very robustly achieve the same (or even better efficiency) without using any randomization. The existence of such protocols is somewhat surprising because conventional wisdom suggested that both robustness and the efficient information dispersion of gossip protocols stem from their use of randomness.

We also show how combining gossip protocols with network coding can drastically improve the throughput in settings where the amount of data to be multicast is much larger than what can be transmitted in one round. While the idea of using network coded gossip is not new analyzing its performance turned out to be very challenging even in the simplest setting. We introduce projection analysis, as a very simple and powerful technique for providing sharp convergence times in all settings considered in the literature. Beyond this we demonstrate how the projection analysis directly extends to highly dynamic networks.

References

1 B. Haeupler, Analyzing network coding gossip made easy, in Proc. of 43rd ACM Symposium on Theory of Computing (STOC'11), pp. 293-302, ACM, 2011. http://dx.doi.org/10.1145/1993636.1993676

2 B. Haeupler, Simple, Fast and Deterministic Gossip and Rumor Spreading, in Proc. of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'13), pp. 705–716, SIAM, 2013. http://knowledgecenter.siam.org/0236-000055/

3.9 Local Algorithms and Large Graph Analysis

Silvio Lattanzi (Google – New York, US)

Joint work of Lattanzi, Silvio; Kiveris, Raimondas; Korula, Nitish; Mirrokni, Vahab; Alvisi, Lorenzo; Clement, Allen; Epasto, Alessandro; Panconesi, Alessandro

The analysis of very large graph is a central problem in computer science. In this talk we analyze this area of research from a practical prospective. We first describe general problems and directions arising from real world problems then we focus on developing solutions for them using local algorithms.

The first problem that we consider is a security problem in social network. In this context we show how truncated random walk techniques can be used to rescue a social network under sybil attack.

Then we consider a reconciliation problem between two social networks. In this problem we assume to have two social networks and our goal is to map the same user across them. In this setting we show that for several social network model if an initial set of mapping is available it is possible to map almost all the nodes in the network using a simple local algorithm.

3.10 Computing Radius and Diameter of Real World Huge Graphs

Andrea Marino (University of Florence, IT)

Joint work of Crescenzi, Pierluigi; Grossi, Roberto; Habib, Michel; Lanzi, Leonardo; Marino, Andrea
Main reference P. Crescenzi, R. Grossi, M. Habib, L. Lanzi, A. Marino, "On computing the diameter of real-world undirected graphs," Theoretical Computer Science, corrected proof, 2012.
URL http://dx.doi.org/10.1016/j.tcs.2012.09.018

The diameter of a graph is the maximum eccentricity among all its nodes, where the eccentricity of a node x is the distance from x to its farthest node and the distance from x to y is the number of edges contained in the shortest path from x to y.

In the context of real-world networks, the textbook method based on performing a breadth-first search from every node of the graph, requires a prohibitive cost of $\mathcal{O}(nm)$ time, where n is the number of nodes and m is the number of edges of the graph: indeed, it is not rare that a real-world graph contains several millions of nodes and several millions of edges. Even more efficient theoretical methods turn out to be too much time consuming.

We have shown a new algorithm for computing the diameter of directed (or undirected) weighted (or unweighted) graphs. Our algorithm dynamically refines a lower and upper bound of the diameter by analysing the eccentricity of the nodes in a specified "good" order. Although its worst-case complexity is $\mathcal{O}(nm)$ time, we have experimentally shown that our algorithm works in $\mathcal{O}(m)$ time in practice, requiring few breadth-first searches to complete its task.

By applying the same approach, we have shown that it is possible to compute efficiently also the radius, that is the minimum eccentricity of the nodes of a graph.

3.11 Towards Truly Distributed Matrix Computations

Gerhard Niederbrucker (Universität Wien, AT)

Joint work of Niederbrucker, Gerhard; Straková, Hana; Gansterer, Wilfried N.

Numerical linear algebra kernels are fundamental for computational science. Traditionally, algorithms for matrix computations are studied on reliable parallel systems. In contrast to that, we investigate whether or not we can provide such kernels also on unreliable decentralized distributed systems, like sensor networks or peer-to-peer systems. Moreover, future high performance computing systems are expected to be more distributed than parallel in their nature and hence, it is further expected that fault tolerance already at the algorithmic level will become a necessity.

To address the rising need for fault tolerance at the algorithmic level, we investigate the potential of gossip-based algorithms to robustly provide core computational primitives, e.g., distributed summations. While most existing gossip-based approaches are highly flexible and efficient, a single failure in the system in general immediately translates to wrong results. Concerning this issue, we present a technique which allows for deriving fault tolerant variants of a variety of known algorithms which naturally tolerate a wide range of possible system failures.

3.12 Epidemics in Urban Environments – A Virus' Tale

Adrian Ogierman (Universität Paderborn, DE)

Joint work of Ogierman, Adrian; Elsässer, Robert

We consider a dynamic epidemic process in a certain urban environment. The epidemic is spread among n agents, which move from one location to another according (mostly) to a power law distribution. If two agents meet at some spot, a possible infection may be transmitted from one agent to the other.

For the theoretical part, we analyze two different scenarios. We show that at least a small number of agents remains uninfected and the epidemic is stopped after a logarithmic number of rounds. Then, by adding some countermeasures, the epidemic is stopped after $(\log \log n)^{\mathcal{O}(1)}$ steps affecting at most a polylogarithmic number of agents.

For the experimental part, we compare our model to real world data provided by the RKI. Additionally, we empirically analyze the effects of certain countermeasures as applied in the US against the Influenza Pandemic in 1918–1919. Based on our empirical tests, we show that by utilizing the right parameters – some of them being obtained from real world observations – one can efficiently approximate the course of a disease in real world.

3.13 Rumor Spreading in Models of Social Networks

Konstantinos Panagiotou (LMU München, DE)

Joint work of Fountoulakis. Nikolaos; Panagiotou, Konstantinos; Sauerwald, Thomas
Main reference N. Fountoulakis, K. Panagiotou, T. Sauerwald, "Ultra-fast rumor spreading in social networks," in Proc. of the 23rd Annual ACM-SIAM Symp. on Discrete Algorithms (SODA '12), pp. 1642–1660, SIAM, 2012.
URL http://dl.acm.org/citation.cfm?id=2095246

We analyze the popular push-pull protocol for spreading a rumor in networks. Initially, a single node knows of a rumor. In each succeeding round, every node chooses a random neighbor, and the two nodes share the rumor if one of them is already aware of it. We present the first theoretical analysis of this protocol on random graphs that have a power law degree distribution with an arbitrary exponent $\beta > 2$.

Our main findings reveal a striking dichotomy in the performance of the protocol that depends on the exponent of the power law. More specifically, we show that if $2 < \beta < 3$, then the rumor spreads to almost all nodes in $\mathcal{O}(\log \log n)$ rounds with high probability. On the other hand, if $\beta > 3$, then $\mathcal{O}(\log n)$ rounds are necessary.

We also investigate the asynchronous version of the push-pull protocol, where the nodes do not operate in rounds, but exchange information according to a Poisson process with rate 1. Surprisingly, we are able to show that, if $2 < \beta < 3$, the rumor spreads even in constant time, which is much smaller than the typical distance of two nodes.

3.14 Rumor Spreading in Random Evolving Graphs

Francesco Pasquale (University of Rome "La Sapienza", IT)

Joint work of Clementi, Andrea; Crescenzi, Pierluigi; Doerr, Carola; Fraigniaud, Pierre; Isopi, Marco; Pasquale, Francesco; Panconesi, Alessandro; Silvestri, Riccardo
Main reference A. Clementi, P. Crescenzi, C. Doerr, P. Fraigniaud, M. Isopi, A. Panconesi, F. Pasquale, R. Silvestri, "Rumor Spreading in Random Evolving Graphs," arXiv:1302.3828v1 [cs.DM].
URL http://arxiv.org/abs/1302.3828v1

Randomized gossip is one of the most popular way of disseminating information in large scale networks. This method is appreciated for its simplicity, robustness, and efficiency. In the "Push" protocol, every informed node selects, at every time step (a.k.a. round), one of its neighboring node uniformly at random and forwards the information to this node. This protocol is known to complete information spreading in $\mathcal{O}(\log n)$ time steps with high probability (w.h.p.) in several families of n-node "static" networks. The Push protocol has also been empirically shown to perform well in practice, and, specifically, to be robust against dynamic topological changes.

We analyze the Push protocol in "dynamic" networks. We consider the "edge-Markovian" evolving graph model which captures natural temporal dependencies between the structure of the network at time t, and the one at time $t+1$. Precisely, a non-edge appears with probability p, while an existing edge dies with probability q. In order to fit with real-world traces, we mostly concentrate our study on the case where $p = \Omega(\frac{1}{n})$ and q is constant. We prove that, in this realistic scenario, the Push protocol does perform well, completing information spreading in $\mathcal{O}(\log n)$ time steps w.h.p. Note that this performance holds even when the network is, w.h.p., disconnected at every time step (e.g., when $p \ll \frac{\log n}{n}$). Our result provides the first formal argument demonstrating the robustness of the Push protocol against network changes. We also address other ranges of parameters p and q (e.g., $p+q=1$ with arbitrary p and q, and $p = \frac{1}{n}$ with arbitrary q). Although they do not precisely fit with the measures performed on real-world traces, they can be of independent interest for other settings. The results in these cases confirm the positive impact of dynamism.

3.15 From Unstructured to Structure Epidemics: Past and Future Work

Luis Rodrigues (Technical University – Lisboa, PT)

Joint work of Rodrigues, Luis; Leitao, Joao; Ferreira, Mario; Branco, Miguel

We illustrate some of our research that aims at optimizing practical epidemic information dissemination protocols using different approaches to add structured to an otherwise unstructured dissemination process. Two approaches are illustrated: embedding trees and biasing gossip.

The embedded approach is illustrated with Thicket. One way to efficiently disseminate information in a P2P overlay is to rely on a spanning tree. However, in a tree, interior nodes support a much higher load than leaf nodes. Also, the failure of a single node can break the tree, impairing the reliability of the dissemination protocol. These problems can be addressed by using multiple trees, such that each node is interior in just a few trees and a leaf node

in the remaining; the multiple trees approach allows to achieve load distribution and also to send redundant information for fault-tolerance. Thicket is a decentralized algorithm to efficiently build and maintain such multiple trees over a single unstructured overlay network.

The biasing approach is illustrated with BoundedGossip. Gossip-based protocols are very robust and are able to distribute the load uniformly among all nodes. Furthermore, gossip-protocols circumvent the oscillatory phenomena that are known to occur with other forms of reliable multicast. As a result, they are excellent candidates to support the dissemination of information in large-scale datacenters. However, in this context, topology oblivious approaches may easily saturate the switches in the highest level of the datacenter network fabric. BoundedGossip provides an adequate load distribution among the different layers of the switching fabric of the datacenter, avoiding being a source of network bottlenecks.

We also provide some brief overview of our future research directions.

This work was partially supported by FCT – Fundacao para a Ciencia e a Tecnologia under the projects PEst-OE/EEI/LA0021/2011 and HCPI under the grant PTDC/EIA-EIA/102212/2008.

References

1 M. Branco, J. Leitão, L. Rodrigues. Bounded Gossip: A Gossip Protocol for Large-Scale Datacenters. In *Proceedings of the 28th Symposium On Applied Computing (SAC 2013)*, Coimbra, Portugal. March 1013.

2 Leitão, J. Pereira and L. Rodrigues. HyParView: a membership protocol for reliable gossip-based broadcast. In *Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, Edinburgh, UK, June, 2007.

3 M. Ferreira, J. Leitão, and L. Rodrigues. Thicket: A Protocol for Building and Maintaining Multiple Trees in a P2P Overlay. In *Proceedings of the 29th IEEE Symposium on Reliable Distributed Systems (SRDS)*, New Delhi, India, 31 October–3 November 2010.

3.16 Revealing Epidemic Processes: On the Detection of Peer-to-peer Botnets

Stefan Ruehrup (FZ Telekommunikation Wien, AT)

Joint work of Ruehrup, Stefan; Urbano, Pierfrancesco; Berger, Andreas; D'Alconzo, Alessandro

Main reference S. Ruehrup, P. Urbano, A. Berger, A. D'Alconzo, "Botnet detection revisited: theory and practice of finding malicious P2P networks via Internet connection graphs," 5th IEEE International Traffic Monitoring and Analysis Workshop (TMA'13), 2013.

Structured peer-to-peer networks are designed such that information can be spread fast in the network. Their communication structure often emerges from some kind of epidemic process. Since peer-to-peer networks have been used as command and control structures for botnets (networks of malicious software), it is important for network operators to reveal such structures. For this analysis, network traffic can be represented by a Traffic Dispersion Graph, which contains IP addresses as nodes and edges whenever at least one IP packet is exchanged between the respective hosts. The start point for detection is an infected host, which is trapped by a honeypot in the operator's network. Then the rest of the botnet is most likely connected to this start node.

In theory, structured peer-to-peer networks form dense communities, which can be separated from legitimate background traffic by using community detection methods such as the Louvain method. Our experiments on real DSL network traces show that traffic

dispersion graphs of legitimate traffic often contain densely connected components, which makes separation from malicious peer-to-peer traffic difficult. An alternative is to target the neighborhood of known bots. The comparison of such approaches on real network traces show that a local graph search starting from known bots can reveal members of a Kademlia botnet efficiently, where a separation by high degree, betweenness or page rank alone gives non-satisfactory recall factors.

3.17 Discrete Load Balancing on Arbitrary Network Topologies

Thomas Sauerwald (MPI für Informatik – Saarbrücken, DE)

Joint work of Sauerwald, Thomas; Sun, He
Main reference T. Sauerwald, H. Sun, "Tight Bounds for Randomized Load Balancing on Arbitrary Network Topologies," in Proc. of the 53rd Annual IEEE Symp. on Foundations of Computer Science (FOCS'12), pp. 341–350, IEEE CS, 2012.
URL http://dx.doi.org/10.1109/FOCS.2012.86

We consider the problem of balancing load items (tokens) on networks. Starting with an arbitrary load distribution, we allow in each round nodes to exchange tokens with their neighbors. The goal is to obtain a distribution where all nodes have the same number of tokens. For the continuous case where tokens are arbitrarily divisible, most load balancing schemes correspond to Markov chains whose convergence is fairly well-understood in terms of their spectral gap.

However, in many applications load items cannot be divided arbitrarily often and we need to deal with the discrete case where load is composed of indivisible tokens. In this talk we investigate a natural randomized protocol and demonstrate that there is almost no difference between the discrete and continuous case. Specifically, we show that for any regular network, all nodes have the same number of tokens up to an additive constant in the same number of rounds as in the continuous case.

3.18 From Push&Pull to Pointer Push&Pull

Christian Schindelhauer (Universität Freiburg, DE)

Main reference T. Janson, P. Mahlmann, C. Schindelhauer, "A Self-Stabilizing Locality-Aware Peer-to-Peer Network Combining Random Networks, Search Trees, and DHTs," in Proc. of the 16th Int'l Conf. on Parallel and Distributed Systems (ICPADS'10), pp. 123–130, IEEE, 2010.
URL http://dx.doi.org/10.1109/ICPADS.2010.42

We revisit dynamic graph transformations for establishing low diameter, high expansion random d-regular connected graphs. After presenting some connected areas like Pointer-Jumping in PRAMs and Hot Link Assignment we present the two open questions about the Mixing Time of the Flipper operation and the Pointer- Push&Pull operation.

We survey some recent publications and conclude with the demonstration of Pointer-Push&Pull operations inside an existing Peer-to-Peer network.

3.19 Randomness-Efficient Information Spreading

He Sun (MPI für Informatik – Saarbrücken, DE)

Joint work of Sun, He; Guo, Zeyu

We study the classical rumor spreading problem, which is used to spread information in an unknown network with n nodes. We present the first push-based protocol on a complete graph G of n nodes such that, by using $\mathcal{O}(\log n \log\log n)$ random bits, every node gets informed in $\mathcal{O}(\log n)$ rounds with high probability. Both of runtime and randomness complexity are known to be almost tight. Moreover, for several graph topologies our protocols are as fast as the classical protocol and use $\tilde{O}(\log n)$ random bits in total, in contrast to $\mathcal{O}(n \log^2 n)$ random bits used in the classical rumor spreading protocol. These results together give us almost full understanding of the randomness requirement for this basic epidemic process.

3.20 Gossip Protocols for Renaming and Sorting

Philipp Woelfel (University of Calgary, CA)

Joint work of Giakkoupis, George; Woelfel, Philipp

Gossip protocols have emerged as an important communication paradigm for large networks of distributed systems, such as sensor, peer-to-peer, or mobile ad-hoc networks. In these protocols, every node contacts only a few random nodes in each round and exchanges a small amount of information with them. Such protocols are attractive because they offer reasonable performance and at the same time are simple, scalable, and fault-tolerant, and decentralized. Often, they are designed so that nodes need only little computational power and a small amount of storage space.

In this talk I describe efficient gossip-based algorithms for some fundamental distributed tasks. In every round of these protocols, each of the n nodes exchanges information of size $\mathcal{O}(\log n)$ bits with (at most) one other randomly chosen peer. We first consider the *renaming* problem, that is, to assign distinct IDs from a small name space to all nodes of the network. We propose a simple gossip protocol with logarithmic round complexity and name space $\{1, ..., (1+\epsilon)n\}$, for any fixed $\epsilon > 0$. Then we solve the *tight* renaming problem, where each node obtains a unique ID in $\{1, ..., n\}$, in $\mathcal{O}(\log^2 n)$ rounds.

Next we study the following *sorting* problem: Nodes have consecutive IDs 1 up to n, and they receive numbers as inputs. They then have to exchange those inputs so that in the end the input of rank k is located at the node with ID k. Jelasity and Kermarrec (2006) suggested a simple and natural protocol, where nodes exchange keys with peers chosen uniformly at random, but it is not hard to see that this protocol requires $\omega(n)$ rounds. We prove that the same protocol works in $\mathcal{O}(\log^2 n)$ rounds if peers are chosen according to a non-uniform power law distribution.

3.21 Flooding in Dynamic Graphs with Arbitrary Degree Sequence

Pierluigi Crescenzi (University of Florence, IT)

Joint work of Fraigniaud, Pierre; Baumann, Hervé

In this talk we address the flooding problem in dynamic graphs, where flooding is the basic mechanism in which every node becoming aware of an information at step t forwards this information to all its neighbors at all forthcoming steps $t' > t$. Dynamic graphs are modeled as a sequences of graphs $(G_0, G_1, G_2, \ldots)$ with the same node set $[n]$, such that G_t is drawn independently at random according to some random graph model $\mathcal{G}$, for every $t \geq 0$. The case of a sequence of Erdös-Renyi random graphs (i.e., $\mathcal{G} = \mathcal{G}_{n,p}$) has been extensively studied in the literature. In this talk, we consider a sequence of random graphs $(G_t)_{t\geq 0}$ where every G_t is drawn at random in $\mathcal{G}_{\mathbf{w}}$, the model of random graphs with given expected degree sequence $\mathbf{w}$. We show that our techniques developed in a previous paper, for analyzing flooding in a Markovian sequence of $\mathcal{G}_{n,p}$ graphs, is robust enough to be used also in the general case of $\mathcal{G}_{\mathbf{w}}$ whenever there is mutual independence between consecutive graphs in the sequence. In particular, in the case of power-law degree distributions with intercept α, and exponent β, we prove that flooding in a sequence of graphs drawn from $\mathcal{G}_{\alpha,\beta}$ takes a.s. $O(\log n)$ steps even if, a.s., none of the graphs in the sequence is connected, while in the case of graphs with an arbitrary degree sequence we can prove several bounds, which depend on some specific properties of the degree sequence itself. For instance, if this sequence is specially admissible, then the flooding completion time is bounded by $O(\log n(\frac{1}{\log(1+\tilde{d})} + \frac{1}{w_{\min}}))$, where $w_{\min}$ denotes the smallest value in the degree sequence, and $\tilde{d}$ is the second order average degree.

4 Open Problems

During the seminar two open problem sessions were organized, in which several participants presented open problems within the research topics of the seminar. One of these open problems is close to be solved due to a collaboration between the participants.

Participants

- Chen Avin
Ben Gurion University – Beer Sheva, IL
- Keren Censor-Hillel
MIT – Cambridge, US
- Pierluigi Crescenzi
University of Florence, IT
- Oksana Denysyuk
INESC-ID – Lisboa, PT
- Benjamin Doerr
MPI für Informatik – Saarbrücken, DE
- Carola Doerr
MPI für Informatik – Saarbrücken, DE
- Robert Elsässer
Universität Salzburg, AT
- Pascal Felber
Université de Neuchâtel, CH
- Pierre Fraigniaud
University Paris-Diderot, FR
- Davide Frey
INRIA Bretagne Atlantique – Rennes, FR
- Tobias Friedrich
Universität Jena, DE
- Chryssis Georgiou
University of Cyprus, CY
- George Giakkoupis
INRIA Bretagne Atlantique – Rennes, FR
- Bernard Haeupler
MIT, US
- Hovhannes A. Harutyunyan
Concordia Univ. – Montreal, CA
- Anna Huber
University of Durham, GB
- Amos Korman
LIAFA – Paris, FR
- Silvio Lattanzi
Google – New York, US
- Andrea Marino
University of Florence, IT
- Gerhard Niederbrucker
Universität Wien, AT
- Adrian Ogierman
Universität Paderborn, DE
- Konstantinos Panagiotou
LMU München, DE
- Alessandro Panconesi
University of Rome “La Sapienza”, IT
- Francesco Pasquale
University of Rome “La Sapienza”, IT
- Luis Rodrigues
Technical Univ. – Lisboa, PT
- Stefan Rührup
FZ Telekommunikation Wien, AT
- Thomas Sauerwald
MPI für Informatik – Saarbrücken, DE
- Christian Schindelhauer
Universität Freiburg, DE
- He Sun
MPI für Informatik – Saarbrücken, DE
- Philipp Woelfel
University of Calgary, CA

Report from Dagstuhl Seminar 13051

Software Certification: Methods and Tools

Edited by

Darren Cofer[1], John Hatcliff[2], Michaela Huhn[3], and Mark Lawford[4]

1 **Rockwell Collins – Cedar Rapids, US,** ddcofer@rockwellcollins.com
2 **Kansas State University, US,** hatcliff@cis.ksu.edu
3 **TU Clausthal, DE,** michaela.huhn@tu-clausthal.de
4 **McMaster University – Hamilton, CA,** lawford@McMaster.CA

Abstract

With the pervasive deployment of software in dependable systems used in everyday life, society is increasingly demanding that software used in critical systems must meet minimum safety, security and reliability standards. *Certification* is the procedure by which an authorized person or agency assesses and verifies characteristics of a system or product in accordance with established requirements, standards, or regulations. For software, it encompasses traditional notions of verification, but also includes the evidence, tools, methods, and personnel qualifications that are needed to convince the certification authority that the system or product conforms to the relevant standard. Manufacturers of these systems need consistent and effective guidelines as to what constitutes acceptable evidence of software quality, and how to achieve it.

Compared to process-oriented certification procedures, recent approaches provide evidence for dependability by the thorough evaluation of the product itself and the adequacy, coverage and maturity of design and quality assurance methods. Substantial progress has been made in areas including safety and assurance cases, the conceptual foundation of evidence and formal methods, and tooling for software design and verification. New approaches are necessary to develop holistic and cost-effective methodologies and to provide integrated tool support for creating certifiable software-intensive systems, as well as product-focused approaches to certifying these systems.

Experts from academia and industrial practitioners met in the Dagstuhl Seminar 13051 "Software Certification: Methods and Tools" to discuss and software certification challenges, best practices, and the latest advances in certification technologies in several different software-intensive domains (automotive, aircraft, medical, nuclear, and rail).

Seminar 27. January to 01. February, 2013 – www.dagstuhl.de/13051
1998 ACM Subject Classification D.2.0 Software Engineering / General, D.2.4 Software/Program Verification, D.2.9 Management / Software Quality Assurance, I.6.4 Model Validation and Analysis, K.4.1 Public Policy Issues / Human Safety, K.5.2 Governmental Issues / regulation, K.6.3 Software Management / Software Process
Keywords and phrases dependable systems, safety, security, certification, formal methods, model-driven development, validation & verification, tools
Digital Object Identifier 10.4230/DagRep.3.1.111
Edited in cooperation with Sara Bessling

Software Certification: Methods and Tools, *Dagstuhl Reports*, Vol. 3, Issue 1, pp. 111–148
Editors: Darren Cofer, John Hatcliff, Michaela Huhn, and Mark Lawford

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

1 Executive Summary

Darren Cofer
John Hatcliff
Michaela Huhn
Mark Lawford

Context

An increasingly important requirement for success in many domains is the ability to cost-effectively develop and certify software for critical systems (e.g. pacemakers, health monitoring equipment, core banking applications, financial reporting, nuclear reactors, rail automation and active safety in vehicles etc.). Software errors in each of these domains continue to lead to catastrophic system failures, sometimes resulting in loss of life. A recent report by the U.S. National Academy of Sciences [1], concludes that "new techniques and methods will be required in order to build future software systems to the level of dependability that will be required...In the future, more pervasive deployment of software...could lead to more catastrophic failures unless improvements are made." Thus, society is increasingly demanding that software used in critical systems must meet minimum safety, security and reliability standards. Manufacturers of these systems are in the unenviable position of not having consistent and effective guidelines as to what constitutes acceptable evidence of software quality, and how to achieve it. This drives up the cost of producing these systems without producing a commensurate improvement in dependability.

Multiple trends and activities (a) point to the changing nature of development of certified systems and (b) indicate the need for community-wide efforts to assess and form a vision of the future for development of certified systems.

New and Evolving Standards

To adapt to the significant changes in the role of software in dependable systems and to improve current industrial practice in software engineering, international standards like the IEC 61508 are currently under revision. DO-178C governing certification of software in commercial aircraft has recently been revised to accomodate the use of software technologies such as formal methods and model-based development processes. In several other software-intensive domains new domain-specific standards are being developed.

Process- vs. Product-oriented Certification

In practice, current certification of software-intensive systems is primarily process based. A reliance on process oriented standards has established a certification practice that is dominated by assessing process-related documents and marking off checklists that are derived from the recommendation annexes of the standards or so-called "approved practice in use". Thorough evaluation of the product itself or the adequacy, coverage and maturity of design and quality assurance methods are sometimes neglected because there is currently no fundamental agreement on software engineering principles and product qualities to achieve demonstrably dependable software. An alternative to process-oriented certification regimes is "safety and assurance cases" [7]. In Europe, and particularly in the UK, assurance cases have been adopted as a product centric alternative approach to certification and are widely used in

practice already. Recently the U.S. FDA has issued guidance documents recommending the use of assurance cases in submissions for approval of infusion pumps. However, while assurance cases offer some product oriented focus to certification, the lack of standardization of safety and assurance case arguments has its own pitfalls [8].

Advances in Formal Methods

In academia, research on formal methods has made substantial progress with respect to scalability and coverage recently, e.g. in tool-supported model-based design and code generation, but also in the area of software model checking or timing analysis [2]. Thus, formally assuring safety requirements has become feasible at least on the level of components. Nevertheless, research usually focuses on specific techniques, thereby often neglecting the cross-cutting nature of dependability and the need of providing traceable evidence.

Software Development Trends

Two trends relevant in industrial software development for critical systems are the success of model-based design environments that support automated code generation and the need to integrate pre-developed or Commercial Off The Shelf (COTS) software components: (1) Model based tools facilitate rapid prototyping and validation and verification in earlier design phases than traditional processes, but with a price of higher effort in the design phases performed by well-trained and experienced personnel. Software quality will only benefit from these approaches, if certification procedures are adapted towards a cost-effective assessment on the level of models wherever it is adequate. For instance, if model based tools are supported by V & V tools that perform some verification at design time, how does this affect certification standards that require independent design and V & V teams? (2) Evidence based upon prior usage and operating history are typically key components in making decisions in industry about the "fitness for use" of a pre-developed software application or component. However, platform-specific and environmental constraints on the usage are sometimes not specified in detail which has lead to catastrophic failures in the past.

Community-building Activities

Various community-building organizations are being formed drive research, education, and cross-domain coordination in the area of software certification. For instance, the Software Certification Consortium (SCC) was formed in 2007 as a North American initiative to promote product based software certification. Its members are drawn from regulators, industry and academia. SCC has been successful in highlighting shortcomings in current certification regimes and in providing challenge problems and example certification artifacts to the broader community.

Seminar Topics and Goals

The Dagstuhl Seminar 13051 *Software Certification: Methods and Tools* brought together experts for the purpose of assessing the current state of practice, identifying challenges, promising techniques/methods, and for creating a road map for future research, education, and standards development in the area of certification of software and systems.

The seminar addressed the following topics:

- Identification of the challenges, regulatory bodies, primary certification standards, typical development and certification processes in variety of safety-critical domains including avionics, automotive, medical systems, and rail, as well as cross-cutting aspects of security certification.
- Developing a *rational basis* for the primary activities in certification. This included work on the interrelation between i) how we develop software in a way that facilitates certification; and, ii) how we collect and use evidence about software products to evaluate whether they should or should not be certified for use, and iii) cost-benefits issues in certification.
- Pros and cons of assurance-cases in regulatory regimes, assessing the confidence given by assurance cases, new techniques for presenting assurance case arguments, tools for managing the collection of evidence and organization of arguments for assurance cases, and the relationship between assurance cases and software certification standards such as DO-178C.
- The use of tools and open source infrastructure in certification, along with new approaches and guidelines for qualifying tools for use in development of certified systems.
- The latest advances in relevant formal methods for software verification, and integrating formal method with other quality assurance techniques such as testing in the context of certified system development.
- The increasing use of "systems of systems" in safety-critical domains, and the need for new approaches supporting compositional certification and reuse of components in the context of certified systems.
- The structure, nature, use, of current certification standards, current business models and organizational principles for developing standards, and how these aspects might be evolved to better address the needs of the community.
- Strategies for managing the complexity of software intensive systems, including model based development, refinement-based methodologies, and generative techniques.
- Challenges problems, infrastructure, and pedagogical resources to support research and education for both academia and industry in the area of certified system development.

Seminar Participants and Activities

41 researchers participated in the "Software Certification" Seminar, 21 academic researchers, 10 are affiliated to research institutes and 10 experts from industries proving the strategic relevance of the subject to both, research and practice. With about 40% the portion of North American participants was remarkable high.

The seminar started with an introductory session on Monday morning at which the organizers recapitulated the outline, the objectives, and goals of the seminar. Each participant shortly introduced him/herself, his/her scientific background and personal goals for the seminar week. Then senior experts gave an overview on software certification in different domains, namely the avionic, nuclear, medical devices, automotive, and the rail domain. Monday afternoon ended with a discussion on the major differences and similarities between software assessment in the domains and cross-domain challenges. From Tuesday to Thursday experts presented their work. Panel discussions, challenge problem advertisements as well as working group sessions took place in the afternoons and evenings. A wide range of topics was covered including assurance cases and the fundamentals of how to achieve evidence, tool support to software assessment in the certification process, experience reports and new

methodologies for the medical device domain, model based design methods appropriate to certification, issues in cloud security and security certification, tools and methods for static analysis, formal verification and testing. On Thursday evening we had a fruitful discussion with the participants of the Dagstuhl seminar on "Multicore Enablement for Embedded and Cyber-Physical Systems" organized by Andreas Herkersdorf, Michael Hinchey, and Michael Paulitsch that was held in parallel. Among others the following questions were discussed: What are the requests on predictability that have to be satisfied by multicore architectures to be well suited for dependable systems? What are the compelling cyberphysical dependable applications that need multicore architectures? What mechanisms known from dependable software development may be transferred to multicore architecture design and vice versa? Friday was dedicated to working groups as well as outlining and scheduling post seminar proceedings in which we plan to summarize the state of the art in software certification and the results of the seminar. The areas identified by the plenum to be most relevant for further progress on software certification are:

- Fundamentals on confidence and evidence
- Compositional certification
- Education on dependable systems and certification
- Tool qualification
- Security
- Methods for the development of certifiable software and methods supporting certification

References

1 D. Jackson, M. Thomas, L. Millett. "Software for Dependable Systems: Sufficient Evidence?" Committee on Certifiably Dependable Software Systems, National Research Council, National Academies Press, 2007.

2 M. Huhn, H. Hungar. UML for software safety and certification – Model-based development of safety-critical software-intensive systems. In H. Giese, G. Karsai, E. Lee, B. Rumpe, and B. Schätz (Eds.): Model-Based Engineering of Embedded Real-Time Systems – Int'l Dagstuhl Workshop, Dagstuhl Castle, Germany, November 4–9, 2007. Revised Selected Papers, LNCS 6100, Springer, pp. 203–240. 2011. DOI: 10.1007/978-3-642-16277-0_8

3 M. Huhn, A. Zechner. Analysing Dependability Case Arguments Using Quality Models. In B. Buth, G. Rabe, and T. Seyfarth (Eds.): 28th Int'l. Conf. on Computer Safety, Reliability, and Security (SAFECOMP), LNCS 5775, Springer, pp. 118–131, 2009. DOI: 10.1007/978-3-642-04468-7_11

4 A. Wassyng, T.S.E. Maibaum, M. Lawford, On Software Certification: We Need Product-Focused Approaches. C. Choppy and O. Sokolsky (Eds.): Monterey Workshop 2008, LNCS Vol. 6028, Springer, pp. 250–274, 2010. DOI: 10.1007/978-3-642-12566-9_13

5 J. Hatcliff, M.P.E. Heimdahl, M. Lawford, T.S.E. Maibaum, A. Wassyng, F.L. Wurden. A Software Certification Consortium and its Top 9 Hurdles. In Proc.of the First Workshop on Certification of Safety-Critical Software Controlled Systems (SafeCert 2008), ENTCS, Vol. 238, No. 4, pp. 11–17, 2009. DOI: 10.1016/j.entcs.2009.09.002

6 FDA, "FDA Launches Initiative to Reduce Infusion Pump Risks," News Release, April 23, 2010 (see: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm209042.htm)

7 R. Bloomfield and P. Bishop. Safety and assurance cases: Past, present and possible future – an Adelard perspective. In C. Dale, T. Anderson (Eds.): Making Systems Safer, Proc. of the Eighteenth Safety-Critical Systems Symp., Bristol, UK (February 2010), pp. 51–67.

8 A. Wassyng, T.S.E. Maibaum, M. Lawford and H. Behr. Is there a case against safety cases? Submitted to post-proceedings volume of Monterey 2010 Workshop, to be published in LNCS.

2 Table of Contents

3 Overview of Talks

3.1 Modeling Requirements for Embedded Systems with RDAL

Dominique Blouin (Université de Bretagne Sud, FR)

Joint work of Blouin, Dominique; Turki, Skander, Senn Eric
Main reference D. Blouin, E. Senn, S. Turki, "Defining an annex language to the architecture analysis and design language for requirements engineering activities support," in Proc. of Model-Driven Requirements Engineering Workshop (MoDRE'11), pp. 11–20, IEEE, 2011.
URL http://dx.doi.org/10.1109/MoDRE.2011.6045362

In this talk, I will introduce the Requirements Definition and Analysis Language (RDAL) that we are developing as an annex of the Architecture Analysis and Design Language (AADL). I will present the needs for such language, its main features and show how it can be used to formalize requirements specifications supporting requirements engineering best practices such as those of the FAA Requirements Engineering Management Handbook. I will also present the modeling of a concrete example from the FAA handbook with RDAL, and discuss some modeling issues and inconsistencies of the natural language specification revealed by the process of formalizing the specification. I would like to get inputs on the potential impact of finding these inconsistencies on the system development process and the benefits of the modeling activity taking into account the overhead it implies.

3.2 Technology Infusion Study for DO-333

Darren Cofer (Rockwell Collins – Minneapolis MN, US)

Joint work of Cofer, Darren; Miller, Steven

In 2012, RTCA published DO-178C, DO-278A, and DO-333, which together define a framework for applying formal methods in the certification of airborne and air traffic management systems. However, there remain significant challenges to the successful infusion of formal methods into development and certification workflows in the aviation industry. Under NASA funding, Boeing and Rockwell Collins are developing technical material that will help industry to effectively apply formal methods in a DO-178C/DO-333 context. Our work includes a survey of currently available formal methods and tools that are most relevant for verification and certification of airborne software. We have also developed an integrated case study to demonstrate the use of some of these tools to satisfy DO-178C certification objectives using the augmented guidance in DO-333. The case study is based on a fault-tolerant flight control system than includes requirements and design artifacts specified using PVS, Simulink/Stateflow, and source code. We have verified different aspects of the system design using theorem proving, model checking, and abstract interpretation, and show how various certification objectives are satisfied using these formal techniques. The survey and the case study will form the basis for a new formal methods guidebook, which will be publicly available along with all of the models and verification artifacts. Both the technical descriptions and the examples will be presented at varying levels of detail, suitable for system developers, certifiers, and other stakeholders.

3.3 Integrating Formal Program Verification with Testing

Cyrille Comar (AdaCore, Paris, FR)

Joint work of Comar, Cyrille; Kanig, Johannes; Moy, Yannick

The Hilite project proposes a framework that offers the possibility of using formal and non-formal verification of requirements expressed in the form of subprogram contracts in a DO-178C context. In particular, we explore the conditions for validating an approach mixing formal verification and testing whose goal is to reduce the costs of testing and ease the adoption of formal verification in the industry.

3.4 Functional Safety and Certification of Automotive E/E systems

Mirko Conrad (The MathWorks GmbH – Ismaning, DE)

ISO 26262 "Road vehicles – Functional safety" is a set of safety standards for electrical and/or electronic (E/E) systems installed in series production passenger cars. ISO 26262 constitutes the adaptation of IEC 61508 to comply with needs specific to the application sector of E/E systems within such road vehicles. ISO 26262 which was published in 2011 provides:

- An automotive safety lifecycle incl. means for tailoring the necessary activities
- An automotive-specific risk-based approach to determine integrity levels (Automotive Safety Integrity Levels, ASILs) and uses ASILs to specify applicable requirements to avoid unreasonable residual risk
- Requirements for validation and confirmation measures to ensure a sufficient and acceptable level of safety being achieved
- Requirements for OEM-supplier relationship.

Functional safety as per the standard is defined as "absence of unreasonable risk due to hazards caused by malfunctioning behaviour of E/E systems". ISO 26262 does not have a notion of certification, but it sets out three types of so called confirmation measures (confirmation review, functional safety audit, and functional safety assessment). The standard also provides a process to establish the required confidence in the correct functioning of software tools. This process comprises two steps, tool classification and subsequent tool qualification (if applicable). The talk provides an overview of key concepts of ISO 26262 that are related to the topic area of the seminar.

3.5 Abstraction, Fidelity and (In-)Competence: modelling cyber-physical systems and systems of systems

John S. Fitzgerald (Newcastle University, GB)

Joint work of Fitzgerald, John S.; Larsen, Peter G; Verhoef, Marcel HG; Pierce, K; Gamble, C

Formal model-based methods have been advocated as a means of gaining assurance in the early stages of product development. Formal notations are said to promote careful abstraction, the explicit recording of assumptions, and the elimination of infeasible designs. The range of applications of formal methods, and the capability of analysis techniques, using proof and model-checking for example, have grown in the last quarter century. However, the use of a notation that happens to have formal semantics is not in itself a basis for making substantive judgements about any product that may ultimately be derived from the model. What, then, are the costs, benefits and risks of formal modelling and analysis?

My talk focuses on the new and challenging domains of cyber-physical systems in which there is a close interaction between networked computing devices and the physical environment, and in systems of systems, which are composed of independently owned and managed constituents about which we have limited knowledge and over which we have limited control. What can formal techniques contribute to the assurance of such systems?

I use examples drawn from recent work on the following projects in which formal model-based techniques are developed with the aim of help to de-risk development by eliminating inadequate or infeasible designs at early stages. In both cases, the quality of models depends on balancing abstraction and fidelity while not compromising competence.

- DESTECS (Design Support and Tooling for Embedded Control Software) www.destecs.org defines methods and tools for collaborative modelling and co-simulation for embedded systems, linking discrete-event models of software to continuous-time models of controlled plant and environment. A reconciled operational semantics for simulators in both domains enables systematic exploration of design spaces.
- COMPASS (Comprehensive Modelling for Advanced Systems of Systems) www.compass-research.eu is developing formal modelling techniques for Systems of Systems. This entails providing common semantic bases for a range of aspects including data, functionality, architectural structure, time and mobility.

3.6 What is Mission-Assurance?

Kim R. Fowler (Kansas State University, US)

I will present several case studies, from military, medical, and appliance industries, as examples of dependable systems, which illustrate some aspects of mission-assurance. We, the seminar participants, will discuss what "dependable" and "mission-assurance" really mean in the context of each of these case studies.

I will present and discuss with you some components of development processes that lead towards mission-assurance in systems. My primary conclusion is: Appropriate processes and procedures in design and development lead to dependable systems and mission-assurance, not blind application of all possible techniques.

3.7 A naive look at software certification practices – and proposals for enhancement

Hubert Garavel (INRIA Rhône-Alpes, FR)

In this talk I will provide the feedback of an academic computer scientist confronted to the current multiplicity of software certification standards. The current situation seems far from optimal because of the diversity of vocabulary and concepts, and because key ideas of software engineering and formal methods seem to be missing from current standards. Based on these remarks, some proposals for enhancing certification practices are formulated.

3.8 Bringing evidence-based arguments into practice

Janusz Gorski (Gdansk University of Technology, PL)

Evidence-based arguments have a potential to strengthen trust relationships in different contexts. To support their wider application, several problems have to be solved, including: choosing an adequate argument model, integration of arguments and the supporting evidence, diversity of evidence formats (e.g. text, graphics, video stream, audio), diversity of the evidence repositories, user-friendly interface, argument assessment (including multiple assessments and diverse assessment mechanisms), communication of the argument assessment results, deployment models for supporting tools, information security (in particular, security of an argument and security of the evidence supporting the argument), scalability, version management and so on.

To address this and the related problems we are developing the TRUST-IT methodology [2] and the related platform of software services, called NOR-STA [1]. TRUST-IT is focused on representation and assessment of evidence based arguments and on their possible usage scenarios [10, 8, 7, 5, 6]. The arguments are represented graphically, with the help of NOR-STA software services which are deployed in accordance with the SaaS (Software-As-A-Service) cloud computing model. The "strength" of an argument can be appraised by an independent assessor using the appraisal mechanism based on Dempster-Shafer theory. The use of other, application domain specific argument appraisal mechanisms is also supported. The results of argument assessment are visualized by coloring the argumentation tree which provides for effective and efficient communication and decision making support.

TRUST-IT and NOR-STA have been already used in different application scenarios, including justification of safety, security and privacy properties of IT systems and services developed in four different European research projects (DRIVE, PIPS, ANGEL and DECOS) [4], justification of the trustworthiness of the assessment criteria of web-based sources of medical information (applied by the Health-On-the Net foundation based in Switzerland) [3], and are presently used to support the processes of achieving and assessing standards conformance in different domains, including accreditation standards for hospitals [9], a standard for risk management in outsourcing processes, CAF (Common Assessment Framework) – a set of guidelines for self-assessment and self-improvement of public organizations, and presently we begin to support HACCP (Hazard Analysis and Critical Control Points), a systematic

preventive approach to food safety and allergenic, chemical, and biological hazards. New application scenarios, including parallel monitoring of critical requirements in different sites and automatic assessment of selected claims are under investigation.

References

1 http://www.nor-sta.eu/en
2 http://iag.pg.gda.pl/iag/?s=research&p=trust_cases
3 http://www.hon.ch/Global/copyright.html
4 Górski J, Jarzêbowicz A, Miler J, Witkowicz M, Czyznikiewicz J, Jar P: Supporting Assurance by Evidence-Based Argument Services, SAFECOMP Workshops 2012, Springer-Verlag Berlin, Heidelberg, pp. 417–426
5 Cyra Ł., Górski J., Support for argument structures review and assessment, Reliability Engineering and System Safety, Elsevier, Volume 96, 2011, pp. 26—37
6 Cyra Ł., Górski J., SCF – a Framework Supporting Achieving and Assessing Conformity with Standards, Computer Standards & Interfaces, Elsevier, Volume 33 Issue 1, January, 2011,pp. 80–95
7 Górski J., Jarzebowicz A., Leszczyna R., Miler J., Olszewski M.: Trust case: justifying trust in IT solution, Proc. Safecomp Conference, Reliability Engineering and System Safety, Elsevier, vol. 89/1, 2005, pp. 33–47. 8
8 Górski J.: Trust-IT – a framework for trust cases, Workshop on Assurance Cases for Security – The Metrics Challenge, Proc. of DSN 2007, June 25-28, Edinburgh, UK, 2007, pp. 204–209.
9 Górski J., Jarzêbowicz A., Miler J.,Validation Of Services Supporting Healthcare Standards Conformance, Journal on Metrology and Measurement Systems, vol. XIX, No. 2, 2012, pp. 269–284
10 Górski J., Trust Case – a case for trustworthiness of IT infrastructures, in Cyberspace Security and Defense: Research Issues, NATO Science Series II: Mathematics, Physics and Chemistry, Vol. 196, Springer-Verlag, 2005, pp. 125–142

3.9 Static Analysis of Real-Time Embedded Systems with REK

Arie Gurfinkel (CMU – Pittsburgh PA, US)

Joint work of Gurfinkel, Arie; Chaki, Sagar; Strichman, Ofer; Kong, Soonho
Main reference S. Chaki, A. Gurfinkel, S. Kong, O. Strichman, "Compositional Sequentialization of Periodic Programs," in Proc. of the 14 Int'l Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI'13), LNCS, Vol. 7737, pp. 536–554, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-35873-9_31

Real-Time Embedded Software is an important class of safety-critical software systems. It plays a crucial role in controlling systems ranging from airplanes and cars, to infusion pumps and microwaves. Verifying the correct operation of RTES is an important open problem.

In this presentation, I will describe the START (Static Analysis of Real-Time Embedded Systems) project that I am leading together with Sagar Chaki at the Software Engineering Institute at Carnegie Mellon University. The focus of START is verification of safety properties (e.g., race conditions, mutual exclusion, and deadlocks) of periodic programs scheduled with Rate Monotonic Scheduling (RMS) policy. Such programs are common in automotive and avionics domains. I will describe our experience in building a Bounded Model Checker, called Rek, for programs written for OSEK/VDX operating system, and our experience in using Rek to verify properties of a robotics controller.

3.10 Certification for Medical Devices and Systems: An Overview and Challenges

John Hatcliff (Kansas State University – Manhattan KS, US)

Joint work of Hatcliff, John; Knight, John; Weber, Jens; Heimdahl, Mats

Medical devices and systems have long been an example of a safety-critical domain with many challenges related to risk assessment, regulatory policy, and certification. However, an aging population, innovations in mobile computing devices, increased reliance on integrated systems, and increased importance of storing and leveraging patient data are introducing a variety of strains and pressures on existing certification approaches.

In this talk, we give a summary of important certification-related issues in the medical device domain including the types of products certified in the domain, the most common regularatory processes and agencies, development and verification tools used in the medical device domain, and relevant standards for medical device certification. We conclude with a discussion of important trends and technologies within the medical device space that are giving rise to challenges that need the attention of researchers working in the areas of certification and verification.

3.11 Requirements Specification and Supporting Artifacts for an Open Source Patient-Controlled Analgesic Pump

John Hatcliff (Kansas State University – Manhattan KS, US)

Joint work of Hatcliff, John; Larson, Brian
URL http://info.santoslab.org/research/pca

The dynamic nature of the medical domain is driving a need for continuous innovation and improvement in techniques for developing and assuring medical devices. Unfortunately, research in academia and communication between academics, industrial engineers, and regulatory authorities is hampered by the lack of realistic non-proprietary development artifacts for medical devices.

In this talk, we give an overview of a detailed requirements document for a Patient-Controlled Analgesic (PCA) pump developed under the US NSF's Food and Drug Administration (FDA) Scholar-in-Residence (SIR) program. This 60+ page document follows the methodology outlined in the US Federal Aviation Administrations (FAA) Requirements Engineering Management Handbook (REMH) and includes a domain overview, use cases, statements of safety & security requirements, and formal top-level system architectural description. Based on previous experience with release of a requirements document for a cardiac pacemaker that spawned a number of research and pedagogical activities, we believe that the described PCA requirements document can be an important research enabler within the formal methods and software engineering communities.

3.12 Concerning the implicit DO-178C assurance case

Michael Holloway (NASA Langley ASDC – Hampton, US)

Main reference C.M. Holloway, "Towards Understanding the DO-178C / ED-12C Assurance Case," in Proc. of the IET 7th Int'l Conf. on System Safety, October 2012, Edinburgh, Scotland.
URL http://hdl.handle.net/2060/20120016708

This informal discussion without visual aids describes ongoing work towards identifying and expressing explicitly the arguments contained in, or implied by, DO-178C, which implicitly justify the assumption that the document meets its stated purpose of "providing guidelines for the production of software for airborne systems and equipment that performs its intended function with a level of confidence in safety that complies with airworthiness requirements."

3.13 Software verification in the medical domain

Jozef Hooman (Radboud University Nijmegen, NL)

Joint work of Hooman, Jozef; Mooij, Arjan; Keshishzadeh, Sarmen; Albers, Rob

We present an overview of research activities to improve the verification and validation of medical systems at Philips Healthcare. In particular, we focus on the interventional X-ray systems of Philips. To reduce the test and integration phase of these systems and obtain a more efficient development process, the aim is to detect faults earlier by applying various modeling and analysis techniques. This includes executable models, domain specific languages, and formal methods.

3.14 Bridging the modeling/verification gap

Jerôme Hugues (ISAE – Toulouse, FR)

Model Driven Engineering (MDE) provides an appealing framework for supporting engineering activities, from early design phases to acceptance tests; going through refinement, architectural and functional design down to code generation and V&V efforts. Yet, certification activities may interfere with such process: traceability must be demonstrated, specific verification or validation activities must be performed, some of which are project or domain specific.

In this talk, I present current discussions on the part of the AADL standardization committee to enrich Architecture Description Language with a Constraint language. The objective is to increase the coupling between modeling and verification. By making the verification part of extended semantics rules of an ADL, we control the patterns used to describe the system, ensuring designers respect process requirements, but also integrate V&V as part of the modeling effort. Thus, it provides a lean approach to certification through MDE.

3.15 Opening up the Verification and Validation of Safety-Critical Software

Hardi Hungar (German Aerospace Center – Braunschweig, DE)

Joint work of Hungar, Hardi; Behrens, Marc
Main reference In: M. Huhn, S. Gerken and C. Rudolph (eds.), Proc. ZeMoSS 2013, to appear

Smooth cross-border rail traffic is of important interest to commercial realizations of ETCS (European Train Control System). Starting from the hypothesis that the traditional way of developing software for safety-critical systems might be an obstacle to standardizing rail traffic, the ITEA 2 project openETCS has set out to pursue the idea of transferring an open-source development style to this domain, taking the EVC (European Vital Computer, core of the on-board unit) as a target.

The goal is to formalize the requirements in a functional model, derive, via design models, an implementation, and demonstrate how the verification and validation activities necessary for certifying a resulting product could be performed. All of this is to be done as an open-source project, employing only open-source tools. One of the main motives behind the approach is to use the potential of an open community to detect design and implementation flaws much earlier than the resource-limited inspection in a traditional development setting.

This talk discusses the challenges this new approach faces from the legal requirement of adhering to the standards, mainly the EN 50128 in this case, particularly with respect to verification and validation. This comprises the interpretation and application of the standard throughout all lifecycle phases for a open-source model-based development and qualification issues for personnel and tools.

3.16 Using Code Analysis Tools for Software Safety Certification

Daniel Kaestner (AbsInt – Saarbrücken, DE)

Joint work of Kaestner, Daniel; Ferdinand, Christian
Main reference D. Kästner, C. Ferdinand, "Efficient Verification of Non-Functional Safety Properties by Abstract Interpretation: Timing, Stack Consumption, and Absence of Runtime Errors," in Proc. of the 29th Int'l System Safety Conference (ISSC'11), ISBN 9781618399922. Las Vegas, 2011.

In automotive, railway, avionics and healthcare industries more and more functionality is implemented by embedded software. A failure of safety-critical software may cause high costs or even endanger human beings. Also for applications which are not highly safety-critical, a software failure may necessitate expensive updates. Safety-critical software has to be certified according to the pertinent safety standard to get approved for release.

Contemporary safety standards – including DO-178B, DO-178C, IEC-61508, ISO-26262, and EN-50128 – require the identification of potential functional and non-functional hazards and to demonstrate that the software does not violate the relevant safety goals. To ensure functional program properties, automatic or model-based testing and formal techniques like model checking are becoming more widely used.

For non-functional properties identifying a safe end-of-test criterion is a hard problem since failures usually occur in corner cases and full test coverage cannot be achieved. For some non-functional program properties this problem is solved by abstract interpretation-based

static analysis techniques which provide full control and data coverage and yield provably correct results. Like model checking and theorem proving, abstract interpretation belongs to the formal software verification methods.

This talk focuses on static analyses of worst-case execution time, stack consumption, and runtime errors, which are increasingly adopted by industry in the validation and certification process for safety-critical software. First we will give an overview of the most important safety standards with a focus on the requirements for non-functional software properties. We then explain the methodology of abstract interpretation based analysis tools and discuss the role of formal verification methods in current safety standards. Using tools for certification requires an appropriate tool qualification. We will address each of these topics, report on industrial experience, and address open issues.

3.17 Towards an Effective Safety Demonstration Framework

Peter Karpati (Institute for Energy Technology – Halden, NO)

This talk with introduce our project which aims at assembling evolutionarily adaptable guidelines for an effective safety demonstration framework based on exploring state of the art and state of practice in the field.

3.18 Software Certification: Where is Confidence Won and Lost?

Tim Kelly (University of York, GB)

Given that we cannot prove the safety of software (in a system context) we are forced to wrestle with the issue of confidence in software certification. Some draw confidence from compliance with software assurance standards and believe this is sufficient, yet we don't have consensus in these standards. Some establish confidence through the process of constructing and presenting a software assurance case, but ignore the experience and "body of knowledge" provided by standards. Some (sensibly) use a combination of these approaches. Using our framework of 4+1 principles of software safety, this talk will discuss where and how in current safety-critical software development and assessment approaches confidence is typically won and lost. Based on this assessment, we describe how the activity and structure of an assurance case should best be targeted to explicitly address issues of confidence.

3.19 User Assembled Medical System of Systems

Andrew King (University of Pennsylvania, US)

Joint work of King, Andrew; Lee, Insup; John, Hatcliff
Main reference J. Hatcliff, A. King, I. Lee, A. Macdonald, A. Fernando, M. Robkin, E. Vasserman, S. Weininger, J.M. Goldman, "Rationale and Architecture Principles for Medical Application Platforms," in Proc. of the 2012 IEEE/ACM Third Int'l Conf. on Cyber-Physical Systems (ICCPS'12), pp. 3–12, IEEE/ACM, 2012.
URL http://dx.doi.org/10.1109/ICCPS.2012.9

In safety critical domains, there is a typically a prime contractor that is responsible for integration and system-level verification and validation. In user assembled systems (such as plug and play medical systems), system modules are sold to users directly. These non-technical users then assemble these modules into (possible safety critical) systems of systems towards some purpose. We describe plug and play medical systems and associated certification challenges.

3.20 Three Challenges

John C. Knight (University of Virginia, US)

In this talk, I will briefly summarize: (a) the concerns I have about the dependability of medical systems and why I think medical system dependability is more difficult than other domains, and (b) the concerns I have about standards together with some ideas about a "standard for standards."

3.21 Certification of Medical Device Composition

Brian Larson (Multitude Corp., US)

The Software Certification Consortium seeks to host "mock" certification evaluations of safety-critical systems dependent upon software. My presentation considers what artifacts should be part of submissions for mock certification, particularly evidence from analysis of architecture models of errors and behavior. Only formal proofs of safety and effectiveness of apps and medical devices they control will likely obtain regulatory approval.

3.22 Bayesian Probabilistic Approaches to Confidence are Impossible: The Need for a Baconian Approach (pace Jonathan Cohen)

Tom S. Maibaum (McMaster University – Hamilton, CA)

Many have recognized the need for some notion of confidence in relation to safety and assurance cases. After all, a regulator has to have enough confidence in the case to prover a certification. After all, other domains also use such notions of confidence. These include the legal domain, where (at least in English law) judgements have to be made "on the balance of probabilities" (civil cases) and "beyond reasonable doubt" (criminal cases) are standards of confidence required of judges and juries about the guilt of the accused. Similarly, there is an implicit notion of confidence amongst scientists about the theories they use in their subject. Corroborative experiments raise this level of confidence, whilst negative result may lower the level of confidence (but not necessarily to 0, pace Popper). A number of philosophers/logicians/scientists have attempted to characterize this notion of confidence, including Bacon, and more recently, Carnap, Keynes, etc. More recently, Jonathan Cohen, in The Probable and the Provable, has demonstrated that the concept of probability underlying the concept of confidence, which he claims is Baconian, simply cannot be reduced to conventional Pascalian, frequency of occurrence, notions of probability. The argumentation basis for safety/assurance cases uses confidence as a tenet for the approach. Toulmin's argument schemes present a form of inductive/scientific reasoning with explicit reference to confidence to justify the applications of a scheme. Interestingly enough, a scheme reduces to a deductive rule of inference when there is no uncertainty about its application.

3.23 Software Certification: The Return on Investment?

John McDermid (University of York, GB)

Certification costs money; it also has benefits, perhaps most importantly reduction in risk to system users and third parties, to which we give a value:

- The costs will be in manpower and other resources for testing, code inspections, formal verification, etc.;
- The benefit will be in terms of security breaches avoided, hazardous failures avoided, etc.;
- The value will be in terms of the assets protected (e.g. €Ms), or lives saved (perhaps monetised by multiplying by the VPF (value of preventing a fatality)), etc.

At its simplest, for there to be a positive return on investment (RoI) the value has to outweigh the cost. In practice, there are difficulties in determining RoI, for example mapping benefit to value due to uncertainties about how the software will be used. These difficulties are exacerbated prior to undertaking a software certification activity, e.g.:

- How do we predict costs?
  - How much will it cost to apply the technique to the software?
  - How much will it cost to rectify any flaws identified?
- How do we predict benefit before carrying out the activity, e.g. by simulation?
  - What flaws will we find?

  - What flaws will we find that we would not find by other means?
- How do we predict value, as the benefits are really contingent?
  - Which assets will now be protected, should the attack we know could have been successful, but won't be now, actually occur? How likely is the attack?
  - How many lives will be saved, should the operational scenario we know could have been hazardous, but won't be now, actually arise? How likely is that scenario?

Whilst expressed in theoretical, or economic, terms there is a real practical issue here; when running a project, how much effort should be put into certification, and how do we know when to stop? (It is always possible to do more work and spend more money.) The talk will identify some of the inherent uncertainties in managing software certification and give some sanitised metrics on certification costs (mainly from the safety domain) which illustrate how much RoI can vary. It will also suggest some criteria by which we might judge any scheme to manage software certification so as to deliver positive RoI, and use this to stimulate a debate on how we evaluate the benefit of individual methods used in support of software certification.

3.24 Refinement may help for Certification

Dominique Mery (LORIA – Nancy, FR)

Joint work of Mery, Dominique; Singh, Neeraj Kumar

Formal methods have emerged as an alternative approach to ensuring the quality and correctness of the high confidence critical systems, overcoming limitations of the traditional validation techniques such as simulation and testing. Certification aims at assessing or demonstrating that a system complies with a collection of rules, regulations and standards defining the minimum requirements a system must have to be deployed, operated and dismissed in a safe way. It appears that formal methods may help in the process of certification... but with toil and with tools. We describe a methodology for developing critical systems from requirement analysis to automatic code generation with standard safety assessment approach. This methodology combines the refinement approach with various tools including verification tool, model checker tool, real-time animator and finally, produces the source code into languages using automatic code generation tools. This approach is intended to contribute to further the use of formal techniques for developing critical systems with high integrity and to verify complex properties, which help to discover potential problems. Assessment of the proposed methodology is given through developing a standard case study: the cardiac pacemaker. The pacemaker is a proposed case study of the grand challenge and we analyse the role of the refinement in the methodology, which is to be improved and experimented on other case studies. Finally, the refinement relationship provides a way to play with abstractions of software system under design. The general methodology promotes the use of refinement and the incremental development of models so called abstractions. As in classical engineering approaches, abstractions play a central role and we have provided additional tools for easing the communication between domain experts and method experts.

References

1 D. Méry, N. K. Singh. Real-Time Animation for Formal Specification. In: Complex Systems Design & Management 2010, M. Aiguier, F. Bretaudeau, D. Krob (réd.), Springer, pp. 49–60. Paris, France, octobre 2010.
2 D. Méry, N. K. Singh. Critical systems development methodology using formal techniques. In: 3rd International Symposium on Information and Communication Technology – SoICT 2012, ACM, pp. 3–12. Ha Long, Viet Nam, août 2012.
3 mery:inria-00638473 D. Méry, N. K. Singh. Formalisation of the Heart based on Conduction of Electrical Impulses and Cellular-Automata. In: International Symposium on Foundations of Health Information Engineering and Systems (FHIES, 2011), Z. Liu, A. Wassyng (réd.), UMIST Macau. Johannesburg, South Africa, août 2011. conference 2011.
4 D. Méry, N. K. Singh. Technical Report on Formal Development of Two-Electrode Cardiac Pacing System. Rapport, février 2010.
5 D. Méry, N. K. Singh. Technical Report on Formalisation of the Heart using Analysis of Conduction Time and Velocity of the Electrocardiography and Cellular-Automata. Rapport, août 2011.
6 D. Méry, N. K. Singh. Technical Report on Interpretation of the Electrocardiogram (ECG) Signal using Formal Methods. Rapport, 2011.
7 D. Méry, N. Singh. EB2ALL- The Event-B to C, C++, Java and C# Code Generator. 2011, http://eb2all.loria.fr.

3.25 Certification Challenges for Software With Uncertainty

Richard F. Paige (University of York, GB)

Joint work of Paige, Richard F.; Burton, Frank; Rose, Louis; Kolovos, Dimitrios; Poulding, Simon; Smith, Simon
Main reference F.R. Burton, R.F. Paige, L.M. Rose, D.S. Kolovos, S. Poulding, S. Smith, "Solving Acquisition Problems Using Model-Driven Engineering," in Proc. of the 8th European Conf. on Modelling Foundations and Applications (ECMFA'12), LNCS, Vol. 7349, pp. 428–443, Springer, 2012.
URL http://dx.doi.org/10.1007/978-3-642-31491-9_32

This talk presented several different types of software applications – all considered safety-critical or safety-related – that exhibited uncertainty in some sense. By uncertainty, we meant "openness" or "unpredictability of behaviour", wherein some behaviours were emergent at run-time. One application, based on [**?**], made use of evolutionary algorithms to calculate optimal solutions to decision making problems in the enterprise (e.g., acquisition of new equipment, new training of personnel, development of new policies). The second application exhibited two types of uncertainty: in terms of data sources (which could change at run-time, e.g., switching from a batch or cached data source where data quality was guaranteed to a real-time error-high data source) and in terms of user-programmability (where users could encode new behaviours in the system by developing new execution patterns). The two applications were illustrated and a set of certification challenges were identified. These challenges included gathering of evidence, auditing evidence, assessing the quality of evidence, and process issues (e.g., ensuring that sufficient assurance was obtained for specific types of changes).

3.26 Models and Certification

Andras Pataricza (Budapest Univ. of Technology & Economics, HU)

Model Driven Engineering (MDE) becomes the main trend in critical embedded systems design. The presentation focuses on the fundamental question: How the different models and artifacts can be reused for supporting the certification process? A novel solution for traceability support is presented. The use of ontologies for the formalization of text documents like standards is proposed. Dependability analysis is addressed in the context of MDE. Finally, an approach for the empirical validation of models and underlying assumptions is proposed.

3.27 From Tool Qualification to Tool Chain Design

Jan Philipps (Validas AG – München, DE)

As engineering disciplines become more mature, more emphasis is put not only on the way of working, but also on the tools used. In safety standards, we can observe a similar development. Earlier standards put only little emphasis on tool use, perhaps roughly demanding a separate argument that each tool individually be "fit for use." Recent standards, such as the ISO 26262, take a more holistic viewpoint. Not only tools themselves, but also their use in the development process of the project must be analyzed, risks identified and only if necessary, further tool qualification measures employed.

Qualification measures typically are a combination of extensive tool testing and an analysis of the development process used for the tools. These approaches tend to be rather costly. Some tool vendors give support in the form of so-called qualification kits, which, for instance, include ready-to-run test cases. In practice, however, development tool chains consist not only of commercial tools, but also of open source or bespoke tools. The cost of qualifying all these tools would be prohibitive.

However, the holistic viewpoint taken in the ISO 26262 opens up new possibilities in building trust for tools and tool chains. Instead of just looking at isolated tools, it is possible to reduce qualification demand through tool diversity or through error avoidance and detection mechanisms within the tool chain. This is not unlike systems design, where safety is achieved by a combination of architectural choices and component reliability.

The talk presents this shift from tool qualification to tool chain design and gives some examples from an industrial project.

3.28 Cloud Security: Information Segregation and Data Privacy

Julia Rubin (IBM – Haifa, IL)

As cloud computing gains popularity both in private and enterprise sectors, it is only reasonable that customers expect guarantees for secure care of their data. In this talk, we focus on certification for cloud security, specifically information segregation and data privacy. We motivate the need for such certification and discuss its main differential factors from the more established certification procedures, e.g., those in the domains of safety-critical systems and, recently targeted, sensitive data protection in the healthcare and financial industries.

One of the major challenges for security certification is to verify not only that the system does what it is expected to do, but also that is does not do what it is not expected to. To address this challenge, we propose a specification-based runtime verification approach which relies on model checking, model-based testing, monitoring and run-time data analysis techniques. We believe that this approach is significantly more powerful than conventional testing and more practical than traditional formal verification for verifying security properties of a system.

3.29 Logic and Epistemology in Assurance Cases

John Rushby (SRI – Menlo Park CA, US)

Any assurance case comes down to two kinds of questions: how complete and accurate is my *knowledge* about aspects of the system (e.g., its requirements, environment, implementation, hazards) and how accurate is my *reasoning* about the design of the system, given my knowledge.

The first of these is a form of *epistemology* and requires human experience and insight, but the second can, in principle, be reduced to *logic* and then checked and automated using the methods of formal verification. (There are "inner" epistemic questions here, concerning correctness of the verifier, but we'll postpone those for the time being.)

The distinction between concerns that are epistemic or logical in origin is exactly that underlying the traditional partitioning of assurance into Validation and Verification (V&V). To some extent, it is possible to trade the two kinds of concerns (e.g., strong fault models allow simple fault-tolerant implementations: this reduces logic doubt about correctness of the implementation but increases epistemic doubt about verity of the model).

We propose that reducing epistemic doubt should be the main focus in assurance cases, and discuss ways in which this might be achieved.

3.30 Model-Based Development and Functional Safety

Bernhard Schaetz (fortiss GmbH – München, DE)

Model-based development has demonstrated its benefits in the automotive industry in improving development time and costs. Being a de-facto standard in automotive software development, it has found explicit acknowledgement as relevant technique by ISO 26262.

Its main advantages lie in front-loading of quality assurance techniques and automation of implementation steps. We show how this can also contribute to functional safety by explicating assumptions about platform and environment, by enabling high degree of precision as well as a scalable degree of detail, by supporting in-depth understanding of assumption, by supporting correctness of design and implementation, and by enabling automation of analysis and synthesis.

3.31 Software Cerrtification Challenges in the Nuclear Power Domain

Alan Wassyng (McMaster University – Hamilton, CA)

The current state of (software) certification in the nuclear power industry, its major challenges and a brief comparison with other domains in which systems are safety critical. The primary examples are drawn from a Canadian perspective.

3.32 Certification of Medical Information Systems – A paradigm shift: from devices to systems, from functions to data

Jens H. Weber (University of Victoria, CA)

Medical information systems (MIS) play a pivotal role in modern health care systems. They perform increasingly critical functions with respect to human safety, privacy and security. A significant number of adverse events has been associated with failures in MIS and this has resulted in increasing calls for their certification and regulation. Unfortunately, there is much confusion and little agreement on how to certify these systems. Existing paradigms as for example used for certifying traditional forms of medical devices do not readily apply. My presentation contrasts unique aspects of MIS from those found in other classes of critical computer-based systems. I point out why current certification techniques are insufficient with respect to achieving the overarching certification objectives. I then describe a paradigm shift that is needed in order to arrive a more effective certification program for MIS. Finally, I indicate research challenges and opportunities in this area.

3.33 Software certification in aeronautics

Virginie Wiels (ONERA – Toulouse, FR)

This talk presents an overview of avionics software certification, including the process, the certification authorities, the different criticality levels. It describes the main principles of DO-178, which is the certification standard in this domain. It briefly mention the recent update of DO-178 and in particular the Formal Methods Technical Supplement. It also lists future challenges for software certification.

References

1 Guidance for Using Formal Methods in a Certification Context. Duncan Brown, Hervé Delseny, Kelly Hayhurst, Virginie Wiels. ERTSS 2010 May 2010, Toulouse, France

3.34 Some experience and remarks on security certification at industry

David von Oheimb (Siemens AG – München, DE)

In my overview talk I share my experience with IT security certification at Siemens Corporate Technology. This type of activity is relatively rare in industry, for a number of reasons backed up by several examples. I briefly introduce the Common Criteria (CC), list several types of involvement of our group in their use, at examples like digital tachographs, an airplane SW distribution system, smart card processors, and smart metering gateways.

4 Overview of Working Groups

4.1 Challenges: Compositional Certification

In virtually all domains of modern computing systems, software size and system complexity are growing rapidly. Increased scale and complexity are straining current methods for system development, and in particular, methods for ensuring safety and for certification. A general engineering principle for managing complexity is to (a) decompose a system down into multiple smaller components that can be worked with individually through multiple phases of development, and (b) integrate components in later stages of development to form a complete system. Decomposing systems into components can also lead to cost reductions and decreased development time when components are *reused* across multiple systems. Unfortunately, the effectiveness of these strategies is limited in the context of certified systems, because almost all certification regimes for safety-critical domains certify complete systems – not system components. This state of affairs does not result from a lack of insight or interest on the part of industry or certification authorities. Rather, it is driven by the fact safety issues often arise due to incorrect context assumptions or unanticipated interactions between components, between software and associated hardware, and between a functioning system and its environment. Therefore, reasoning about safety, using current practices, is most easily carried when one has as much context information as possible. Of course, context

information is maximized when working with *all* system components in an integrated state, i.e., when working with the system as it will actually be deployed.

A seminar working group considered the issues related to component-wise development in the context of certified systems. In particular, the group was interested in the issue of *compositional certification:* a process by which individual components could be certified so that when components where assembled into complete systems, certification activities and arguments would not require a full assessment of all components implementations but instead could rely, to a large extent, the certification outcomes of the individual components. Such a process would allow both component implementations *and their certification artifacts/results* to be reused across different system implementations.

4.1.1 Current practices and trends motivating the need for compositional approaches to certification

In addition to the general challenge of dealing with systems of increasing size and complexity, the working group noted several important trends in system development that are motivating the need for compositional approaches to certification.

The notion of an integrated "system of systems" (SoS) is a category of systems that are increasingly prevalent and particularly challenging to certify. While definitions vary, a SoS is generally understood to be a collection of systems that can each function as a stand-alone system in some capacity but are integrated typically by some network-centric architecture to achieve new mission capabilities not obtainable from the individual systems. Small to mid-scale examples include modern automotive and avionics systems, where many microcontrollers, sensors, and actuators interconnect via a communications infrastructure that allows information from each set of sensors to calculate actions of actuators across the entire system. In larger examples, military command and industrial control systems are increasingly moving from stand-alone, monolithic designs to integrated platforms.

The concept of *software product line engineering* has proven to be very effective in a number of large organizations. Product line engineering is applied when one has a family of similar systems. An effort is made to (a) identify functionality that is common across multiple systems within the family, (b) design and implement components that provide that functionality, and (c) systematically design systems so that common components can be reused across multiple systems within the family. The end result is that costs are saved and development time is decreased by reusing components across multiple systems. While the product line concept is most often applied within a single organization, component reuse across organizations is facilitated when component interfaces are clearly defined or standardized and when a commodity market develops for component implementations.

Architecting *computational platforms* is becoming a common approach for achieving reuse. In this approach, a run-time environment, common services, and frequently used application components are shared between many stakeholders. To develop an application, one need not develop system functionality from the ground up; instead, one focuses on developing application logic using the shared services and application building blocks to produce an application that executes in the provided run-time environment. Smart phone platforms such as the iPhone and Android one of the most prominent examples in the consumer space. A platform approach can also encourage innovation since, by removing the need to build many parts of the system, it allows more people with less capital to enter the market and contribute ideas.

While integrated systems in general are not common in the medical device space as they are in e.g., avionics, automotive, etc., several talks in this seminar have presented research

results related to the concept of a *medical application platform* (MAP) [4]. A MAP is a safety- and security- critical real-time computing platform for (a) integrating heterogeneous devices, medical IT systems, and information displays via a communication infrastructure and (b) hosting application programs (i.e., *apps*) that provide medical utility via the ability to both acquire information from and update/control integrated devices, IT systems, and displays.

In summary, the compositional and/or component-wise development present in all of these types of systems mentioned above cannot currently be adequately aligned with existing certification regimes due to their lack of support for compositional certification arguments.

4.1.2 What would we like to be able to do that we can't do now?

The group identified several specific capabilities, technologies, and products that it believed would be important for moving toward more compositional approaches to building safety- and security-critical systems.

- (pre)-Certifiable high-assurance platforms that organizations can purchase and re-use.
- Better technologies (or more widely applied technologies) for ensuring partitioning in terms of space and time. This helps ensure that composition will not negatively impact safety.
- Approaches for making compositional safety arguments
- Formal capture of functional and non-functional properties to enable mechanical reasoning (e.g., checking interface compatibility, checking interface compliance) either a priori or at run-time/composition time.

4.1.3 What are the barriers (technological, standards, regulatory policy, political, social) that are hindering an advance toward a solution?

Regarding technology barriers, there is a need for better methods for dealing with *emergent behaviors* – behaviors that are not present or at least cannot be understood well in individual components but only arise when components are integrated. Specifically, better techniques are needed for:

- a priori recognition (or the possibility of) hazardous emergent behaviors – in essence, there is a need to develop compositional approaches for hazard analysis and risk assessment,
- engineering principles for minimizing unanticipated interactions and emergent behaviors,
- better and more systematic post-deployment means of detecting harmful emergent behaviors (you don't know what you don't know, so at least try to check after the system is deployed if something is going wrong).

Proposition compositional development relies on well-defined and precisely specified interfaces. There is need for better methods of (a) capturing interface properties that specify *non-functional* "contracts" including quality of service (QoS) and real-time properties, (b) capturing effects that a component can have with the environment (to detect possible interactions through the environment). As an example of the later property, magnetic resonance imaging (MRI) machines may impact the environment by causing interference with wireless communication of other medical devices in close proximity. Better technology is also needed for formally capturing modal/state behavior and security policies of components. In the vision for medical application platforms introduced above, system composition/integration in a health-care delivery organization – after system platforms and associated apps and devices are certified and deployed. Therefore, there is a need for technology that would allow the

run-time system of the medical platform itself to dynamically moderate composition so as to only enable those compositions that will produce a safe system.

Regarding business barriers, it was noted that some companies prefer proprietary systems so that they can lock small manufactures out of the market. Thus, while society as a whole would benefit from more open, component-based approaches, many companies are not interested in defining the interfacing standards necessary for achieving this vision. Moreover, systems composed from heterogeneous components (components from different manufacturers) often lead to questions about liability, i.e., which manufacturer is liable when a system fails.

Regarding regulatory barriers, allowance for compositional approaches in existing certification and regulatory guidelines is minimal. In the medical space for example, there are currently no guidelines for how systems following the notion of "medical application platform" should be regulated. To some degree this is justified because the community has not yet arrived at a convincing approach for demonstrating safety in the compositional setting. In the medical space, interoperability and compositional approaches are also hampered by the lack of widely implemented interoperability standards. In avionics, Integrated Modular Avionics (IMA) [2] provides some notion of reuse, but only supports static configuration; that is, changing configurations requires new certification. Seminar attendees shared the perspective that some in the broader community believe that IMA will reduce costs, but others say that the use of IMA increases the difficult of making safety arguments – to the extent that the cost reductions achieved are not as significant as they are "advertised" to be.

4.1.4 What evidence is there that success is possible?

Participants shared several notions of reuse or compositionality in certified systems.

- In the security domain, one of the original motivations for the Multiple Independent Levels of Security (MILS) architecture [1] was to promote a commodity market of reusable components for security systems to be certified according to various protection profiles within the Common Criteria [3].
- In the avionics domain, guidance for Integrated Modular Avionics (DO-297) describes how a particular architecture supports flexibility, modularity, reusability and interoperability.
- ISO 26262, which addresses functional safety for road vehicles, includes the notion of "safety element out of context" (SEooC) which allows the statement of assumptions and guarantees for a particular element (informally, a sub-system) whose development is being carried out by, e.g., a sub-contractor.
- The FAA Advisory Circular on Reusable Software Components (AC 20-148) provides a means for reusable software component (RSC) developers, integrators, and applicants to gain: FAA "acceptance" of a software component that may be only a part of an airborne system's software applications and intended functions, and credit for the reuse of a software component in follow-on systems and certification projects, including "full credit" or "partial credit" for compliance to the objectives of RTCA/DO-178B, Software Considerations in Airborne Systems and Equipment Certification.
- UK Defense Standard 23-09, standardizes interfaces within the Generic Vehicle Architecture, an open source architecture specification with well-defined interfaces that aims to encourage reuse in military ground vehicles.
- The cross-domain functional safety standard IEC 61508, allows the notion of a "pre-certified" component that can be integrated to a 61508-certified system while reusing some of the certification effort. For example, a Green Hills produces a a pre-certified IEC 61508 Safety Integrity Level 3 real-time operating system kernel – the Platform for Industrial Safety (PIS). IEC 61508 is an international standard for the functional safety

of electrical/electronic, programmable electronic systems (PES) and is well established in the industrial process control and automation industry. Because IEC 61508 serves as the meta-standard for a range of industries and published standards, the Platform for Industrial Safety is directly applicable to railway (CENELEC EN 50128), medical (IEC 60601), nuclear (IEC 61513), process control (IEC 61511), and automotive (ISO 26262). Such pre-certified components are documented in 61508 by a "Safety Manual".

These existing approaches could be considered to be precursors to an eventual more robust and rigorous approach to compositional certification. Techniques such as self-verifying software, runtime verification, proof-carrying code, model composition focused on system verification, modular assurance cases [5] may be key enabling technologies.

References

1. C. Boettcher, R. DeLong, J. Rushby, W. Sifre. The MILS Component Integration Approach to Secure Information Sharing. Presented at the 27th IEEE/AIAA Digital Avionics Systems Conference (DASC), St. Paul MN, October 2008.
2. P. Conmy, M. Nicholson, J. McDermid. Safety assurance contracts for integrated modular avionics. Proceedings of the 8th Australian workshop on Safety critical systems and software, Volume 33, pp. 69–78, 2003.
3. R. DeLong, J. Rushby. A common criteria authoring environment supporting composition. Proceedings of the 8th International Common Criteria Conference, 2007.
4. J. Hatcliff, A. King, I. Lee, A. Macdonald, A. Fernando, M. Robkin, E. Vasserman, S. Weininger, J. Goldman. Rationale and Architecture Principles for Medical Application Platforms, Proceedings of the 2012 International Conference on Cyber-Physical Systems, pp. 3–12, April, 2012.
5. T. Kelly, S. Bates. The Costs, Benefits, and Risks Associated With Pattern-Based and Modular Safety Case Development. Proceedings of the UK MoD Equipment Safety Assurance Symposium 2005, October 2005.

4.2 Challenges: Education and Challenge Problems

The seminar included significant discussions on education related to certification, and in particular, on the use of challenge problems and realistic case studies in education.

In prepared remarks given at the Information Technology and Innovation Foundation, April 12, 2012, Washington, DC, Thomas Kalil listed several properties/goals of Grand Challenges: (1) they can potentially have a major impact in the domain, (2) they should be ambitious but achievable, (3) they should be compelling and motivating, (4) they should be focused – must know when they have been achieved, and (5) they should drive innovation and advance technology. In addition, we note that challenge problems should be "research intensive". That is, if you look at the definition of a grand challenge it should be obvious that the emphasis is on the research involved in the challenge.

The community associated with this Dagstuhl Seminar has created several challenge problems in the medical domain (e.g., the Pacemaker, and Patient-Controlled Analgesic Pump described in talk abstracts appearing earlier in this report). For example, Boston Scientific (through Brian Larson) released into the public domain a sanitized requirements document for a 10 year old pacemaker. Brian also worked with an Electrical and Computing Engineering Capstone class at University of Minnesota to design a hardware reference platform. Mark Lawford manufactured and sold 50 modded units. The specific challenge was to use the natural language requirements document as the basis for a formal approach to building a

pacemaker Currently more than ten different prominent institutions tackling the challenge in coursework. Over twenty research papers have been published that use pacemaker artifacts. A Dagstuhl Seminar on the Pacemaker Challenge has been approved for Feb 2–7, 2014.

Assessing the pacemaker challenge against Kalil's criteria for research challenges, it has all the attributes of a Grand Challenge but one: it is not focused enough. There was a lack of specificity regarding the research objectives and expected artifacts to be produced. The lack of focus has led to the effort being less of a driver for true innovations in verification techniques. Teams tended to just take small bites or slices of the problem and address issues that existing tools were already able to handle well. We must make sure not to repeat that mistake in future challenges. Another issues is that the challenge does not run itself – there are real resourcing issues. Experts really need to be available to provide background domain knowledge, occasional support for the hardware, and evaluation of proposed solutions.

However, as a driver for innovative pedagogy, the Pacemaker Challenge has been unexpectedly successful. For example, the capstone undergraduate class at McMaster (run by Mark Lawford), the third year class on software development for ECE and Mechatronics Engineers (Alan Wassyng) at McMaster, courses at the University of Pennsylvania, etc. all use the Pacemaker materials. In addition, the SCORE student competition activity at the International Conference in Software Engineering in 2009, the Pacemaker artifacts were used by four 4 groups, one of which won the prize for best use of formal methods.

There are major components of these grand challenges that are reflective of state of the art and do not require the final research breakthrough that is being sought in the Grand Challenge. They typically involve hot topics in their specific domains. Tremendous effort is being expended in tackling these problems, and we can borrow from that effort to reinforce our education components in these domains. We think we can bridge the gap between state of the art/practice and general practice by disseminating these problems with guidance material.

The audiences who might benefit from an increased emphasis on *educational* material associated with challenges like Pacemaker are quite diverse: (a) students in safety/security-critical systems, (b) industry engineers who we want to orient to "new/improved" techniques that we are proposing regulatory / certification agencies, and certification community members working on research topics to advance the state of the art in certification / regulatory science.

If one focuses on the pedagogical/education space with challenges problems, there slightly different goals to consider. The problem should be complex enough to require sound techniques for their solution, when supporting projects the problem should be complex enough to need groups rather than individuals, to add realism it is ideal to have both hardware and software components, the challenge should be "solvable" in a span of time related to semester length or academic year length (e.g., 8 months for a year-long capstone project, and can be scaled down so that a subset that can be solved in four months), and finally, there needs to be appropriate domain background material, requirements, etc.

What are some specific things that should be offered:

- Requirements specification – should this be natural language, formal, or both?
- Dependable hardware platform at reasonable cost
- Hardware manual(s)
- Clear goals
- Guidance for development of solutions
- Mechanisms for evaluating solutions
- Domain specific background material
- Certification related background material

- Details on a slice through the system
- Wiki with FAQ and discussion groups
- Competitions

This type of challenge problem activity has several benefits to students. They get to tackle a project that is representative of "real" projects, they (hopefully) get excellent support material, they get guidance "best practices" and on how to "do it right" for a realistic problem, and they are typically excited to have it appear on their resumes. Pre-packaged educational material for challenge problems also has numerous benefits for instructors. The details of a "case study" is given to them with all supporting material including availability of a hardware platform. They get support through a Wiki. They get the benefit of cross-fertilization of a common problem tackled in many countries When the supplied pedagogical material includes evaluation criteria and guides, they instructors do not need to work so hard on this themselves. Finally, as a challenge problem's use increases, "frequently asked questions" and Wiki contents accumulate which provides more resources for future issues.

4.3 Challenges: Security

IT security is of increasing importance in various areas as electronic commerce and governance, any kind of access control systems, privacy of personal and in particular medical data, and information segregation in the cloud. IT-security is a cross-cutting concern in most systems. Moreover, with the tighter integration of formerly stand-alone systems to systems of systems and the enhancement of functionality, functional safety of systems relies on IT security. Incidents emphasizing the vulnerability of safety-critical systems by security attacks have been reported e.g. from the transportation domain and medical systems.

Whereas safety considerations on software have a tradition since decades as they were required with the first usage of programmed functionality in safety-critical systems, the need for specific and systematic IT-security analysis and design methods has been recognized later. Initially a lot of methods have been transferred from safety, in particular approaches to assure functional correctness apply for both areas. However, the requirement to specifically identify potential loopholes usable by malicious attackers is a distinctive feature of security analysis. Standards, in particular the Common Criteria, are widely accepted nowadays. Due to practitioners, the standards have proven beneficial, but but need to be applied in a sensible way.

The working group stated the following *general* challenges of software certification for IT security: How can manufacturers actually prove that their products are secure, rather than just declaring it? How can the proliferation of profound new security techniques into industrial practice be accelerated? Due to short life cycles of many applications and and missing compositionality, security is approved only for a short period of time. Thus approaches to continuous (re)certification are needed. Many standards can be improved in order to be to more prescriptive, more up to date, and document best practices. Security and safety concerns shall be unified to support correct-by-design, safe-by-design, and secure-by-design.

As a particular issue, that shall be addressed in further research, *compositionality in security analysis and design* was identified: Obviously assembling secure components is not sufficient to produce a secure system since security – as safety – is an emergent system property. Layering in the architectural design and considering security a first class issue may help, but will not solve all issues. Further classification is needed in order to identify architectural patterns that support particular security properties.

Awareness and education is still an issue, because even software developers who graduated recently are not aware of security and do not know about methods to design and implement secure software. Even in large IT projects an (independent) security expert in the design team is often missing.

Last but not least the perception of IT security in the public audience will be a key factor not only for future research activities on IT security, but even more for the success or failure of several sectors within IT industries like social media or the cloud that rely on commonly accepted security policies.

4.4 Challenges: Tool Qualification

Tool qualification is the process by which certification credit may be claimed for the use of a software tool. The purpose of tool qualification is to provide sufficient confidence in the tool functionality so that its output may be trusted. Tool qualification is, therefore, a significant aspect of any certification effort. The tool qualification working group attempted to identify initial needs and challenges related to use of software tools in certification efforts.

One way that we may obtain confidence in the output of a software tool is to provide some independent verification of the tool output. This may be accomplish by a manual review (if feasible) or by using an independent tool of equivalent functionality and comparing the outputs. Qualification is required whenever a software tool is used to eliminate, reduce, or automate a software life cycle process without the tool's output being verified.

Many different types of tools may be used in a software development process that could impact the correctness of the software. For example:

- Requirements engineering tools for eliciting, capturing, and specifying requirements
- Traceability tools for managing the design rationale and connections between software life cycle data
- Design tools for transforming requirements and constraints into design models
- Transformation tools such as code generators and compilers
- Test generation tools for producing test cases from requirements specifications
- Verification tools for analyzing design models or code to determine compliance with requirements
- Tools for generating and checking software configuration data

However, tools may be broadly categorized for qualification purposes (as in DO-178B) according to how they may impact software correctness:

- Tools that could introduce a fault into the software (development or transformation tools)
- Tools that could fail to detect an error already present in the software (verification tools)

Tools may be further categorized according to the criticality of the product software that they are generating or verifying.

In general, the qualification requirements for a development or transformation tool are much more stringent than for a verification tool. A reasonable approach in some cases is to qualify an independent verification tool to check the output of an unqualified development tool. For example, a complex tool may be needed to generate a schedule; however, tool for checking the correctness of a given schedule is relatively simple.

4.4.1 Needs

The working group identified a number of needs in the area of tool qualification.

Software tools are used in development processes to automate life cycle activities that are complex and error-prone if performed by humans. The use of such tools should, in principle, be encouraged from a certification perspective to provide confidence in the correctness of the software product. Therefore, we should avoid unnecessary barriers to tool qualification which may inadvertently reduce the use of tools that would otherwise enhance software quality and confidence.

Most software tools are not used in isolation, but are used as part of a complex tool chain requiring significant integration effort. In general, these tools have been produced by different organizations. We need to develop better and more reliable methods for integrating tools from different vendors (including university tools, open source tools, and commercial tools).

A given software tool may be used in different application domains having very different requirements for both certification and tool qualification. Furthermore, the methods and standards for tool development varies across domains. Consistent qualification requirements across different domains would simplify the process.

4.4.2 Barriers

The working group discussed several barriers that may inhibit tool qualification today.

Complexity is a barrier to tool qualification that may manifest itself in several ways. The input/output space of the tool may be complex, as is the case for compilers and code generators. For other tools, such as model checkers, the results produced are difficult verify. Another source of complexity may be the algorithm used by tools, as in the case of transformation tools that rely on artificial intelligence techniques, evolutionary methods, or non-deterministic algorithms.

The platform on which a tool executes may present barriers to qualification. For example, the use of a virtual machine (VM) may introduce uncertainties in how a tool executes that may be hidden from view, or may vary when a different VM is used. Reliance on COTS libraries outside the scope of the software tool itself is another source of uncertainty if these libraries are not completely identified and specified.

Many valuable and effective software tools are developed using less than rigorous tool development processes. This is true for many tools developed in the university environment where the research agenda is a higher priority that qualification requirements. This can be an impediment the subsequent qualification of the tool for use in a certification process.

4.4.3 Roadmap

A number of steps were discussed that could be part of a roadmap for improving the tool qualification process.

Improved methods for tool analysis could be developed. For example, tool chain integration has been identified as a significant issue. Perhaps a HAZOP-type approach could be used to assess the potential errors introduced in integration. Another approach would be to identify common patterns of tool errors and way to control or avoid this errors.

In the area of tool construction, new methods for providing evidence of correct tool operation could be developed. For example, a tool could provide evidence as part of its installation or at runtime or correct operation. Another approach is to focus efforts on tool architectures based on the idea of a complex (but unqualified) transformation that is checked by a simpler (qualified) verification tool.

Another new and promising approach to tool qualification is found in the use of formal methods to verify software tools. The CompCert compiler project (http://compcert.inria.fr/) is an example of this approach.

A final need is a thorough comparison of the qualification viewpoints and demands of the different domains. This should include an assessment of the point of view of different regulatory bodies.

Several useful references for qualification in the avionics and nuclear domains were identified:

- *DO-330/ED-215: Benefits of the New Tool Qualification Document*
  http://www.adacore.com/knowledge/technical-papers/do-330-ed-215-benefits-of-the-new-tool-qualification-document/
- *Licensing of safety critical software for nuclear reactors: Common position of seven European nuclear regulators and authorised technical support organisations*
  http://www.hse.gov.uk/nuclear/software.pdf

4.5 Intellectual Basis for Certification & Confidence

The current intellectual basis for certification of software intensive systems and how regulators gain confidence in the safety and reliability of systems in largely based upon the application of process oriented standards. Currently there is growing consensus that the current status quo has to change. A working group at the seminar considered the related questions (i) what should form the intellectual basis of certification? And (ii) How do we gain sufficient confidence in software intensive systems?

4.5.1 What is happening in today's world and into the future that is driving the need for change?

- There is a lack of repeatability of certification results. Currently the outcome of a certification is overly dependent upon the evaluator(s), or, in some cases similar cases before the same evaluator results in different outcomes.
- Regulators and developers do not know how to evaluate different types of evidence and understand how they may be combined to achieve a desired level of confidence in the system. Further they do not know what evidence to collect or do not agree about relevant values of particular types of evidence. There are strong disagreements even among recognized experts on these points.
- There is unnecessary variation across application domains in terms of both practice and regulation. The goals are largely the same in all sectors – achieving tolerable risk in a software intensive system – yet there is wide variation on what constitutes acceptable evidence in different domains. This makes it more difficult to share expertise in both development and evaluation as well as methods and tools.
- Often developers and certifiers feel that there is wasted effort that does not add value since they are unsure sure how things contribute to adding confidence. There is a desire to eliminate busy work and focus efforts on those aspects of development and evaluation that make a real contribution achieving tolerable risk. Currently it is unclear how to determine whether effort is wasted or not.
- There is a gross association between Dependability Assurance Levels (DALs) or Safety Integrity Levels (SILs) to moderate confidence. It is not clear that such course grained classification of systems and their respective appropriate evidence is the best method.

- The types of systems we now want to build are different than they were in the past. They are now continually evolving, adaptive systems that may be rapidly reconfigurable during operations.
- It is not always clear how much confidence is "enough". There are sometimes technical issues in making this determination, but there are others such as the societal acceptance of the trade offs involved in the issues of risks vs. confidence vs. benefit vs. cost. It is possible to be overly cautiousness because of fear of the unknown or on the other hand be too quick to accept what may be a largely unknown risk.
- The difference between normal engineering (e.g. development of a minor revision of a well understood product) vs. radical design (development of a novel, first of a kind product) is not always understood. In the current practice of the engineering of software intensive systems, there is often no clear basis for distinguishing between them.
- The amount of safety-critical software is increasing substantially. Functions for which people were once responsible are being transferred to software systems. This may results in an inadequate ability to control the rate of change in requirements.
- In order to reduce costs there is trend towards the increased use of commercial software and hardware in critical systems. The incorporation of Software Of Unknown Providence (SOUP) into critical systems imposes further challenges for certification. This is but one example of how systems may be evolving in ways that are antithetical to safety-critical use.

To summarize, what would we like to be able to do that we cannot do now with the current basis of certification is to solve the problems posed above in a systematic way.

4.5.2 The Barriers

What are the barriers (technological, standards, regulatory policy, political, social) that are hindering an advance toward a solution? Commercial interests dictate that in order to preserve intellectual property such as trade secrets, companies are reluctant to provide full source code and system documentation, preferring "black box" certification. Further entrenched cadres are invested in the current practice viewing their knowledge of how to navigate current opaque certification regime as a competitive advantage and a barrier to entry for start-up competitors. There is little historic data available on certification. Correlating the data that does exist with causal factors of success/failure is very difficult.

There is a widely held idea in academia that verification is equal to certification. This demonstrates a lack of understanding of industry practice by academics resulting in research into and teaching of methods and tools that do not scale to industry problems. The view of the working group was that currently there is inadequate education for a basis for certification. Even what we know works well is not known widely enough and taught even less. The result is that we are graduating engineers that are not recognizing that confidence is an issue that needs to be addressed. When it is address there was concern in the working group that we are often using the wrong mathematics for reasoning about confidence. A view was expressed that quantitative arguments about the reliability of software intensive systems is misleading since it is extremely difficult to get accurate failure rates for software.

4.5.3 Evidence that success is possible

What elements of potential solutions exist in research, existing standards, existing technology? Some domains, such as aerospace, achieve good results although the reasons for the success are not necessarily known and the costs are often high. There is growing recognition of

the problems involved in the certification of software intensive systems and this is resulting in improvements in the capabilities of tools to support aspects of certification. Domains that previously have not had knowledge in the development and certification of software intensive systems are gaining it and there are some signs of movement from the past process-orientation certification regimes to product-orientation methods. There is a current trend towards company standards including explicit confidence arguments and a number of case studies are being performed. There is work being done of the development of reference designs to normalize engineering in areas such as infusion pumps and there are efforts to provide engineering cookbooks such as the FAA Requirements Engineering Management Handbook.

4.6 Methods for Developing Certifiable Systems and Methods of Certifying Systems

As part of the seminar a working group discussed methods of developing and certifying systems. This section summarizes their findings.

4.6.1 Need

What is happening in today's world or in the future that is driving the need for change?

We are wanting to build increasingly complex and interconnected systems (and systems-of-systems) and are using them in contexts that previously would be considered untenable or unsuitable for software. Our aspirations and reach, as engineers, is growing. The only real tools that we as software engineers have to work with are abstraction and decomposition, though it is questionable whether such reductionist approaches will apply to ultra-large-scale and complex systems. We must fundamentally use models (whether formal or structured) to manage growing complexity, and certification processes must reflect this reality, as well as the new forms of analysis that are available and go beyond what testing can feasibly achieve.

What would we like to be able to do that we can't do now?

- We would like to front-load our analysis and not have to wait until we have a prototype or completely build a system, which may be too late to realize that we are building a system that cannot be certified.
- Replace parts of testing with more rigorous formal analysis/exhaustive analysis and have the results of exhaustive analysis be considered as valid evidence by certifiers.
- Qualifying advanced model-based development tools (which include non-traditional programming constructs – e.g., rule-based programming). Why do we need to develop and use models (of any kind) to help develop certifiable systems? They are useful in particular for evolution.

4.6.2 Barriers

What are the barriers (technological, standards, regulatory policy, political, social) that are hindering an advance toward a solution?

Regulatory policy is conservative. Standards evolve slowly and do not necessarily reflect what can be achieved with formal methods or model-based design. This may be a positive thing, given that engineers may be concerned with safety, but having ways to evolve standards that can take into account proven forms of analysis and tooling would be beneficial though we can argue about what "proven" means.

A sociological problem associated with formal methods exists. They have been oversold, misused and misunderstood. This has resulted in bad experiences in critical systems development – typically leading to either hiding use of formal methods ("under the hood") or not using them at all. Certification is costly to begin with – what effect does formal methods or model-based development have on this cost, especially taking into account the cost of adopting formal methods or model-based development. This links back to John McDermid's talk on ROI. There has been failures in standardization. For example, UML is conceptually a successful standard but implementation wise (e.g., in terms of XMI/export/input) it is a failure. XMI evolved too slowly and tool vendors did not implement it correctly (or added their own "flavour").

4.6.3 Evidence that success is possible

What elements of potential solutions exist in research, existing standards, existing technology? There are success stories in applying MBD and Formal methods in different application domains such as nuclear, aerospace and medical devices. We hope to summarize these experiences in the post seminar publication.

4.6.4 Envisioned solution

How would the world (or at least the context of certified systems) change if the challenge could be overcome? What form would solution(s) to the challenge take? new technology? new standards? new methodologies / principled approaches?

Best practice guides to using formal methods or model-based development for building certified systems are critical. This clearly needs some success stories and pilot experiments, of which there are many (at least in the academic literature), but these probably need to be reformulated and expressed differently to get the value proposition to industry across more clearly. Standard interfaces that connect current practice (e.g., modeling in CSV) to new practices (e.g., modeling tools) are also needed.

4.6.5 Roadmap to a solution / research agenda

What concrete steps might the community take to move forward on this challenge?

- Developing methods of analysis for model consistency and coverage
- Clarifying the notions of abstraction used in different models and how to communicate abstractions across domains
- Gateways between methods and tools and improved traceability
- Scale: Your method should scale and you need to provide indicators to decide when it won't
- Education: Being aware of the notion of "logical proof" already at high school
- Assessment examples (rather like Tim Kelly's presentation comparing standards) that show how specific MBD or formal methods can produce evidence that is "at least as convincing as" what is indicated in various standards. For example, take DO178C MBD annex and actually compare/assess a specific MBD approach against it to see what constraints need to be applied to the approach, and whether the approach can produce sufficiently compelling evidence.
- If you are an academic with a specific MBD/formal technique, working well on large problems, consider how far are you from something that can be certified.
- A "how-to" guide for building a formal method/MBD technique that can be certified/qualified.

Participants

- Dominique Blouin
Université de Bretagne Sud, FR
- Darren Cofer
Rockwell Collins – Cedar Rapids, US
- Cyrille Comar
AdaCore, Paris, FR
- Mirko Conrad
The MathWorks GmbH – Ismaning, DE
- John S. Fitzgerald
Newcastle University, GB
- Kim R. Fowler
Kansas State University, US
- Hubert Garavel
INRIA Rhône-Alpes, FR
- Janusz Górski
Gdansk Univ. of Technology, PL
- Arie Gurfinkel
CMU – Pittsburgh, US
- John Hatcliff
Kansas State University, US
- Mats P. E. Heimdahl
University of Minnesota, US
- Constance L. Heitmeyer
Naval Res. – Washington, US
- Michael Holloway
NASA Langley ASDC – Hampton, US
- Jozef Hooman
Radboud Univ. Nijmegen, NL
- Jérôme Hugues
ISAE – Toulouse, FR
- Michaela Huhn
TU Clausthal, DE
- Hardi Hungar
German Aerospace Center – Braunschweig, DE
- Daniel Kästner
AbsInt – Saarbrücken, DE
- Peter Karpati
Institute for Energy Technology – Halden, NO
- Vikash Katta
Institute for Energy Technology – Halden, NO
- Tim Kelly
University of York, GB
- Andrew King
University of Pennsylvania, US
- John C. Knight
University of Virginia, US
- Brian Larson
Multitude Corp., US
- Mark Lawford
McMaster Univ. – Hamilton, CA
- Dominik Mader
Berner & Mattner Systemtechnik – Berlin, DE
- Tom S. Maibaum
McMaster Univ. – Hamilton, CA
- John McDermid
University of York, GB
- Dominique Méry
LORIA – Nancy, FR
- Frank Ortmeier
Universität Magdeburg, DE
- Richard F. Paige
University of York, GB
- Andras Pataricza
Budapest Univ. of Technology & Economics, HU
- Jan Philipps
Validas AG – München, DE
- Robby
Kansas State University, US
- Julia Rubin
IBM – Haifa, IL
- John Rushby
SRI – Menlo Park, US
- Bernhard Schätz
fortiss GmbH – München, DE
- David von Oheimb
Siemens AG – München, DE
- Alan Wassyng
McMaster Univ. – Hamilton, CA
- Jens H. Weber
University of Victoria, CA
- Virginie Wiels
ONERA – Toulouse, FR

Report from Dagstuhl Seminar 13052

Multicore Enablement for Embedded and Cyber Physical Systems

Edited by

Andreas Herkersdorf[1] and Michael Paulitsch[2]

1 **TU München, DE,** herkersdorf@tum.de
2 **EADS – München, DE,** michael.paulitsch@eads.net

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13052 "Multicore Enablement for Embedded and Cyber Physical Systems". During the seminar the participants from industry and academia actively discussed chances and problems of multicore processors in embedded in cyber-physical systems. The focus of the seminar was on the exchange of experiences and discussion of the challenges of reusable and transferable multicore technologies. Those were covered in the individual talks and plenum discussions. Beside that, working groups have been formed to discuss and present important topics in detail, which are also part of this report.

Seminar 27. January to 01. February, 2013 – www.dagstuhl.de/13052
1998 ACM Subject Classification C.1.4 Parallel Architectures
Keywords and phrases Multicore, hardware, software, platforms, embedded systems, security, real-time, safety, cyber physical systems
Digital Object Identifier 10.4230/DagRep.3.1.149
Edited in cooperation with Stefan Wallentowitz

1 Executive Summary

Andreas Herkersdorf
Michael G. Hinchey
Michael Paulitsch

Multicore processors are a key enabling technology for solving grand societal challenges of the coming decades. Secure and ecological mobility, geographic coverage of high-tech healthcare, sustainable energy generation, distribution and management, and in general the development of our digitized society impose compute performance requirements on distributed embedded and cyber physical IT equipment which makes multicore technology indispensable. All leading processor vendors – ARM, Freescale, IBM, Infineon, Intel, MIPS, Nvidia – follow a strictly multicore-oriented strategy. Due to the paradigm shift from exploiting instruction level to process level parallelism, multicore processors are superior over single-core representatives with respect to computing performance and energy efficiency. Prerequisite is, processes can be balanced among parallel cores such that the nominally available computing performance can be utilized effectively, and cores can be set into sleep mode or power gated when not busy. As of today, the ability to efficient utilize the available resources depends to a large extent on the aptitude of experienced programmers and the inherent ability of being able to parallelize the computing problem.

Embedded and Cyber Physical Systems exhibit demands for "non-functional requirements", such as low(est) power and energy dissipation, reliability, availability and security, real-time and cost constraints, which are typically not found to the same extent in general purpose computing applications. The enablement of multicore technology for embedded and cyber-physical markets imposes serious challenges to industry and academia which can easily overwhelm the capabilities and capacities of individual corporations or even consortia. Industry and university research in Europe recognized early and invested significantly into the establishment of multicore know-how and competences. Examples of related projects at EU level and in Germany are: RECOMP – Reduced Certification Costs Using Trusted Multicore Platforms, ACROSS – ARTEMIS CROSS-Domain Architecture, SPES 2020 – Software Plattform Embedded Systems 2020, Cesar – Cost-efficient methods and processes for safety relevant embedded systems, MERASA – Multicore Execution of Hard Real-Time Applications Supporting Analysability (see Relationship to other seminars and projects for a more complete listing), and ARAMiS – Automotive, Railway and Avionics Multicore Systems.

The seminar brought together leading industry and university research groups from different fields of embedded system design and application development, multicore architecture and hardware/software design methodology & tools. The main objective of the seminar was on reporting experiences and discussing challenges of reusable and transferable multicore technologies among participants representing different application markets and scientific backgrounds. The technical focus of the agenda was on:

- Generic hardware/software building blocks for real-time performance, dependability, functional safety and security for embedded systems built around enhanced standard multicore solutions.
- System modeling, design and validation methods and tools for such platforms.

The seminar established new and strengthened existing ties between players and networks in the area of multicore and embedded technologies. Topical working groups were formed on the following topics:

- Specification & Interference
- Industrial Perspective on MultiCore Motivations and Challenges
- Certification of Safety-Critical Multicore Systems: Challenges and Solutions
- Network-on-Chip – Dependability and Security Aspects
- Multicore Ecosystem
- Secure Elements in future embedded multicore systems

The working groups compiled summaries reflecting the status and outlook on the respective topic. These summaries can be found in the sequel of this report.

2 Table of Contents

3 Overview of Talks

3.1 The ARTEMIS ACROSS project

Christian El Salloum (TU Wien, AT)

The European ARTEMIS ACROSS project aims to overcome the limitations of existing Multi-Processor Systemson-a-Chip (MPSoC) architectures with respect to safety-critical applications. MPSoCs have a tremendous potential in the domain of embedded systems considering their enormous computational capacity and energy efficiency. However, the currently existing MPSoC architectures have significant limitations with respect to safety-critical applications. These limitations include difficulties in the certification process due to the high complexity of MPSoCs, the lacking temporal determinism and problems related to error propagation between subsystems. These limitations become even more severe, when subsystems of different criticality levels have to be integrated on the same computational platform. Examples of such mixed-criticality integration are found in the avionics and automotive industry with their desire to integrate safety-critical, mission critical and non-critical subsystems on the same platform in order to minimize size, weight, power and cost. The main objective of ACROSS is to develop a new generation of multicore processors designed specially for safety-critical embedded systems; the ACROSS MPSoC. This talk will show how the ACROSS MPSoC overcomes the limitations of existing MPSoC architectures in order to make the multi-core technology available to the safety-critical domain.

3.2 A Model-based Approach for Optimizing Existing Real-Time Software on Multicore Processors

Michael Deubzer (Timing Architects Embedded Systems GmbH, DE)

In many upcoming real-time system projects multicore processors are an integral part of the roadmap. Till today, software in the embedded industry has been mostly developed for single-core processor systems and represents a high investment for a company. To protect their investment those companies are now faced with the challenge to migrate the software to multicore processor systems. The approach presented in this talk describes a methodology to migrate single-core software to multicore processor systems by regrouping of functions and adding a hardware abstraction layer. The methodology is applied on the architectural granularity of functions which are directly called in tasks of a multitasking system and described by dataflow, execution time and resource demands. In the first step a partitioning heuristic groups functions, considering the dataflow and execution time demands, to tasks and allocates them to cores. In real-time systems two major requirements have to be guaranteed, namely coherency and consistency. This is solved by evaluating data dependencies and creating a middle-ware which copies shared data items of functions to a local memory and writes back data after execution. In order to avoid conflicts at the access to shared resources, a protection mechanism is applied which manages exclusive access and therefore limits the degree of interference between applications. For the timing evaluation of a certain allocation

of tasks to cores, an explorative simulation is applied and timing constraints are checked. This process has been automated and is used for a multi-objective optimization algorithm, searching for best partitioning of functions and allocation of tasks in terms of predefined criteria like reaction times, memory usage or bus traffic. By splitting tasks in smaller subtasks and allocating those to cores a set of software allocations is generated and solutions, fitting the system requirements and configuration best, can be selected.

3.3 IDAMC – A manycore architecture for mixed critical applications

Rolf Ernst (TU Braunschweig, DE)

Joint work of Ernst, Rolf; Jonas Diemer; Philip Axer
URL http://www.ida.ing.tu-bs.de/en/research/projects/aramis/

Mixed critical systems integrate application functions of different safety and time criticality. For such systems, safety standards require that the critical functions must adhere to the reliability requirements and are not be affected by the non-critical functions ("freedom from interference") while the less critical tasks shall be implemented with maximum efficiency. This talk will present a many-core architecture based on a Network-on-chip that provides core isolation and supports dynamic flow control for bounded timing interference. A formal timing model allows formal verification of performance constraints. First results will be presented.

3.4 Commerical Challenges of MultiCores in Automotive Domain

Glenn Farrall (Infineon – Bristol, GB)

While there are many challenging aspects to the deployment of Multicore devices, this talk takes a step back and considers some of the basic implementation issues of Multicore SoCs.

As well as additional die area – many characteristics of devices targeting the Automotive market (especially safety requirements) add to the power budget working against the major driver of multicore usage, namely higher performance per $ and per Watt. This talk briefly covers several of these issues.

3.5 Timing Predictability of Multi-Core Processors

Christian Ferdinand (AbsInt – Saarbrücken, DE)

All contemporary safety standards require to demonstrate the availability of sufficient resources to sustain correct functioning of the system. This includes determining safe upper bounds on the worst-case execution and response time of real-time tasks. In mixed-criticality

systems the entire system is subject to the highest occurring safety integrity level unless the independence of all safety functions can be demonstrated in the spatial and temporal domain. These requirements are imposed, e.g., by DO-178B, DO-178C, ISO-26262, IEC-61508, and EN-50128. Since FDA regulations and the German Medizinproduktegesetz require to take into account the state of the art they also pertain to software for medical devices. Spatial independence can be ensured by using partitioned operating systems, or can be proven by static analysis tools which, e.g. can demonstrate the absence of stack overflows or other runtime errors. However, many multi-core processors exhibit characteristics that make it difficult or even impossible to ascertain predictable performance: it may be hard to ensure freedom of interference and to determine safe worst-case execution time bounds. We give an overview of hardware features leading to interference and predictability problems, shows examples of predictability-oriented multi-core configurations, and describe a tool-based methodologies to ensure the correct timing behavior.

3.6 Road to the use of multicore processors in space systems

Massimo Ferraguto (Space Systems Finland Ltd – Espoo, FI)

The use of multicore in the space domain can be beneficial in terms of greater processing capability (concentration of multiple functions in one single computer, with partitioning by criticality level and/or function; more payload data processing on-board), weight, power and fuel reduction which ultimately lead to longer lifetime and cost efficiency.

The European space industry is developing the enabling technology to reach the necessary readiness level to be able to use multicore processors in real space missions. The main enabling technologies considered and under development include: multicore processors (Leon 4, etc.), Time and Space Partitioning approach of the integrated Modular Avionics for Space (started from single-core and inspired from the ARINC 653 standard), hypervisor technology (XtratuM, etc.) and SW architecture (SAVOIR-IMA). In particular the Time and Space partitioning of resources is considered to be an essential driver to ensure the predictability needed for critical missions.

3.7 Analysis of Embedded Software for Multicore in the Automotive Domain

Steffen Goerzig (Daimler AG – Böblingen, DE)

Multicore technology promises reduced energy consumption, reduced package dimensions, and higher performance. But the road to multicore is covered with hazards – race hazards. This is especially true for most of the software as it will be ported rather than re-implemented for multicore platforms. The talk presents current approaches to avoid race hazards in embedded software including technology transfer from academia to industry. First results of automotive applications are shown.

3.8 Multi core – Single bus

Rene Graf (Siemens AG – Nürnberg, DE)

The use of multi core processors in embedded systems is essential in future designs. Though, system architects have to think about the implications porting a former single core application to a multi core hardware, since most of the peripherals will still remain a single resource.

Even if the application can be easily splitted into independent parts, which work on different peripherals, the bus, e.g. PCI, between processor cores and peripherals will become a bottle neck. Assuming one part of the distributed application running with real time conditions, the latencies that were met on a single core processor may not be met any more due to the interfering bus accesses of the other parts.

The modeling and analysis of such a system in the early phase of development can help to find these implications both in a qualitative and quantitative manner.

The method to analyse those systems is described using a real system with a multi core processor and different peripherals, which are connected by a single PCI bus. Finally, the simulation results are compared with real measurement figures.

3.9 Distilling Programs for Multicore Architectures

Geoff Hamilton (Dublin City University, IE)

The proliferation of increasingly parallel architectures will have a significant impact on software developers; they can no longer develop software for a sequential architecture and expect performance to improve as the underlying architecture becomes faster. There is therefore a need to develop software that harnesses the power of parallel architectures directly. The development of parallel software is inherently more difficult than the development of sequential software; parallelization of programs by hand is very difficult, tedious and error-prone. By automatically introducing parallelism into programs, the programmer can be freed from explicitly implementing parallelism and can therefore concentrate on algorithmic issues. However, producing automatically parallelized code which is comparable in performance to code which has been parallelized by hand is very difficult, particularly for imperative programming languages.

There has been a recent upsurge of interest in the parallelization of functional programming languages. Functional programs are claimed to be better suited to parallelization than their sequential counterparts for a couple of reasons. Firstly, computations do not involve a shared state, which is problematic for parallel implementations; secondly, execution orders are solely constrained by data dependencies, as opposed to the unnecessary dependencies caused by statement sequences. Also, it is claimed that functional programs are easier to analyze and more amenable to transformation. However, functional programs also have the disadvantage that expressions are often combined using intermediate data structures, which would result in costly inter-process communication if these expressions were to be evaluated in parallel.

In this work, we show how our own program transformation algorithm which we call distillation can be used to transform programs into a form which makes functional programs

more amenable to parallelization and execution on multicore architectures. Distillation is a very powerful source-to-source program transformation algorithm for removing intermediate data structures which can achieve superlinear improvement in the run-time of sequential programs. Programs produced by transformation are in a specialised form called distilled form in which most functions are tail recursive and there are very few intermediate data structures. We show how distillation can be used to convert programs defined over sequential data structures to equivalent programs defined over data structures which are more easily partitioned to facilitate parallel execution. We then show how the resulting programs can be parallelised using Glasgow Parallel Haskell. We argue that this has the advantage over alternative techniques that fewer intermediate data structures are created in the resulting programs, so they can be executed more efficiently.

3.10 Necessity for & Feasibility of a Multicore Ecosystem

Andreas Herkersdorf (Technische Universität München, DE)

Multicore technology overcomes the bottleneck of sequential task execution and provides superior processing performance and power efficiency compared to sophisticated single-core ancestors. Multicore technology also lets industry and academia face entirely new challenges with respect to coping with system complexity. For the time being, the efficient utilization of vast amounts of parallel processing resources relies predominantly on the skills of expert programmers and system architects, but isn't yet accessible for the broad community of software engineers. In the field of embedded and cyber physical systems, multicore processors must satisfy tough requirements with respect to real-time, power efficiency, reliability, safety and security. Methods for multicore system modeling, verification and software debugging, if existing, are specific to an individual processors, but not generically applicable to classes of multicore systems (Would we need a Sync-Point as an enhancement to a Break-Point?).

Finding generic, flexible and scalable solutions to these problems in order to enable multicore on an even broader scale for embedded systems applications may be beyond the skills and capacities of individual enterprises. Therefore, the "Working Group Multicore" within the Bavarian ICT Innovation Cluster BICCNet proposed establishing a research and development network to jointly tackle these challenges. Through mutual exchange of knowledge and (partially) providing access to solutions in the specific domain of expertise of partners, a multicore ecosystem would gradually evolve. My objective for this Dagstuhl seminar is to stimulate discussions on the feasibility of such an ecosystem, to hear what reservations industry might have and how approaches for an initial instantiation could look like.

3.11 "Heterogeneous Multiprocessing or Just a Bunch of Coprocessors?" – The case for unified programmability

Enno Luebbers (Intel GmbH – Feldkirchen, DE)

Heterogeneous multiprocessor systems combine general-purpose processors with specialized (and thus highly efficient) processing units for the acceleration of application-dependent functionality. The increased overall efficiency, however, comes at the cost of programmability, as usually, the application developer needs to be an expert in the respective programming models for the individual acceleratos (FPGAs, GPUs, DSPs, ...) in order to exploit the heterogeneous elements to their full potential.

Many approaches exist to cover hetereogeneous elements with new or extended existing programming languages and models. In the face of upcoming challenges in embedded systems like openness, extensibility, safety and security requirements and ever-increasing complexity, the questions rises whether there is actually a silver bullet, at least for certain application domains, or if we should look at different programming models which at least allow the integration of heterogeneous parts, developed by domain experts, into heterogeneous applications that fulfill the promise in terms of efficiency that heterogeneous platforms have made.

3.12 Fault-Tolerant Time-Triggered Communication Infrastructure for Multi-Processor Systems-on-a-Chip

Roman Obermaisser (Universität Siegen, DE)

The ongoing technological advances in the semiconductor industry make MPSoCs more attractive, because uniprocessor solutions do not scale satisfactorily with increasing transistor counts. However, higher integration causes more sensitivity w.r.t. energy variations which requires new fault-tolerance measures to overcome the transient fault rates that have significantly increased. In the transient tolerant time-triggered system-on-chip architecture, fault-tolerance mechanisms for application components, communication interfaces and the time-triggered network-on-a-chip are introduced. In addition, a fault injection framework was developed to compare state-of-the-art integrated architectures (e.g., hypervisors such as XtratuM) and the transient tolerant time-triggered system-on-chip architecture. Experiment evaluations have provided evidence for the reliability of the architecture in the presence of soft-errors.

3.13 Multi-Core in Avionics – On Problems and One Technical Approach Monitoring-Based Shared Resource Separation for Commercial Multi-Core System-On-Chip

Michael Paulitsch (EADS – München, DE)

Multi-core computer architectures are the first choice in consumer electronics. Their performance and power efficiency are also attractive features for safety-critical applications, as in avionics. But increased integration and optimizations for average case performance poses challenges when deploying them for such domains. In the Dagstuhl presentation, we first present visions and problems of multi-core processors. In an exemplary approach towards solving a specific solution, we focus on the problems of temporal indeterminism and fault containment introduced by shared resources such as network on chip and shared memory. Pursuing previous work that quantified the impact of concurrent usage of shared resources, targeting the integration of mixed-criticality applications on the same platform, we propose a partitioning approach to control those interferences. We present a partitioning concept, which is further used to develop a modified worst-case analysis for multi-core processors. For evaluation we use representative benchmarks of the EEMBC Autobench benchmark suite on the Freescale 8-core PowerPC P4080.

3.14 Talk on ARTEMIS Project RECOMP

Michael Paulitsch (EADS – München, DE)

"RECOMP" stands for Reduced Certification Costs Using Trusted Multi-core Platforms and is a European funded project from ARTEMIS JOINT UNDERTAKING (JU). The project started April 1st of 2010 and has a duration of 36 months.

RECOMP research project pretend to form a joint European task force contributing to the European Standard Reference Technology Platform for enabling cost-efficient certification and re-certification of safety-critical systems and mixed-criticality systems, i.e. systems containing safety-critical and non-safety-critical components. The aim is establish methods, tools and platforms for enabling cost-efficient (re-)certification of safety-critical and mixed-criticality systems. Applications addressed are automotive, aerospace, industrial control systems, and lifts and transportation systems.

RECOMP recognizes the fact that the increasing processing power of embedded systems is mainly provided by increasing the number of processing cores. The increased numbers of cores is a design challenge in the safety-critical area, as there are no established approaches to achieve certification. At the same time there is an increased need for flexibility in the products in the safety-critical market. This need for flexibility puts new requirements on the customization and the upgradability of both the non- safety-critical and safety-critical parts. The difficulty with this is the large cost in both effort and money of the re-certification of the modified software-

3.15 Talk on German Project ARAMiS – Automotive, Railway, and Avionics Multi-Core Systems

Michael Paulitsch (EADS – München, DE)

ARAMiS (Automotive, Railway and Avionic Multicore Systems) aims at the development of concepts for multi-core processors in automotive, railway and avionics to reach a gain in safety, comfort and efficiency. In current aircrafts or cars only single-core processors are used since only their functionality can be certified according to domain specific safety-standards. But these singe-core architectures cannot reach the performance needed for future applications and are getting obsolete. To develop efficient multi-core architectures several research institutions and manufacturers from the automotive, railway, and avionics domain, their suppliers as well as hardware and software producers are working together in the ARAMiS project.

The project involves the following partners: AbsInt, Airbus, Audi, BMW, Bosch, Cassidian, Continental, Daimler, Diehl, EADS, Freescale, Infineon, Intel, Liebherr, OpenSynergy, Symta Vision, Vector, Wind River and various research institutions (Technische Universität München, Technische Universität Braunschweig, Karlsruher Institut für Technologie, Universität Stuttgart, Technische Universiät Kaiserslautern, Christian-Albrechts-Universität zu Kiel, Universität Paderborn, Fraunhofer IESE AISEC, Offis, Fortiss).

3.16 Fine grained process migration for MPSoCs

Sri Parameswaran (UNSW – Syndey, AU

Process migration (PM) is a method used in Multi-Processor System on Chips (MPSoCs) to improve reliability, reduce thermal hotspots and balance loads. However, existing PM approaches are limited by coarse granularity (i.e. can only switch at application or operating systems boundaries), and thus respond slowly. Such slow response does not allow for fine control over temperature, nor does it allow frequent migration which is necessary in certain systems.

In this work, we showcase Thor, an approach which is a fine-grained reliable PM scheme, for Embedded MPSoCs, to overcome the limitations of existing PM approaches. Our approach leverages custom instructions to integrate a base processor architecture, with PM functionality. We have proposed three schemes, Thor-BM (migration at basic block boundaries), Thor-BM/CR (migration at basic block boundaries with checkpoint and recovery), and ThorIM/CR (migration at instruction level with checkpoint and recovery). Our main motivation is to realize a fine-grained PM approach beneath the OS level within the local processor architecture. Performing locally, and without the use of the OS, results in short initiation Time. If such a scheme were to be implemented, then this would let Dynamic Thermal Management techniques (especially those with policies relying on frequent task migrations, improve overall performance while maintaining low peak temperatures. Such a scheme would allow for a fast process migration in the presence of faults. It is also possible for fast load balancing scenarios to take place. Our experiments show that the execution time overhead is less than 2%, while the additional area cost and power consumption costs are approximately 50% (excluding main memories, which if taken into account would substantially decrease this overhead). The average migration time cost is just 289 cycles.

3.17 Task Mapping for Manycore-based Embedded Real-Time Systems

Stefan M. Petters (ISEP-IPP – Porto, PT)

Joint work of Nikolic, Boirslav; Petters, Stefan M.;
Main reference B. Nikolic, S.M. Petters, "Application Mapping In NoC-Based Many-Cores," Technical Report, HURRAY-TR-121201, 2012.
URL http://www.cister.isep.ipp.pt/docs/733/

Manycores-based processors are clearly on the agenda for their eventual deployment in embedded systems. While the widespread usage of manycores is still a few years into the future, it is worth spending now some effort in considering implications and requirements for this deployment. A current research activity in the area of operating systems for such processors is the Barrelfish OS. Barrelfish operates on a limited migrative model, where a task can only migrate within a subset of cores. Within the talk, various challenges in the mapping process have been discussed, when it comes to the communication within the dispatcher entities of an application, as well as between applications. In particular the notion of proxies to simplify the analysis process of the communication delay has been introduced. Since an exhaustive search for feasible mappings is of exponential complexity, the talk presented a heuristic, which allowed via one parameter to control the mapping complexity and via another parameter, the greediness of high priority applications when it comes to chip area. A set of experiments showcase the complexity of the approach, as well as the impact of the two parameters on the mapping result.

3.18 Sustainable Development of Software in the Multi-Core Age

Matthias Pruksch (sepp.med – Röttenbach, DE)

In order to benefit from progress in hardware, software development faces a paradigm change to parallelism. The impact is even more important, since life-cycle of software is much longer than that of hardware: investments in software last for decades. In addition, software now plays a crucial role for system design. Model based methods for development and quality assurance show the prospect to master the increasing complexity of such systems. Notably, if you are working in a context of products that have to be certified.

3.19 Chances and risks for security in Multicore processors

Georg Sigl (Technische Universität München, DE)

The main security challenges in future systems in the application areas such as automotive, transport, industry automation, health care, smart grid are:

- Security for 10 and more (30) years.
- Secure autonomous interaction of heterogeneous machines (M2M).
- Protection against manipulation and misuse.

- Fulfilling security requirements while keeping real time requirements.
- Consider resource limitations.
- Managing increasing complexity in embedded systems.
- Protection of intellectual property (hardware and software) in embedded systems against counterfeiting.
- Support of adaptation of cyber physical systems through securely adaptable embedded systems.

In order to fulfill these requirements in the future, secure elements will be integrated in the systems in order to provide security services while still being resistant even against hardware attacks. Hardware attacks can be classified into probing attacks, which try to extract information out of a chip by probing internal signals or by forcing values. Side channel attacks observe power consumption or electromagnetic radiation and fault attacks inject faults into the chip, e.g. through spikes on the current supply or light. Examples for systems with secure elements are cars with secure elements as investigated in the German funded project SEIS, the smart meter gateway solution as specified by the German BSI (Bundesamt für Sicherheit in der Informationstechnik) or mobile phones with up to three secure elements, the SIM a secure element soldered in the phone for NFC payment and a SD-card with secure element provided by a bank.

In multicore systems we have alternative solutions for solving security problems which are nowadays solved with a secure element. In order to detect fault attacks a very proper means is redundancy, which can be implemented in multicore systems very well. Security critical tasks can be parallelized and the results can be checked afterwards for correctness. A very good countermeasure against side channel attacks is randomization, which increases the effort for the attacker to observe critical operations. In multicore systems the execution of tasks is randomized by default and can be even increased by actively assigning parts of tasks to different core s with changing degrees of parallelization. Another way to counteract side channel attacks is the implementation of secret sharing schemes which avoid the use of a complete secret key on one core but distribute it to many cores. Separation is another important security measure, which may be easier on multicores as long as the underlying architecture supports that. If there are too many shared resources this could however be also a security risk. With multicores it may be even possible to assign the role of a secure element flexibly to one or more of the standard cores, which the responsibility to monitor the behavior of the system and to provide security services.

Multicores enable creation of much more complex systems than today's processors. Increasing complexity usually increases the risk of vulnerabilities like denial of service, undetected malware, or buffer overflows. The reason for these vulnerabilities is resource sharing, lack of monitoring or control and badly separated software. Another risk is sid e channel attacks which may be executed on multicore systems through software which is executed on the same system on chip.

Overall multicore may offers opportunities to improve the security of embedded systems, where we have currently a lot of open and un solved problems.

3.20 Isolation of Cores to Support Development of Mixed Critical Systems

Claus Stellwag (Elektrobit Automotive – Erlangen, DE)

The major issue when using multi-core controller in embedded devices is the huge amount of legacy code developed in former times. This code is normally well suited ("proven in use") but not aware of execution parallelism and therefore cause problems when being executed on multi-core controller. Reasons for this behavior are: fast interrupt locks, implicit communication (e.g. activation order of threads) or cooperative scheduling (thus avoiding the use of explicit locks for shared resources) among others. Additionally new requirements have to be considered in the design (e.g. minimize energy consumption) and the use of standards is forced (AUTOSAR). If safety relevant software has to be executed on the same device with standard QM (or legacy) software the designer has to make sure that the safety part is not influenced from non-safety parts.

One common idea when migrating software from single to multi-core is to perform a redesign and split up threads to the different cores. Experience showed that this is possible but causes lot of work, e.g. all implicit communication or locking have to be considered and sometimes must be made explicit. Also the proof that the new partitioning of the redesign is free of errors is difficult - especially from the safety viewpoint. Therefore a different approach is presented here.

Central idea is to focuses on the separation and isolation of cores. This means that existing applications should not be split but kept together. Still multiple applications can run on one device, but each one mapped to its own core. Communication between the cores shall be minimized in order to get the maximum performance. Safety applications get with this mapping their own core and (if the hardware supports isolation) can be completely protected from non-safety software. One controller family which supports this is the Infineon AURIX. On this chip the cores and their RAM can be configured in way that only read but no write access from other cores is possible.

As an example just consider an AUTOSAR system where normally the safety and non-safety applications (SWCs) can be separated, but all the basic software is shared. This means that all basic software must be developed according the related safety standard. If isolation is used one core can use a standard AUTOSAR system with all the non-safety code. The other core can be used for safety related software only. If required a core-to-core communication module can offer an exchange mechanism between the cores. The safety core might also need some basic software, but typically the amount of such modules for safety is very low. In general it is always better to limit the amount of safety related software to handle the complexity. Within one EU funded project (RECOMP) Elektrobit implemented this isolation approach for the AURIX. The implementation was used by Delphi (automotive supplier) to demonstrate an electrical steering column lock. The demonstrator showed that the isolation method works.

3.21 Safe(r) Loop Computations on Multi-Cores

Jürgen Teich (Universität Erlangen-Nürnberg, DE)

Main reference J. Teich, W. Schröder-Preikschat, A. Herkersdorf, "Invasive Computing - Common Terms and Granularity of Invasion," arXiv:1304.6067v1 [cs.OS]
URL http://arxiv.org/abs/1304.6067v1

The necessity of satisfaction of non-functional constraints such as guaranteed data processing throughputs, deadline reactive processing or safety properties on the correctness of computational results is of utmost importance for the successful introduction of multi-core technology in many future embedded system products.

In this visionary introduction, we treat the problem of architectures, methods and tools that allow a developer to specify a certain safety level for a quite general and important class of loop computations. Loop programs are known to be quite amenable to parallel processing and are typically also quite scalable. However, no existing work is known to us how to make loop computations safe so to guarantee the correctness of the corresponding computed results of a loop program at run-time.

In this realm, we propose first ideas how, dependent on a specified safety level, the core allocation might be properly controlled for allowing concepts such as DMR and TMR known for single processor systems to loop computations on multi-cores including the way how deterministic voting may be efficiently implemented on a class of domain-specific multi-core architectures called tightly-coupled processor arrays (TCPAs).

We conclude how these concepts of redundant in-sync loop computations might be nicely supported by the recent parallel computing concept of invasive computing.

3.22 parMERASA- Multi-Core Execution of Parallelised Hard Real-Time Applications

Theo Ungerer (Universität Augsburg, DE)

URL http://www.parmerasa.eu/

Providing higher performance than state-of-the-art embedded processors can deliver today will increase safety, comfort, number and quality of services, while also lowering emissions as well as fuel demands for automotive, avionic and automation applications. Such a demand for increased computational performance is widespread among European key industries. Engineers who design hard real-time embedded systems in such embedded domains express a need for several times the performance available today while keeping safety as major criterion. A breakthrough in performance is expected by parallelising hard real-time applications and running them on an embedded multi-core processor, which enables combining the requirements for high-performance with time-predictable execution.

The talk will discuss preliminary results of the EC FP-7 project parMERASA (Multi-Core Execution of Parallelised Hard Real-Time Applications Supporting Analysability, started Oct. 1, 2011). The project targets timing analysable systems of parallel hard real-time applications running on a scalable and predictable multi-core processor with up to 64 cores. We target in particular future complex control algorithms by parallelising hard real-time

application programs to run on multi-/many-core processors. Application companies of avionics, automotive, and construction machinery domains cooperate with tool developers and multi-core architects to reach the project objectives.

3.23 OpTiMSoC – An Open Source Experimentation Platform for Multicore

Stefan Wallentowitz (Technische Universität München, DE)

Joint work of Stefan Wallentowitz, Philipp Wagner, Michael Tempelmeier, Thomas Wild, Andreas Herkersdorf
Main reference S. Wallentowitz, P. Wagner, M. Tempelmeier, T. Wild, A. Herkersdorf, "Open Tiled Manycore System-on-Chip," arXiv:1304.5081v1 [cs.AR].
URL http://arxiv.org/abs/1304.5081v1

Future System-on-Chip will integrate an increasing amount of processing elements. Tiled manycore System-on-Chip are a promising approach to organize future platforms. In such platforms a Network-on-Chip connects "tiles" of processing elements, memories and I/O, often as a mesh. For example Intel has presented the "Single Chip Cloud Computer".

Research of such platforms and especially prototyping rely on building such a platform or is bound to simulation. At LIS we develop a framework for prototyping of such tiled manycore System-on-Chip: Open Tiled Manycore System-on-Chip (OpTiMSoC). It is based on open source components and itself freely available. Essential elements and target platforms are part of the library and a platform generator tool is in development to allow for fast creation and implementation of platforms.

This talk gives an overview of OpTiMSoC and presents the current status and roadmap of the project.

3.24 Efficient observation of Multicore SoCs

Alexander Weiss (Accemic GmbH & Co. KG – Kiefersfelden, DE)

Comprehensive observability of multicore System-on-Chip (SoC) is the basis for efficient debugging, especially for the analysis of root causes of non-deterministic failures. Furthermore, it is also important for the detection of race conditions, the measurement of WCET, cache and memory layout optimization as well as different kinds of coverage measurements. Solutions for multicore SoC observation can be rated by the completeness of accessible information, which includes executed instructions, clock cycle accurateness, data access (value, address, direction), cache and bus operations. This information should be captured in parallel for multiple CPUs and other bus masters. Observation should be long-time, non-intrusive, available in real-time and applicable for mass-produced SoCs. By using multicore SoCs the traditional computation-centric observation strategy of single core SoCs has to be amended by communication-centric observation. Today's solutions are software instrumentation and embedded trace, both with limitations especially in the multicore area. The bottleneck is the required bandwidth for trace data output, which increases superlinear with the number of CPU cores. With hidICE a new observer based approach for full visibility of internal states

of multicore SoCs was developed. The fundamental idea behind hidICE is to equip the SoC with a facility that allows the synchronization with an external emulator. Thus, only data is communicated from the SoC to the emulator that is required to reconstruct all internal data and the program flow. All other system responses are defined by the program code. The advantage of this approach is the fact, that in most cases the bandwidth required for synchronization is significantly lower than the bandwidth required for traditional trace data output. The emulation provides full access to all internal CPU and bus states, including CPU register trace or bus trace with the deep view as known from a logic state analyzer. The hidICE approach was evaluated for different CPU architectures, including a LEON3 based multicore system. This implementation requires a very low gate count and provides full, real-time, continuous, and concurrent observation of all CPU cores. Another challenge in multicore SoC observation is an efficient strategy to handle the huge amount of trace data, accessible from multicore SoCs. The traditional offline analysis has to be complemented and replaced by online analysis approaches, such as runtime verification.

3.25 Many cores – many problems

Reinhard Wilhelm (Universität des Saarlandes, DE)

The embedded-systems industry goes multi-core. The main reason is the good performance per consumed-energy ratio. For time-critical systems, this transition is problematic. No sound and efficient method for the verification of timing constraints of embedded systems executed on multi-core platforms exists. This problem is difficult and can not solved by the established methods due to the interference on shared resources. There is a proposed method for timing analysis: 1. analyze an application by established single-core methods assuming that the access to shared resources happens instantaneously, then 2. add a safe bound on the delays of all accesses as indicated by a given abstraction of the access behavior of all co-running applications.

Current core designs do not allow this approach since they are not timing compositional. In addition, the currently used abstractions seem to be too coarse as to allow precise bounding of access delays. The MULCORS report, Use of Multicore Processors in airborne systems, submitted by Thales Avionics to the European Aviation Safety Agency is considered a document of capitulation in the face of the problem. It recommends to use measurement-based, unsound methods for timing verification of safety-critical and mixed-critical avionics systems implemented on COTS multi-core architectures. They ignore both the state of the art in single-core timing analysis as well as fundamental problems of measurement-based approaches. For example, they recommend using corrective factors to unsafe measurement-based WCET estimations based on an identification of worst-case perturbances, ignoring the high likelihood of domino effects on such architectures.

3.26 High-level Simulation-based Design Space Exploration on Multicore Virtual Platforms

Thomas Wild (Technische Universität München, DE)

For an efficient simulation based performance assessment, instruction set simulators (ISSs) cannot meet the simulation performance requirements to be applied in the exploration of the macroarchitecture of multicore systems-on-chip (MPSoCs). In order to speed up the exploration it is essential to use more abstract simulation models, which limit the number of simulation events and yet deliver accurate and indicative insights into the architecture to assist the designer in finding an optimal solution.

This talk presents McSim, a SystemC based high level simulation tool, which allows the use of trace driven as well as compiled binary level simulation models. In a trace model, internal structure and functionality of a processor core are abstracted in time interleaved reads and writes, which are replayed and superimposed on the shared resources of the MPSoC architecture. A binary simulation model includes in addition the execution of the functionality, allowing the simulation of flexible input stimuli for the processor. Experiments with various benchmark applications from MiBench and MediaBench show - compared to SimpleScalar ISS – an increase of the simulation performance of a factor of up to 200 and 30, respectively. The maximum error of the two models relative to SimpleScalar are 15% and 10%. Architecture exploration with McSim consists of two phases: A one time generation of either the trace or the binary level model (which includes running the ISS). During the actual exploration phase, in each iteration the trace and binary level models are executed thus profiting from the acceleration compared to the ISS. The adaptation of HW/SW architecture and of task mapping/scheduling for the solution alternative under investigation is done via XML files.

4 Working Groups

4.1 Specification & Interference

Claus Stellwag, Michael Deubzer, and Glenn Farrall

The workgroup focuses on the topics specification and interference but during the discussions also other topics were touched. The following topics (A, B, C) were considered to be most critical for a wide use of multi-core architectures in cyber physical systems.

A) Application Specification

It is of central importance that an application specification exists which details different aspects of the application design. Especially the document should:

1. include all requirements include non-functional ones (e.g. requirements on timing).
2. enable legacy code integration and shall explain how this is performed/which rules are applied.
3. help to avoid and safely bound interferences on shared resources.

Furthermore, as a base a well-defined software component model in needed. Basic characteristics shall include dependencies between components and interface to e.g. global variables which are accessed by the component. An example for such a model can be found in AUTOSAR. Depending on the design stage the information will be updated or extended. Example: During top down design the timing constraints are given and budgets are assigned to software components, which can be later verified when the executable of the application is available (e.g. by static analysis or tests on the final hardware). Standard component models talk about functional interactions. But the models shall also be able to include meta information e.g. by specifying memory interference, WCET, options to map the components to different cores, or energy behavior. This would greatly simplify the composition of applications in multicore architectures because it is assumed that the component developer and the integrator of the system are different persons. This way also parts of the design knowledge can be saved within the model and later reused e.g. to find the best mapping of components to cores.

Requirements shall contain all necessary properties like timing constraints, safety requirements, data loss constraints, and criticality information.

It would be very helpful to use specification guidelines which help the software architect to structure the specifications and to make sure that no important aspects gets lost. Such guidelines should

- be built upon a well-defined model of computation (interacting components)
- make all component dependencies explicit
- classify variables into configuration, start-up and volatile
- provide memory partitioning on an appropriate level of abstraction (local to process, shared between subsets of processes, global)

A compiler can map this to NUMA shared memory architectures. Depending on the real needs security aspects shall be listed. (E.g. a security architecture could add monitors)

B) Legacy Code

There is no easy way to migrate existing code to multicore controllers. Most code has implicit assumptions about the single core behavior, e.g. implicit communication between cooperative threads. Ideally depending on the original design the effort of the changes can be estimated. Unfortunately in many cases the know-how of this original design is lost (e.g. responsible people have left or retired) and therefore the fear of changes is huge. (“don’t touch it, it’s a piece of art”). The following four levels have been discussed (effort grows from low to high)

1. Small changes are often possible because their impact can be foreseen. (E.g. software unit internal behavior which is not visible from outside). Additional abstraction layers and interfaces to shared resources help to limit the degree of interference between software components.
2. Mapping of existing SW to cores. This approach maps parts of the existing software to separate cores and tries to minimize (or even omit) the communication between cores and avoid sharing resources with other cores. If resources are shared (e.g. I/O) it has to be handled in special way. Since the software runs only on a single core most of the old behavior can be established without modifications (e.g. implicit communication will work since all relevant communication processes are on the same core)
3. Redesign. Completely redesign (central) parts of the application based on a clear model of computation and interaction that exposes (at least) the required degree of concurrency.

4. New development. Some industries (e.g. signal processing, mobile communications) have to rewrite completely the application in a well-defined way, e.g. by model of computation that reveals concurrency and interaction (e.g. data flow models). In these domains, clearly defined design flows are available for multiprocessor platforms that optimize and guarantee timing, energy, and memory.

C) Virtualization

Virtualization could help to migrate software to multicore architectures if it would talk about real-time and not virtual time. In cyber physical systems programs typically have dependencies to real-time, this is different to IT systems (e.g. servers) where the timing within a virtualized software is not that important. Practically virtualization is quite established in PC like environments, but only very rarely used in embedded systems.

As an example: just consider an algorithm which performs an analog to digital conversation. This piece of software must know when it is time to convert the next analog value. If this time is missed there will be no result. Supporting real-time in virtualized environments is currently no considered in existing solutions and might also impact the used hardware (e.g. when peripherals are virtualized).

4.2 Industrial Perspective on MultiCore Motivations and Challenges

Glenn Farrall, Christian Ferdinand, Massimo Ferraguto, Steffen Görzig, Michael Paulitsch, Matthias Pruksch, Claus Stellwag, Sergey Tverdyshev, and Alexander Weiss

There were two questions asked of industrial participants to the "Multicore Enablement for Embedded and Cyber Physical Systems" seminar.

For each question there is a summary of the response, and then for completeness the full set of text is provided.

Q1: What is the major motivation for *using* multicore devices — both today and in the future? (where using can be "implement a product with" or "supply tools and/or services to support")

There were two main answers to the first question. This is not surprising since the motivations for multicore usage have high commonality and are aligned with the realities working against performance of single core devices increasing ad infinitum.

Availability Lack of availability of single core devices was one strong answer. There is a clear expectation that eventually there will only be multicore devices available (above a certain performance threshold) — this answer has been summarised as TINA (There Is No Alternative)!

Derived from product requirements The drive for new features, or step changes in capability demand the extra performance or the extra performance per unit (of power, weight, volume, etc.) that multicore devices offer. As should be obvious from the considered markets (those with Cyber Physical aspects) included in features required are safety and availability, and these also are attractively enhanced with multicore devices.

Collated feedback to question 1 with attribution (in alphabetical order):

Christian Ferdinand (AbsInt) Even for highly safety-critical applications, multi-core processors seem to promise better performance. Therefore, many of AbsInt's potential customers are exploring the possibilities. Typical COTS multi-core processors use shared memory and shared memory buses/networks. Such architectures introduce a high degree of resource sharing that would not be there in a distributed memory architecture. This additional resource sharing can lead to large interference effects between cores. This complexity has created a new demand for timing verification tool support not only for the highest criticality classes.

Massimo Ferraguto (Space Systems Finland) The use of multicore in the space domain can be beneficial in terms of greater processing capability (concentration of multiple functions in one single computer, with partitioning by criticality level and/or function; more payload data processing on-board), weight, power and fuel reduction which ultimately lead to longer lifetime and cost efficiency.

Steffen Görzig (Daimler) TINA (There Is No Alternative). When you can only buy multicore processors in the market, your only choice is to switch cores off...

Michael Paulitsch (EADS)

- Quest for more compute power due to new application demands.
- Wish to reduce size, weight and power of computing for improved performance of aircraft and improved environmental performance.
- Tighter integration leads to the need of more powerful central compute platforms
- wish to use COTS chips for reasons like possible lower cost (COTS are likely multicore devices or SoC)

Matthias Pruksch (sepp.med) Customers, who are the customers of sepp.med customers, demand ever increasing value of products in terms of quality, functionality and interoperability. To achieve this, many products gain from smart acquisition, control and handling of information. Savings in space, weight and power as well as added value by cyber physical systems are just two important drivers to name. Multicore devices offer the prospect to sustain the increasing demand for computing and communication performance. The fundamental switch to multicore architectures lead to a paradigm shift in software development and has a tremendous impact on the installed base of legacy software and how new software has to be written in order to be sustainable. Specifically, regulated domains like medical or avionics are challenged by current multicore implementations that neither fulfil stringent requirements for predictability and absence of interference, nor enable certified legacy software to be migrated without expensive reengineering and costly re-certification. sepp.med is highly interested to provide qualified services for their customers and partners from consulting, over development and quality assurance up to certification.

Claus Stellwag (Electrobit Automotive) The increasing complexity and the rise of new functions (Automotive: e.g. advanced driver assistance functions which build the base for autonomous cars in the future) requires a lot more performance. Typical embedded controllers with only some megahertz will not be able to solve this performance gap. On the other hand the embedded environment (no active cooling of controllers, EMV, etc.) forbids the usage of standard (PC like) processors. So the only way to deliver the required performance and keeping the price affordable are multicore controllers.

Sergey Tverdyshev (SYSGO) The competition on the market drives companies to innovate. This innovation includes developing new functionalities, higher utilisation of available resources, sinking costs. Probably these are the main three reasons driving industry to

adopt multicore microprocessors.
SYSGO is one of the leading embedded RTOS providers provides and is constantly researching and improving support for multi-core systems. The following are the some of the reasons behind this work:

- Customer demands for support multicore systems including system-level architectural support (e.g. AMP, SMP)
- OS support to maximise utilisation of HW resources
- Increasing RTOS/OS performance which is transparent to the user
- Hope to increase dependability for high-criticality systems with multicore

Alexander Weiss (Accemic)

- traditional computation-centric observation has to be amended by communication-centric observation
- multicore possible increases the amount of non-deterministic failures
- traditional storage and offline analysis of trace data gets more and more limited, online trace data decompression and computation (for run-time verification, WCET, race conditions, profiling, ...) seems to be the next step in tool evolution → new generation of tools are required

Q2: What would be a key enabler to making their usage easier, or more prolific or perhaps more profitable?

The second question had a much wider range of answers — as the challenges are not as neatly encompassed as the first question on motivations was by physics.

- The timing behaviour of multicores is consistently raised as a challenge which needs to be made easier to cope with. So mechanisms for interference reduction (or elimination as a goal) are definitely required for enhanced usage in Avionics and other safety critical domains.
- Documentation of existing behaviour (especially of interference or arbitration conditions) is also clearly in a poor state today and improvements in this area would help (or at least add confidence) to any safety case made on COTS based systems.
- Architectures that would allow software to port seamlessly from singlecore to multicore (of any number) are also on the wish list. This has a clear economic advantage — most systems are not created from scratch, but involve reuse of existing code. The investment in this existing code can be very significant and very few products will start from a clean sheet. So either an architecture or tooling to enable this migration would enhance deployment.

Collated feedback to question 2 with attribution (in alphabetical order):

Christian Ferdinand (AbsInt) Interferences complicate timing verification. The adoption of multi-core processors in highly safety-critical applications could be helped by providing support for processor configurations that reduce interferences and a clear documentation of the resource conflict resolution mechanisms

Massimo Ferraguto (Space Systems Finland) The main enabling technologies considered include: multicore processors (Leon 4, etc.), Time and Space Partitioning approach of the integrated Modular Avionics for Space (started from single-core and inspired from the ARINC 653 standard), hypervisor technology (XtratuM, etc.) and SW architecture (SAVOIR-IMA). In particular the Time and Space partitioning of resources is considered to be an essential driver to ensure the predictability needed for critical missions.

Steffen Görzig (Daimler) Methods and tools to migrate old software to multicore without adding errors.

Michael Paulitsch (EADS) Helping would be anything that addresses the limiting points below

- tight integration of function blocks on SoC with limited ability to control and monitor their behaviour
- complexity of SoC and fear of design faults and possible limitations of mitigation or getting detailed design info to argue correctness
- complexity of SoC and ability to holistically understand it
- gap between average and worst-case performance increases
- less ability to control at SoC system level (guaranteed switch off cores, controlled access to shared resources ...)
- increasing gap between COTS SoC design environment and avionics design environment

Making use easier:

- having WCET in mind (does not necessarily have to be the centre of focus)
- access to essential details of chip design

Matthias Pruksch (sepp.med) First objective is to get legacy programmes running without touching the code. Therefore, a mandatory key enabler is access to sustainable multicore devices that provide predictability without sacrificing performance: lockstep mode is no solution. This enables the consolidation of previously separate devices of mixed criticality into one multicore system. Second objective is to define requirements and design guidelines for sustainable software development, e.g. scalability by number of cores. This means, by the increasing number of cores, performance improvements can be realized. Integrated development environments and tool chains are needed to tackle the complexity in terms of an increasing number of software components and constraints and to support design space exploration. Model driven development (correctness by design, formal methods ...) and model based testing show the prospect to handle those aspects. In our opinion, solutions will bring us to the next level of system creation towards hardware/software co-design, with the benefit of faster time-to-market, improved reuse and conservation of resources.

Claus Stellwag (Electrobit Automotive) Key enabler will be the handling of the new complexity (real parallelism). Devices with a clear partitioning approach can help as well as good tools. Some areas need to be developed further (e.g. shared resources and the access times to it).

Sergey Tverdyshev (SYSGO) However, there are obstacles which hinder wide adoption of multi-core in safety critical domains:

- The state-of-the-art COTS multi-core design is driven by cost reduction and average performance optimisation. This lead to the gap between worst and best performance is increasing and predictability of the system behaviour is decreasing. Lesser predictability mitigates advantages of MC in medium to high critical systems.
- The state-of-the-art COTS multi-core architecture focus on “simply adding” new cores. This increases communication load on interconnects and peripheral devices making them the truly bottlenecks for safety and sometimes security.
- The lack of in-depth documentation on COTS hardware makes it impossible to mitigate HW deficits in software on OS, middleware, RTE, or application levels.
- The lack of acceptance by certification agencies

The current situation in *safety* critical area is similar to a round-dance around a huge-camp fire where dancers see or believe to see something very precious in the middle and

while dancing trying to figure how to get it out with being burned.
Interestingly the security critical domains are not that hard affected by safety issues. In theses domains MC is widely used (at least in prototypes), especially in the systems without physical access to the hardware for attackers. The most of the problems are similar to single core systems, e.g. lack endurance/evidences that the produced COTS hardware is indeed the one which is described in documentation and does not contain malicious changes.

Alexander Weiss (Accemic) All the observation issues are very important. Not only the common instruction and data traces, also

- cycle accurate trace options for all devices
- trace of all bus masters (not only CPUs)
- easy differentiation on data trace between read and write access without the need for computing the instruction trace in parallel (ARM!)
- access to all information to observe scenarios as listed in Michael's paper [1, table 1]
- high bandwidth trace ports (in combination with smart port replacement approaches)

References

1 O. Kotaba, M. Paulitsch, J. Nowotsch, S. Petters, H. Theiling, *Multicore In Real-Time Systems – Temporal Isolation Challenges Due To Shared Resource* Workshop on Industry-Driven Approaches for Cost-effective Certification of Safety-Critical, Mixed-Criticality Systems (WICERT); part of DATE 2013. Grenoble, France. 2013

4.3 Certification of Safety-Critical Multicore Systems: Challenges and Solutions

Stefan M. Petters and Rene Graf

Certification in general is the process of a certification agent attesting to certain properties. Depending on the domain, the procedures vary substantially. In most cases certification applies to a complete system, which includes hardware, software and tools to build and validate the artefact. However, for example in the automotive domain, the concept of "certification-out-of-context" implies that some components may be certified in isolation. In general the current landscape in the area of systems certification within the automotive domain, industry automation and aerospace domains is dominated by a very conservative mind-set. Anecdotal evidence suggests that at the time of writing at least one dual-core system is in the certification process and the positive outcome of the process is not yet assured. This is driven by a number of reasons. Most notably there is, for obvious reasons, non-existent to negligible experience with the certification of multicore systems. This is exacerbated by the lack of answers to some of the technological challenges and consequently the lack of a reasonable safety argument. In particular the technological challenges with the deployment of multicores often presents substantial hurdles for the migration of legacy systems on Commercial-of-the-Shelf (COTS) based multicore platforms. The latter is driven primarily by the changed paradigm of performance gains which are now achieved via parallelism. This parallelism is in most cases not exposed in legacy systems. Even legacy applications using multiple threads are not verified running on a multicore processor.

One of the challenges identified within the working group is the increase of complexity in multicore systems. On the one hand this stems from the concurrency, as well as the fine grained resource sharing present in multicore, on the other hand the integration of additional components into the die which were previously accessible individually. Both concerns mean the traditional divide and conquer approach of looking at individual components and reasoning about the interaction between components in the analysis and safety argument are no longer possible. In order to mitigate the issue, it is of paramount importance to identify all resource sharing whether it's explicitly or implicitly shared. Once this identification is complete, platforms need to provide isolation as far as reasonably possible. The isolation properties can be achieved by either separation of duplication of the respective shared resource. For example, for caches, it would be of large benefit to avoid shared caches but rather have the possibility to partition this or use it as scratchpad, which is at least in some COTS processors available. However, shared caches or scratchpad memories may be needed in real SMP applications, where communication and synchronisation of threads need to be performed across several cores. In general it is also preferable to have the ability to control parameters and access to the shared resources in a fine grained manner.

For those resources, where strict isolation is not possible, like for example the memory interface, it is important to assure a fair and/or deterministic arbitration scheme. This would enable to reason about the maximal interference suffered by the core under investigation that is caused by applications running on the other cores. To validate the interference the platform needs to expose enough observability. With current technology, the integration of non-intrusive debug interfaces on individual cores and shared resources presents the most promising path to achieve the required observability.

While there are a number of academic approaches available to provide predictable and analysable HW platforms, their deployment in real-certified systems is very limited. This is driven by the differences at both ends of the spectrum, where current certification practice is focused on single core platforms while industrial pre-development is eyeing the computing capabilities of modern COTS based multicore systems, to implement a plethora of new system functionalities. The COTS issues can be mitigated to some degree by a working close communication with the HW vendors. However, such a communication base is not achieved on short notice but needs to be developed over years. The alternative is an individual development of a platform, however, this is neither a likely path to a high performant computer system, nor appears to be a cost effective solution.

Furthermore, the presence of dynamic and potentially non-controllable features in many COTS present further issues. Examples for such features are the NMI interrupt in Intel based systems or commonly used clock throttling for temperature and/or power management. Similar to the resource sharing, the identification of such features requires great care to be taken. Secondly, if such mechanisms may not be switched off in a guaranteed manner or provide no way to reason about their behaviour and impact it will make the deployment of such an architecture likely impossible in a certified environment.

Besides the HW architecture features discussed above, the software stack running on top of a multicore platform has the potential to mitigate some of the issues occurring on the HW platform. This reaches from the application driven pipelining of image processing segments to concentrate memory accesses at the start and end of the execution and thus avoiding interference on the shared read and write paths. On the operating system level tasks may be pinned to specific cores and thus avoiding the indeterminism of migrating applications.

One fundamental asset in the certification process of new HW/SW platforms would be to obtain a set of guidelines identifying "best practice" in the area. However, such guidelines

would naturally have to be developed by the certification agencies and are usually a product of experience. Consequently the process faces a similar paradigm shift, as the introduction of the first processors in certified systems delivered, with little clear idea on how to approach the problem. The first system mentioned in the introduction is certainly a step in this direction, even in this case though the process cannot be successfully completed, as it means first testing the waters and then refining with future attempts.

4.4 Network-on-Chip – Dependability and Security Aspects

Roman Obermaisser, Christian El Salloum, Theo Ungerer, and Thomas Wild

A major requirements for NoCs in embedded systems in *predictability*. Techniques for predictability range from static scheduling (e.g., time-triggered) to dynamic scheduling (e.g., priority-based). Also, NoCs provide solutions for monitoring and enforcement of resource budgets (e.g., AEthereal). Predictable application behavior for a given NoC also requires suitable modeling and timing analysis techniques.

The second challenge for NoCs is *composability*. Composability refers to a framework that supports the integration and reuse of independently developed components in order to increase the level of abstraction in the design process. Prior services of components must not be invalidated by integration, which is facilitated by temporal and spatial isolation based on precise interface specifications. In addition, the goal of composability is to avoid unintended emerging side effects at the system-level. Of particular importance in NoCs is deadlock freedom. Deadlock freedom can be ensured by isolation (i.e., control of dependencies), suitable routing strategies without deadlocks and formal analysis methods for routing cycles. Furthermore, resource guarantees such as bandwidth, jitter and latency must be maintained upon component integration. A key mechanism are specifications with explicit resource and memory requirements.

A third challenge for NoCs is *fault-tolerance* and *robustness* to support the reliable operation in the presence of faults. NoCs need to support the provision of an acceptable level of service on an MPSoC despite the occurrence of transient and permanent hardware faults of resources. For permanent hardware faults, important techniques are active redundancy or migration and reconfiguration of services exploiting spare resources. Transient hardware faults require the recovery in predictable time with state recovery. In mixed-criticality systems, containment of design faults is the primary concern, which is supported by NoCs with strong temporal and spatial isolation. The error detection mechanisms for operational faults and design faults can be based on a priori knowledge, information redundancy, analytical redundancy or replication. Recent advances in fault-tolerance of NoCs focus on proactive fault-tolerance. For example, wear-out specific scheduling takes into account temperature variations or increasing fault-rates.

Security is rapidly gaining significance in the field of embedded systems. In particular in safety-critical systems, security has to be considered as a safety aspect in scenarios where a malicious attack can lead to unspecified system behavior with catastrophic consequences (e.g., sabotage or terroristic attacks). The NoC in a multi-core processor provides the perfect opportunity to implement security mechanisms directly in hardware in order to enforce specific inter-core security policies. Considering the role that a NoC has, namely establishing communication among the individual cores and other entities like on-chip device controllers, a security-enabled NoC should establish the following properties:

- Authenticity of the sender: A receiving core on the NoC should be able to reliably determine the core from which a message was sent. It should be not possible for any core to forge the sender address without being detected.
- Message integrity: It should be not possible for any core to modify, delete or duplicate the messages sent by any other core. In a real-time system, the integrity of a message does not only depend on the message content, but also on the timing of the message. Therefore the integrity requirement has to be extended, such that it also should be not possible for any core to change the timing of messages that where sent by any other core.
- Message confidentiality: Only the intended receivers of a message should be allowed to read the message contents.
- Availability of guaranteed communication resources: In a (hard) real-time system, guaranteed communication resources have to be given to the individual cores, in order to assure that end-to-end deadlines are always met. From a security perspective it must be prevented, that the behavior of a malicious core (e.g., due to a compromised program running on a core) can lead to a violation of such guarantees (i.e. Denial-of-Service DoS attacks).
- Prevention of side-channel attacks: For some applications, it is required that there are no other possible means of communication than the explicitly defined communication channels. Such other means of communication are called side channels. An example of a side channel can be found in a NoC where the temporal properties of a given communication channel depend (even slightly) on the communication activities on another communication channel. Imagine two malicious nodes located on two explicitly defined distinct communication channels. Since the communication channels are defined as distinct, it should be not possible to leak confidential information from one channel to the other. The problem in a NoC that is not free from temporal interference, is that the two malicious cores can use that interference to illegally exchange information in a Morse-Code like manner. One core can induce a specific temporal interference pattern (by sending a pattern of channel-local messages) which can be observed and interpreted by the core on the other channel.

An example of a NoC satisfying all the above stated requirements is the time-triggered NoC (TTNoC) in the ACROSS architecture. In ACROSS the cores have no direct access to the NoC, but only via the Trusted Interface Sub System (TISS) which acts as a guardian. The TISS stores a statically defined time-triggered message schedule, which holds for each message the sent instance as well as the route and the set of receivers. Thus, the time-triggered schedule holds the entire topology which defines to which receivers a given message will be forwarded. All other cores will never see the message. The statically defined topology ensures message integrity and authenticity as well as confidentiality. Furthermore, the TISS ensures that messages are only sent according to the time-triggered schedule such that there is absolutely no temporal interference among different messages. Thereby the TTNoC establishes availability of the communication resources and prevents hidden side-channel attacks.

Adaptiveness is a challenge for NoCs to support system evolution, context adaptation and resource variation. In long-lived systems, the integration of new components, services and resources is needed to cope with changed application requirements. Technique for adaptiveness include predictable, fault-tolerant and secure configuration of the NoC. A prerequisite are standardized interfaces supporting configuration. Recent techniques for self-optimization in NoCs are a promising approach to autonomously and continuously adapted to the application behavior and resource availability. The extension of these feedback techniques for safety-critical embedded systems is a future research challenge.

4.5 Multicore Ecosystem

Andreas Herkersdorf, Johan Lilius, Massimo Ferraguto, Christian Thiel, Stefan Wallentowitz, and Thomas Wild

Multicore as an ICT Key Technology

Multicore processors are a key technology for coping with the important challenges our society will face in the upcoming decades. Secure and sustainable mobility, comprehensive healthcare, universal power management and the development of a digital society pose great demands on a distributed and powerful information and communication technology (ICT). These demands on embedded and cyber physical systems can only be met with multicore processors. All leading processor vendors – Intel, IBM, ARM, Nvidia, Freescale, Infineon, MIPS, TI – pursue a multicore architecture strategy. Such multicore processors are superior to their single-core ancestors with respect to processing performance and power efficiency, as they can execute different tasks concurrently on less complicated but parallel processor cores. On the other hand, industry and academia are facing entirely new challenges with respect to system complexity. The efficient utilization of parallel processing resources currently relies predominantly on the individual skills of the programmers. In the field of embedded and cyber physical systems, multicore processors must adhere to much stronger demands of real-time, power efficiency, reliability, safety and security when compared to standard desktop machines. Furthermore, multicore-enabled test and debugging tools are often missing along with universal methods for the modeling, design and validation of system issues. Various industrial and academic institutions in Europe have identified the relevance of multicore as a key ICT technology from the very beginning, and have established a competitive knowledge base for multicore technologies. However, finding flexible and scalable solutions for non-functional requirements, performance and power efficiency in increasingly demanding embedded system applications will soon be beyond the capabilities of large-scale enterprises or even networks of companies.

Roles and Benefits of a Multicore Ecosystem

The "Working Group Multicore" within the Bavarian ICT Innovation Cluster BICCNet proposes the establishment of a research and development network to jointly tackle these challenges [1]. In particular, topics such as parallelization support for non-functional requirements, migration of existing software and the development of sophisticated tools for debugging, testing and validation of multicore systems need to be addressed jointly. By achieving their individual goals, partners in this research network will also contribute to the above-mentioned topics. The results from publicly funded projects can be designed with compatibly in mind by using standard interfaces, and are available to all partners, allowing a growing *multicore ecosystem* to develop. Along with software and hardware components, this ecosystem will also contain models, methods and tools for multicore solutions; to the advantage of all contributing partners.

A multicore ecosystem can have a number of beneficial aspects. It can be an innovation ecosystem, where the idea is to encourage the interaction among of the actors to create new innovations, or its goal can be to create new business. In the first the main goal is the creation of new ideas, while in the latter the goal is to create new economic value. In

addition, the mere information exchange among different players and recognizing, who can bring what asset to the table and looks for filling what gaps in the own portfolio, may bring together new partners and represents a value by itself.

An ecosystem is often recognised post-facto, when one realises that there are strong activities around an issue. In the area of multi-core, the classical example of a business ecosystem is the ARM ecosystem, that has grown around ARM processors. Other examples of ecosystems are e.g. the activities around the AUTOSAR standard (which maybe is not as clearly identified as an ecosystem yet), or the activities around the eclipse tools (for which there is not necessarily big economic gain for the participants). Characteristic for these ecosystems is that there is a central entity around which the actors of the ecosystems place themselves to achieve added value.

In order to get started, the basic set of entities for a multicore ecosystem could center around could be a set of tools or hardware and software intellectual property building blocks (such as, e.g., elements of the AUTOSAR stack) that are either difficult to obtain (portability, licensing), very expensive to buy, and would be too complex to build oneself. Identifying such a set of tools and building blocks, and providing them for use to the community could be an interesting foundation to start ecosystem building activities. Providing this set of tools as open-source is crucial, since as noted by Riehle [3] this provides an avenue not only for users of the software, but also for system integrators and other actors to increase profit. Finally participation in the further development of the tools also becomes crucial for companies, as this will allow them to participate in the decision processes and influence the tool evolution.

As a proposal the OpTiMSoC [2] tools could form a starting point.

Many-core Monday

An ecosystem needs a platform for interaction and for attracting new participants. One such platform is the regular BICCnet AK Multicore meetings in Munich. Another interesting concept is *Mobile Monday*[1]. This is an open community platform of mobile industry visionaries, developers and influential individuals fostering brand neutral cooperation and cross-border P2P business opportunities through live networking events to demo products, share ideas and discuss trends from both local and global markets. Mobile Monday started as an informal gathering in Helsinki, with the aim of bringing together persons in the mobile industry. Initially it was just a group of people inviting friends and colleagues to an informal drink in a bar on Mondays. Often there were one or two presentations about something interesting, but the emphasis was on informal discussion and networking. The movement has grown and is organised into chapters that have organised events in over 140 cities worldwide.

Open Innovation

Open Innovation is an idea promoted by Henry Chesbrough [4], where companies use both internal and external ideas to create new products. An open innovation ecosystem consists of a group of actors that share both risk and rewards, creating growth for everybody. Central in this idea is that it is possible to build on top of other ideas. In the multicore area, open innovation could help create larger toolflows if tool vendors would make their tools interoperable, and would build new tools based on these toolflows. Open innovation is promoted by many large companies, and e.g. Nokia has been working successfully with a number of Universities (EPFL Lausanne, Berkley, Aalto), by forming “lablets” small research

[1] http://www.mobilemonday.net/ – “Mobile Monday”

groups at the University campus. This makes it easy for the industrial and the academic researchers to interact.

References

1 A. Herkersdorf et al., *Relevanz eines Multicore-Ökosystems für künftige Embedded Systems*, BICC-net, 2011, http://www.bicc-net.de/nachrichten/artikel/multicore-oekosystem/.
2 S. Wallentowitz, P. Wagner, M. Tempelmeier, T. Wild, A. Herkersdorf, *Open Tiled Manycore System-on-Chip*, arXiv:1304.5081, http://arxiv.org/abs/1304.5081
3 D. Riehle, *The economic motivation of open source software: Stakeholder perspectives*, Computer, vol. 40, no. 4, pp. 25–32, 2007.
4 H. W. Chesbrough, *Open Innovation*, Harvard Business Press, 2006.

4.6 Secure Elements in future embedded multicore systems

Georg Sigl, Sri Paramareswaran, Michael Paulitsch, Stefan M. Petters, Matthias Pruksch, Sergey Tverdyshev, and Stefan Wallentowitz

There is currently a trend that a specifically designed hardware attack resistant very well isolated secure element is integrated in systems on chip (SoC). The purpose of this secure element is:

- Integrity check of software that is executed on the system during boot
- Remote Attestation to confirm to a remote party that integrity of system is ok
- Provide identity and authentication for communication with other parties (PKI)
- Access control to resources and configuration registers of the SoC
- Key storage and secure memory

One example where these features are used is the boot process in a system. Normally in a microprocessor system using secure boot technology as specified by the Trusted Computing Group, the boot code calculates a hash value of its own executable, sends this to the secure element where the value is compared to an expected value. Then the secure element would recognize if the boot code has been changed. After that the boot core calculates the hash value of configuration code, e.g. BIOS, and afterwards of the operating system kernel and sends these values to the secure element as well for comparison. After completion of this process the secure element can confirm that the system has been started with trusted software. In a multicore embedded system this start up procedure is executed by one core (core 0) which then starts the boot process of other cores.

This security solution is based on the assumption that no hardware attacks are performed on the system on chip and a software security layer, e.g. a hypervisor, provides sufficient security and separation services for the applications running on the multicore system. For a better separation of cores a new architecture with an encryption and decryption (D/E) unit at each core could be helpful as shown in Figure 1. All data leaving a core would be encrypted and could be put into common memories without the chance to access it from other cores unless the correct key is known.

If we assume hardware attacks on a multicore system we currently have no hardware means in the multicore to detect them. The only solution in a system with a secure element is to move all security critical tasks into the secure element and establish an end-to-end

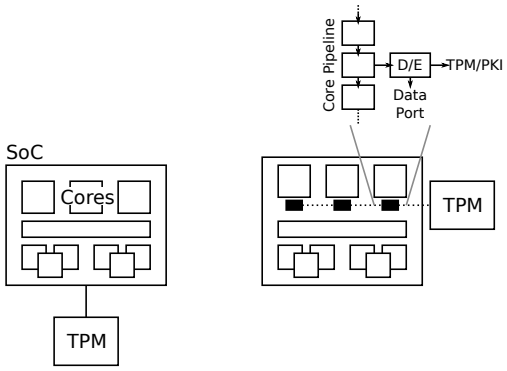


Figure 1 An encryption and decryption unit at each core can help with better separation of cores.

protected communication channel between the secure element and the remote application requesting the security service. If we have security services, which need a high bandwidth such as car2car communication with hundreds of signatures to be computed within seconds, there may be a need to integrate many secure elements as well. Otherwise there may be a bandwidth problem in an architecture as shown above.

There are both chances and risks in multicore systems concerning security (see presentation of Georg Sigl). One example where we see even both advantages and disadvantages in multicore systems is the chance to implement monitoring services in multicores. One core could be used to monitor the behavior of others in order to detect misbehavior created by malware running. On the other side the monitoring could be used to perform side channel attacks with a much better measurement accuracy compared to an external measurement of the cache behavior, e.g..

A very good countermeasure against many attacks, such as side channel attacks, is randomization. Multicores give plenty of opportunity for randomization, which is exactly the most severe concern of safety-critical system design engineers. Designers and certification authorities insist in deterministic behavior for these systems in order to determine, e.g., worst case execution times and to guarantee certain timings. As a solution to resolve this conflict for secure safety-critical systems, it would be very interesting to investigate implementations, which accept random behavior and still guarantee a timely execution with high probability. The project Proartis[2] goes into this direction. Synergies between this approach and the needs and solutions developed in the security domain could be a very interesting research direction and may be also a topic for a future Dagstuhl seminar.

[2] http://www.proartis-project.eu/

4.7 Inter-seminar workgroup: Software Certification & Multicore Processing

Michael Paulitsch

Multicore processing for safety-critical and security-relevant and safe deployment strongly depends on the ability to certify software running on multicore processors in the system context. The workshop "Multicore Enablement for Embedded and Cyber Physical Systems" has been incidentally running in parallel to the workshop "Software Certification: Methods and Tools (Seminar 13051)". Both groups realized the link between the two topics and organized an open common exchange and discussion session that addressed both topics in some detail. The common session increased the understanding of each other's workshop topic and made participants realize the complexity of certification involving multicore processors. An exemplary common observation of both seminar participant groups was the ever increasing gap and wishes of simplicity of processing in certified safety-critical environments for deterministic execution of critical software and the increasing modern multicore processor complexity.

Participants

- Michael Deubzer
Timing Architects Embedded Systems GmbH, DE
- Christian El Salloum
TU Wien, AT
- Rolf Ernst
TU Braunschweig, DE
- Glenn Farrall
Infineon – Bristol, GB
- Christian Ferdinand
AbsInt – Saarbrücken, DE
- Massimo Ferraguto
Space Syst. Finland Ltd – Espoo, FI
- Steffen Görzig
Daimler AG – Böblingen, DE
- René Graf
Siemens AG – Nürnberg, DE
- David Gregg
Trinity College Dublin, IE
- Geoff Hamilton
Dublin City University, IE
- Andreas Herkersdorf
TU München, DE
- Johan Lilius
Abo Akademi University, FI
- Enno Lübbers
Intel GmbH – Feldkirchen, DE
- Roman Obermaisser
Univ. Siegen – Feldkirchen, DE
- Sri Parameswaran
UNSW – Sydney, AU
- Michael Paulitsch
EADS – München, DE
- Stefan M. Petters
ISEP-IPP – Porto, PT
- Matthias Pruksch
sepp.med – Röttenbach, DE
- Georg Sigl
TU München, DE
- Claus Stellwag
Elektrobit Automotive – Erlangen, DE
- Jürgen Teich
Univ. Erlangen-Nürnberg, DE
- Christian Thiel
BICCnet – München, DE
- Lothar Thiele
ETH Zürich, CH
- Sergey Tverdyshev
Sysgo AG – Mainz, DE
- Theo Ungerer
Universität Augsburg, DE
- Stefan Wallentowitz
TU München, DE
- Alexander Weiss
Accemic GmbH & Co. KG – Kiefersfelden, DE
- Thomas Wild
TU München, DE
- Reinhard Wilhelm
Universität des Saarlandes, DE