Franz Baader (Ed.)

LNCS 4533

Term Rewriting and Applications

18th International Conference, RTA 2007
Paris, France, June 2007
Proceedings

Springer

Franz Baader (Ed.)

Term Rewriting and Applications

18th International Conference, RTA 2007
Paris, France, June 26-28, 2007
Proceedings

Springer

Volume Editor

Franz Baader
TU Dresden
Theoretical Computer Science
01062 Dresden, Germany
E-mail: baader@tcs.inf.tu-dresden.de

Library of Congress Control Number: 2007929743

CR Subject Classification (1998): F.4, F.3.2, D.3, I.2.2-3, I.1

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

ISSN 0302-9743
ISBN-10 3-540-73447-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73447-5 Springer Berlin Heidelberg New York

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12086580 06/3180 5 4 3 2 1 0

Lecture Notes in Computer Science 4533

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

Preface

This volume contains the papers presented at the 18th International Conference on Rewriting Techniques and Applications (RTA 2007), which was held during June 26–28, 2007, on the campus of the Conservatoire National des Arts et Métiers (CNAM) in Paris, France.

RTA is the major forum for the presentation of research on all aspects of rewriting. Previous RTA conferences were held in Dijon (1985), Bordeaux (1987), Chapel Hill (1989), Como (1991), Montreal (1993), Kaiserslautern (1995), Rutgers (1996), Sitges (1997), Tsukuba (1998), Trento (1999), Norwich (2000), Utrecht (2001), Copenhagen (2002), Valencia (2003), Aachen (2004), Nara (2005), and Seattle (2006).

For RTA 2007, 24 regular papers and 3 system descriptions were accepted for publication out of 69 submissions. Each submission was reviewed by at least three expert reviewers, and an electronic Program Committee (PC) meeting was held on the Internet, using Andrei Voronkov's *EasyChair* system. The reviews were written by the 14 PC members and 131 additional reviewers, who are listed in these proceedings. I would like to thank the PC members and the additional reviewers for doing such a great job in writing high-quality reviews in time and participating in the electronic PC discussion.

The RTA programme also included three invited talks, by Xavier Leroy (Formal Verification of an Optimizing Compiler), Robert Nieuwenhuis (Challenges in Satisfiability Modulo Theories), and Frank Pfenning (On a Logical Foundation for Explicit Substitutions). The talk by Frank Pfenning was a joint invited talk of RTA and the collocated Eighth International Conference on Typed Lambda Calculi and Applications (TLCA 2007).

The RTA PC decided to award a prize of 1,000 euro for the best paper to the article "On Linear Combinations of λ-Terms" by Lionel Vaux. Moreover, several travel grants could be given to students.

RTA 2007 was held as part of the Federated Conference on Rewriting, Deduction, and Programming (RDP), together with the following events:

- The Eighth International Conference on Typed Lambda Calculi and Applications (TLCA 2007)
- The colloquium From Type Theory to Morphologic Complexity in honor of Giuseppe Longo
- The workshop on Higher Order Rewriting (HOR)
- The workshop on Proof Assistants and Types in Education (PATE)
- The workshop on Rule-Based Programming (RULE)
- The workshop on Security and Rewriting Techniques (SecReT)
- The workshop on Unification (UNIF)

- The workshop on Functional and (Constraint) Logic Programmming (WFLP)
- The workshop on Reduction Strategies in Rewriting and Programming (WRS)
- The workshop on Termination (WST)

Many people helped to make RTA 2007 a success. In particular, I would like to thank the Conference Chairs, Ralf Treinen and Xavier Urbain, as well as the rest of the local organization team, and the sponsors of RDP 2007:

- The Conservatoire des Arts et Métiers (CNAM)
- The Centre National de la Recherche Scientifique (CNRS)
- The École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise (ENSIEE)
- The GDR Informatique Mathematique
- The Institut National de Recherche en Informatique et Automatique (INRIA) unit Futurs
- The Région Île de France

Barbara Morawska and the *EasyChair* system helped to produce the camera-ready copy of these proceedings.

April 2007 Franz Baader

Conference Organization

Program Chair

Franz Baader	Dresden

Conference Chairs

Ralf Treinen	Cachan
Xavier Urbain	Paris

Publicity Chair

Ashish Tiwari	Menlo Park

Program Committee

Alessandro Armando	Genova
Roberto Di Cosmo	Paris
Jürgen Giesl	Aachen
Deepak Kapur	Albuquerque
Hélène Kirchner	Nancy
Barbara König	Duisburg
Salvador Lucas	Valencia
Narciso Martí-Oliet	Madrid
Tobias Nipkow	Munich
Femke van Raamsdonk	Amsterdam
Aaron Stump	St. Louis
Sophie Tison	Lille
Ralf Treinen	Cachan

Local Organization

Antonio Bucciarelli	Paris
Vincent Padovani	Paris
Ralf Treinen	Cachan
Xavier Urbain	Paris

RTA Steering Committee

Jürgen Giesl	Aachen
Delia Kesner	Paris

Vincent van Oostrom	Utrecht
Ashish Tiwari	Menlo Park
Maribel Fernández	London
Bernhard Gramlich	Vienna

External Reviewers

Andreas Abel
Beatriz Alarcón
María Alpuente
Roberto M. Amadio
Oana Andrei
Sergio Antoy
Takahito Aoto
Zena Ariola
Pablo Arrighi
Emilie Balland
Alexandru Berlea
Frederic Blanqui
Stefan Blom
Eduardo Bonelli
Iovka Boneva
Guillaume Bonfante
H.J. Sander Bruggink
Roberto Bruni
Antonio Bucciarelli
Wilfried Buchholz
Guillaume Burel
Sergiu Bursuc
Anne-Cécile Caron
Yannick Chevalier
Adam Chlipala
Manuel Clavel
Dario Colazzo
Evelyne Contejean
Andrea Corradini
Rene David
Philippe Devienne
Kevin Donnelly
Gilles Dowek
Francisco Durán
Rachid Echahed
Santiago Escobar
Jerôme Euzenat
Stephan Falke
Emmanuel Filiot
Thomas Genet
Alfons Geser
Jean Goubault-Larrecq
Bernhard Gramlich
Yves Guiraud
Raúl Gutiérrez
Peter Habermehl
Florian Haftmann
Michael Hanus
Tobias Heindel
Joe Hendrix
Miki Hermann
Thomas Hildebrandt
Clement Houtmann
Samuel Hym
Florent Jacquemard
Deepak Kapur
Benny George Kenkireth
Delia Kesner
Jeroen Ketema
Zurab Khasidashvili
Konstantin Korovin
Sava Krstic
Yves Lafont
Francois Lamarche
Julia Lawall
Aurélien Lemay
Joachim Niehren
Stéphane Lengrand
Christof Loeding
Denis Lugiez
Ian Mackie
Yitzhak Mandelbaum
Jacopo Mantovani
Claude Marché
Ralph Matthes
Francois Metayer

Sebastian Mödersheim
Georg Moser
Leonardo de Moura
Paliath Narendran
Enrica Nicolini
Joachim Niehren
Karl-Heinz Niggl
Thomas Noll
Albert Oliveras
Mizuhito Ogawa
Hitoshi Ohsaki
Vincent van Oostrom
Peter Padawitz
Vincent Padovani
Miguel Palomino
Ricardo Peña
Detlef Plump
François Pottier
Pierre Rety
Frank Raiser
Silvio Ranise
Didier Remy
Christian Retoré
Adrian Riesco
Mario Rodriguez-Artalejo
Yves Roos
Grigore Roşu
David Sabel
Masahiko Sakai
Manfred Schmidt-Schauß
Aleksy Schubert
Traian Serbanuta
Jakob Grue Simonsen
Isabelle Simplot-Ryl
Sergei Soloviev
Mark-Oliver Stehr
Toshinori Takai
Jean-Marc Talbot
Alwen Tiu
Marc Tommasi
Xavier Urbain
Christian Urban
Rafael del Vado Virseda
Alberto Verdejo
René Vestergaard
Germán Vidal
Alicia Villanueva
Eelco Visser
Fer-Jan de Vries
Roel de Vrijer
Johannes Waldmann
Edwin Westbrook
Hongwei Xi
Hans Zantema
Francesco Zappa Nardelli

Printing: Mercedes-Druck, Berlin
Binding: Stein + Lehmann, Berlin

Lecture Notes in Computer Science

For information about Vols. 1–4490

please contact your bookseller or Springer

Vol. 4600: H. Comon-Lundh, C. Kirchner, H. Kirchner, Rewriting, Computation and Proof. XVI, 273 pages. 2007.

Vol. 4595: D. Bošnački, S. Edelkamp (Eds.), Model Checking Software. X, 285 pages. 2007.

Vol. 4592: Z. Kedad, N. Lammari, E. Métais, F. Meziane, Y. Rezgui (Eds.), Natural Language Processing and Information Systems. XIV, 442 pages. 2007.

Vol. 4591: J. Davies, J. Gibbons (Eds.), Integrated Formal Methods. IX, 660 pages. 2007.

Vol. 4590: W. Damm, H. Hermanns (Eds.), Computer Aided Verification. XV, 562 pages. 2007.

Vol. 4589: J. Münch, P. Abrahamsson (Eds.), Product-Focused Software Process Improvement. XII, 414 pages. 2007.

Vol. 4588: T. Harju, J. Karhumäki, A. Lepistö (Eds.), Developments in Language Theory. XI, 423 pages. 2007.

Vol. 4587: R. Cooper, J. Kennedy (Eds.), Data Management. XIII, 259 pages. 2007.

Vol. 4586: J. Pieprzyk, H. Ghodosi, E. Dawson (Eds.), Information Security and Privacy. XIV, 476 pages. 2007.

Vol. 4584: N. Karssemeijer, B. Lelieveldt (Eds.), Information Processing in Medical Imaging. XX, 777 pages. 2007.

Vol. 4583: S.R. Della Rocca (Ed.), Typed Lambda Calculi and Applications. X, 397 pages. 2007.

Vol. 4582: J. Lopez, P. Samarati, J.L. Ferrer (Eds.), Public Key Infrastructure. XI, 375 pages. 2007.

Vol. 4581: A. Petrenko, M. Veanes, J. Tretmans, W. Grieskamp (Eds.), Testing of Software and Communicating Systems. XII, 379 pages. 2007.

Vol. 4578: F. Masulli, S. Mitra, G. Pasi (Eds.), Fuzzy Logic and Applications. XVIII, 693 pages. 2007. (Sublibrary LNAI).

Vol. 4577: N. Sebe, Y. Liu, Y. Zhuang (Eds.), Multimedia Content Analysis and Mining. XIII, 513 pages. 2007.

Vol. 4576: D. Leivant, R. de Queiroz (Eds.), Logic, Language, Information, and Computation. X, 363 pages. 2007.

Vol. 4574: J. Derrick, J. Vain (Eds.), Formal Techniques for Networked and Distributed Systems – FORTE 2007. XI, 375 pages. 2007.

Vol. 4573: M. Kauers, M. Kerber, R. Miner, W. Windsteiger (Eds.), Towards Mechanized Mathematical Assistants. XIII, 407 pages. 2007. (Sublibrary LNAI).

Vol. 4572: F. Stajano, C. Meadows, S. Capkun, T. Moore (Eds.), Security and Privacy in Ad-hoc and Sensor Networks. X, 247 pages. 2007.

Vol. 4570: H.G. Okuno, M. Ali (Eds.), New Trends in Applied Artificial Intelligence. XXI, 1194 pages. 2007. (Sublibrary LNAI).

Vol. 4569: A. Butz, B. Fisher, A. Krüger, P. Olivier, S. Owada (Eds.), Smart Graphics. IX, 237 pages. 2007.

Vol. 4566: M.J Dainoff (Ed.), Ergonomics and Health Aspects of Work with Computers. XVIII, 390 pages. 2007.

Vol. 4565: D.D. Schmorrow, L.M. Reeves (Eds.), Foundations of Augmented Cognition. XIX, 450 pages. 2007. (Sublibrary LNAI).

Vol. 4564: D. Schuler (Ed.), Online Communities and Social Computing. XVII, 520 pages. 2007.

Vol. 4563: R. Shumaker (Ed.), Virtual Reality. XXII, 762 pages. 2007.

Vol. 4561: V.G. Duffy (Ed.), Digital Human Modeling. XXIII, 1068 pages. 2007.

Vol. 4560: N. Aykin (Ed.), Usability and Internationalization, Part II. XVIII, 576 pages. 2007.

Vol. 4559: N. Aykin (Ed.), Usability and Internationalization, Part I. XVIII, 661 pages. 2007.

Vol. 4549: J. Aspnes, C. Scheideler, A. Arora, S. Madden (Eds.), Distributed Computing in Sensor Systems. XIII, 417 pages. 2007.

Vol. 4548: N. Olivetti (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. X, 245 pages. 2007. (Sublibrary LNAI).

Vol. 4547: C. Carlet, B. Sunar (Eds.), Arithmetic of Finite Fields. XI, 355 pages. 2007.

Vol. 4546: J. Kleijn, A. Yakovlev (Eds.), Petri Nets and Other Models of Concurrency – ICATPN 2007. XI, 515 pages. 2007.

Vol. 4545: H. Anai, K. Horimoto, T. Kutsia (Eds.), Algebraic Biology. XIII, 379 pages. 2007.

Vol. 4544: S. Cohen-Boulakia, V. Tannen (Eds.), Data Integration in the Life Sciences. XI, 282 pages. 2007. (Sublibrary LNBI).

Vol. 4543: A.K. Bandara, M. Burgess (Eds.), Inter-Domain Management. XII, 237 pages. 2007.

Vol. 4542: P. Sawyer, B. Paech, P. Heymans (Eds.), Requirements Engineering: Foundation for Software Quality. IX, 384 pages. 2007.

Vol. 4541: T. Okadome, T. Yamazaki, M. Makhtari (Eds.), Pervasive Computing for Quality of Life Enhancement. IX, 248 pages. 2007.

Vol. 4539: N.H. Bshouty, C. Gentile (Eds.), Learning Theory. XII, 634 pages. 2007. (Sublibrary LNAI).

Vol. 4538: F. Escolano, M. Vento (Eds.), Graph-Based Representations in Pattern Recognition. XII, 416 pages. 2007.

Vol. 4537: K.C.-C. Chang, W. Wang, L. Chen, C.A. Ellis, C.-H. Hsu, A.C. Tsoi, H. Wang (Eds.), Advances in Web and Network Technologies, and Information Management. XXIII, 707 pages. 2007.

Vol. 4536: G. Concas, E. Damiani, M. Scotto, G. Succi (Eds.), Agile Processes in Software Engineering and Extreme Programming. XV, 276 pages. 2007.

Vol. 4534: I. Tomkos, F. Neri, J. Solé Pareta, X. Masip Bruin, S. Sánchez Lopez (Eds.), Optical Network Design and Modeling. XI, 460 pages. 2007.

Vol. 4533: F. Baader (Ed.), Term Rewriting and Applications. XII, 419 pages. 2007.

Vol. 4531: J. Indulska, K. Raymond (Eds.), Distributed Applications and Interoperable Systems. XI, 337 pages. 2007.

Vol. 4530: D.H. Akehurst, R. Vogel, R.F. Paige (Eds.), Model Driven Architecture- Foundations and Applications. X, 219 pages. 2007.

Vol. 4529: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), Foundations of Fuzzy Logic and Soft Computing. XIX, 830 pages. 2007. (Sublibrary LNAI).

Vol. 4528: J. Mira, J.R. Álvarez (Eds.), Nature Inspired Problem-Solving Methods in Knowledge Engineering, Part II. XXII, 650 pages. 2007.

Vol. 4527: J. Mira, J.R. Álvarez (Eds.), Bio-inspired Modeling of Cognitive Tasks, Part I. XXII, 630 pages. 2007.

Vol. 4526: M. Malek, M. Reitenspieß, A. van Moorsel (Eds.), Service Availability. X, 155 pages. 2007.

Vol. 4525: C. Demetrescu (Ed.), Experimental Algorithms. XIII, 448 pages. 2007.

Vol. 4524: M. Marchiori, J.Z. Pan, C.d.S. Marie (Eds.), Web Reasoning and Rule Systems. XI, 382 pages. 2007.

Vol. 4523: Y.-H. Lee, H.-N. Kim, J. Kim, Y. Park, L.T. Yang, S.W. Kim (Eds.), Embedded Software and Systems. XIX, 829 pages. 2007.

Vol. 4522: B.K. Ersbøll, K.S. Pedersen (Eds.), Image Analysis. XVIII, 989 pages. 2007.

Vol. 4521: J. Katz, M. Yung (Eds.), Applied Cryptography and Network Security. XIII, 498 pages. 2007.

Vol. 4519: E. Franconi, M. Kifer, W. May (Eds.), The Semantic Web: Research and Applications. XVIII, 830 pages. 2007.

Vol. 4517: F. Boavida, E. Monteiro, S. Mascolo, Y. Koucheryavy (Eds.), Wired/Wireless Internet Communications. XIV, 382 pages. 2007.

Vol. 4516: L. Mason, T. Drwiega, J. Yan (Eds.), Managing Traffic Performance in Converged Networks. XXIII, 1191 pages. 2007.

Vol. 4515: M. Naor (Ed.), Advances in Cryptology - EUROCRYPT 2007. XIII, 591 pages. 2007.

Vol. 4514: S.N. Artemov, A. Nerode (Eds.), Logical Foundations of Computer Science. XI, 513 pages. 2007.

Vol. 4513: M. Fischetti, D.P. Williamson (Eds.), Integer Programming and Combinatorial Optimization. IX, 500 pages. 2007.

Vol. 4511: C. Conati, K. McCoy, G. Paliouras (Eds.), User Modeling 2007. XVI, 487 pages. 2007. (Sublibrary LNAI).

Vol. 4510: P. Van Hentenryck, L. Wolsey (Eds.), Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. X, 391 pages. 2007.

Vol. 4509: Z. Kobti, D. Wu (Eds.), Advances in Artificial Intelligence. XII, 552 pages. 2007. (Sublibrary LNAI).

Vol. 4508: M.-Y. Kao, X.-Y. Li (Eds.), Algorithmic Aspects in Information and Management. VIII, 428 pages. 2007.

Vol. 4507: F. Sandoval, A. Prieto, J. Cabestany, M. Graña (Eds.), Computational and Ambient Intelligence. XXVI, 1167 pages. 2007.

Vol. 4506: D. Zeng, I. Gotham, K. Komatsu, C. Lynch, M. Thurmond, D. Madigan, B. Lober, J. Kvach, H. Chen (Eds.), Intelligence and Security Informatics: Biosurveillance. XI, 234 pages. 2007.

Vol. 4505: G. Dong, X. Lin, W. Wang, Y. Yang, J.X. Yu (Eds.), Advances in Data and Web Management. XXII, 896 pages. 2007.

Vol. 4504: J. Huang, R. Kowalczyk, Z. Maamar, D. Martin, I. Müller, S. Stoutenburg, K.P. Sycara (Eds.), Service-Oriented Computing: Agents, Semantics, and Engineering. X, 175 pages. 2007.

Vol. 4501: J. Marques-Silva, K.A. Sakallah (Eds.), Theory and Applications of Satisfiability Testing – SAT 2007. XI, 384 pages. 2007.

Vol. 4500: N. Streitz, A. Kameas, I. Mavrommati (Eds.), The Disappearing Computer. XVIII, 304 pages. 2007.

Vol. 4499: Y.Q. Shi (Ed.), Transactions on Data Hiding and Multimedia Security II. IX, 117 pages. 2007.

Vol. 4498: N. Abdennahder, F. Kordon (Eds.), Reliable Software Technologies – Ada Europe 2007. XII, 247 pages. 2007.

Vol. 4497: S.B. Cooper, B. Löwe, A. Sorbi (Eds.), Computation and Logic in the Real World. XVIII, 826 pages. 2007.

Vol. 4496: N.T. Nguyen, A. Grzech, R.J. Howlett, L.C. Jain (Eds.), Agent and Multi-Agent Systems: Technologies and Applications. XXI, 1046 pages. 2007. (Sublibrary LNAI).

Vol. 4495: J. Krogstie, A. Opdahl, G. Sindre (Eds.), Advanced Information Systems Engineering. XVI, 606 pages. 2007.

Vol. 4494: H. Jin, O.F. Rana, Y. Pan, V.K. Prasanna (Eds.), Algorithms and Architectures for Parallel Processing. XIV, 508 pages. 2007.

Vol. 4493: D. Liu, S. Fei, Z. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part III. XXVI, 1215 pages. 2007.

Vol. 4492: D. Liu, S. Fei, Z. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part II. XXVII, 1321 pages. 2007.

Vol. 4491: D. Liu, S. Fei, Z.-G. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part I. LIV, 1365 pages. 2007.

Table of Contents

Formal Verification of an Optimizing Compiler 1
Xavier Leroy

Challenges in Satisfiability Modulo Theories 2
Robert Nieuwenhuis, Albert Oliveras, Enric Rodríguez-Carbonell, and Albert Rubio

On a Logical Foundation for Explicit Substitutions 19
Frank Pfenning

Intruders with Caps 20
Siva Anantharaman, Paliath Narendran, and Michael Rusinowitch

Tom: Piggybacking Rewriting on Java 36
Emilie Balland, Paul Brauner, Radu Kopetz, Pierre-Etienne Moreau, and Antoine Reilles

Rewriting Approximations for Fast Prototyping of Static Analyzers 48
Yohan Boichut, Thomas Genet, Thomas Jensen, and Luka Le Roux

Determining Unify-Stable Presentations 63
Thierry Boy de la Tour and Mnacho Echenim

Confluence of Pattern-Based Calculi 78
Horatiu Cirstea and Germain Faure

A Simple Proof That Super-Consistency Implies Cut Elimination 93
Gilles Dowek and Olivier Hermant

Bottom-Up Rewriting Is Inverse Recognizability Preserving 107
Irène Durand and Géraud Sénizergues

Adjunction for Garbage Collection with Application to Graph Rewriting 122
Dominique Duval, Rachid Echahed, and Frederic Prost

Non Strict Confluent Rewrite Systems for Data-Structures with Pointers 137
Rachid Echahed and Nicolas Peltier

Symbolic Model Checking of Infinite-State Systems Using Narrowing ... 153
Santiago Escobar and José Meseguer

Delayed Substitutions 169
José Espírito Santo

Innermost-Reachability and Innermost-Joinability Are Decidable for Shallow Term Rewrite Systems 184
Guillem Godoy and Eduard Huntingford

Termination of Rewriting with Right-Flat Rules 200
Guillem Godoy, Eduard Huntingford, and Ashish Tiwari

Abstract Critical Pairs and Confluence of Arbitrary Binary Relations ... 214
Rémy Haemmerlé and François Fages

On the Completeness of Context-Sensitive Order-Sorted Specifications 229
Joe Hendrix and José Meseguer

KOOL: An Application of Rewriting Logic to Language Prototyping and Analysis 246
Mark Hills and Grigore Roşu

Simple Proofs of Characterizing Strong Normalization for Explicit Substitution Calculi 257
Kentaro Kikuchi

Proving Termination of Rewrite Systems Using Bounds 273
Martin Korp and Aart Middeldorp

Sequence Unification Through Currying 288
Temur Kutsia, Jordi Levy, and Mateu Villaret

The Termination Competition 303
Claude Marché and Hans Zantema

Random Descent 314
Vincent van Oostrom

Correctness of Copy in Calculi with Letrec 329
Manfred Schmidt-Schauß

A Characterization of Medial as Rewriting Rule 344
Lutz Straßburger

The Maximum Length of Mu-Reduction in Lambda Mu-Calculus 359
Makoto Tatsuta

On Linear Combinations of λ-Terms 374
Lionel Vaux

Satisfying KBO Constraints 389
Harald Zankl and Aart Middeldorp

Termination by Quasi-periodic Interpretations 404
Hans Zantema and Johannes Waldmann

Author Index 419

Formal Verification of an Optimizing Compiler

Xavier Leroy

INRIA Rocquencourt
Domaine de Voluceau, B.P. 105, 78153 Le Chesnay, France
Xavier.Leroy@inria.fr

Programmers naturally expect that compilers and other code generation tools produce executable code that behaves as prescribed by source programs. However, compilers are complex programs that perform many subtle transformations. Bugs in compilers do happen and can lead to silently producing incorrect executable code from a correct source program. This is a significant concern in the context of high-assurance software that has been verified (at the source level) using formal methods (static analysis, model checking, program proof, etc): any bug in the compiler can potentially invalidate the guarantees so painfully established by the use of formal methods.

There are several ways to generate confidence in the compilation process, including translation validation and proof-carrying code. This talk focuses on applying program proof technology to the compiler itself, in order to prove a semantic preservation theorem for every pass of the compiler. We present preliminary results from the Compcert experiment: the development and proof of correctness of a moderately-optimizing compiler for a large subset of the C language. The proof of correctness is mechanized using the Coq proof assistant. Moreover, most of the compiler itself is written directly in the functional subset of the Coq specification language, from which executable Caml code is automatically extracted.

The preliminary results are encouraging and suggest two directions for long-term research. One is the formal verification of other tools (code generators, static analyzers, provers, ...) involved in the production and certification of high-assurance software. The other is the systematic use of proof assistants to mechanize programming language semantics, type systems, program transformations and related formal systems.

References

1. Bertot, Y., Grégoire, B., Leroy, X.: A structured approach to proving compiler optimizations based on dataflow analysis. In: Filliâtre, J.-C., Paulin-Mohring, C., Werner, B. (eds.) TYPES 2004. LNCS, vol. 3839, pp. 66–81. Springer, Heidelberg (2006)
2. Blazy, S., Dargaye, Z., Leroy, X.: Formal verification of a C compiler front-end. In: Misra, J., Nipkow, T., Sekerinski, E. (eds.) FM 2006. LNCS, vol. 4085, pp. 460–475. Springer, Heidelberg (2006)
3. Leroy, X.: Formal certification of a compiler back-end, or: programming a compiler with a proof assistant. In: 33rd symposium Principles of Programming Languages, pp. 42–54. ACM Press, New York (2006)

F. Baader (Ed.): RTA 2007, LNCS 4533, p. 1, 2007.

Challenges in Satisfiability Modulo Theories

Robert Nieuwenhuis, Albert Oliveras,
Enric Rodríguez-Carbonell, and Albert Rubio*

Abstract. Here we give a short overview of the DPLL(T) approach to Satisfiability Modulo Theories (SMT), which is at the basis of current state-of-the-art SMT systems. After that, we provide a documented list of theoretical and practical current challenges related to SMT, including some new ideas to exploit SAT techniques in Constraint Programming.

1 Introduction

Propositional satisfiability checkers (SAT solvers) are currently being applied in more and more contexts, including hardware and software verification, in Operations Research (planning, scheduling), as well as in Biology, Linguistics and Medicine. Most SAT solvers are based on the Davis-Putnam-Logemann-Loveland (DPLL) procedure [DP60, DLL62]. The performance of DPLL-based SAT solvers has improved spectacularly in the last years, due to better implementation techniques and conceptual enhancements such as *backjumping*, *conflict-driven lemma learning* and *restarts* [MSS99, MMZ+01, ES03]. However, some practical problems are more naturally expressed in logics that are more expressive than propositional logic.

For example, for timed automata, a good choice is *difference logic*, where formulas contain atoms of the form $a - b \leq k$, which are interpreted with respect to a background theory T of the integers, rationals or reals. Similarly, for the verification of pipelined microprocessors it is convenient to consider a logic of *Equality with Uninterpreted Functions (EUF)*, where the background theory T specifies a congruence [BD94]. To mention just one other example, the conditions arising from program verification usually involve arrays, lists and other data structures, so it becomes very natural to consider satisfiability problems *modulo* the theory T of these data structures. In such applications, problems may contain thousands of clauses like

$$p \quad \vee \quad \neg q \quad \vee \quad a = f(b-c) \quad \vee \quad read(s,\, f(b-c)\,) = d \quad \vee \quad a - g(c) \leq 7$$

containing purely propositional atoms as well as atoms over (combined) theories. This is known as the *Satisfiability Modulo Theories* (SMT) problem for a theory T: given a formula F, determine whether F is T-satisfiable, i.e., whether there exists a model of T that is also a model of F.

* Technical Univ. of Catalonia, Barcelona. All authors partially supported by Spanish Min. of Educ. and Science through the LogicTools project (TIN2004-03382) and Intel Corp. Research Grant: "SMT Solvers for High-Level Hardware Verification".

F. Baader (Ed.): RTA 2007, LNCS 4533, pp. 2–18, 2007.

SMT has become an extremely active area of research. A rapidly growing library of benchmarks for SMT with a formal syntax and semantics exists [RT03], as well as a yearly SMT competition (both SMT-LIB and SMT-COMP are easily found on the web).

The DPLL(T) approach to SMT is based on a general DPLL(X) engine, whose parameter X can be instantiated with specialized solvers $Solver_T$ for given theories T, thus producing a system DPLL(T). Once the DPLL(X) engine has been implemented, new theories can be dealt with by simply plugging in new theory solvers. These solvers must only be able to deal with *conjunctions* of theory literals and conform to a minimal and simple set of additional requirements.

In Sections 2, 3 and 4 of this paper, by means of a rewrite-rule-based framework called Abstract DPLL we first give a brief overview of DPLL, SMT, and the DPLL(T) approach to SMT (we refer to [NOT06] for all details and references). In Section 5 we describe a number of theoretical and practical challenges in SMT. Extensions for handling new theories and applications, including optimization and constraint programming are discussed, as well as for first-order theorem proving. Other challenges involve the design of efficient data structures and algorithms for implementing certain key parts of SMT solvers. All of them are closely related to the area of rewriting.

2 The DPLL Procedure

Let P be a fixed finite set of propositional symbols. If $p \in P$, then p and $\neg p$ are *literals* of P. The *negation* of a literal l, written $\neg l$, denotes $\neg p$ if l is p, and p if l is $\neg p$. A *clause* is a disjunction of literals $l_1 \vee \ldots \vee l_n$. A *unit clause* is a clause consisting of a single literal. A (finite, non-empty, CNF) *formula* is a conjunction of one or more clauses $C_1 \wedge \ldots \wedge C_n$. When it leads to no ambiguities, we sometimes also write such a formula in set notation $\{C_1, \ldots, C_n\}$ or simply replace $\wedge$ connectives by commas.

A (partial truth) *assignment* M is a set of literals such that $\{p, \neg p\} \subseteq M$ for no p. A literal l is *true* in M if $l \in M$, it is *false* in M if $\neg l \in M$, and l is *undefined* in M otherwise. M is *total* over P if no literal of P is undefined in M. A clause C is true in M if at least one of its literals is true in M. It is false in M if all its literals are false in M, and it is undefined in M otherwise. A formula F is true in M, or *satisfied* by M, denoted $M \models F$, if all its clauses are true in M. In that case, M is called a *model* of F. If F has no models then it is called *unsatisfiable*. If F and F' are formulas, we write $F \models F'$ if F' is true in all models of F. Then we say that F' *is entailed by* F, or is a *logical consequence* of F.

In what follows, (possibly subscripted or primed) lowercase l *always* denotes literals. Similarly C and D always denote clauses, F and G denote formulas, and M and N are assignments. If C is a clause $l_1 \vee \ldots \vee l_n$, we sometimes write $\neg C$ to denote the formula $\neg l_1 \wedge \ldots \wedge \neg l_n$.

Here a DPLL procedure is modeled by a transition relation over states. A state is either *FailState* or a pair $M \parallel F$, where F is a finite set of clauses and M is a sequence of literals that is seen as a partial assignment. Some literals l