

Tero Harju  
Juhani Karhumäki  
Arto Lepistö (Eds.)

LNCS 4588

# Developments in Language Theory

11th International Conference, DLT 2007  
Turku, Finland, July 2007  
Proceedings



Springer

Tero Harju Juhani Karhumäki  
Arto Lepistö (Eds.)

# Developments in Language Theory

11th International Conference, DLT 2007  
Turku, Finland, July 3-6, 2007  
Proceedings



Springer

## Volume Editors

Tero Harju  
Department of Mathematics  
University of Turku  
20014 Turku, Finland  
E-mail: tero.harju@utu.fi

Juhani Karhumäki  
Department of Mathematics  
University of Turku  
20014 Turku, Finland  
E-mail: karhumak@utu.fi

Arto Lepistö  
Department of Mathematics  
University of Turku  
20014 Turku, Finland  
E-mail: alepisto@utu.fi

Library of Congress Control Number: 2007929029

CR Subject Classification (1998): F.4.3, F.4.2, F.4, F.3, F.1, G.2

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

ISSN	0302-9743
ISBN-10	3-540-73207-1 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-73207-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper      SPIN: 12080298      06/3180      5 4 3 2 1 0

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Switzerland*

John C. Mitchell

*Stanford University, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

Oscar Nierstrasz

*University of Bern, Switzerland*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*University of Dortmund, Germany*

Madhu Sudan

*Massachusetts Institute of Technology, MA, USA*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Moshe Y. Vardi

*Rice University, Houston, TX, USA*

Gerhard Weikum

*Max-Planck Institute of Computer Science, Saarbruecken, Germany*

# Lecture Notes in Computer Science

For information about Vols. 1–4467

please contact your bookseller or Springer

Vol. 4600: H. Comon-Lundh, C. Kirchner, H. Kirchner (Eds.), *Rewriting, Computation and Proof*. XVIII, 273 pages. 2007.

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

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

Vol. 4581: A. Petrenko, M. Veanes, J. Tretmans, W. Grieskamp (Eds.), *Testing of Software and Communicating Systems*. XII, 379 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. 4569: A. Butz, B. Fisher, A. Krüger, P. Olivier, S. Owada (Eds.), *Smart Graphics*. IX, 237 pages. 2007.

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

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

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

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. 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. Reitenspiß, 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, 497 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. 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.
- Vol. 4490: Y. Shi, G.D. van Albada, J. Dongarra, P.M.A. Sloot (Eds.), *Computational Science – ICCS 2007*, Part IV. XXXVII, 1211 pages. 2007.
- Vol. 4489: Y. Shi, G.D. van Albada, J. Dongarra, P.M.A. Sloot (Eds.), *Computational Science – ICCS 2007*, Part III. XXXVII, 1257 pages. 2007.
- Vol. 4488: Y. Shi, G.D. van Albada, J. Dongarra, P.M.A. Sloot (Eds.), *Computational Science – ICCS 2007*, Part II. XXXV, 1251 pages. 2007.
- Vol. 4487: Y. Shi, G.D. van Albada, J. Dongarra, P.M.A. Sloot (Eds.), *Computational Science – ICCS 2007*, Part I. LXXXI, 1275 pages. 2007.
- Vol. 4486: M. Bernardo, J. Hillston (Eds.), *Formal Methods for Performance Evaluation*. VII, 469 pages. 2007.
- Vol. 4485: F. Sgallari, A. Murli, N. Paragios (Eds.), *Scale Space and Variational Methods in Computer Vision*. XV, 931 pages. 2007.
- Vol. 4484: J.-Y. Cai, S.B. Cooper, H. Zhu (Eds.), *Theory and Applications of Models of Computation*. XIII, 772 pages. 2007.
- Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), *Logic Programming and Nonmonotonic Reasoning*. IX, 327 pages. 2007. (Sublibrary LNAI).
- Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), *Rough Sets, Fuzzy Sets, Data Mining and Granular Computing*. XIV, 585 pages. 2007. (Sublibrary LNAI).
- Vol. 4481: J. Yao, P. Lingras, W.-Z. Wu, M. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), *Rough Sets and Knowledge Technology*. XIV, 576 pages. 2007. (Sublibrary LNAI).
- Vol. 4480: A. LaMarca, M. Langheinrich, K.N. Truong (Eds.), *Pervasive Computing*. XIII, 369 pages. 2007.
- Vol. 4479: I.F. Akyildiz, R. Sivakumar, E. Ekici, J.C.d. Oliveira, J. McNair (Eds.), *NETWORKING 2007. Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*. XXVII, 1252 pages. 2007.
- Vol. 4478: J. Martí, J.M. Benedí, A.M. Mendonça, J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Part II. XXVII, 657 pages. 2007.
- Vol. 4477: J. Martí, J.M. Benedí, A.M. Mendonça, J. Serrat (Eds.), *Pattern Recognition and Image Analysis*, Part I. XXVII, 625 pages. 2007.
- Vol. 4476: V. Gorodetsky, C. Zhang, V.A. Skormin, L. Cao (Eds.), *Autonomous Intelligent Systems: Multi-Agents and Data Mining*. XIII, 323 pages. 2007. (Sublibrary LNAI).
- Vol. 4475: P. Crescenzi, G. Prencipe, G. Pucci (Eds.), *Fun with Algorithms*. X, 273 pages. 2007.
- Vol. 4474: G. Prencipe, S. Zaks (Eds.), *Structural Information and Communication Complexity*. XI, 342 pages. 2007.
- Vol. 4472: M. Haindl, J. Kittler, F. Roli (Eds.), *Multiple Classifier Systems*. XI, 524 pages. 2007.
- Vol. 4471: P. Cesar, K. Chorianopoulos, J.F. Jensen (Eds.), *Interactive TV: a Shared Experience*. XIII, 236 pages. 2007.
- Vol. 4470: Q. Wang, D. Pfahl, D.M. Raffo (Eds.), *Software Process Dynamics and Agility*. XI, 346 pages. 2007.
- Vol. 4469: K.-c. Hui, Z. Pan, R.C.-k. Chung, C.C.L. Wang, X. Jin, S. Göbel, E.C.-L. Li (Eds.), *Technologies for E-Learning and Digital Entertainment*. XVIII, 974 pages. 2007.
- Vol. 4468: M.M. Bonsangue, E.B. Johnsen (Eds.), *Formal Methods for Open Object-Based Distributed Systems*. X, 317 pages. 2007.

# Preface

The 11th International Conference on Developments in Language Theory (DLT 2007) was held at the University of Turku, Finland, July 3–6, 2007. This was the second time DLT took place in Turku. Indeed, the very first meeting was organized in Turku in 1993. Consequent meetings were held in Magdeburg (1995), Thessaloniki (1997), Aachen (1999), Vienna (2001), Kyoto (2002), Szeged (2003), Auckland (2004), Palermo (2005), and Santa Barbara (2006). The conference series is held under the auspices of the European Association for Theoretical Computer Science.

The DLT meeting can be viewed as the main conference on automata and formal language theory. The current topics of the conference include the following: grammars, acceptors and transducers for strings, trees, graphs and arrays, efficient text algorithms, algebraic theories for automata and languages, combinatorial and algebraic properties of words and languages, variable-length codes, symbolic dynamics, decision problems, relations to complexity theory and logic, picture description and analysis, polyominoes and bidimensional patterns, cryptography, concurrency, bio-inspired computing, quantum computing. This volume of *Lecture Notes in Computer Science* contains the papers that were presented at DLT 2007, including the abstracts or full papers of the six invited speakers Volker Diekert (Stuttgart), Thomas Henzinger (Lausanne), Michal Kunc (Brno), Ming Li (Waterloo), Jacques Sakarovitch (Paris), and Kai Salomaa (Kingston)

For the conference, 32 contributed papers were selected from a record-breaking 74 submissions. We warmly thank the authors of the papers, the members of the Program Committee, who faced many hard decisions, and the reviewers of the submitted papers for their valuable work. All these efforts were the basis of the success of the conference. In particular, we are very thankful to the invited speakers of the conference. Finally, we thank the Organizing Committee for its splendid work and also the members of the Steering Committee.

Finally, we wish to thank the support of the conference sponsors: The Academy of Finland, The Finnish Cultural Foundation, the Finnish Academy of Science and Letters / Vilho, Yrjö and Kalle Väisälä Foundation, the City of Turku, the University of Turku, the Turku Centre for Computer Science, and Centro Hotel.

April 2007

Tero Harju  
Juhani Karhumäki  
Arto Lepistö

# Organization

DLT 2007 was organized by the Department of Mathematics and Centre for Fundamentals of Computing and Discrete Mathematics (FUNDIM), University of Turku.

## Invited Speakers

Volker Diekert (Stuttgart, Germany)  
Thomas Henzinger (Lausanne, Switzerland)  
Michal Kunc (Brno, Czech Republic)  
Ming Li (Waterloo, Canada)  
Jacques Sakarovitch (Paris, France)  
Kai Salomaa (Kingston, Canada)

## Program Committee

Cristian Calude (Auckland, New Zealand)  
Julien Cassaigne (Marseille, France)  
Christian Choffrut (Paris, France)  
Jürgen Dassow (Magdeburg, Germany)  
Bruno Durand (Marseille, France)  
Massimiliano Goldwurm (Milan, Italy)  
Tero Harju, Co-chair (Turku, Finland)  
Juraj Hromkovič (Zurich, Switzerland)  
Oscar Ibarra (Santa Barbara, USA)  
Lucian Ilie (London, Canada)  
Masami Ito (Kyoto, Japan)  
Nataša Jonoska (Tampa Bay, USA)  
Juhani Karhumäki, Co-chair (Turku, Finland)  
Wojciech Rytter (Warsaw, Poland)  
Mikhail Volkov (Ekaterinburg, Russia)  
Klaus Wagner (Würzburg, Germany)

## Organizing Committee

Juhani Karhumäki (Co-chair)  
Vesa Halava (Co-chair)  
Elena Czeizler  
Eugen Czeizler  
Eero Lehtonen  
Arto Lepistö  
Alexander Okhotin  
Petri Salmela



## Referees

F. Ablayev	I. Honkala	F. Otto
R. Alur	J. Honkala	B. Palano
A. Angeleska	J. Hromkovič	G. Paun
M. Babenko	O. Ibarra	D. Perrin
A. Ballier	L. Ilie	G. Pighizzini
L. Bienvenu	M. Ito	J.-E. Pin
H.-J. Böckenhauer	E. Jeandel	J. Pirnot
P. Boldi	N. Jonoska	W. Plandowski
L. Breveglieri	J. Justin	R. Radicioni
H. Buhrman	C. Kapoutsis	B. Ravikumar
C. Calude	J. Karhumäki	C. Reutenauer
O. Carton	J. Kari	G. Richomme
J. Cassaigne	L. Kari	A.E. Romashchenko
A. Cherubini	D. Kephart	W. Rytter
C. Choffrut	R. Kralovic	K. Saari
M. Chrobak	D. Krob	P. Salmela
B. Courcelle	M. Kutrib	A. Salomaa
M. Crochemore	S. Lasota	K. Salomaa
F. D'Alessandro	M. Latteux	N. Santean
J. Dassow	A. Lepistö	M. Santha
F. Drewes	F. Levé	S. Seibert
E. Duchi	G. Lischke	V. Selivanov
B. Durand	K. Lodya	A. Shen
Z. Fülöp	M. Lohrey	A. Staninska
E. Formenti	S. Lombardy	J.-M. Talbot
A. Frid	V. Lonati	D. Tamascelli
P. Gastin	P. Madhusudan	N. Tran
M. Goldwurm	P. Massazza	S. Travers
P. Goralčík	B. Masson	R. Treinen
J. Gruska	C. Mereghetti	G. Vaszil
Y. Guesnet	T. Momke	M. Volkov
V. Halava	B. Nagy	K. Wagner
T. Harju	E. Ochmański	H.-C. Yen
M. Hirvensalo	A. Okhotin	S. Yu
M. Holzer	N. Ollinger	

## Sponsoring Institutions

The Academy of Finland

The Finnish Cultural Foundation

Finnish Academy of Science and Letters / Vilho, Yrjö and Kalle Väisälä Fund

Nokia Foundation

Turku University Foundation

Centro Hotel

# Table of Contents

---

## I Invited Talks

---

On First-Order Fragments for Words and Mazurkiewicz Traces . . . . .	1
<i>Volker Diekert and Manfred Kufleitner</i>	
Quantitative Generalizations of Languages . . . . .	20
<i>Thomas A. Henzinger</i>	
What Do We Know About Language Equations? . . . . .	23
<i>Michal Kunc</i>	
Information Distance and Applications . . . . .	28
<i>Ming Li</i>	
Finite Automata and the Writing of Numbers . . . . .	29
<i>Jacques Sakarovitch</i>	
Descriptive Complexity of Nondeterministic Finite Automata . . . . .	31
<i>Kai Salomaa</i>	

---

## II Contributed Papers

---

From Determinism to Non-determinism in Recognizable Two-dimensional Languages . . . . .	36
<i>Marcella Anselmo, Dora Giammarresi, and Maria Madonia</i>	
Coding Partitions: Regularity, Maximality and Global Ambiguity . . . . .	48
<i>Marie-Pierre Béal, Fabio Burderi, and Antonio Restivo</i>	
Multi-letter Reversible and Quantum Finite Automata . . . . .	60
<i>Aleksandrs Belovs, Ansis Rosmanis, and Juris Smotrovs</i>	
Approximability and Non-approximability Results in Computing the Mean Speedup of Trace Monoids . . . . .	72
<i>Alberto Bertoni and Roberto Radicioni</i>	
The Dynamics of Cellular Automata in Shift-Invariant Topologies . . . . .	84
<i>Laurent Bienvenu and Mathieu Sablik</i>	
Two Element Unavoidable Sets of Partial Words . . . . .	96
<i>Francine Blanchet-Sadri, Naomi C. Brownstein, and Justin Palumbo</i>	

Hairpin Finite Automata.....	108
<i>Henning Bordihn, Markus Holzer, and Martin Kutrib</i>	
Characterizing Reduction Graphs for Gene Assembly in Ciliates .....	120
<i>Robert Brijder and Hendrik Jan Hoogeboom</i>	
2-Visibly Pushdown Automata.....	132
<i>Dario Carotenuto, Aniello Murano, and Adriano Peron</i>	
An Efficient Computation of the Equation $\mathbb{K}$ -Automaton of a Regular $\mathbb{K}$ -Expression .....	145
<i>Jean-Marc Champarnaud, Faissal Ouardi, and Djelloul Ziadi</i>	
An Extension of Newton's Method to $\omega$ -Continuous Semirings .....	157
<i>Javier Esparza, Stefan Kiefer, and Michael Luttenberger</i>	
Non-constructive Methods for Finite Probabilistic Automata .....	169
<i>Rūsiņš Freivalds</i>	
The Unambiguity of Segmented Morphisms .....	181
<i>Dominik D. Freydenberger and Daniel Reidenbach</i>	
Commutation of Binary Factorial Languages .....	193
<i>Anna E. Frid</i>	
Inapproximability of Nondeterministic State and Transition Complexity Assuming $\mathbf{P} \neq \mathbf{NP}$ .....	205
<i>Hermann Gruber and Markus Holzer</i>	
State Complexity of Union and Intersection of Finite Languages .....	217
<i>Yo-Sub Han and Kai Salomaa</i>	
Bisimulation Minimisation for Weighted Tree Automata .....	229
<i>Johanna Högberg, Andreas Maletti, and Jonathan May</i>	
Conjunctive Grammars Can Generate Non-regular Unary Languages ...	242
<i>Artur Jež</i>	
Deterministic Blow-Ups of Minimal Nondeterministic Finite Automata over a Fixed Alphabet .....	254
<i>Jozef Jirásek, Galina Jirásková, and Alexander Szabari</i>	
Reduced Languages as $\omega$ -Generators.....	266
<i>Sandrine Julia and Tran Vinh Duc</i>	
Avoiding Approximate Squares .....	278
<i>Dalia Krieger, Pascal Ochem, Narad Rampersad, and Jeffrey Shallit</i>	
Duplication Roots.....	290
<i>Peter Leupold</i>	

Complexity Theory for Splicing Systems . . . . .	300
<i>Remco Loos and Mitsunori Ogihara</i>	
Descriptive Complexity of Bounded Context-Free Languages . . . . .	312
<i>Andreas Malcher and Giovanni Pighizzini</i>	
Definable Transductions and Weighted Logics for Texts . . . . .	324
<i>Christian Mathissen</i>	
A Star Operation for Star-Free Trace Languages . . . . .	337
<i>Edward Ochmański and Krystyna Stawikowska</i>	
Finite Automata on Unranked and Unordered DAGs . . . . .	346
<i>Lutz Priese</i>	
On Almost Periodicity Criteria for Morphic Sequences in Some Particular Cases . . . . .	361
<i>Yuri Pritykin</i>	
A Local Balance Property of Episturmian Words . . . . .	371
<i>Gwénaél Richomme</i>	
Suffix Automata and Standard Sturmian Words . . . . .	382
<i>Marinella Sciortino and Luca Q. Zamboni</i>	
Fine Hierarchy of Regular Aperiodic $\omega$ -Languages . . . . .	399
<i>Victor L. Selivanov</i>	
On Transition Minimality of Bideterministic Automata . . . . .	411
<i>Hellis Tamm</i>	
<b>Author Index</b> . . . . .	423

# On First-Order Fragments for Words and Mazurkiewicz Traces

## A Survey

Volker Diekert<sup>1</sup> and Manfred Kufleitner<sup>2</sup>

<sup>1</sup> Universität Stuttgart, FMI, Germany  
`diekert@fmi.uni-stuttgart.de`

<sup>2</sup> Université Bordeaux 1, LaBRI, France  
`manfred.kufleitner@labri.fr`

**Abstract.** We summarize several characterizations, inclusions, and separations on fragments of first-order logic over words and Mazurkiewicz traces. The results concerning Mazurkiewicz traces can be seen as generalizations of those for words. It turns out that over traces it is crucial, how easy concurrency can be expressed. Since there is no concurrency in words, this distinction does not occur there. In general, the possibility of expressing concurrency also increases the complexity of the satisfiability problem.

In the last section we prove an algebraic and a language theoretic characterization of the fragment  $\Sigma_2[E]$  over traces. Over words the relation  $E$  is simply the order of the positions. The algebraic characterization yields decidability of the membership problem for this fragment. For words this result is well-known, but although our proof works in a more general setting it is quite simple and direct. An essential step in the proof consists of showing that every homomorphism from a free monoid to a finite aperiodic monoid  $M$  admits a factorization forest of finite height. We include a simple proof that the height is bounded by  $3|M|$ .

## 1 Introduction

The concept of partially commutative free monoids has first been considered by Cartier and Foata [1]. Later Keller and Mazurkiewicz used them as a model for concurrent systems and Mazurkiewicz established the notion of *trace monoids* for these structures [16,19,20]. Since then the elements of partially commutative monoids are called *Mazurkiewicz traces*. Many aspects of traces and trace languages have been researched, see *The Book of Traces* [7] for an overview.

Over words it has turned out that finite monoids are a powerful technique to refine the class of recognizable languages [9]. For fragments of first-order logic, in many cases it is a characterization in terms of algebra which leads to decidability of the membership problem. For example, on the algebraic side first-order logics as well as temporal logics corresponds to aperiodic monoids, see e.g. [12]. The probably most interesting fragment of them is given by the variety **DA**. It admits many different characterizations, which led to the title *Diamonds are Forever* in

[30]. One of the purposes of this paper is to survey the situation over words and Mazurkiewicz traces.

Words can be seen as a special case of Mazurkiewicz traces and the corresponding results for words have been known before their generalizations to traces. Since over words we do not have any concurrency the situation is more complex for traces, and therefore not all word results remain valid for traces. It turns out that for traces the distinction between so-called dependence graphs and partial orders is rather crucial. Over words, both notions coincide.

The paper is organized as follows. In Section 2 we introduce Mazurkiewicz traces using a graph theoretic approach since this directly translates into the logic setting. After that we present further notions used in this paper which include the definition of fragments of first-order logic and temporal logic, some language operations, and the connections to finite monoids. In Section 3 we give several characterizations of languages whose syntactic monoid is aperiodic or in the variety **DA**. In a second part of this section we describe the alternation hierarchy of first-order logic using language operations. Section 4 contains some ideas and approaches revealing how concurrency increases the expressive power of logical fragments and in Section 5 we present some results showing that in general, concurrency also increases the complexity of the satisfiability problem.

Finally, in Section 6 we give a self-contained proof of a language theoretic and an algebraic characterization of the fragment  $\Sigma_2$  over traces. The algebraic characterization yields decidability of the membership problem for this fragment. For words this result is well-known, but although our proof works in a more general setting it is quite simple and direct. A main tool in this proof are factorization forests. We give a simple and essentially self-contained proof for Simon's theorem on factorization forests in the special case of finite aperiodic monoids  $M$ . Our proof can be generalized to arbitrary monoids and still yields that the height of the factorization forests is bounded by  $3|M|$ . The previously published bound was  $7|M|$ , see [2]. After having completed our paper we learned that the bound  $3|M|$  has been stated in the Technical Report [3], too.

## 2 Preliminaries

### Words and Mazurkiewicz Traces

A *dependence alphabet* is a pair  $(\Gamma, D)$  where the alphabet  $\Gamma$  is a finite set (of actions) and the *dependence relation*  $D \subseteq \Gamma \times \Gamma$  is reflexive and symmetric. The *independence relation*  $I$  is the complement of  $D$ . A *Mazurkiewicz trace* is an isomorphism class of a node-labeled directed acyclic graph  $t = [V, E, \lambda]$ , where  $V$  is a finite set of vertices labeled by  $\lambda : V \rightarrow \Gamma$  and  $E \subseteq (V \times V) \setminus \text{id}_V$  is the edge relation such that for any two different vertices  $x, y \in V$  we have either  $(x, y) \in E$  or  $(y, x) \in E$ .

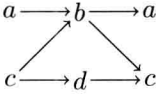
We call  $[V, E, \lambda]$  a *dependence graph*. By  $<$  we mean the transitive closure of  $E$ . We write  $x \parallel y$  if  $x \neq y$  and the vertices  $x$  and  $y$  are incomparable with respect to  $<$ . In this case we say that  $x$  and  $y$  are *independent* or *concurrent*. Node labeled graphs  $(V, E, \lambda)$  and  $(V', E', \lambda')$  are isomorphic if and only if the

corresponding labeled *partial orders*  $(V, <, \lambda)$  and  $(V', <', \lambda')$  are isomorphic. The transitive reduction of a trace is called the *Hasse diagram*.

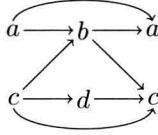
For  $D = \Gamma \times \Gamma$  we obtain *words*. The vertices in words are linearly ordered and the relations  $E$  and  $<$  are identical. Let  $t_1 = [V_1, E_1, \lambda_1]$  and  $t_2 = [V_2, E_2, \lambda_2]$  be traces. Then we define the concatenation of  $t_1$  and  $t_2$  to be  $t_1 \cdot t_2 = [V, \leq, \lambda]$  where  $V = V_1 \cup V_2$  is a disjoint union,  $\lambda = \lambda_1 \cup \lambda_2$ , and  $E = E_1 \cup E_2 \cup \{(x, y) \in V_1 \times V_2 \mid (\lambda(x), \lambda(y)) \in D\}$ . The set  $\mathbb{M}$  of traces becomes a monoid with the empty trace  $1 = (\emptyset, \emptyset, \emptyset)$  as unit. It is generated by  $\Gamma$ , where a letter  $a$  is viewed as a graph with a single vertex labeled by  $a$ . Thus, we obtain a canonical surjective homomorphism  $\pi : \Gamma^* \rightarrow \mathbb{M}$ . The effect of the mapping  $\pi$  can be made explicit as follows. We start with a word  $w = a_1 \cdots a_n$  where all  $a_x$  are letters in  $\Gamma$ . Each  $x$  is viewed as an element in  $\{1, \dots, n\}$  with label  $\lambda(x) = a_x$ . We draw an arc from  $x$  to  $y$  if and only if both,  $x < y$  and  $(a_x, a_y) \in D$ . This dependence graph is  $\pi(w)$ . Note that  $\mathbb{M}$  is also canonically isomorphic to the quotient monoid  $\Gamma^* / \{ab = ba \mid (a, b) \in I\}$ . By abuse of notation we often identify a trace  $t$  and its word representatives  $w \in \pi^{-1}(t)$ .

*Example 1.* Let  $(\Gamma, D) = a - b - c - d$  where self-loops are omitted. Consider the trace  $t = acdbca$ . We have  $acdbca = cabadc$  in  $\mathbb{M}$ . The trace  $t$  has the following graphical presentations:

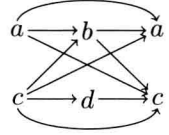
Hasse diagram:



Dependence graph  $E$ :



Partial order  $<$ :



In  $t$ , the node labeled with  $d$  is concurrent to all nodes labeled with  $a$  or  $b$ .  $\square$

There is a basic observation which holds for all  $t \in \mathbb{M}$  and all vertices  $x, y$  of  $t$ :

$$(x, y) \in E \Leftrightarrow (x, y) \in E^+ \wedge (\lambda(x), \lambda(y)) \in D \quad (1)$$

$$(x, y) \in E^+ \Leftrightarrow \exists x_1 \cdots \exists x_{|r|} : \left\{ \begin{array}{l} x_{|r|} = y \wedge (x, x_1) \in E \wedge \\ \bigwedge_{1 \leq i < |r|} (x_i, x_{i+1}) \in E \cup \text{id}_V \end{array} \right\} \quad (2)$$

This shows that traces can be either represented by their dependence graphs or as a partial order without losing any information. There are some standard notations we adopt here. By  $\text{alph}(t)$  we denote the *alphabet* of a trace  $t$ , i.e., the set of letters occurring as labels of some position. By  $|t|$  we denote the *length* of a trace, i.e., the number of vertices of  $t$ . A *trace language*  $L$  is a subset of  $\mathbb{M}$ .

## First-Order Logic and Temporal Logic

The syntax of first-order logic formulas  $\text{FO}[E]$  is built upon atomic formulas of type

$$\top, \lambda(x) = a, \text{ and } (x, y) \in E,$$

where  $\top$  means *true*,  $x, y$  are variables and  $a \in \Gamma$  is a letter. If  $\varphi, \psi$  are first-order formulas, then  $\neg\varphi, \varphi \vee \psi, \exists x \varphi$  are first-order formulas, too. We use the usual shortcuts as  $\perp = \neg\top$  meaning *false*,  $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$ , and  $\forall x \varphi = \neg\exists x \neg\varphi$ . Note that  $x = y$  can be expressed by

$$\bigvee_{a \in \Gamma} (\lambda(x) = a \wedge \lambda(y) = a) \wedge (x, y) \notin E \wedge (y, x) \notin E$$

We let  $\text{FO}^m[E]$  be the set of all formulas with at most  $m$  different names for variables. There are completely analogous definitions for the first-order logic  $\text{FO}[<]$ . The only difference is that instead of  $(x, y) \in E$  we have an atomic predicate  $x < y$ .

Given  $\varphi \in \text{FO}[E] \cup \text{FO}[<]$  the semantics is defined as usual [32]. In particular, if all free variables in  $\varphi$  belong to a set  $\{x_1, \dots, x_m\}$ , then for all  $t \in \mathbb{M}$  and all  $x_1, \dots, x_m \in t$  we write  $t, x_1, \dots, x_m \models \varphi$  if  $t$  satisfies  $\varphi(x_1, \dots, x_m)$ . We identify formulas by semantic equivalence (over finite traces). Hence, if  $\varphi$  and  $\psi$  are formulas with  $m$  free variables, then we write  $\varphi = \psi$  as soon as  $t, x_1, \dots, x_m \models (\varphi \leftrightarrow \psi)$  for all  $t \in \mathbb{M}$  and all  $x_1, \dots, x_m \in t$ . Due to (1) we have that  $\text{FO}^m[E]$  is a fragment of  $\text{FO}^m[<]$ . A *first-order sentence* is a formula in  $\text{FO}[E]$  or  $\text{FO}[<]$  without free variables. For a first-order sentence  $\varphi$  we define  $L(\varphi) = \{t \in \mathbb{M} \mid t \models \varphi\}$ . A trace language  $L \subseteq \mathbb{M}$  is called *first-order definable* if  $L = L(\varphi)$  for some first-order sentence  $\varphi$  and we let  $\text{FO}(\mathbb{M}) = \{L(\varphi) \mid \varphi \in \text{FO}[E]\}$ . We do not write  $\text{FO}[E](\mathbb{M})$ , because  $\text{FO}(\mathbb{M}) = \{L(\varphi) \mid \varphi \in \text{FO}[<]\}$  as well, due to (2). So, in first-order it is not necessary to distinguish between  $E$  and  $<$ . However, for subclasses of  $\text{FO}$  we need this distinction. We define the following classes for  $E' = E$  and  $E' = <$ , respectively.

The fragment  $\Sigma_n[E']$  contains all formulas in prenex normal form with  $n$  blocks of alternating quantifiers starting with a block of existential quantifiers whereas in  $\Pi_n[E']$  formulas start with a block of universal quantifiers. According to our convention to identify equivalent formulas, it makes sense to write e.g.  $\varphi \in \Sigma_n[E'] \Leftrightarrow \neg\varphi \in \Pi_n[E']$ . Although in general the transitive closure of binary relations is not expressible in first-order logic, we have  $\bigcup_{0 \leq n} \Sigma_n[E] = \text{FO}[<]$  due to the following observation obtained from (1) and (2):

$$\Sigma_n[E] \subseteq \Sigma_n[<] \subseteq \Sigma_{n+1}[E]$$

For  $E' = E$  and  $E' = <$  we define the following language classes:

- $\text{FO}^m[E'](\mathbb{M}) = \{L(\varphi) \mid \varphi \in \text{FO}^m[E']\}$ .
- $\Sigma_n[E'](\mathbb{M}) = \{L(\varphi) \mid \varphi \in \Sigma_n[E']\}$ .
- $\Pi_n[E'](\mathbb{M}) = \{L(\varphi) \mid \varphi \in \Pi_n[E']\}$ .
- $\Delta_n[E'](\mathbb{M}) = \Sigma_n[E'](\mathbb{M}) \cap \Pi_n[E'](\mathbb{M})$ .

Now,  $\text{FO}^m[E'](\mathbb{M})$  and  $\Delta_n[E'](\mathbb{M})$  are Boolean algebras and  $\Sigma_n[E'](\mathbb{M})$  and  $\Pi_n[E'](\mathbb{M})$  are closed under union and intersection.

Local temporal logic formulas are defined by first-order formulas having at most one free variable. In this paper we focus on unary operators and local semantics. In temporal logic we write  $a(x)$  for the atomic formula  $\lambda(x) = a$ .



Inductively, we define  $\text{SF}\varphi(x)$  (*Strict Future*),  $\text{SP}\varphi(x)$  (*Strict Past*),  $\text{M}\varphi(x)$  (*soMewhere*),  $\text{Eco}\varphi(x)$  (*Exists concurrently*) as follows.

$$\begin{aligned}\text{SF}\varphi(x) &= \exists y : x < y \wedge \varphi(y) \\ \text{SP}\varphi(x) &= \exists y : y < x \wedge \varphi(y) \\ \text{M}\varphi(x) &= \exists y : \varphi(y) \\ \text{Eco}\varphi(x) &= \exists y : x \parallel y \wedge \varphi(y)\end{aligned}$$

It is common to write  $\varphi$  instead of  $\varphi(x)$ . Let  $\mathcal{C}$  be a subset of temporal operators from the set above, then  $\text{TL}[\mathcal{C}]$  means the formulas where all operators are from  $\mathcal{C}$ . In order to pass to languages we would like to define  $L(\varphi) \subseteq \mathbb{M}$ , even if  $\varphi$  has a free variable. There is however no canonical choice, so we use an existential variant; and we define here:

$$L_{\exists}(\varphi) = \{ t \in \mathbb{M} \mid \exists x \in t : t, x \models \varphi \} = L(\text{M}\varphi).$$

Define  $\text{TL}[\mathcal{C}](\mathbb{M})$  as the Boolean closure of languages defined by  $L_{\exists}(\varphi)$  with  $\varphi \in \text{TL}[\mathcal{C}]$ .

## Languages and Language Operations

We now define some operations on classes of languages that are used to describe the expressive power of logical fragments. Let  $\mathcal{V}$  be a class of trace languages. By  $\mathbb{B}(\mathcal{V})$  we denote the Boolean closure of  $\mathcal{V}$ . A language  $L$  is a *monomial* over  $\mathcal{V}$  of *degree*  $m$  if there exist  $n \leq m$ ,  $a_i \in \Gamma$  and  $L_i \in \mathcal{V}$  with

$$L = L_0 a_1 L_1 \cdots a_n L_n$$

Note that the degree of a monomial is not unique. A finite union of monomials over  $\mathcal{V}$  is called a *polynomial* over  $\mathcal{V}$ . A polynomial has *degree*  $m$  if it can be written as a union of monomials of degree  $m$ . The class of all polynomials over  $\mathcal{V}$  is denoted by  $\text{Pol}(\mathcal{V})$ . The class  $\text{Pol}(\mathcal{V})$  is often called the *polynomial closure* or the closure under product and union of the class  $\mathcal{V}$ . By  $\text{co-Pol}(\mathcal{V})$  we denote the class of languages  $L$  such that  $\mathbb{M} \setminus L \in \text{Pol}(\mathcal{V})$ . If we speak of monomials and polynomials without referring to some class  $\mathcal{V}$  then we mean monomials and polynomials over  $\mathcal{A} = \{ A^* \mid A \subseteq \Gamma \}$ , respectively. In particular,  $\text{Pol} = \text{Pol}(\mathcal{A})$  and  $\text{co-Pol} = \text{co-Pol}(\mathcal{A})$ . For example, if  $A, B \subseteq \Gamma$  then  $A^* B^* \in \text{Pol}$  since

$$A^* B^* = A^* \cup \bigcup_{b \in B} A^* b B^*$$

The class of *star-free* languages  $\text{SF}$  is the closure of the empty set under Boolean operations and polynomials. If  $\mathcal{V}$  is a class of word languages then  $\text{UPol}(\mathcal{V})$  consists of the word languages that are disjoint finite unions of unambiguous monomials. A monomial  $L_0 a_1 L_1 \cdots a_n L_n$  is *unambiguous* if every  $w \in L_0 a_1 L_1 \cdots a_n L_n$  has a unique factorization  $w = w_0 a_0 w_1 \cdots a_n w_n$  with  $w_i \in L_i$ . A similar language operation is  $\mathbb{B}\text{-UPol}$ . By  $\mathbb{B}\text{-UPol}(\mathcal{V})$  we denote the closure of  $\mathcal{V}$  under Boolean