Lecture Notes in Computer Science 7024

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

Roberto Grossi Fabrizio Sebastiani
Fabrizio Silvestri (Eds.)

String Processing and Information Retrieval

18th International Symposium, SPIRE 2011
Pisa, Italy, October 17-21, 2011
Proceedings

Springer

Volume Editors

Roberto Grossi
Università di Pisa
Dipartimento di Informatica
Largo Bruno Pontecorvo 3, 56127 Pisa, Italy
E-mail: grossi@di.unipi.it

Fabrizio Sebastiani
Fabrizio Silvestri
Istituto di Scienza e Tecnologia dell'Informazione
"Alessandro Faedo"
Consiglio Nazionale delle Ricerche
Area della Ricerca di Pisa
Via Giuseppe Moruzzi 1
56124 Pisa, Italy
E-mail:{fabrizio.sebastiani; fabrizio.silvestri}@isti.cnr.it

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-24582-4 e-ISBN 978-3-642-24583-1
DOI 10.1007/978-3-642-24583-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011937624

CR Subject Classification (1998): H.3, J.3, H.2.8, I.5, I.2.7, H.4

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

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In the 18 years since its inauguration back in 1993 the International Symposium on String Processing and Information Retrieval (SPIRE) has become the reference meeting for the interdisciplinary community of researchers whose activity lies at the crossroads of string processing and information retrieval. This volume contains the proceedings of SPIRE 2011, the 18th symposium in the series. The first four events concentrated mainly on string processing, and were held in South America under the title "South American Workshop on String Processing" (WSP) in 1993 (Belo Horizonte, Brazil), 1995 (Valparaiso, Chile), 1996 (Recife, Brazil), and 1997 (Valparaiso, Chile). WSP was renamed SPIRE in 1998 (Santa Cruz, Bolivia) when the scope of the event was broadened to include information retrieval. The change was motivated by the increasing relevance of information retrieval and its close interrelationship with the general area of string processing. From 1999 to 2007, the venue of SPIRE alternated between South / Latin America (odd years) and Europe (even years), with Cancun, Mexico in 1999; A Coruña, Spain in 2000; Laguna de San Rafael, Chile in 2001; Lisbon, Portugal in 2002; Manaus, Brazil in 2003; Padova, Italy in 2004; Buenos Aires, Argentina in 2005; Glasgow, UK in 2006; and Santiago, Chile in 2007. This pattern was broken when SPIRE 2008 was held in Melbourne, Australia, but it was restarted in 2009 when the venue was Saariselkä, Finland, followed by Los Cabos, Mexico in 2010.

The SPIRE 2011 call for papers resulted in the submission of 102 papers. Each submitted paper was reviewed by three of the 64 members of the Program Committee, who eventually engaged in discussions coordinated by the two PC Chairmen in cases of lack of consensus. We believe this resulted in a very accurate selection of the truly best submitted papers. As a result, 30 long papers and 10 short papers were accepted and have been published in these proceedings.

The dense program of SPIRE 2011 started on October 17 with four tutorials providing in-depth coverage of both introductory as well as advanced topics in string processing ("Introduction to Sequence Learning", by Corinna Cortes and Mehryar Mohri, and "Space-Efficient Data Structure", by Francisco Claude and Gonzalo Navarro) and information retrieval ("Introduction to Web Retrieval" by Ricardo Baeza-Yates, and "Computational Geography", by Vanessa Murdock and Gary Gale). The main conference featured keynote speeches by Erik Demaine and Abdur Chowdhury, plus the presentations of the 30 full papers and 10 short papers. Following the main conference, on October 21, SPIRE 2011 hosted two workshops, i.e., the Workshop on the Algorithmic Analysis of Biological Data (WAABD 2011) and the Workshop on Compression, Text, and Algorithms (WCTA 2011). A Best Paper Award and a Best Student Paper Award were also assigned, each consisting of a check of 1000 EUR and sponsored by Google and NoemaLife, respectively.

We would like to take the opportunity to thank Google, NoemaLife, Microsoft Research, the Department for Information and Communication Technologies of the Italian National Council of Research, the Italian Association for Automatic Computation (AICA), Yahoo! Research, Twitter, and the ASSETS project. All of them provided generous sponsorship, which allowed the organizers to keep the registration fees as low as possible and thus to enhance participation.

We would also like to thank everybody involved in making SPIRE 2011 such an exciting event. Specifically, we would like to thank all conference, tutorial, and workshop participants and presenters, who provided a fascinating one-week program of high-quality presentations and intensive discussions. Thanks also to all the members of the Program Committee and to the additional reviewers, who went to great lengths to ensure the high quality of this conference, and to the coordinator of the SPIRE Steering Committee, Ricardo Baeza-Yates, who provided assistance and guidance in the organization.

Furthermore, we would like to thank all the members of the local organizing team at the Italian National Council of Research and at the University of Pisa. Particularly, we would like to thank Andrea Esuli who acted as Tutorials Chair, Nadia Pisanti who acted as Workshops Chair, our Webmaster Stefano Baccianella, Catherine Bosio and Giulio Galesi who gave us support in local arrangements, Beatrice Rapisarda who designed the official poster of the symposium, and all the student volunteers. They all made a tremendous effort to make sure that this event was exciting and enjoyable. It is due to them that the organization of SPIRE 2011 was not just hard work, but also a pleasure.

October 2011

Roberto Grossi
Fabrizio Sebastiani
Fabrizio Silvestri

Organization

Program Committee

Omar Alonso	Microsoft
Gianni Amati	Fondazione Ugo Bordoni
Amihood Amir	Bar-Ilan University and Johns Hopkins University
Leif Azzopardi	University of Glasgow
Rolf Backofen	Albert-Ludwigs-University Freiburg
Ricardo Baeza-Yates	Yahoo! Research
Alvaro Barreiro	University of A Coruña
Philip Bille	Technical University of Denmark
Paolo Boldi	Università degli Studi di Milano
Danny Breslauer	University of Haifa
Edgar Chavez	Universidad Michoacana
Charles Clarke	University of Waterloo
Maxime Crochemore	King's College London and Université Paris
Brian Davison	Lehigh University
Nadia El-Mabrouk	University of Montreal
Paolo Ferragina	University of Pisa
Frantisek Franek	McMaster University
Leszek Gasieniec	University of Liverpool
Dora Giammarresi	University of Rome "Tor Vergata"
Nazli Goharian	Georgetown University
Gregory Grefenstette	Exalead
Roberto Grossi	Università di Pisa
Concettina Guerra	University of Padova and Georgia Tech
Antonio Gulli	Microsoft
Jan Holub	Czech Technical University in Prague
Heikki Hyyrö	University of Tampere
Lucian Ilie	University of Western Ontario
Costas Iliopoulos	King's College London
Shunsuke Inenaga	Kyushu University
Shen Jialie	Singapore Management University
Jaap Kamps	University of Amsterdam
Takuya Kida	Hokkaido University
Marcin Kubica	Warsaw University
Gregory Kucherov	CNRS/LIGM
Mounia Lalmas	University of Glasgow
Moshe Lewenstein	Bar Ilan University
Alistair Moffat	University of Melbourne
Laurent Mouchard	University of Rouen

Gonzalo Navarro	University of Chile
Wolfgang Nejdl	L3S and University of Hannover
Iadh Ounis	University of Glasgow
Laxmi Parida	IBM Research
Kunsoo Park	Seoul National University
Marco Pellegrini	Institute for Informatics and Telematics of C.N.R.
Pierre Peterlongo	INRIA Rennes-Bretagne-Atlantique
Andrea Pietracaprina	University of Padova
Ely Porat	Bar-Ilan University
Venkatesh Raman	The Institute of Mathematical Sciences
Horacio Rodriguez	Universitat Politècnica de Catalunya
Marie-France Sagot	Université de Lyon
Cenk Sahinalp	Simon Fraser University
Leena Salmela	University of Helsinki
Jeanette Schmidt	Stanford University
Fabrizio Sebastiani	ISTI - CNR
Fabrizio Silvestri	ISTI - CNR
Steven Skiena	Stony Brook University
Dina Sokol	Brooklyn College of the City University of New York
Jens Stoye	Bielefeld University
Torsten Suel	Yahoo! Research
Fabio Vandin	Brown University
Stéphane Vialette	Université Paris-Est
Alain Viari	INRIA
Jeff Vitter	University of Kansas
Oren Weimann	Weizmann Institute of Science
Le Zhao	CMU
Nivio Ziviani	Federal University of Minas Gerais

Additional Reviewers

Andonov, Rumen	Dan, Ovidiu
Antoniou, Pavlos	David, Julien
Arroyuelo, Diego	Epifanio, Chiara
Badkobeh, Golnaz	Fernandes, David
Bernhardt, Daniel	Fertin, Guillaume
Blanco, Roi	Fonseca, Paulo
Bressan, Marco	Franek, Frantisek
Canovas, Rodrigo	Frigeri, Achille
Ceccarelli, Diego	Galle, Matthias
Chikhi, Rayan	Gerlach, Wolfgang
Claude, Francisco	Giraud, Mathieu
Constant, Matthieu	Gurevich, Maxim
Costa, Fabrizio	Hamel, Sylvie
Dai, Na	Hegerty, Ian

Husemann, Peter
Jahn, Katharina
Jaroš, Jakub
Jiang, Minghui
Karenos, Kyriakos
Kopelowitz, Tsvi
Levy, Avivit
Lonati, Violetta
Losada, David
Manzini, Giovanni
Markowetz, Alexander
Martinez-Prieto, Miguel A.
Menezes, Guilherme
Möhl, Mathias
Nanni, Mirco
Nardini, Franco Maria
Noe, Laurent
Peterlongo, Pierre
Pinkas, Benny
Pissis, Solon
Prochazka, Petr
Puglisi, Simon
Qi, Xiaoguang
Radoszewski, Jakub
Rojas, Pablo
Rosone, Giovanna
Russo, Luis M.S.
Santos, Rodrygo
Satti, Srinivasa Rao
Schmidt, Jeanette
Silva, Altigran
Silvestri, Francesco
Starikovskaya, Tatiana
Tannier, Eric
Tischler, German
Tolomei, Gabriele
Utro, Filippo
Veloso, Adriano
Venturini, Rossano
Vigna, Sebastiano
Vildhøj, Hjalte Wedel
Walen, Tomasz
Will, Sebastian
Wittler, Roland
Xu, Bojian
Xue, Zhenzhen
Yan, Hao
Yin, Dawei
Yorukoglu, Deniz
Zelikovitz, Sarah

Table of Contents