

Lecture Notes in Computer Science

730

Edited by G. Goos and J. Hartmanis

Advisory Board: W. Brauer D. Gries J. Stoer



David B. Lomet (Ed.)

Foundations of Data Organization and Algorithms

4th International Conference, FODO '93
Chicago, Illinois, USA, October 13-15, 1993
Proceedings

Springer-Verlag

Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Series Editors

Gerhard Goos
Universität Karlsruhe
Postfach 69 80
Vincenz-Priessnitz-Straße 1
D-76131 Karlsruhe, Germany

Juris Hartmanis
Cornell University
Department of Computer Science
4130 Upson Hall
Ithaca, NY 14853, USA

Volume Editor

David B. Lomet
Digital Equipment Corporation, Cambridge Research Lab
One Kendall Square, Building 700, Cambridge, MA 02139, USA

CR Subject Classification (1991): E.1-2, F.2.2, H.2-5

ISBN 3-540-57301-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-57301-1 Springer-Verlag New York Berlin Heidelberg

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-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1993
Printed in Germany

Typesetting: Camera-ready by author
Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
45/3140-543210 - Printed on acid-free paper

External Referees

Berndt Amann	INRIA, France
Masatoshi Arikawa	Kyoto University, Japan
Fritz Augenstein	University of Freiburg, Germany
Guangyi Bai	Kyushu University, Japan
Anthony Berglas	University of Queensland, Australia
Andrew Black	Digital Equipment Corp., USA
Yuri Breitbart	University of Kentucky, USA
Jae-Woo Chang	Chun-Pook University, Korea
Andrew E. Deacon	ETH Zurich, Switzerland
Sam DeFazio	Digital Equipment Corp., USA
David DeWitt	University of Wisconsin, USA
Gisbert Droege	ETH Zurich, Switzerland
Juergen Eckerle	University of Freiburg, Germany
Doo-Hun Eum	Duk-Sung Women's University, Korea
Cathy Ewald	University of Queensland, Australia
Jay Feenan	Digital Equipment Corp., USA
Tetsuya Furukawa	Kyushu University, Japan
Stephane Grumbach	INRIA France
Alois Heinz	University of Freiburg, Germany
Edmund Ihler	University of Freiburg, Germany
Yannis Ioannidis	University of Wisconsin
Bin Jiang	Union Bank of Switzerland
Yoki Kusumi	Matsushita Electric, Ltd., Japan
Steve Langdon	Digital Equipment Corp., USA
Dick Loveland	Digital Equipment Corp., USA
Xuemin Lin	University of Queensland, Australia
Bo-jiang Liu	Osaka University, Japan
Akifumi Makinouchi	Kyushu University, Japan
Moira Norrie	ETH Zurich, Switzerland
Beng-Chin Ooi	National University of Singapore
Young-Chul Park	Kyung-Pook University, Korea
Werner Schaad	ETH Zurich, Switzerland
Craig Schaffert	Digital Equipment Corp., USA
Juergen Schoening	University of Freiburg, Germany
Stefan Schroedl	University of Freiburg, Germany
Sven Schuierer	University of Freiburg, Germany
David Shertleff	Digital Equipment Corp., USA
Amit P. Sheth	Bellcore, USA
Ian Smith	Digital Equipment Corp., USA
Kazutoshi Sumiya	Matsushita Electric, Ltd., Japan
Wolfgang Weck	ETH Zurich, Switzerland
Peter Zabback	ETH Zurich, Switzerland
Yanchun Zhang	University of Queensland, Australia
Xiaofang Zhou	University of Queensland, Australia

Message of the General Chairman

It has been my great pleasure to organize The Fourth International Conference on Data Organization and Algorithms (FODO'93) in Evanston. FODO has from its inception been an international conference. First in Warsaw (1981), then in Kyoto (1985), and Paris (1989), the conference has brought researchers together to share ideas and friendship. The 1981 conference concentrated on the consecutive retrieval property; the following meetings broadened the focus to include general access methods. FODO'93 reflects the maturing of the database field which has been driven by the enormous growth in the range of applications for database systems. The "non-standard" applications of the not-so-distant past, such as hypertext, multimedia, scientific, and engineering databases, now provide some of the central motivation for the advances in hardware technology and data organization and algorithms.

We were very fortunate to have a group of dedicated people who have worked enthusiastically to make this conference happen. I would like to thank first the steering committee, and Yahiko Kambayashi in particular, for keeping the spirit of FODO alive. My sincere appreciation goes to David Lomet and Katsumi Tanaka, the Program Chairs, and their Program Committees for an outstanding job in selecting high quality papers. Thanks too, to Edward Omiecinski for publicity, to Kathi Davis for registration, Aris Ouksel for finances, Chris Clifton for local arrangements, and to our international coordinators.

The proceedings bring together twenty-two contributed papers as well as invited talks by David Maier, Yahiko Kambayashi and C. Mohan. The talks by Maier and Kambayashi present stimulating ideas on data organizations for new hardware technology, specifically continuous-media and flash memory. The talk by Mohan presents an integrated view of many of the implementation issues of large databases. I thank them for their fine contributions. Additional highlights are the panels, "Highways and Jukeboxes" and "Scientific Databases". I thank Margaret Murphy, our panel chair, for her help in organizing these. The conference has been preceded by tutorials on "Mobile Computing and Database Problems" by Tomasz Imielinski and "Audio/Video Databases" by Simon Gibbs and Christian Breiteneder. Tutorials and panels complement the paper sessions by presenting challenges for data organization and algorithms in new domains.

I would like to express my gratitude to the industrial and university sponsors without whom FODO'93 could not have happened: DEC, HP, Kyoto University, Matsushita and Northwestern University.

I hope that FODO will continue to serve the database community as a conference at the forefront between emerging technology and fundamental data organizations and algorithms.

Peter Scheuermann
FODO'93 General Chair

Message from the Program Committee Chairs

FODO'93 continues the FODO tradition of focusing on the data organizations and how systems use them to support data management. This year's call-for-papers explicitly emphasized not only traditional concerns but also called attention to physical storage and new application areas. The number of submissions and the quality of the accepted papers document the progress within the FODO scope, including these new areas.

The 1993 FODO program is a strong one. This reflects well on the authors of the papers. FODO has successfully attracted strong participants in the data organizations and algorithms community. The review process scoring was based on the standards of the VLDB and SIGMOD conferences. The scores for the accepted FODO papers are comparable to those of these pre-eminent database conferences. Some of the papers received outstanding scores. We wish to thank all authors for their submissions as it is the quality of the technical program that determines the value of a conference.

The strength of this year's program is a tribute to the program committees. The ability of conferences like FODO to attract quality submissions depends on the quality and credibility of the program committee. This year's FODO committees were outstanding. It is a pleasure to report also on the excellent cooperation and communication between the committees. We wish to thank the distinguished database experts who served on the FODO program committees for their efforts. The quality of the submissions made for tough choices.

The technical program is both focused on data organizations and quite diverse in coverage. The session on access methods applies these methods to new forms of data and system configurations. The text retrieval section contrasts signature and index trie methods. Multimedia considerations are addressed from the user interface level to data storage. The physical storage session deals directly with disk storage considerations. The query processing session emphasizes data organization and management aspects. A "new directions" session captures some of the more recent trends, OO, hypertext, multiversioned objects. The industrial papers session is of special interest, with two papers describing work that commercial enterprises expect to impact the marketplace. We hope you learn from and enjoy this fine program.

Our message would be incomplete without further acknowledgements. Thanks particularly to conference chair Peter Scheuermann, whose vision and determination brought FODO'93 to fruition. Peter deserves credit also for recruiting our outstanding invited speakers. Thanks also to Michelle Gillespie of DEC's Cambridge Research Lab, who formatted the proceedings front material and handled PC correspondence.

David Lomet and Katsumi Tanaka
FODO'93 Program Co-chairs

FODO'93 Conference

General Chairman: Peter Scheuermann
Northwestern University

Steering Committee: Claude Delobel, Altair, France
Sakti Ghosh, IBM, USA
Yahiko Kambayashi, Kyoto University, Japan

ORGANIZING COMMITTEE:

American Coordinator: Clement Yu, University of Illinois at Chicago, USA
Asian Coordinator: Yoshifumi Masunaga, ULIS, Japan
European Coordinator: Gerhard Weikum, ETH, Switzerland
Treasurer: M. Aris Ouksef, University of Illinois at Chicago, USA
Publicity: Edward Omiecinski, Georgia Institute of Technology, USA
Local Arrangements: Chris Clifton, Northwestern University, USA
Registration: Kathy Hogshead Davis, Northern Illinois University, USA

SPONSORED BY: Northwestern University

IN COOPERATION WITH:

Association for Computing Machinery (ACM)
Special Interest Group on Management of Data (SIGMOD)
Institute of Electrical and Electronic Engineers (IEEE)
Technical Committee on Data Engineering (TCDE)

SUPPORTED BY:

Digital Equipment Corporation
Hewlett-Packard Labs
Kyoto University
Matsushita Electric, Ltd.

FODO'93 Program Committee

America/Europe Committee

David Lomet (chair)
DEC Cambridge Research Lab
USA

Stavros Christodoulakis
Technical University of Crete
GREECE

Jim Gray
DEC San Francisco Systems Center
USA

Laura Haas
IBM Almaden Research Center
USA

Rivka Ladin
DEC Cambridge Research Lab
USA

Witold Litwin
University of Paris
FRANCE

J. Eliot Moss
University of Massachusetts, Amherst
USA

Frank Olken
Lawrence Berkeley Laboratory
USA

Thomas Ottmann
University of Freiburg
GERMANY

M.V. Ramakrishna
Michigan State University
USA

Betty Salzberg
Northeastern University
USA

Hans Schek
ETH Zurich
SWITZERLAND

Asia/Pacific Committee

Katsumi Tanaka (chair)
Kobe University
JAPAN

David Abel
CSIRO
AUSTRALIA

Hideto Ikeda
Hiroshima University
JAPAN

Yoshihiko Imai
Matsushita Electric, Ltd.
JAPAN

Yasushi Kiyoki
Tsukuba University
JAPAN

Hongjun Lu
National University of Singapore
SINGAPORE

Akifumi Makinouchi
Kyushu University
JAPAN

Shojiro Nishio
Osaka University
JAPAN

A. Ohori
OKI Electric
JAPAN

Maria E. Orlowska
University of Queensland
AUSTRALIA

Kyu-Young Whang
KAIST
KOREA

Masatoshi Yoshikawa
Kyoto Sangyo University
JAPAN

Table of Contents

Invited Talk

Storage System Architectures for Continuous Media Data <i>David Maier, Jonathan Walpole and Richard Staehli</i>	1
--	---

Session: Multimedia

Physical Storage Organizations for Time-Dependent Multimedia Data <i>Huang-Jen Chen and T.D.C. Little</i>	19
User Interface Management System Embedded in a Multimedia Document Editor Framework <i>Takashi Ohtsu and Michael A. Harrison</i>	35
Management of Physical Replicas in Parallel Multimedia Information Systems <i>Shahram Ghandeharizadeh and Cyrus Shahabi</i>	51

Session: Access Methods

Efficient Similarity Search in Sequence Databases <i>Rakesh Agrawal, Christos Faloutsos and Arun Swami</i>	69
Filtered Hashing <i>Ilsoo Ahn</i>	85
Design and Implementation of DDH: A Distributed Dynamic Hashing Algorithm <i>Robert Devine</i>	101

Session: Text Processing

Performance Analysis of Superimposing-Coded Signature Files <i>Sam Y. Sung</i>	115
Trie Methods for Representing Text <i>T. H. Merrett and Heping Shang</i>	130
Estimation of False Drops in Set-valued Object Retrieval with Signature Files <i>Hiroyuki Kitagawa, Yoshiaki Fukushima, Yoshiharu Ishikawa and Nobuo Ohbo</i>	146

Invited Talk

Data Structure and Algorithms for New Hardware Technology <i>Yahiko Kambayashi, Hiroki Takakura and Shintaro Meki</i>	164
--	-----

Session: Query Processing

Evaluation of Upper Bounds and Least Nodes as Database Operations <i>Joachim Biskup and Holger Stiefeling</i>	197
A C++ Binding for Penguin: a System for Data Sharing among Heterogeneous Object Models <i>Arthur M. Keller and Catherine Hamon</i>	215
Dynamic Selectivity Estimation for Multidimensional Queries <i>William I. Grosky, Junping Sun and Farshad Fotouhi</i>	231
Reducing Page Thrashing in Recursive Query Processing <i>Rakesh Agrawal and Jerry Kiernan</i>	247

Panel

Highways and Jukeboxes: A Revolution for Data Structures?264
Witold Litwin (chair), Tomasz Imielinski, David Lomet and Marie-Ann Neimat

Session: Industrial Papers

The Sybase Replication Server Project
Yongdon Wang, Jane Chiao and Vasu Nori265
Analogical Inferred Compensation Method for Heterogeneous Database Schema
Mismatch and its Application to Hypermedia Cooperative Work System
Shigeru Shimada, Toshihisa Aoshima and Tetsuzo Uehara271

Invited Talk

A Survey of DBMS Research Issues in Supporting Very Large Tables
C. Mohan279

Session: Physical Storage

Partition-Based Clustering in Object Bases: From Theory to Practice
Carsten Gerlhof, Alfons Kemper, Christoph Kilger, Guido Moerkotte301
An Algorithm for the Implementation of Replicated Tables
Brahma Dathan317
Improved Traditional Mirror
Cyril U. Orji, Mark A. Weiss and Jon A. Solworth329
Adaptive Load Balancing in Disk Arrays
Peter Scheuermann, Gerhard Weikum and Peter Zabback345

Panel

Scientific Databases: Challenges to the Database Community.....361
*Robert Grossman (chair), Yannis Ioannidis, Edward May, Sakti Pramanik
and Ruth Pordes*

Session: New Directions

An Object-Oriented Office Space Description Model and an Office View
Management Mechanism for Distributed Office Environment
Hideyuki Takada and Yahiko Kambayashi.....362
Stamp Locking Method for Multiversion Composite Objects
Wojciech Cellary and Waldemar Wiczerzycki378
Queries on Structures in Hypertext
Tatsuo Minohara, Ryuichi Watanabe and Mario Tokoro394

Author Index:412