

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

339

M. Rafanelli J. C. Klensin
P. Svensson (Eds.)

Statistical and Scientific Database Management

Fourth International Working Conference SSDBM
Rome, Italy, June 21–23, 1988
Proceedings



Springer-Verlag

Berlin Heidelberg New York London Paris Tokyo

Editorial Board

D. Barstow W. Brauer P. Brinch Hansen D. Gries D. Luckham
C. Moler A. Pnueli G. Seegmüller J. Stoer N. Wirth

Editors

Maurizio Rafanelli
I.A.S.I. – CNR
Viale Manzoni, 30, I-00185 Roma, Italy

John C. Klensin
Massachusetts Institute of Technology,
Room 20A–226
77 Massachusetts Avenue, Cambridge, MA 02139, USA

Per Svensson
Swedish Defence Research Establishment
S-10254 Stockholm, Sweden

CR Subject Classification (1987): H, G.3, I.3.5, J.2–3

ISBN 3-540-50575-X Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-50575-X 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 other ways, and storage in data banks. Duplication of this publication or parts thereof is only permitted under the provisions of the German Copyright Law of September 9, 1965, in its version of June 24, 1985, and a copyright fee must always be paid. Violations fall under the prosecution act of the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1989
Printed in Germany

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2145/3140-543210

FOREWORD

The Fourth International Working Conference on Statistical and Scientific Data Base Management (IV SSDBM) held on June 21 - 23, 1988 in Rome, Italy, continued the series of conferences initiated in California in December 1981 by the Lawrence Berkeley Laboratory (LBL) of the University of California. The purpose of this conference was to bring together database researchers, users and system builders, working in this specific field, to discuss the particular points of interest, to propose new solutions to the problems of the domain and to expand the topics of the previous conferences, both from the theoretical and from the applicational point of view.

Much to our pleasure we succeeded in transforming the workshop into a working conference in order to provide an environment in which those who are interested in this field may freely exchange their ideas. At the same time we intend to select the contributions very carefully in order to obtain a better quality and scientific respectability.

Why "working conference" ?

"Working" : If on the one hand the transformation of the "working groups" into "open panels" requested a greater effort of preparation, on the other hand more qualified and *participating* people joined the open panels. In each of the three scheduled panels a chair discussed the problems regarding a selected topic (presenting a paper included in the proceedings) and gave proposals and suggestions about the possible solutions. Subsequently, he opened a discussion with some panelists previously chosen from the experts of the area and also with the audience in the conference hall.

"Conference" : A rigorous selection has been made of the submitted papers. This fact requested a stronger engagement from all: each paper has been referred by three to six members (we have sent papers of European authors to three overseas members and viceversa, and, to obtain a more uniform and impartial evaluation, a subset of the papers were sent both to three European and to three overseas members).

The structure of this conference, based on the experience of the previous events, included four scientific sessions in which eleven selected scientific papers were presented and discussed by the authors. These papers dealt with the following topics: knowledge base and expert system, data model, natural language processing, query language, time performance, user interface, heterogeneous data classification, storage constraints, automatic drawing, ranges and trackers, and arithmetic coding.

Also two special sessions had been scheduled, in which six other scientific papers, containing important points of originality, were presented and discussed as work in progress. In particular, the topics of these papers are geographical data modeling, spatial database queries, user interface in an Object Oriented SDB, interpretation of queries, graphical query languages, and knowledge browsing - front ends.

The following paper by Maurizio Rafanelli provides both an introduction on the general concepts helpful to people outside the field and a survey of all the papers in these proceedings.

The conference also had three invited papers regarding topics of particular interest such as "Temporal Data", "Statistical Data Management Requirements" and "Knowledge Based Decision Support Systems", included in this volume.

Furthermore, as already said, in three open panels relevant topics have been discussed regarding statistical relational model, geographical database systems and statistical information, and database management systems for statistical and scientific applications. Papers by the chairmen, contributions by the panelists and a summary of the respective discussions which followed the presentations of the chairmen are included in this volume, too.

As General Chairman and Program Chairmen of the conference, we wish to thank the Istituto di Analisi dei Sistemi ed Informatica for the financial and organizational support given to the conference. Our thanks also go to Arie Shoshani for his valuable suggestions, as well as to Roger Cubitt for his useful advice and for the immediate (and fundamental) financial support which proved to be essential for the start of the conference. At last, we are also grateful to Andrew Westlake for the circulation of the final program in the Statistical Software Newsletters journal. Nearly one thousand people in more than seventy countries of the world were informed.

We would like to extend our thanks to all the authors who submitted papers for consideration; we are particularly pleased by the fact that among them there are researchers from countries which had not been represented in previous workshops.

We repeat our thanks to all the Program Committee Members (R.A. Becker, F.Y. Chin, R.E. Cubitt, H. Hinterberger, N.C. Lauro, H. Lutz, F.M. Malvestuto, M. McLeish, Z. Michalewicz, R.J. Muller, S.B. Navathe, S. Nordbotten, F. Olken, G. Ozsoyoglu, F.L. Ricci, H. Sato, A. Shoshani, B. Sundgren, Y. Vassiliou and K.M. Wittkowski) and the other referees (T. Carey, K.W. Chen, R. Dawson, D. Defays, G. Di Battista, P. Di Felice, G. Gambosi, S. Knudsen, R. Marti, J.C. Nordbotten, H. Pilat, D. Rotem, T. Saydam, A. Segev, M. Theodoridou, Y. Yeorgavoudakis, R. Zahir) for their effective and efficient work.

At last, we would like to thank all the sponsors of the conference, that is the Universities of Rome "La Sapienza" and of Naples, ISTAT (Central Institute of Statistics), ENEA (National Committee for Research and Development of Nuclear Energy and Alternative Sources), IASC (International Association for Statistical Computing), the Swedish Defense Research Establishment, EUROSTAT (Commission of the European Communities Statistics Office) and IBM.

Our thanks also go to ACM (association for Computing Machinery) and to IEEE (the Institute of Electrical and Electronics Engineers - Middle and South Italian Section) for their cooperation and support, and the Italian Society of Statistics and the Ministry of Scientific and Technological Research for the high patronage given to the conference.

On a more personal level we would like to thank Nella Ricci for the most valuable cooperation given both in the conference organization and to us personally.

John C. Klensin
Co-chairman

Maurizio Rafanelli
General Chairman

Per Svensson
Co-chairman

CONTENTS

Research Topics in Statistical and Scientific Database Management: The IV SSDBM Maurizio RAFANELLI	1
---	---

INVITED LECTURES

Statistical Data Management Requirements and the SQL Standards - An Evolving Comparison John C. KLENSIN and Roselyn M> ROMBERG	19
The Representation of a Temporal Data Model in the Relational Environment Arie SEGEV and Arie SHOSHANI	39
Knowledge Based Support for the Management of Statistical Databases Knut M. WITTKOWSKI	62

CONTRIBUTED PAPERS

A Study of Modified Interpolation Search in Compressed, Fully Transposed, Ordered Files Magnus ANDERSSON and Per SVENSSON	72
Design of Statistical Information Media: Time Performances and Storage Constraints Giulio BARCAROLI, Giuseppe DI BATTISTA, Ennio FORTUNATO and Claudio LEPORELLI	93
Handling Time in Query Languages Mostafa A. BASSIOUNI and M. LLEWELLYN	105
Software and Hardware Enhancement of Arithmetic Coding Mostafa A. BASSIOUNI, N. RANGANATHAN and Amar MUKHERJEE . .	120
Orthogonal Range Retrieval using Bucket Address Hashing C.C. CHANG and C.Y. CHEN	133

Automatic Drawing of Statistical Diagrams Giuseppe DI BATTISTA	141
The Classification Problem with Semantically Heterogeneous Data Francesco M. MALVESTUTO and Carla ZUFFADA	157
Database Issues for a Veterinary Medical Expert System Mary McLEISH, Matthews CECILE and Alex LOPEZ-SUAREZ	177
Ranges and Trackers in Statistical Databases Zbigniew MICHAŁEWICZ and Keh-Wei CHEN	193
A Data Model, Knowledge Base, and Natural Language Processing for Sharing a Large Statistical Database Hidetō SATO	207
Precision-Time Tradeoffs: A Paradigm for Processing Statistical Queries on Databases Jaideep SRIVASTAVA and Doron ROTEM	226
 <i>WORK IN PROGRESS PAPERS</i>	
Interpretation of Statistical Queries to Relational Databases Alessandro D'ATRI and Fabrizio L. RICCI	246
GQL, a Graphical Query Language for Semantic Databases Manoochehr AZMOODEH and Hongbo DU	259
A Conceptual Model for the Representation of Statistical Data in Geographical Information Systems Giorgio GAMBOSI, Enrico NARDELLI and Maurizio TALAMO	278
Spatial Data Base Queries: Relational Algebra Versus Computational Geometry Robert LAURINI and Françoise MILLERET	291
Design of the User-interface for an Object-oriented Statistical Database Erik MALMBORG	314

Knowledge Browsing - Front Ends to Statistical Databases

Geoffrey A. STEPHENSON327

OPEN PANELS

Statistical Relational Model

Sakti P. GHOSH (chairman)338

Panelists: Lawrence McNAMEE, Francesco MALVESTUTO, Knut
WITTKOWSKI*Contribution paper:* A Model of Summary Data and its Applications
in Statistical DatabasesMeng Chang CHEN, Lawrence McNAMEE, Michel
MELKANOFF356

Database Management Systems for Statistical and Scientific Applications:

Are Commercially Available DBMS Good Enough?

Per Svensson (chairman)388

Panelists: Martin PODEHL, Geoffrey STEPHENSON, Michael
CAWSON

Geographical Database Systems and Statistical Information

Andrew WESTLAKE (chairman)420

Panelists: Giorgio GAMBOSI, David RHIND*Contribution paper:* The Analysis of Geographical Data: Data Rich,
Technology Adequate, Theory Poor

David RHIND, Stan OPENSHAW, Nick GREEN427