Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

420

Z. Michalewicz (Ed.)

Statistical and Scientific Database Management

Fifth International Conference, V SSDBM
Charlotte, N.C., USA, April 3–5, 1990
Proceedings

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo Hong Kong

Editor
Zbigniew Michalewicz
Department of Computer Science
The University of North Carolina at Charlotte
Charlotte, N. C. 28223, USA

CR Subject Classification (1987): H, G.3, I.3.5, J.2, E.2, E.5

ISBN 3-540-52342-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-52342-1 Springer-Verlag New York Berlin Heidelberg

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2145/3140-543210 – Printed on acid-free paper

PREFACE

This volume contains papers which have been selected for presentation at the Fifth International Conference on Statistical and Scientific Databases (V SSDBM), held in Charlotte, North Carolina, April 3–5, 1990. The Conference was hosted by UNC-Charlotte and sponsored by: UNC-Charlotte; National Aeronautics and Space Administration; Statistical Office of the European Communities—EUROSTAT (Luxembourg); Ente Nazionale Energie Alternative (Italy); Statistics Sweden; Microelectronics Center of North Carolina; International Association for Statistical Computing; Department of Statistics (New Zealand); and Istituto di Analisi dei Sistemi ed Informatica del CNR (Italy).

The Conference has been organized by the Department of Computer Science of the University of North Carolina at Charlotte with the following Organizing Committee: K.-W. Chen (UNC-C); S. Chen (UNC-C); L. Groves (VUW-Wellington); R. Lejk (UNC-C); Z. Michalewicz (Conference Chairman, UNC-C); Z. Ras (UNC-C); and M. Zemankova (NSF and UT-Knoxville).

The V SSDBM continued the series of conferences started nine years ago in California (1981 and 1983), then in Europe (Luxembourg, 1986, and Rome, 1988). The purpose of this conference was to bring together database researchers, users, and system builders, working in this specific area of activity, to discuss the particular issues of interest, to propose new solutions to the problems of the area, and to extend the themes of the previous conferences, both from the theoretical and from the application point of view.

The contributed papers have been selected by the following Program Committee: S. Chen (UNC-C); R. Cubitt (EUROSTAT, Luxembourg); S. Ghosh (IBM, San Jose); H. Hinterberger (Program Co-chairman, ETH Zurich, Switzerland); J. Klensin (MIT); F. Malvestuto (ENEA, Italy); M. McLeish (Guelph, Canada); Z. Michalewicz (UNC-C); G. Ozsoyoglu (Case Western Reserve University); M. Rafanelli (IASI-CNR, Italy); D. Rotem (Program Co-chairman, Lawrence Berkeley Laboratory); Y. Sato (Tokyo, Japan); A. Shoshani (Lawrence Berkeley Laboratory); B. Sundgren (Stockholm, Sweden); P. Svenson (Stockholm, Sweden); and A. Westlake (University of London, UK). All accepted papers were reviewed by at least three, and, in many cases, by seven members of the Program Committee. Only 13 papers were accepted for presentation. These papers cover a wide area of research for statistical and scientific databases: object oriented database systems, semantic modelling, deductive mathematical databases, security of statistical databases, implementational issues for scientific databases, temporal summary table management, graphical and visual interfaces for statistical databases, query optimization, distributed databases, and economic and geographical databases.

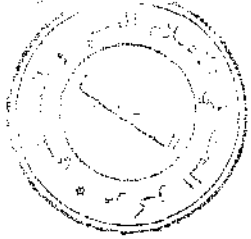
There were also two panel sessions: the first one on *Scientific Data Management for Human Genome Applications*, chaired by A. Shoshani (Lawrence Berkeley Laboratory), the other one on *Expert Statistical Systems*, chaired by Roger Cubitt (EUROSTAT). Additionally, there was a presentation *On the NSF Scientific Database Workshop* by James C. French, Anita K. Jones, and John L. Pfaltz (Institute of Parallel Computation). The reports from the panels and the presentation will appear in the IEEE Data Engineering Bulletin.

Finally, I wish to express my thanks to Won Kim (Microelectronics and Computer Corporation); Doron Rotem (Lawrence Berkeley Laboratory); Andrew Westlake (University of London); and Maria Zemankova (NSF and UT-Knoxville), who gave invited talks at the V SSDBM. I would like also to express my appreciation to all V SSDBM sponsors, to all who submitted papers for presentation at the conference and in this proceedings, to the V SSDBM Organizing Committee, and to all of those who contributed to the symposium program—in particular to Doron Rotem (Lawrence Berkeley

Laboratory) and Hans Hinterberger (ETH Zurich) who served as program co-chairmen. In addition, I would like to thank M. Hadzikadic (UNC-C), F. Olken (Lawrence Berkeley Laboratory), L. M. Pfau (ETH Zurich, Switzerland) and A. Segev (Lawrence Berkeley Laboratory) for their help in refereeing papers. I would like to acknowledge the many useful comments I received from Arie Shoshani (Lawrence Berkeley Laboratory), Roger Cubitt (EUROSTAT), Zbigniew Ras (UNC-C), and John Klensin (MIT). My further thanks go to Hans Wössner at Springer-Verlag, without whose help the present volume could not have been completed. On a more personal level, my special thanks goes to my wife Eva for her enthusiasm and help in organizing the conference.

Charlotte, N.C., November 20th, 1989

Zbigniew Michalewicz
V SSDBM Chairman



CONTENTS

OBJECT-ORIENTED APPROACH TO MANAGING STATISTICAL AND SCIENTIFIC DATABASES

Won Kim

Microelectronics and Computer Technology Corporation
3500 West Balcones Center Drive
Austin, Texas 78759

The surging interest in object-oriented programming has given rise to a flurry of recent research and development efforts in object-oriented databases. In this paper we discuss capabilities and limitations of object-oriented database systems that support a set of core object-oriented modeling concepts, and illustrate those aspects of core object-oriented database systems that are particularly appropriate for managing large statistical and scientific databases.

1. INTRODUCTION

The rapidly increasing use of object-oriented concepts in various components of software technology has naturally necessitated object-oriented database systems. An object-oriented database system is a persistent and sharable repository and manager of an object-oriented database; an object-oriented database is a collection of objects defined by an object-oriented data model, that is, objects that capture the semantics of objects supported in object-oriented programming. Object-oriented concepts form a good basis for a rich data model for the next-generation database applications, such as computer-aided design and engineering systems, artificial intelligence, office information systems, and advanced human interface systems. The reason is that object-oriented concepts already include data modeling concepts found in conventional database languages, such as grouping objects into a class (corresponding to grouping records into a record type, or tuples into a relation), aggregation relationships between an object and objects it consists of (nested objects), and generalization relationships between a class of objects and classes of objects specialized from it. The fact that an object-oriented data model includes the aggregation and generalization relationships means that an object-oriented database system provides a user interface for the definition and manipulation of the relationships among objects. This in turn means that application programmers need not explicitly program and manage these relationships. For example, the aggregation relationship is the basis of the recursive definition of a complex object in terms of other objects. This makes it possible to successively refine the representation of complex objects often found in such applications as computer-aided design and engineering and compound documents.

In this paper, we provide a definition of an object-oriented data model and object-oriented database systems in Sections 2 and 3, respectively (this part of the paper is largely taken from a technical article this author has prepared for publication in a trade journal in early 1990.) Then in Sections 4 and 5, we discuss capabilities and limitations of an object-oriented data model in managing statistical and scientific databases. Section 8 closes the paper with a historical perspective on object-oriented database systems.

2. AN OBJECT-ORIENTED DATA MODEL

Despite the high level of activities aimed at the potential marketplace for object-oriented database systems, currently there is no standard object-oriented data model, and thus no standard object-oriented database language. However, if one examines existing object-oriented programming languages, knowledge repre-