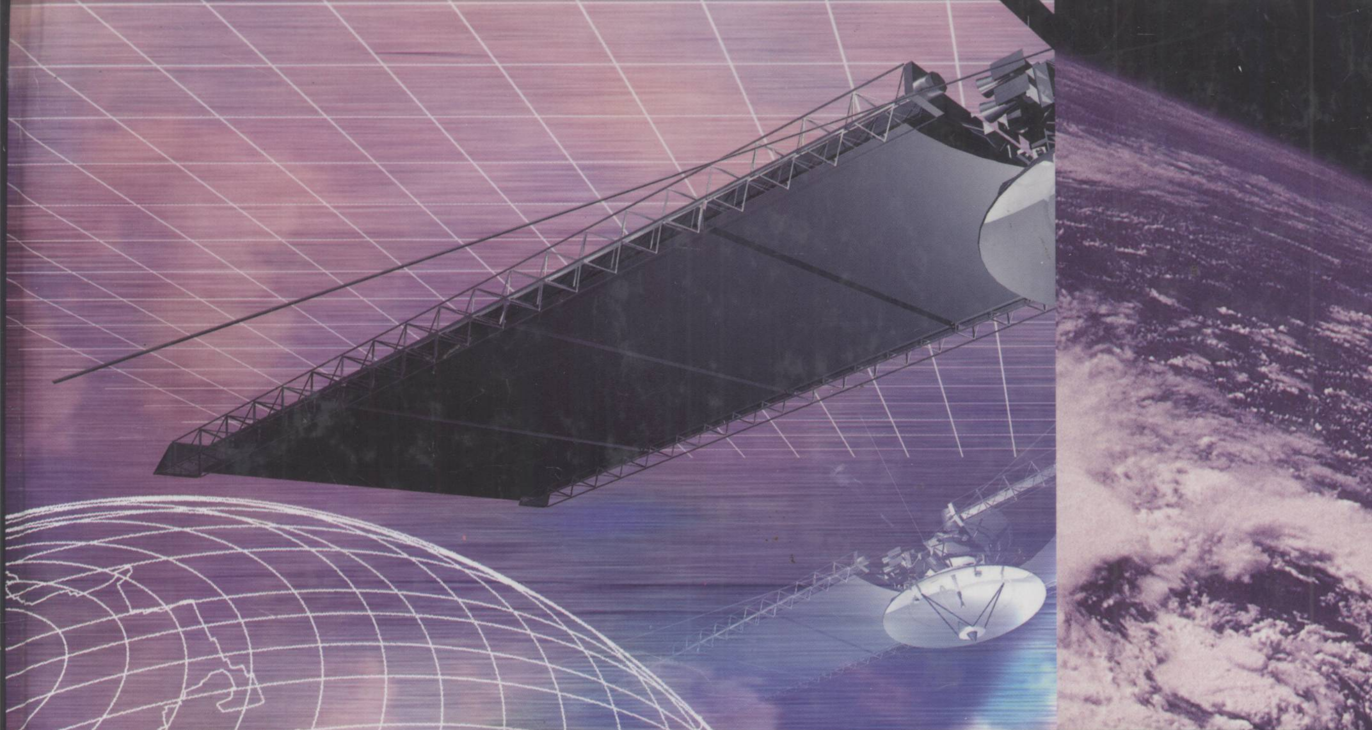
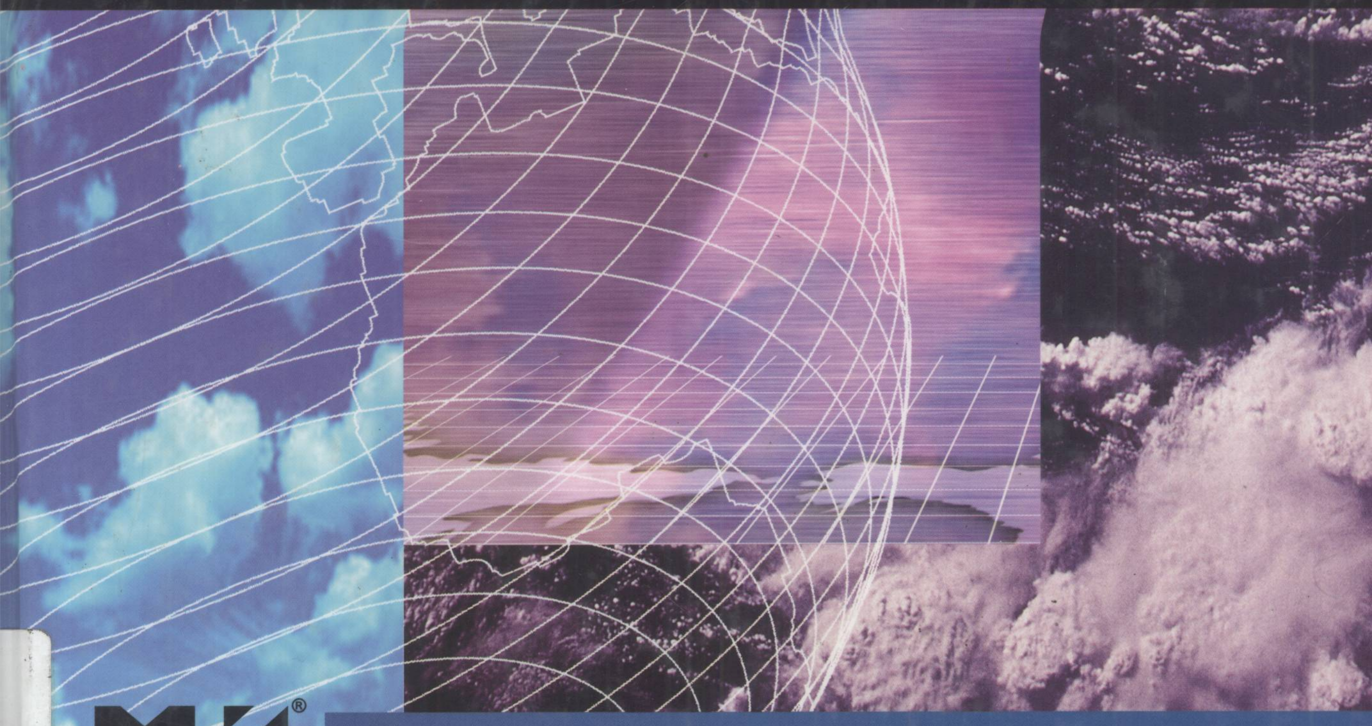


MOVING OBJECTS DATABASES



MORGAN KAUFMANN

Ralf Hartmut Güting and Markus Schneider

Moving Objects Databases

Ralf Hartmut Güting
Hagen, Germany

Markus Schneider
Gainesville, Florida, USA

ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
MORGAN KAUFMANN PUBLISHERS IS AN IMPRINT OF ELSEVIER

MORGAN KAUFMANN PUBLISHERS

Publisher	Diane D. Cerra
Publishing Services Manager	Simon Crump
Project Manager	Andrew Therriault
Editorial Assistant	Asma Stephan
Production Assistant	Melinda Ritchie
Cover Design	Yvo Riezebos Design
Composition	Multiscience Press, Inc.
Copyeditor	Multiscience Press, Inc.
Proofreader	Multiscience Press, Inc.
Indexer	Multiscience Press, Inc.
Interior printer	The Maple-Vail Book Manufacturing Group
Cover printer	Phoenix Color Corp.

Morgan Kaufmann Publishers is an imprint of Elsevier.
500 Sansome Street, Suite 400, San Francisco, CA 94111

This book is printed on acid-free paper.

Library of Congress Cataloging-in-Publication Data
Application submitted.

ISBN-13: 978-0-12-088799-6

ISBN-10: 0-12-088799-1

Moving Objects Databases

The Morgan Kaufmann Series in Data Management Systems

Series Editor: Jim Gray, Microsoft Research

Moving Objects Databases
Ralf Hartmut Güting and Markus Schneider

Foundations of Multidimensional and Metric Data Structures
Hanan Samet

Joe Celko's SQL Programming Style
Joe Celko

Data Mining, Second Edition: Concepts and Techniques
Ian Witten and Eibe Frank

Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration
Earl Cox

Data Modeling Essentials, Third Edition
Graeme C. Simsion and Graham C. Witt

Location-Based Services
Jochen Schiller and Agnès Voisard

Database Modeling with Microsft® Visio for Enterprise Architects
Terry Halpin, Ken Evans, Patrick Hallock, Bill Maclean

Designing Data-Intensive Web Applications
Stephano Ceri, Piero Fraternali, Aldo Bongio, Marco Brambilla, Sara Comai, and Maristella Matera

Mining the Web: Discovering Knowledge from Hypertext Data
Soumen Chakrabarti

Advanced SQL: 1999—Understanding Object-Relational and Other Advanced Features
Jim Melton

Database Tuning: Principles, Experiments, and Troubleshooting Techniques
Dennis Shasha and Philippe Bonnet

SQL:1999—Understanding Relational Language Components
Jim Melton and Alan R. Simon

Information Visualization in Data Mining and Knowledge Discovery
Edited by Usama Fayyad, Georges G. Grinstein, and Andreas Wierse

Transactional Information Systems: Theory, Algorithms, and Practice of Concurrency Control and Recovery
Gerhard Weikum and Gottfried Vossen

Spatial Databases: With Application to GIS
Philippe Rigaux, Michel Scholl, and Agnes Voisard

Information Modeling and Relational Databases: From Conceptual Analysis to Logical Design
Terry Halpin

Component Database Systems
Edited by Klaus R. Dittrich and Andreas Geppert

Managing Reference Data in Enterprise Databases: Binding Corporate Data to the Wider World
Malcolm Chisholm

Data Mining: Concepts and Techniques
Jiawei Han and Micheline Kamber

Understanding SQL and Java Together: A Guide to SQLJ, JDBC, and Related Technologies
Jim Melton and Andrew Eisenberg

Database: Principles, Programming, and Performance, Second Edition
Patrick and Elizabeth O'Neil

The Object Data Standard: ODMG 3.0
Edited by R. G. G. Cattell and Douglas K. Barry

Data on the Web: From Relations to Semistructured Data and XML
Serge Abiteboul, Peter Buneman, and Dan Suciu

Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations
Ian Witten and Eibe Frank

Joe Celko's SQL for Smarties: Advanced SQL Programming, Third Edition
Joe Celko

Joe Celko's Data and Databases: Concepts in Practice
Joe Celko

Developing Time-Oriented Database Applications in SQL
Richard T. Snodgrass

Web Farming for the Data Warehouse
Richard D. Hackathorn

Database Modeling & Design, Third Edition
Toby J. Teorey

Management of Heterogeneous and Autonomous Database Systems
Edited by Ahmed Elmagarmid, Marek Rusinkiewicz, and Amit Sheth

Object-Relational DBMSs: Tracking the Next Great Wave, Second Edition
Michael Stonebraker and Paul Brown,with Dorothy Moore

A Complete Guide to DB2 Universal Database
Don Chamberlin

Universal Database Management: A Guide to Object/Relational Technology
Cynthia Maro Saracco

Readings in Database Systems, Third Edition
Edited by Michael Stonebraker and Joseph M. Hellerstein

Understanding SQL's Stored Procedures: A Complete Guide to SQL/PSM
Jim Melton

Principles of Multimedia Database Systems
V. S. Subrahmanian

Principles of Database Query Processing for Advanced Applications
Clement T. Yu and Weiyi Meng

Advanced Database Systems
Carlo Zaniolo, Stefano Ceri, Christos Faloutsos, Richard T. Snodgrass, V. S. Subrahmanian, and Roberto Zicari

Principles of Transaction Processing
Philip A. Bernstein and Eric Newcomer

Using the New DB2: IBMs Object-Relational Database System
Don Chamberlin

Distributed Algorithms
Nancy A. Lynch

Active Database Systems: Triggers and Rules For Advanced Database Processing
Edited by Jennifer Widom and Stefano Ceri

Migrating Legacy Systems: Gateways, Interfaces, & the Incremental Approach
Michael L. Brodie and Michael Stonebraker

Atomic Transactions
Nancy Lynch, Michael Merritt, William Weihl, and Alan Fekete

Query Processing for Advanced Database Systems
Edited by Johann Christoph Freytag, David Maier, and Gottfried Vossen

Transaction Processing: Concepts and Techniques
Jim Gray and Andreas Reuter

Building an Object-Oriented Database System: The Story of O_2
Edited by François Bancilhon, Claude Delobel, and Paris Kanellakis

Database Transaction Models for Advanced Applications
Edited by Ahmed K. Elmagarmid

A Guide to Developing Client/Server SQL Applications
Setrag Khoshafian, Arvola Chan, Anna Wong, and Harry K. T. Wong

The Benchmark Handbook for Database and Transaction Processing Systems, Second Edition
Edited by Jim Gray

Camelot and Avalon: A Distributed Transaction Facility
Edited by Jeffrey L. Eppinger, Lily B. Mummert, and Alfred Z. Spector

Readings in Object-Oriented Database Systems
Edited by Stanley B. Zdonik and David Maier

To Edith, Nils David, and Helge Jonathan

R. H. G.

To Annette, Florian Markus, and Tim Christopher,
and my parents Hans and Christel Schneider

M. S.

Foreword

How can you represent a hurricane in a database? How can you represent an ocean's waves and currents? How can you represent a fleet of ships operating in that environment? If you have a representation, can it answer the interesting and difficult questions, such as what is the average current and windspeed experienced by these ships? These are challenging spatial database problems—but they are *static*. Now consider the more realistic problem when the hurricane, currents, and ships are moving. How would you structure and index the database and what kind of query language would you provide to answer questions such as: How long will it take each rescue ship and helicopter to arrive at the scene of the emergency?

With the advent of wireless and mobile computing, RFIDs, and sensor networks, it seems that every problem we encounter forces us to deal with objects moving in the 4D universe of space and time. Manufacturing, environmental monitoring, transportation and distribution, emergency services, and telecommunications all have challenging problems in representing and querying databases describing moving objects.

These data representation and data query problems were intractable 15 years ago—each had to be handled manually or approximated. But there has been huge progress in spatial databases, temporal databases, database indexing, and data querying over the last decade. We have learned how to represent 3D objects in compact ways, and have learned how to represent and reason about time and dynamics. The last five years have seen a synthesis of these spatial and temporal ideas into spatio-temporal data types and methods.

This book represents a milestone in that synthesis of temporal and spatial database concepts and techniques. It unifies and organizes the existing research into a coherent whole; and it also presents substantial new results and approaches in many areas. In each case it begins with what is known, then it introduces the new concepts in an abstract and general model, and then it translates the ideas into a pragmatic representation of data structures or SQL-like query language extensions. As such, the book makes both an excellent text and an excellent reference. It also takes you to the frontiers of our understanding, so it is a great point of departure for a new researcher who wants to advance this field.

Ralf Güting and Markus Schneider have been leaders in the unification of spatio-temporal databases. I have been working in an area peripheral to spatio-temporal databases and so was aware of some of the advances. But, in reading the preprint, I found myself saying "Aha!" more than once as each chapter progressed. As Alan Kay observed, "Perspective is worth 100 points of IQ"—meaning that if you look at a problem in the right way, it is easy to understand the problem and the solution. This book has lots of "perspective." I am now eager to apply this perspective to my problems.

Jim Gray
Microsoft Research

Preface

This book is about *moving objects databases*, a relatively new research area that has received a lot of interest in recent years. The general goal of this research is to allow one to represent moving entities in databases and to ask queries about such movements. Moving entities could be cars, trucks, aircraft, ships, mobile phone users, terrorists, or polar bears. For these examples, usually only the time-dependent position in space is relevant, not the extent; hence, we can characterize them as *moving points*. However, there are also moving entities with an extent—for example, hurricanes, forest fires, oil spills, armies, epidemic diseases, and so forth. These we would characterize as *moving regions*.

Moving objects are essentially time-dependent geometries, and database support for them is a specific flavor of *spatio-temporal databases*. The term *moving objects* emphasizes the fact that geometries can now change continuously, in contrast to most of the earlier work on spatio-temporal databases that supported only discrete changes. Note that discrete changes are a special case of continuous developments, and moving objects databases support both.

Some of the interest in this field is spurred by current trends in consumer electronics. Wireless network–enabled and position-aware (i.e., GPS equipped) devices, such as personal digital assistants, on-board units in vehicles, or even mobile phones, have become relatively cheap and are predicted to be in widespread use in the near future. This will lead to many kinds of new applications, such as location-based services. Similarly, big retail companies are moving toward tracking their products by indoor location devices (e.g. using RFID tags). Both trends mean that a huge volume of movement information (sometimes called trajectories) will become available and will need to be managed and analyzed in database systems.

The focus of the book is on the underlying database technology to support such applications. Extending database technology to deal with moving objects means—as for many other nonstandard database applications—to provide facilities in a DBMS data model for describing such entities and to extend the query language by constructs for analyzing them (e.g., for formulating predicates about them). Second, it means that the implementation of a DBMS must be extended. The two major strategies for this are: to build a layer on top of an existing DBMS and so to map moving object representations and predicates to existing facilities of the DBMS or to actually extend the DBMS by providing data structures for moving objects, methods for evaluating operations, specialized indexes and join algorithms, and so forth.

There are two major ways of looking at moving objects in databases: to be interested in maintaining continuously information about the current position and predict near future positions; and to consider whole histories of movements to be stored in the database and to ask queries for any time in the past or possibly the future (if we allow "histories" to include the future). This book treats both perspectives in depth.

Purpose and Audience

This book has three major purposes:

- It is intended as a textbook for teaching graduate students or advanced undergraduates.
- By providing a clear and concise presentation of the major concepts and results in the new field of moving objects databases, it should be interesting to researchers who want to get access to the area.
- Domain experts from industry, such as spatial data analysts, GIS experts, and software developers can obtain an introduction into state-of-the-art research on moving objects databases.

Structure and Organization

Chapter 1 gives an introduction to the theme of the book and briefly describes the essential features of database management systems, spatial database systems, and temporal database systems. It then shows that traditional database technology is unable to model and implement moving objects and that new concepts and techniques are needed. A number of applications are listed that demonstrate the necessity of concepts for moving objects.

Chapter 2 gives an overview of two precursors of today's spatio-temporal models. The first approach is based on a temporal extension of simplicial complexes. The second approach is event-based.

Chapter 3 deals with location management, which is related to current and near-future movements. The most prominent model is the MOST model, whose components and main concepts are described in detail. Additional topics are the query language FTL, based on temporal logic; the problem of balancing update cost and imprecision; and the uncertainty of the trajectory of a moving object.

Chapter 4 describes the authors' approach for modeling and querying histories of movement or evolutions of spatial objects over time. This concept is based on abstract data types that can be embedded as attribute types in database schemas. At two different abstraction levels, we present an abstract model with spatio-temporal data types and operations as well as a discrete model providing finite representations for the types of the abstract model. Additionally, the abstract model is enhanced by a concept of spatio-temporal predicates that can be employed for the formulation of spatio-temporal selections and joins.

Chapter 5 presents data structures and algorithms for the data types and operations of the abstract model. Data structures are designed in the context of, and according to, the needs of database systems. Algorithms are designed for operations on temporal data types and for lifted operations. Lifted operations on moving objects are operations that are derived from static operations on corresponding spatial objects in a certain systematic way.

Chapter 6 deals with the constraint approach for moving objects. As an abstract model for geometric entities in a *k*-dimensional space, the constraint approach uses infinite relations. A corresponding discrete model is based on constraint relations. Finally, an implementation of the discrete model is described.

Chapter 7 focuses on spatio-temporal indexing. Starting with some needed geometric preliminaries, which make this book independent from other textbooks, requirements for indexing moving objects are introduced. Then, index structures are presented for current and near-future movements as well as for histories of movement.

Chapter 8 provides a brief overview of further research areas in moving objects databases that are beyond the scope of this book.

The text of this book has been used as course material at Fernuniversität Hagen and is designed for self-study. Each chapter provides exercises within the text. The reader is encouraged to work on these exercises when they occur. Solutions can be found at the end of the book. Further exercises are provided at the end of each chapter; these can be used for homework assignments. Each chapter also provides bibliographic notes explaining the sources for the presented material and related work.

Prerequisites and Possible Courses

This book should be accessible to anyone with a general background on the concepts of database systems. A deeper knowledge on the implementation of database systems is helpful but not required. A background on spatial and temporal database systems is useful but not needed; brief introductions to these fields are provided in Chapter 1.

At Fernuniversität Hagen, the material of the book is used in two ways:

- The complete text is used for a one-semester graduate course, covering modeling and implementation issues.

- A short course, focusing only on the modeling aspects, consists of Chapters 1 through 4.

At the University of Florida, the text of the book is currently used for a one-semester, graduate research class in seminar form. After a broad introduction into the field by the instructor, graduate students are assigned chapters for extensive individual study. The bibliographic notes at the end of the chapters lead them to further material. Students have to give a detailed presentation about their assigned topic and write a report. In addition, the class performs a prototypical implementation project together.

Other selections of subsets for short courses are possible:

- A short course on modeling might consist of Chapters 1 through 4, but replace Section 4.4 with Sections 6.1 and 6.2, to include the constraint model.
- The approaches of Chapters 3, 4, and 6 are alternative—hence, can be read in any order or be omitted. To get a balanced view we recommend including at least Chapters 1 and 3 and Sections 4.1 through 4.3.
- Chapter 5 describes the implementation for the model of Chapter 4 (Sections 4.1 through 4.3).
- Section 7.3 describes indexing for current and near-future movement—hence, is related to the model in Chapter 3. Section 7.4 covers indexing of trajectories or histories of movement and is related to the model in Chapter 4. Nevertheless, one can study the indexing problems without detailed knowledge of these chapters.

Acknowledgments

Writing this book is in some way a result of our research work on moving objects databases, and we are most grateful to several of our colleagues with whom we had the pleasure to work. The CHOROCHRONOS project, funded by the European Union in the late 1990s, has been very influential. This was led by Timos Sellis and Manolis Koubarakis; we thank them for their great leadership and organization of the project and for inviting us to participate initially. We greatly enjoyed meetings and discussions with them and the other members of the project—namely, Andrew Frank, Stéphane Grumbach, Christian Jensen, Nikos Lorentzos, Yannis Manolopoulos, Enrico Nardelli, Barbara Pernici, Hans-Jörg Schek, Michel Scholl, Babis Theodoulidis, and Nectaria Tryfona.

The presentation of our own approach in Chapters 4 and 5 is based on joint work and papers with our coauthors, and we owe a great thanks to Martin Erwig, Luca Forlizzi, Enrico Nardelli, Christian Jensen, Jose Antonio Cotelo Lema, Mike Böhlen, Nikos Lorentzos, and Michalis Vazirgiannis. It was a pleasure to collaborate with them and we are proud of the outcome of our joint work.

Special thanks go to Christian Jensen for continued support and advice and to Ouri Wolfson for many interesting discussions and his great hospitality on the occasion of Ralf's visit to Chicago.

Both of us thank the research agencies partially supporting our work: for Ralf this is the DFG (Deutsche Forschungsgemeinschaft), grant Gu293/8-1; and for Markus this is the NSF (National Science Foundation), grant NSF-CAREER-IIS-0347574.

Our book proposal received a warm welcome and later enthusiastic support by Jim Gray and the editors at Morgan Kaufmann Publishers, Rick Adams and Diane Cerra. We are grateful for that and were impressed by the speed and professionalism by which they set up a board of expert reviewers. We thank the reviewers, Ouri Wolfson, Shashi Shekhar, Vassilis Tsotras, Max Egenhofer, Cyrus Shahabi, and Agnès Voisard, for providing careful and critical reviews of the book proposal in an amazingly short period of time. Thanks to all of them for supporting the project and providing many useful suggestions. Some of these suggestions we will only be able to realize in an expanded second edition . . . Thanks also to Ouri Wolfson, Vassilis Tsotras, and Agnès Voisard for providing detailed second-round reviews on selected chapters.

We thank the research assistants, Dirk Ansorge, Helge Bals, Thomas Behr, and Markus Spiekermann at Fernuniversität Hagen, for reading the course material and providing corrections and especially for developing the homework exercises; they did a great job.

We also thank the students at Fernuniversität Hagen who worked through the course and provided feedback and corrections. Here we mention especially Matthias Ruth, Stefan Schmidt, and Marten Jung.

Ralf Hartmut Güting
Hagen, Germany

Markus Schneider
Gainesville, Florida, USA

February, 2005

Contents