Preface

The papers in this volume were presented at IDEAS'99, the third International Database Engineering and Applications Symposium at Concordia University, Montreal, Canada, on August 2–4, 1999. IDEAS is an annual symposium dedicated to the engineering and application aspects of databases. This is in response to the increasing demands placed on related theories and technologies resulting from the widening scope of database solutions within all forms of enterprise. All submitted papers were reviewed by at least three members of the program committee followed by on-line discussions. The program committee selected 34 papers as regular contributions from 90 submissions. In addition, it selected 17 submissions to be included as short papers. IDEAS’99 invited talks by M. Tamer Ozsu (University of Alberta), Sham B. Navathe (Georgia Tech) and a keynote address by Jeffrey D. Ullman (Stanford University). Summaries of these are included in these proceedings.

Ozsu, in the opening invited talk entitled Issues in Multimedia Data Management, explores the ubiquitous multimedia in evolving computer application and examines their data management needs and functionality.

The second session on Query Optimization consists of three papers. Oommen and Thiyagarajah in Query Result Size Estimation Using a Novel Histogram-like Technique attempts to approximate the density of domain values. In Semantic Query Optimization for ODMG--Databases van Bommel discusses a graphical optimization technique, whereas Che exploits the knowledge of DTD of structured documents in the third paper in this session. Bhowmick et al analyse Cost-Benefit Analysis of Web Bag in A Web Warehouse.

Liu, Ling, and Guan in the first paper of the session on Multi-DB record the attempts to integrate the semistructured objects across the Web and explore their semantics. In A Schema Analysis and Reconciliation Tool Environment for Heterogeneous Databases Castano and Antonellis present a tool to support heterogeneous schema analysis. Jautzy on the other hand, takes an approach of adding generic MDBS management facilities into general purpose OO programming language.

The short paper session on Modelling features papers on current topics in OODBMS, GIS, Heterogeneous DBMS and the Web. In Interactive Design of Object-Oriented Schema van Bommel describes a graphical schema design and semantic capture tool. Nicolle, Cullot and Yetongnon use description logic formalism to provide formal specifications of abstract metatypes. GIS interoperability by dynamically resolving semantic conflicts is the topic of the paper ISIS: A Semantic Mediation Model and an Agent Based Architecture for GIS Interoperability by Leclercq, Benslimane and Yetongnon. Loeser and Ritter propose an integrated Web management component based on ORDBS. Kim and Chang discuss efficient retrieval of content based iconic images in their paper SR Schemes: Spatial-match Iconic Image Representations. A semi-dynamic construction method in a slowly changing system is proposed by Hochin and Tsuji in the paper A Method of Constructing Dynamic Schema Representing the Structure of Semistructured Data.

Wietrzyk and Orgun propose an application of clustering to dynamically reorganize physical database in the first paper of the session on Application of Clustering. On the other hand, in the paper New GDM-based Declustering Methods for Parallel Range Queries Kuo, Winslett and Cho give an algorithm for declustering to parallel disks. Merkl and Rauber examine a system of self-organizing distributed maps in their paper: Self-Organization of Distributed Document
Archives. A technique to apply statistical feature reduction in filtering out articles is the topic of the paper *Feature Reduction and Database Maintenance in NETNEWS Classification* by Hsu and Lang.

In the keynote session entitled *Some Advances in Data-Mining Techniques* Ullman discusses the work underway in the area of querying unstructured information and the work being done at Stanford University in this area.

OLAP and Data Mining continue to be the topics in current database research. The first paper in this session by Rajamani et al examines the performance issues of data mining using the Apriori algorithm. In *Management of Multidimensional Aggregates for Efficient Online Analytical Processing*, Albrecht et al describe the management of multidimensional aggregates via the experimental CUBESTAR OLAP server. Markl, Ramsak and Bayer note the flexibility of multidimensional hierarchical clustering. They claim that clustering improves aggregation queries while not having to pay the penalty of materialization. In their paper, *A Parallel Scalable Infrastructure for OLAP and Data Mining* Goil and Choudhary address the scalability and integration problems while trying to maintain high performance via parallel processing.

The session on View Maintenance contains four papers. Alhajj and Elnagar in *A Model for Deferred View Maintenance* present a deferred OO view maintenance model. In *Incremental Maintenance of Nested Relational Views* Liu, Vincent and Mohania delve into the evaluation and update in nested database. Zhang and Rundensteiner claim a guaranteed concurrent evolution and maintenance of view defined over distributed information sources. The last paper in this stream also by the Rundensteiner team proposes strategies for view adaptation following view synchronization.

Five presentations are featured in the short paper session on Evaluation and Analysis. The first paper by Chu and Wong tackles the hot topic of data mining via an algorithm to discover clustering patterns from numerical data. In *Hash-based Symmetric Data Structure and Join Algorithm for OLAP Applications* Toyama and Ohara apply the fluent join to cluster records on the joining attributes. Supporting topological and direction queries on spatial data is the topic of the paper by Safar and Shahahbi. Signature indexing to provide light weight search tree while eliminating the cross talk of signature from different attributes is examined in *A New Access Index for Fast Execution of Conjunctive Queries over Text Data* by Yang and Singhal. The final paper in this session is on concurrent maintenance of view that allows readers to continue read operations while updates to the views are underway.

In the first paper in the session on Principles of Databases Liu and Yu considers the theory of database queries and develop a translation algorithm. In *Between Well-Founded Semantics and Stable Model Semantics* Luong defines new semantics for negated Datalog programs. Ng proposes lexical ordered FDs in *Lexicographically Ordered Functional Dependencies and Their Application to Temporal Relations*. Montesi, Bagnato and Dallera introduce an evolution graph for termination analysis in active databases.

Navathe’s invited talk is on the complexity of managing the genome data.

A probabilistic model for developing a multi-agent-reasoning system is introduced by Butz and Wong in the first paper on Transaction and Recovery. Bodorik et al investigate an efficient locking service for transaction processing. Requirements for transaction agents are the focus of the paper *Transactional Coordination Agents for Composite Systems* by Schuldt, Schek and
Alonso, Tang and Xing introduce supporting the ad-hoc recovery concept in practical applications.

McClatchey et al in their presentation *The Role of Meta-Objects and Self-Description in an Engineering Data Warehouse* report on a repository to capture engineering data which support interoperability. In the paper *Towards an Active Real-Time Database-Based Architecture for Real-Time Database and Standardized Components*, Munnich et al present an architectural model that separates control and data flow. Zhang and Harder report on the lessons learned in attempted extensibility. In the paper *Scalability of an Object Descriptor Architecture OODBMS*, Yu, Lee and Olson investigate the scalability issue in OODBMS.

The first paper from the short paper session on Transaction Processing and Performance is by Lee and Park on a site-locking concurrency control method. Performance evaluation in a DBMS wherein arrays are first class data types is the topic of the presentation by Widmann and Baumann. An OODBMS called EyeDB that is targeted for the genome project is introduced by Viara. Kang and Kim report on their investigation of *Performance Tradeoffs in Real-Time Transaction Processing with Client Data Caching*. In the paper *Measuring the Performance of Database Object Horizontal Fragmentation Schemes*, Ezeife and Zheng provide a technique for measuring the performance of fragmentation in a DDBMS. Ferreira and Busichia report on a modularization scheme for constructing a flexible information system.

The last session of this symposium deals with DB Engineering and User Issues. In the presentation, *An Implementation Architecture for Orthogonally Persistent Deductive Object-Oriented Database Systems* Ashrafuzzaman and Kusalik propose a novel architectural solution for implementing orthogonal persistence. An interactive customizable interface is the topic of the paper by Cruz and James. Ng and Muntz consider a paradigm to exploit user-defined functions in ORDBMS query optimization.

This completes the synopsis of the articles in this collection. We wish to thank all the authors for their contributions, and the invited speakers for agreeing to address the IDEAS symposium. Thanks, too to the program committee and the external reviewers whose efforts helped to make IDEAS’99 possible.

Bipin C. Desai, Gösta Grahne
Concordia University, Montreal