

CONTENTS

Invited Talks

<i>Innovation in Database Management: Computer Science vs. Engineering</i>	2
<i>Kenneth R. Jacobs (Oracle)</i>	
<i>The Microsoft Repository</i>	3
<i>Philip A. Bernstein, Brian Harry, Paul Sanders, David Shutt, Jason Zander (Microsoft)</i>	
<i>Multidimensional Access Methods: Trees Have Grown Everywhere</i>	13
<i>Timos Sellis (NTUA), Nick Roussopoulos, Christos Faloutsos (University of Maryland)</i>	

Research Sessions

Session: Data Warehousing

<i>Incremental Organization for Data Recording and Warehousing</i>	16
<i>H. V. Jagadish, P. P. S. Narayan, S. Seshadri, S. Sudarshan, Rama Kanneganti</i>	
<i>Multiple-View Self-Maintenance in Data Warehousing Environments</i>	26
<i>Nam Huyn</i>	
<i>Recovering Information from Summary Data</i>	36
<i>Christos Faloutsos, H.V. Jagadish, N.D. Sidiropoulos</i>	

Session: Novel Data Types

<i>A Language for Manipulating Arrays</i>	46
<i>Arunprasad P. Marathe, Kenneth Salem</i>	
<i>Implementing Abstract Objects with Inheritance in Datalog-neg</i>	56
<i>Hasan M. Jamil</i>	
<i>The Case for Enhanced Abstract Data Types</i>	66
<i>Praveen Seshadri, Miron Livny, Raghu Ramakrishnan</i>	

Session: Transactions and Reliability

<i>Integrating Reliable Memory in Databases (best paper award)</i>	76
<i>Wee Teck Ng, Peter M. Chen</i>	
<i>Logical and Physical Versioning in Main Memory Databases</i>	86
<i>Philip Bohannon, Dennis Leinbaugh, Rajeev Rastogi, S. Seshadri, Avi Silberschatz, S. Sudarshan</i>	
<i>Using Versions in Update Transactions : Application to Integrity Checking</i>	96
<i>Francois Llirbat, Eric Simon, Dimitri Tombroff</i>	

Session: Multidimensional Databases

A Foundation for Multi-Dimensional Databases	106
<i>Marc Gyssens, Laks V.S. Lakshmanan</i>	
Fast Computation of Sparse Datacubes	116
<i>Kenneth A. Ross, Divesh Srivastava</i>	
Data Warehouse Configuration	126
<i>Dimitri Theodoratos, Timos Sellis</i>	

Session: Database Design

Algorithms for Materialized View Design in Data Warehousing Environment	136
<i>Jian Yang, Kamalakar Karlapalem and Qing Li</i>	
An Efficient, Cost-Driven Index Selection Tool for Microsoft SQL Server	146
<i>Surajit Chaudhuri, Vivek Narasayya</i>	
Materialized View Selection in a Multidimensional Database	156
<i>Elena Baralis, Stefano Paraboschi, Ernest Teniente</i>	

Session: Data Mining

Efficient Construction of Regression Trees with Range and Region Splitting	166
<i>Yasuhiko Morimoto, Hiromu Ishii, Shinichi Morishita</i>	
Parallel Algorithms for High-dimensional Proximity Joins	176
<i>John C. Shafer, Rakesh Agrawal</i>	
STING: A Statistical Information Grid Approach to Spatial Data Mining	186
<i>Wei Wang, Jiong Yang, Richard Muntz</i>	

Session: Web Databases

Merging Ranks from Heterogeneous Internet Sources	196
<i>Luis Gravano, Hector Garcia-Molina</i>	
To Weave the Web	206
<i>Paolo Atzeni, Giansalvatore Mecca, Paolo Merialdo</i>	
Using Probabilistic Information in Data Integration	216
<i>Daniela Florescu, Daphne Koller, Alon Levy, Avi Pfeffer</i>	

Session: Tertiary Storage

On-Demand Data Elevation in a Hierarchical Multimedia Storage Server	226
<i>Peter Triantafillou, Thomas Papadakis</i>	
Principles of Optimally Placing Data in Tertiary Storage Libraries	236
<i>Stavros Christodoulakis, Peter Triantafillou, Fenia A. Zioga</i>	
Vertical Data Migration in Large Near-Line Document Archives Based on Markov-Chain Predictions	246
<i>Achim Kraiss, Gerhard Weikum</i>	

Session: Heterogeneous Databases

Describing and Using the Query Capabilities of Heterogeneous Sources	256
<i>Vasilis Vassalos, Yannis Papakonstantinou</i>	
Don't Scrap It, Wrap It! A Wrapper Architecture for Legacy Data Sources	266
<i>Mary Tork Roth, Peter Schwarz</i>	
Optimizing Queries Across Diverse Data Sources	276
<i>Laura M. Haas, Donald Kossmann, Edward L. Wimmers, Jun Yang</i>	

Session: Query Optimization

Optimizing Queries with Universal Quantification in Object-Oriented and Object-Relational Databases	286
<i>Jens Claussen, Alfons Kemper, Guido Moekotte, Klaus Peithner</i>	
Parallel Query Scheduling and Optimization with Time- and Space-Shared Resources	296
<i>Minos N. Garofalakis, Yannis E. Ioannidis</i>	
The Complexity of Transformation-Based Join Enumeration	306
<i>Arjan Pellenkoft, Cesar A. Galindo Legaria, Martin Kersten</i>	

Session: Distributed Databases

1-Safe Algorithms for Symmetric Site Configurations	316
<i>Rune Humberstad, Maitrayi Sabaratnam, Svein-Olav Hvasshovd, Oystein Torbjornsen</i>	
Adaptive Data Broadcast in Hybrid Networks	326
<i>Konstantinos Stathatos, Nick Roussopoulos, John S. Baras</i>	
Finding Data in the Neighborhood	336
<i>Andre Eickler, Alfons Kemper, Donald Kossmann</i>	

Session: Database Algorithms

A One-Pass Algorithm for Accurately Estimating Quantiles for Disk-Resident Data	346
<i>Khaled Alsabti, Sanjay Ranka, Vineet Singh</i>	
Concurrent Garbage Collection in O2	356
<i>Marcin Skubiszewski, Patrick Valduriez</i>	
Garbage Collection in Object Oriented Databases Using Transactional Cyclic Reference Counting	366
<i>S. Ashwin, Prasan Roy, S. Seshadri, Avi Silberschatz, S. Sudarshan</i>	

Session: Sorts and Joins

Dynamic Memory Adjustment for External Mergesort	376
<i>Weiye Zhang , Per-Ake Larson</i>	
Evaluation of Main Memory Join Algorithms for Joins with Subset Comparison Join Predicates .	386
<i>Sven Helmer, Guido Moerkotte</i>	
Spatial Joins Using R-trees: Breadth-First Traversal with Global Optimizations	396
<i>Yun-Wu Huang, Ning Jing, Elke A. Rundensteiner</i>	

Session: Spatial Access Methods

A Generic Approach to Bulk Loading Multidimensional Index Structures	406
<i>Jochen van den Bercken, Bernhard Seeger, Peter Widmayer</i>	
A Region Splitting Strategy for Physical Database Design of Multidimensional File Organizations	416
<i>Jong-Hak Lee, Young-Koo Lee, Kyu-Young Whang, and Il-Yeol Song</i>	
M-tree: An Efficient Access Method for Similarity Search in Metric Spaces	426
<i>Paolo Ciaccia, Marco Patella, Pavel Zezula</i>	

Session: Querying and Browsing

DataGuides: Enabling Query Formulation and Optimization in Semistructured Databases	436
<i>Roy Goldman, Jennifer Widom</i>	
Using Taxonomy, Discriminants, and Signatures for Navigating in Text Databases	446
<i>Soumen Chakrabarti, Byron Dom, Rakesh Agrawal, Prabhakar Raghavan</i>	
Towards an ODMG-Compliant Visual Object Query Language	456
<i>Manoj Chavda, Peter T. Wood</i>	

Session: Query Processing

Fast Incremental Maintenance of Approximate Histograms	466
<i>Phillip B. Gibbons, Yossi Matias, Viswanath Poosala</i>	
Groupwise Processing of Relational Queries	476
<i>Damianos Chatziantoniou, Kenneth A. Ross</i>	
Selectivity Estimation Without the Attribute Value Independence Assumption	486
<i>Viswanath Poosala, Yannis E. Ioannidis</i>	

Session: Multimedia Databases

Effective Memory Use in a Media Server	496
<i>Edward Chang, Hector Garcia-Molina</i>	
Efficient User-Adaptable Similarity Search in Large Multimedia Databases	506
<i>Thomas Seidl, Hans-Peter Kriegel</i>	
Resource Scheduling in Enhanced Pay-Per-View Continuous Media Databases	516
<i>Minos N. Garofalakis, Banu Özden, Avi Silberschatz</i>	

Industrial/Commercial Sessions

Session: System Experiences

Integrating SQL Databases with Content-specific Search Engines	528
<i>Stefan Dessloch, Nelson Mattos (IBM)</i>	
Facilitating Multimedia Database Exploration through Visual Interfaces and Perpetual Query Reformulations	538
<i>Wen-Syan Li, K. Selcuk Candan, Kyoji Hirata, Yoshinori Hara (NEC)</i>	
Geo/Environmental and Medical Data Management in the RasDaMan System	548
<i>Peter Baumann, Paula Furtado, Roland Ritsch, Norbert Widmann (FORWISS)</i>	

Session: Application Challenges

GTE SuperPages: Using IR Techniques for Searching Complex Objects	553
<i>Steven Whitehead, Himanshu Sinha, Michael Murphy (GTE)</i>	
Data Manager for Evolvable Real-time Command and Control Systems	558
<i>Eric Hughes, Roman Ginis, Bhavani Thuraisingham, Peter Krupp, John Maurer (MITRE)</i>	
Mining Insurance Data at Swiss Life	562
<i>Jörg-Uwe Kietz, Ulrich Reimer, Martin Staudt (Swiss Life)</i>	

Session: Research Visions

The Network as a Global Database: Challenges of Interoperability, Proactivity, Interactiveness, Legacy	567
<i>Peter C. Lockemann, Ulrike Kölsch, Arne Koschel, Ralf Kramer, Ralf Nikolai, Mechtild Wallrath, Hans-Dirk Walter (FZI Karlsruhe)</i>	
Distributed Processing over Stand-alone Systems and Applications	575
<i>Gustavo Alonso, Claus Hagen, Hans-Jörg Schek, Markus Tresch, (ETH Zürich)</i>	
Critical Database Technologies for High Energy Physics	580
<i>David M. Malon, Edward N. May (Argonne National Laboratory)</i>	

Session: Industry Solutions

Caprera: An Activity Framework for Transaction Processing on Wide-Area Networks	585
<i>Suresh Kumar, Eng-Kee Kwang (Tactica), Divyakant Agrawal (UC Santa Barbara)</i>	
The Oracle Universal Server Buffer Manager	590
<i>W. Bridge, A. Joshi, M. Keihl, T. Lahiri, J. Loaiza, N. Macnaughton (Oracle)</i>	
Efficient Testing of High Performance Transaction Processing Systems	595
<i>Dennis Wildfogel (Tandem), Ramana Yerneni (Stanford University)</i>	
Author Index	599