

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Alex Gray Keith Jeffery Jianhua Shao (Eds.)

Sharing Data, Information and Knowledge

25th British National Conference on Databases, BNCOD 25
Cardiff, UK, July 7-10, 2008
Proceedings



Springer

Volume Editors

Alex Gray
Jianhua Shao
Cardiff University, School of Computer Science
5 The Parade, Cardiff CF24 3AA, UK
E-mail: {w.a.gray, j.shao}@cs.cf.ac.uk

Keith Jeffery
Rutherford Appleton Laboratory, Science and Technology Facilities Council
Harwell Science and Innovation Campus, Didcot OX11 0QX, UK
E-mail: kgj@rl.ac.uk

Library of Congress Control Number: 2008930155

CR Subject Classification (1998): H.2, H.3, H.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN	0302-9743
ISBN-10	3-540-70503-1 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-70503-1 Springer Berlin Heidelberg New York

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 any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media
springer.com

© Springer-Verlag Berlin Heidelberg 2008
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12322770 06/3180 5 4 3 2 1 0

Preface

Since 1981, the British National Conferences on Databases (BNCOD) have provided a forum for database researchers to report the latest progress and explore new ideas. Over the last 28 years, BNCOD has evolved from a predominantly national conference into one that is truly international, attracting research contributions from all over the world.

This volume contains the proceedings of BNCOD 2008. We received 45 submissions from 22 countries. Each paper was reviewed by three referees, and 14 full papers and 7 posters were accepted. All the research papers and posters are included in this volume, and they are organized into five sections: data mining and privacy, data integration, stream and event data processing, query processing and optimization, and posters.

The keynote was delivered by Monica Marinucci, EMEA Programme Director for Oracle in R&D. She has been involved in various advanced developments concerning Oracle, and participated in EC-funded projects as an expert, especially the CHALLENGERS special support action to propose the future of grid computing. In her keynote presentation, she addressed the audience on the topic of the power of data, emphasizing that the ability to store, handle, manipulate, distribute and replicate data and information can provide a tremendous asset to organizations. She also explored some of the latest directions and developments in the database field, and described how Oracle contributes to them partnering up with other leading organizations in different sectors.

BNCOD 2008 marked a special occasion in the BNCOD history – it was the 25th conference in the BNCOD series (BNCOD was not held in 1987, 1997 and 1999 when VLDB and ICDE were held in the UK). To mark this 25th anniversary, an International Colloquium on Advances in Database Research was held as part of the BNCOD main conference. Leading researchers were invited to the colloquium, and they presented and shared their latest research with the audience. The invited papers or the abstracts of the talks presented at the colloquium are included in this volume.

Two workshops were held alongside BNCOD 2008. The Workshop on Teaching, Learning and Assessment of Databases continued to address, as in the previous years, the issues concerning the educational aspects of databases. The papers from this workshop were published separately, and were not included in this volume. The Workshop on Biodiversity Informatics: Challenges in Modelling and Managing Biodiversity Knowledge aimed at advancing understanding in the challenges involved in this important, multi-disciplinary research area. The two best papers from this workshop are included in this volume.

Finally, we would like to thank all the authors for contributing their papers to BNCOD 2008, the referees for their effort in reviewing the papers, the EPSRC for their support of the International Colloquium on Advances in Database

Research, the Welsh Assembly Government for hosting the reception, and the Organizing Committee at Cardiff University for making this conference possible.

July 2008

Alex Gray
Keith Jeffery
Jianhua Shao

Conference Committees

Programme Committee

Alex Gray (Chair)	Cardiff University, UK
David Bell	Queen's University, Belfast, UK
Stefan Böttcher	University of Paderborn, Germany
Peter Buneman	University of Edinburgh, UK
Sharma Chakravarthy	The University of Texas at Arlington, USA
Richard Cooper	University of Glasgow, UK
Alfredo Cuzzocrea	Universita della Calabria, Italy
Barry Eaglestone	University of Sheffield, UK
Suzanne Embury	University of Manchester, UK
A. Fernandes	University of Manchester, UK
Mary Garvey	University of Wolverhampton, UK
Georg Gottlob	University of Oxford, UK
Jun Hong	Queen's University, Belfast, UK
Ela Hunt	Swiss Federal Institute of Technology, Switzerland
Wendy Ivins	Cardiff University, UK
Mike Jackson	Birmingham City University, UK
Anne James	Coventry University, UK
Keith Jeffery	Rutherford Appleton Laboratory, UK
Andrew Jones	Cardiff University, UK
Graham Kemp	Chalmers University, Sweden
Jessie Kennedy	Napier University, UK
Kavin Lu	Brunel University, UK
L MacKinnon	University of Abertay Dundee, UK
Nigel Martin	Birkbeck, University of London, UK
Ken Moody	University of Cambridge, UK
David Nelson	University of Sunderland, UK
M Norrie	ETH Zurich, Switzerland
Werner Nutt	The Free University of Bozen-Bolzano, Italy
Norman Paton	University of Manchester, UK
Alex Poulouvassilis	Birkbeck, University of London, UK
Mark Roantree	Dublin City University, Ireland
Alan Sexton	University of Birmingham, UK
Jianhua Shao	Cardiff University, UK
Paul Watson	Newcastle University, UK
Richard White	Cardiff University, UK
John Wilson	University of Strathclyde, UK
Jian Yang	Macquarie University, Australia

Steering Committee

Lachlan Mackinnon (Chair)	University of Abertay Dundee, UK
Richard Cooper	University of Glasgow, UK
Barry Eaglestone	University of Sheffield, UK
Jun Hong	Queen's University, Belfast, UK
Anne James	Coventry University, UK
Keith Jeffery	Rutherford Appleton Laboratory, UK
David Nelson	University of Sunderland, UK
Alex Poulouvassilis	Birkbeck, University of London, UK
Jianhua Shao	Cardiff University, UK

Organizing Committee

Nick Fiddian (Chair)	Cardiff University
Mikhaila Burgess	Cardiff University
Wendy Ivins	Cardiff University
Andrew Jones	Cardiff University
Jianhua Shao	Cardiff University
Gareth Shercliff	Cardiff University
Alysia Skilton	Cardiff University
Richard White	Cardiff University

25th Anniversary Colloquium Organization

Keith Jeffery	Rutherford Appleton Laboratory, UK
Jianhua Shao	Cardiff University, UK

Additional Referees

Mikhaila Burgess	Iadh Ounis
Grigorios Loukides	Punitha Swamy

Table of Contents

Keynote

The Power of Data	1
<i>Monica Marinucci</i>	

Data Mining and Privacy

Efficient Mining of Frequent Itemsets from Data Streams	2
<i>Carson Kai-Sang Leung and Dale A. Brajczuk</i>	
An Empirical Study of Utility Measures for k -Anonymisation	15
<i>Grigorios Loukides and Jianhua Shao</i>	
HLS: Tunable Mining of Approximate Functional Dependencies	28
<i>Jeremy T. Engle and Edward L. Robertson</i>	
Sentence Ordering for Coherent Multi-document Summary Generation	40
<i>C.R. Chowdary and P. Sreenivasa Kumar</i>	

Data Integration

Schema Matching across Query Interfaces on the Deep Web	51
<i>Zhongtian He, Jun Hong, and David Bell</i>	
A Generic Data Level Implementation of ModelGen	63
<i>Andrew Smith and Peter McBrien</i>	
Reconciling Inconsistent Data in Probabilistic XML Data Integration ...	75
<i>Tadeusz Pankowski</i>	

Stream and Event Data Processing

A Semantics for a Query Language over Sensors, Streams and Relations	87
<i>Christian Y.A. Brenninkmeijer, Ixent Galpin, Alvaro A.A. Fernandes, and Norman W. Paton</i>	
Load Shedding in MavStream: Analysis, Implementation, and Evaluation	100
<i>Balakumar Kendai and Sharma Chakravarthy</i>	
Event-Driven Database Information Sharing	113
<i>Luis Vargas, Jean Bacon, and Ken Moody</i>	

Query Processing and Optimisation

Smooth Interpolating Histograms with Error Guarantees	126
<i>Thomas Neumann and Sebastian Michel</i>	
Virtual Forced Splitting, Demotion and the BV-Tree	139
<i>Alan P. Sexton and Richard Swinbank</i>	
A Functional Data Model Approach to Querying RDF/RDFS Data	153
<i>João Martins, Rui Nunes, Merja Karjalainen, and Graham J.L. Kemp</i>	
Ranking for Approximated XQuery Full-TextQueries	165
<i>Giacomo Buratti and Danilo Montesi</i>	

Poster Papers

Role Based Access to Support Collaboration in Healthcare	177
<i>Alysia Skilton, W. Alex Gray, Omnia Allam, Dave Morry, and Hazel Bailey</i>	
A Peer-to-Peer Database Server	181
<i>John Colquhoun and Paul Watson</i>	
Checking the Integrity Constraints of Mobile Databases with Three-Level Model	185
<i>Hamidah Ibrahim, Zarina Dzolkhifli, and Praveen Madiraju</i>	
Finding Data Resources in a Virtual Observatory Using SKOS Vocabularies	189
<i>Alasdair J.G. Gray, Norman Gray, and Iadh Ounis</i>	
Progressive Ranking for Efficient Keyword Search over Relational Databases	193
<i>Guoliang Li, Jianhua Feng, Feng Lin, and Lizhu Zhou</i>	
Semantic Matching for the Medical Domain	198
<i>Jetendr Shamdasani, Peter Bloodsworth, and Richard McClatchey</i>	
Towards the Automatic Generation of Analytical End-User Tools Metadata for Data Warehouses	203
<i>Jesús Pardillo, Jose-Norberto Mazón, and Juan Trujillo</i>	

25th Anniversary Colloquium on Advances in Database Research

The Hyperdatabase Project – From the Vision to Realizations	207
<i>Hans-Jörg Schek and Heiko Schuld</i>	

From Schema and Model Translation to a Model Management System	227
<i>Paolo Atzeni, Luigi Bellomarini, Francesca Bugiotti, and Giorgio Gianforme</i>	
XtreemOS: Towards a Grid-Enabled Linux-Based Operating System	241
<i>Domenico Laforenza</i>	
High-Assurance Integrity Techniques for Databases	244
<i>Elisa Bertino, Chenyun Dai, Hyo-Sang Lim, and Dan Lin</i>	
Towards General Temporal Aggregation	257
<i>Michael H. Böhlen, Johann Gamper, and Christian S. Jensen</i>	
 Best Papers from the Workshop on Biodiversity Informatics: Challenges in Modelling and Managing Biodiversity Knowledge	
Distributed Systems and Automated Biodiversity Informatics: Genomic Analysis and Geographic Visualization of Disease Evolution	270
<i>Andrew W. Hill and Robert P. Guralnick</i>	
Visualisation to Aid Biodiversity Studies through Accurate Taxonomic Reconciliation	280
<i>Martin Graham, Paul Craig, and Jessie Kennedy</i>	
Author Index	293