Preface

This year marked the coming of age of the British National Conference on Databases with its 21st conference held at Heriot-Watt University, Edinburgh, in July 2004. To mark the occasion the general theme of the conference was "When Data Is Key", reflecting not only the traditional key awarded on a 21st birthday, but also the ever-growing importance of electronic data management in every aspect of our modern lives. The conference was run as part of DAMMS (Data Analysis, Manipulation, Management and Storage) Week, which included a number of co-located and complementary conferences and workshops, including the 2nd Workshop on Teaching, Learning and Assessment in Databases (TLAD2), the BNCOD BioInformatics Workshop, and the 1st International Conference on the Future of Consumer Insight Developments in Retail Banking. The aim of this co-location was to develop synergies between the teaching, research and commercial communities involved in all aspects of database activities, and to use BNCOD as a focus for future synergies and developments within these communities.

Although this is entitled the British National Conference on Databases, BNCOD has always had an international focus, and this year more than most, with the majority of the papers submitted and accepted coming from outwith the UK. We were fortunate in attracting over 70 paper submissions for the conference, of which 18 were accepted, reflecting a stringent and quality-driven peer-review process and maintaining the reputation of BNCOD as a conference reporting high-quality research to the community. However, in our aim to extend the inclusivity of the conference, we also introduced an application paper stream to encourage submissions less focused on research (this stream appears in Vol. 2 of the proceedings), and at the end of this volume you will find the three best papers from TLAD2, as identified by the reviewers for that workshop.

The conference itself reflected the ever-growing diversity of the database community, with papers reflecting traditional database research concerns, through XML to multimedia, but we have strongly resisted the temptation to diversify away from our core concerns. In particular, the importance of database research was reflected in the subjects raised by the two keynote speakers. Both of these subjects are of massive international significance and have achieved the kind of universal importance that permits them to be identified with the definite article – The GRID and The Web.

Our first keynote speaker, Domenico Laforenza, is Technology Director of the Information Science and Technologies Institute (ISTI) of the Italian National Research Council (CNR). An active researcher and promoter of high-performance computing and GRID technologies, he has been deeply involved at both national and European level in the development of GRID infrastructures in Europe and the oversight of future directions for that development. His keynote address "Towards a Next Generation Grid" considered the developments in GRID research and technology that have brought us to the present situation, and elaborated on Next Generation Grid(s) in 2005–2010, and the research opportunities that will result from them, with a particular emphasis on the convergence of multi-disciplinary research.

Our second keynote speaker, Michael Wilson, is Manager of the W3C Office in the UK and Ireland, and a Member of the EU IST programme Advisory Group (ISTAG). His role as Manager in W3C is to help achieve the goal of leading the Web to full potential through ensuring the interoperability of different proprietary systems. As an active researcher in knowledge engineering, HCI, multimedia and VR, he has the background and pedigree to undertake this daunting and complex task. His keynote address "The Future of the World Wide Web" brought together a number of key concepts for this conference, the use of Web services and the Semantic Web, and the relationship between Web and GRID technologies, and he also introduced the issue of trust as a key concept in the future development of all these services.

As identified earlier, we were fortunate in attracting high-quality papers across a range of related topics. The first set of contributed papers were concerned with the processing of queries applied to data streams. Qingchun Jiang and Sharma Chakravarthy introduced novel scheduling strategies to minimize tuple latency and total memory requirements, and presented experimental results on their efficiency and effectiveness. Dan Olteanu, Tim Furche and Francois Bry were concerned with an evaluator (SPEX) for querying XML streams for which the complexity results are polynomial in the sizes of the data and queries, compared to most other methods, which are exponential in the size of the queries.

In the area of integrating data from a set of heterogeneous databases, one problem that needs to be dealt with is that of multiple query languages. Damir Becarevic and Mark Roantree described the EQL language for querying OO and object-relational database schemas in a database- and platform-independent manner. A system that is being developed to perform data integration based on a schema transformation approach is called AutoMed; Lucas Zamboulis described how the integration of XML data sources is accomplished through graph restructuring of their schemas in AutoMed. Kajal Claypool and Elke Rundensteiner focused on the problem of change over time in a database and its effect on the mapping of one data model to another.

In data analytics/manipulations one type of query that would be very useful for data mining is the path query, which determines the relationships between entities in a database. Rachel Hamill and Nigel Martin described one way of extending conventional database technology to support this type of query. Weifeng Chen and Kevin Lu discussed how agent technology has been integrated with conventional data mining techniques to realize a system for financial data mining. In order to limit the damage caused by malicious attacks on a database, Indrakshi Ray et al. considered two techniques for rapid damage assessment and proposed a new one that improves on them.

XML was the focus of much of the work here. With it there was interest in new query languages and query processing techniques. G. Subramanyam and P. Sreenivasa Kumar presented a technique for multiple structural joins and showed that it performs better than existing join algorithms. Jung Kee Park and Hyunchul Kang were concerned with the use of caches to provide efficient support for XML queries, and focused on the problem of XML updates and the effect on cache answerability. Alison Cawsey et al. discussed the use of transformation constraints placed on XML documents by the information provider to constrain transformations applied to it, e.g., in personalization. On a different tack Richard Wheeldon, Mark

Levene and Kevin Keenoy were concerned with keyword search in relational databases and presented a novel algorithm for join discovery.

At the interface with the user, the automatic generation of data entry interfaces to databases usually follows a simplistic approach with limited constraints on data captured. However, Alan Cannon et al. described a semi-automatic tool for generating interfaces, which uses a domain ontology to reflect the semantics of the data, and improve the quality of the captured data. Linas Bukauskas and Michel Bohlen addressed the problem of scalability of visualization systems for large databases when data is extracted from a database and stored in a scene tree. They introduced two new data structures that improve scalability by eliminating the data bottleneck. Images may be stored in a multimedia database as sequences of editing operations. Leonard Brown and Le Gruenwald presented algorithms for performing color-based similarity searches on sets of images stored in this way.

The final session was on spatial databases and there were three papers on this theme. The performance of the M-tree, a data structure to support access to spatial data, depends on the degree of overlap between spatial regions. Alan Sexton and Richard Swinbank presented a new bulk loading algorithm to improve the performance of the M-tree and introduce a variant of it called the SM-tree. Another spatial data problem, the Obstructed Nearest Neighbour, is concerned with finding the nearest neighbours to a point in the presence of obstacles. Chengyi Xia, David Hsu and Anthony Tung presented an efficient algorithm for solving this problem. Despite much research on spatio-temporal data types, data models and query languages, little work has been done on complete spatio-temporal database systems. Tony Griffiths et al. addressed the latter problem with work on support for database programming in the spatio-temporal OODBMS Tripod.

Acknowledgements

We are deeply indebted to a number of different people and organizations without whom BNCOD21 would not have been possible. In particular we would like to thank:

- The members of the programme committee who gave willingly of their time and expertise and were most helpful, especially in having to deal with more papers than anticipated, and in managing to keep to the tight deadlines imposed on us.
- The invited speakers and authors who presented papers at BNCOD as well as those who submitted papers but were not successful. Without you there would have been no conference.
- The sponsors and exhibitors at the workshops and conferences during the whole week; your contributions, be they financial or in time and service, helped to enhance the quality of the experience for the delegates.
- The organizing committee who worked hard behind the scenes to ensure the success of the conference.

May 2004   Howard Williams and Lachlan MacKinnon

Conference Committees

Organizing Committee

L.M. MacKinnon	(Chair)	Heriot-Watt University
A.G. Burger	(BioInformatics Workshop)	Heriot-Watt University
E. Dempster	(Logistics)	Heriot-Watt University
R.G. Dewar	(Publicity, Website)	Heriot-Watt University
P.M.D. Gray	(Competition)	Aberdeen University
D. Lanc	(Data Financials Workshop)	Heriot-Watt University
P.W.Trinder	(Posters, Tutorials)	Heriot-Watt University

Programme Committee

Howard Williams (Chair)	Heriot-Watt University
David Bell	Queen's University, Belfast
Albert Burger	Heriot-Watt University
Richard Connor	Strathclyde University
Richard Cooper	University of Glasgow
Barry Eaglestone	University of Sheffield
Wenfei Fan	University of Edinburgh
Alvaro Fernandes	University of Manchester
Alex Gray	University of Wales, Cardiff
Peter Gray	University of Aberdeen
Mike Jackson	University of Wolverhampton
Anne James	Coventry University
Keith Jeffery	CLRC, Rutherford Appleton Laboratory
Jessie Kennedy	Napier University
Brian Lings	University of Exeter
Lachlan MacKinnon	Heriot-Watt University
Nigel Martin	Birkbeck College, University of London
Ken Moody	University of Cambridge
Werner Nutt	Heriot-Watt University
Norman Paton	University of Manchester
Alex Poulovassilis	Birkbeck College, University of London
Stratis Viglas	University of Edinburgh
Peter Wood	Birkbeck College, University of London

Steering Committee

Brian Lings (Chair)	University of Exeter
Barry Eaglestone	University of Sheffield
Alex Gray	University of Wales, Cardiff
Anne James	Coventry University
Keith Jeffery	CLRC, Rutherford Appleton Laboratory
Roger Johnson	Birkbeck College, University of London