modelling; extend a data modelling methodology to include dynamics.

Creasy, P.N. and Moulin, B. "Approaches to Data Conceptual Modelling," Proc. 6th Annual Workshop on Conceptual Graphs, Binghamton, 1991, pp.387-399.

Moulin, B. and Creasy, P.N. "Extending the Conceptual Graphs Approach for Data Conceptual Modelling," Data & Knowledge Engineering, 8(3), pp.251-258, 1992.

Creasy, P.N. "A Graphical Data Conceptual Modelling Language," Advances in Artificial Intelligence Research, Vol.2, pp.87-100, 1992.

Campbell, L. & Creasy, P.N. "A Conceptual Graphs Approach to Information Systems Design," 7th Annual Workshop on Conceptual Graphs, Las Cruces, 1992, pp.193-202.

Creasy, P.N. & Moulin, B. "Adding Semantics to Semantic Data Models." Current Directions in Conceptual Structures Research, 1992, pp.189-200.

Parallel Computation in Distributed Database Systems

Principal Investigators: Professor Maria Orlowska, Dr Yanchun Zhang, Dr Robert Colomb

Research Students: Mr Ok Cho, Mr Nirad Sharma, Mr Xiaofang Zhou

The purpose of this project is to improve the performance (response time) of very large distributed database systems by using general-purpose multiprocessor computers to process time-consuming relational database operations. We plan to design efficient parallel algorithms for relational database operations on different parallel architectures which include: Massively parallel SIMD computers; MIMD parallel computer; Transputer-based other topological structures; LAN-connected workstations. We plan to investigate parallel query optimisation using parallel processing servers. The main results in the current state of this project are: the analysis and characterization of data skew problems; the develop of a parallel join algorithm and its tests on the MASPAR, and a parallel computation of a transitive closure algorithm.

Zhang, Y. & Orlowska, M.E. "Parallel Processing for the Full Reduction of a Chain Query in Distributed Databases", Information Systems, 18(3), pp.181-187, 1993.

Zhou, X.F., Zhang, Y.C., & Orlowska, M.E. "A Parallel Transitive Closure Algorithm for SIMD Meshes", ACSC'16, 17(1), Brisbane, pp.143-151, 1993.

Zhou, X.F., Orlowska, M.E. & Zhang, Y.C. "Transitive Closure Evaluation of Database Relations on Massively Parallel Computer MasPar MP-1", Joint Symposium on Parallel Processing, Tokyo, Japan, 1993.

Zhou, X.F., & Orlowska, M.E. "A New Dynamic Approach for Handling Data Skew Problems in Parallel Hash Join Computation", TENCON'93, Beijing, China, 1993.

Active Databases

Principal Investigator: Dr John Harrison

A database can be enhanced by adding support for production rules resulting in what is called an "active" database. Each production rule is a declarative representation of the expression: "when an event E occurs then perform action A if condition C is true". Active databases offer a uniform mechanism for supporting materialized view maintenance, integrity constraints, database security, alerters and triggers. If extended to support deductive rules, an active database can react to changes to derived data as well as stored data. Our research addresses active and deductive databases and also their integration. An active, deductive database system is currently under development. The event detector and condition evaluator have been implemented.

Harrison, J. V. & Dietrich, S. W., "Towards an Incremental Condition Evaluation Strategy for Active Deductive Databases", Databases '92, Melbourne, Australia, 1992, pp.81-95.

Harrison, J. V. & Dietrich, S. W., "Maintaining Materialized Views in Deductive Databases: An Update Propagation Approach", Joint Int. Conf. & Symposium on Logic Programming - Deductive Database Workshop, Washington DC, 1992, pp.56-65.

Harrison, J. V. & Dietrich, S. W., "Integrating Active and Deductive Rules", 1st Int. Workshop on Rules in Database Systems, Edinburgh, 1993.

propositional rule-based system into a decision table and further into a decision tree, which gives great advantages in performance and also allows much easier analysis of some important properties of the system. Further, deductive database can be profitably implemented using a main-memory database system, which in its turn can be accelerated with fine-grained parallel hardware. The project includes research into transformation technology and into algorithms and architectures for main-memory implementations.

It is also possible to use deductive database technology to implement repositories for CASE tools, and to specify complex data structures. Projects are under way in both areas, the latter with reference to specification of complex data structures in an open distributed computing environment.

Colomb, R.M. (1992) "A Join Algorithm for a Main-Memory Database Operating on a Fine-Grained Parallel Computer" 3rd Australian Database Conference, Melbourne.

Colomb, R.M. (1992) "Computational Stability of Expert Systems" Expert Systems With Applications, 5(2/3), pp.411-419.

Graphical Conceptual Schema Languages

Principal Investigators: Dr Peter Creasy, Dr Wolfgang Hesse (Marburg University, Germany), Dr Bernard Moulin (Laval University, Canada)

Research Students: Mr Alistair Barros, Ms Linda Campbell, Mr Gerard Ellis

A number of projects are being carried out under the general heading of graphical conceptual schema languages. In general the goal is to find representations to capture as much of the semantics of the UoD as possible. The research involves two modelling tools: NIAM and conceptual graphs. Some of the projects are being carried out with personnel from other universities. Most fact-based language conceptual schema languages with a graphical representation, use only a few validation constructions (e.g. mandatory role, uniqueness). While we can represent what we regard as some of the more important constraints, there is still a large class which cannot be represented. In addition subtype definitions and derivation rules usually have no formal representation. Most of the presentations in the integrity constraint area use first order predicate logic, a language unfamiliar to most analysts. Existential graphs have been developed in a number of forms which correspond to propositional, first order and higher order logics. The graphs provide a very readable logic representation in graphical form which can be readily integrated with graphical knowledge representation languages. The project investigates the incorporation of existential graphs into fact-based languages, to permit subtype definitions, derivation rules and, in particular, constraints to be formally graphically represented. Techniques for efficient computation of the constraints is also being investigated, as well as the use of the extended language as the IRDS conceptual schema language. NIAM seems to have lost out as a modelling language standard to ER. In part, this is due to its semantic richness, which has the disadvantage of complicating most conceptual schema diagrams. The project investigates a two level approach to NIAM. This approach can be extended to make NIAM an object oriented language. Conceptual graphs are a powerful logic based too used in a wide range of computing disciplines. The project has investigated ways of using the graphs (with extensions) in conceptual modelling. Proposals have been made to use the graphs as a core conceptual modelling language. This is similar to the recent ANSI proposals for using the graphs for the IRDS conceptual schema. Mapping between ER, NIAM and the graphs and their use in schema integration have been shown. This project is part of the world-wide PEIRCE project which is developing a conceptual graphs workbench. The descriptive and proscriptive modelling of a Universe of Discourse (ie an application area) is related to static and dynamic aspects. The static aspect is concerned with the structure and integrity of data at any given state. The dynamic aspect is concerned with the processes (generally defined at various levels of abstraction) which act on data, and events which change data from one state to another. Clearly, consideration of data, process and event perspectives for any information system is necessary. Whilst the data perspective is fundamental (since the other perspective act on it), the degree of emphasis on any one perspective depends on the nature of the Universe of Discourse. The project proposes to: determine various classes of problems encountered in data modelling and database design which require support for dynamics; provide a theoretical foundation for inclusion of dynamics with data

Efficient Test for the Valid Distribution of Knowledge Databases", Int. J. Software Engineering & Knowledge Engineering, 2(4), pp.589-609,1992.

Lin, X.M., Orlowska, M.E. & Zhang, Y.C. "A Graph Based Cluster Approach for Vertical Partitioning in Database Design", Data & Knowledge Engineering, 1993.

Zhang, Y. and Orlowska, M.E. "The effect of unary inclusion dependencies on relational database design", Int. J. Computers & Mathematics with Applications, 24(3), pp.49-59, 1992.

Zhang, Y. and Orlowska, M.E. "An improvement on the automatic tool for relational database design", Information Systems, 15(5), pp.647-651, 1990.

Ewald, C.A. & Orlowska, M.E. "A Procedural Approach to Schema Evolution", 5th CAiSE'93, Springer- Verlag LNCS 685, Paris, 1993.

Information Systems Engineering

Principal Investigator: Dr Terry Halpin, Dr Robert Colomb

Research Students: Mr Anthony Berglas, Mr Anthony Bloesch, Ms Linda Campbell, Mr Jon McCormack, Mr John McMillan, Mr Peter Ritson

The primary objectives of this project are to extend conceptual modelling methods for information systems and to provide automated support for information systems engineering. An object-role modelling method based on a formalization of extended NIAM has been developed, together with an associated language (FORML) and a CASE prototype - this part of the research was carried out jointly with ServerWare (USA). Algorithms have been specified to map conceptual schemas to internal and external schemas, including many constraints ignored by typical approaches. Substantial work has been completed on schema optimisation, and automated support for natural verbalization and extended subtyping. An extensible and efficient theorem prover (Genesis) has been developed for a variety of classical and non-classical logics, and its application to database modelling is being explored (e.g. schema transformation and temporal modelling). Related research is currently under way on abstraction mechanisms, process-event modelling, dynamic rules, re-engineering, method extensions, conceptual query languages, metamodelling, object-oriented aspects, and model management systems.

Bloesch, A. & Halpin, T.A. 1993, ‘An Object-oriented Approach to Extensible Parsing', Proc. ACSC'16, 15(1), Brisbane, Australia, pp.345-51.

Campbell, L. & Halpin, T.A. 1993, ‘Automated Support for Conceptual to External Mapping', Proc. 4th European Workshop on Next Generation CASE Tools, Paris.

Halpin, T.A. 1991b, ‘A fact-oriented approach to schema transformation', Proc. MFDBS-91 Int. Conf. on Mathematical Fundamentals of Database & Knowledge Base Systems, Spinger Verlag LNCS 495, Rostock, Germany.

Halpin, T.A. 1991c, ‘WISE: a Workbench for Information System Engineering', Proc. 2nd European Workshop on Next Generation of CASE Tools, Trondheim.

Halpin. T.A. & Orlowska, M.E. 1992, ‘Fact-Oriented Modelling for Data Analysis', J. Information Systems, 2(2), pp.1-23, Blackwell Scientific, Oxford.

Halpin, T.A. & McCormack, J.I. 1992, ‘Automated Validation of Conceptual Schema Constraints', 4th CAiSE'92, Springer-Verlag LNCS 593, Manchester, pp.445-62.

Halpin, T.A. 1992b, "Fact-Oriented Schema Optimization", Proc. CISMOD'92, Bangalore.

McCormack, J.I., Halpin, T.A. & Ritson, P.R. 1993, ‘Automated mapping of conceptual schemas to relational schemas', 5th CAiSE'93, Springer-Verlag LNCS 685, Paris.

Berglas, A. "The Intelligent Developer - towards the automated implementation of information systems." In M.E. Orlowska and M. Papazoglou, editors, Advances in Database Research. World Scientific, Singapore, 1993.

Applications of Deductive Database

Principal Investigator: Dr Robert Colomb

Deductive database is a logical framework for many applications in database and knowledge-based systems. In particular, it is possible to transform a

Another aspect in this project in the concurrency control in MDBS. There are a number of proposed approaches to the autonomous concurrency control, independent from global concurrency protocols. The group is investigating a system's performance improvements. We have constructed a hybrid approach which has a simple global concurrency control like the optimistic method, and has less blocking phenomenon at local sites. The notion of a transaction in "pure" multidatabase system is proposed to be modified, especially when the results from the first part of the project will enhance an enforcement of external constraints. This project is planned for next few years, with the intention to use a simulation as a method for assessing different theoretical solutions.

Finally, there is some work in progress on query optimisation. We have show that some techniques from the centralised cases of relational databases can be extended to cover aspects in multidatabase systems.

Cardiff, J. & Orlowska, M.E. "A geometric approach to semantic query optimisation in relational databases", Proc. Methodologies for Intelligent Systems, 5, Elsevier Science Publishing Co. Inc., pp.126- 134, 1990.

Zhang, Y.C. & Orlowska, M.E. "On Optimistic Concurrency Control in Multidatabase Systems", Proc. Int. Conf. on Parallel and Distributed Systems, Taiwan, pp.456-465, 1992.

Zhang, Y.C. & Orlowska, M.E. "A Hybrid Concurrency Control Approach in Heterogeneous Distributed Database Systems", TENCON'93, Beijing, China, 1993.

Orlowska, M.E. & Zhang, Y.C. "A Linear Model of Distributed Query Execution Strategies", Proc. 2nd Far-East Workshop on Distributed Database Systems, Kyoto, Japan, pp.82-93, 1992.

Jia, X.H. & Zhang, Y.C. "An Optimistic Algorithm for Concurrent Accesses to a Single Replicated File", Information Processing, Vol.1, pp.365-371, 1992.

Design Issues in Distributed Database Systems

Principal Investigators: Professor Maria Orlowska, Dr Yanchun Zhang, Dr Xuemin Lin, Dr Robert Colomb, Mr Jacek Szymczyk

Research Students: Ms Cathy Ewald, Mr Mohsen Rohani

The title of this project is intended to cover a broad class of related problems such as: conceptual schema integration (partial or complete), schema evolution, data fragmentation and the allocation to network sites. Database design may be viewed as an evolutionary process, since a schema may be constructed by repeatedly adding single facts. Similarly, schemata may be integrated by adding one, fact by fact to another. The aim of our research is the development of a unified methodology for database design and integration. This is expected to take the form of a merge operator which takes two correct, well formed, homogeneous conceptual schema diagrams as input and combines them to produce a correct well formed integrated schema. Our current work concentrates on fact-based conceptual schemata designed using the NIAM methodology. It is expected that the procedures developed will be extended to deal with the integration of multiple heterogeneous database schemata. Procedures for insertion and deletion of facts have already been developed. We will also construct other operators for schema evolution, including the insertion, deletion and modification of constraints. Techniques for improving the schema, including the elimination of derivable facts, are under consideration.

This project is also an investigation of the problem of allocation of database fragments to a network, so that the overall communication cost for processing a given set of transactions is minimized. We have developed two data allocation algorithms. The first is a data allocation algorithm with a simple strategy to process the given transactions. Several interesting optimality estimation results of this algorithm are also presented, while the computational intractability of the complete optimisation with respect to a simple given strategy has been shown. Second, a framework of a dynamic data allocation algorithm has been proposed. This starts from an initial data allocation and then iteratively and greedily, reduces the overall communication cost by either adding or dropping a copy of a file based on an existing optimizer - a collection of distributed query optimisation algorithms. This iteration is guaranteed to converge to a local optimal data allocation.

Zhang, Y.C & Orlowska, M.E. "A new polynomial time algorithm for BCNF relational database design, Information Systems, 17(2), pp.185-192, 1992.

Zhang, Y.C., Orlowska, M.E. & Colomb, R.M. "An

Database Research at the University of Queensland

Maria E Orlowska
Department of Computer Science
The University of Queensland
St Lucia Queensland 4072 Australia

Email: maria@cs.uq.oz.au

Introduction

The University of Queensland - a leading Australian university - is located on a scenic bend of the Brisbane River. Its Computer Science Department is one of the largest in the country, with a full-time academic staff of 27 - including five full professors - 30 research staff, 23 support staff and 50 part-time tutors. There are currently over 50 PhD students undertaking research. In 1988 the Australian government awarded the Department the status of Key Centre for Teaching and Research in Software Technology. Three important research centres exist within the Department: the Distributed Systems Technology Centre which performs research in distributed architectures, distributed databases, software tools, management, security and performance; the Centre of Expertise in Distributed Information Systems which conducts research into the design and implementation of distributed information systems; and the Software Verification Research Centre which works on improved methods and tools of industrial significance for developing verified software. Obviously, research is an essential part of the Department's activity and is concentrated in the areas of: database systems, information systems, distributed systems, programming languages, algorithm design, software verification research, cognitive science, and system design.

This report presents the activities of the Database/ Information Systems group, led by Professor Maria Orlowska, of the Computer Science Department. The main areas of research are: distributed database systems, multidatabase systems, conceptual issues in information systems engineering, transaction management in distributed environments, applications of deductive databases, parallel computation and performance of database systems and active databases.

A brief description of major projects, followed by a list of selected recent publications is presented in the sections below.

Transaction Management in a Distributed Environment

Principal Investigators: Professor Maria Orlowska, Dr Yanchun Zhang, Dr Dean Kuo, Mr Jacek Szymczyk, Dr Xuemin Lin

Research Students: Mr John Gilmore, Mr Andrew Goodchild, Mr Ralf Muhlberger, Mr Darryn Reid, Mr George Semeczko

In this project the major focus is on Multidatabase Systems (MDBS) consisting of a number of independently designed databases. The fundamental problem here is establishment of methods or at least some justified groundwork for the conceptual analysis of systems involved in the interoperability. We investigate the interdatabase dependencies, their specification and relationship to the major most frequently used transactions. To maintain, i.e. to enforce the global consistency of a multidatabase system, without - as we assume - existence of a global conceptual schema, the analysis of interdatabase dependencies is necessary. The problem is hard, it can not be fully automated. The progress achieved is very promising; two papers have been submitted. The project is scheduled for the next 2-3 years.