CSIR

our future through science

Technologies for conceptual modelling and intelligent query formulation

R ALBERTS[1], K BRITZ[1], A GERBER[1], K HALLAND[1,2], T MEYER[1], L PRETORIUS[1,2]
(1) Knowledge Systems Group, Meraka Institute, CSIR, Pretoria, Gauteng, South Africa
(2) School of Computing, University of South Africa, Pretoria, South Africa

BACKGROUND

- Improvements in hardware technology in the last decade have opened the way to a new generation of information systems involving complex objects that may integrate components of various kinds: (hyper)text, images, graphics, video, and audio
- The web and digital libraries are becoming the main source of this kind of information for a potentially enormous community of users
- Research and development in this area is of strategic importance, as witnessed by the amount of effort that is converging on the field.

EXISTING PROBLEMS

- Two major problems that arise while designing and managing information systems are conceptual modelling and query management
- In this context, these problems become more complex and challenging, and require new ideas, technologies and tools to be solved.

CONCEPTUAL MODELLING

- In most information system design methodologies - and in particular when designing these new generation systems - the conceptual modelling of the application domain plays an important role
- This phase must be preceded by an accurate collection of requirements, both with respect to the structure of the data and with respect to the operations that need to be performed on them.

GOOD CONCEPTUAL MODELS

- Good conceptual data models put their emphasis on the correct and semantically rich representation of complex relations that may exist between data
- Good conceptual models allow for an abstract representation of data that resembles the way the data is actually perceived and used in the real world, thus shortening the semantic gap between the domain and its representation.

QUERY MANAGEMENT

- The problems of information access and query management in the new generation of information systems do not appear to be close to a solution, because they involve the content of complex documents, that is inherently difficult to understand and model
- The availability of large amounts of data that are heterogeneous in structure and origin has made this problem still more pressing: in this case an excess of information can be equivalent to an absence of information
- It is therefore necessary to use tools that organise data into intelligible and easily-accessible structures and return answers at various levels of detail to support analysis and decision-support.

AIM

The aim of the project is to devise and evaluate algorithms, methodologies, techniques and interaction paradigms to build a tool for conceptual modelling and query management of complex data repositories based on a framework with solid formal foundations. This will contribute to a better understanding of the cross-fertilisation between knowledge representation and database technologies.

OBJECTIVES

- To design and implement a tool supporting the design, management and storage of conceptual schemas and ontologies, connected in the background with a description logic inference server. This is based on a formal framework and a methodology for conceptual modelling borrowed from the most recent ideas from description logic research
- To design and implement a tool for query management extending the above, connected in the background with a description logic inference server extended with query processing capabilities. This is based on a formal framework for query management related to the conceptual modelling

methodology previously established, in agreement with the recent research results on query management paradigms that exploit the knowledge in the ontology.

RESEARCH IN PROGRESS

- In the first strand, a methodology for conceptual design will be devised and a tool supporting it will be implemented
- In the second strand, the query management problem will be studied - in the presence of the previously devised conceptual model: a global framework will be introduced, various basic algorithms for the different tasks involved in query management will be considered, and a tool, integrated with the conceptual modelling tool, will be implemented.

CURRENT STATUS

A number of interesting ideas have been published, various partial prototypes have been implemented, but a set of tools based on a coherent global framework is still lacking.

OUTCOMES OF THE PROJECT

- The establishment of a common framework with solid formal foundation for conceptual modelling and query management
- The design and implementation of tools that demonstrate the validity of these ideas.

Intelligent query answering is indispensable in any fieldwork situation, where relevant information has to be identified from vast information resources.