GLOBAL INTERNET BUSINESS

Update on Best Practice ...
User Experience ...
Oracle Products & Futures

www.eoug.com

CONFERENCE PROGRAMME

EOUG Conference & Exhibition 2000
Madrid / Spain, June 21st - 23rd

European Oracle User Group

EOUG

In Cooperation with Oracle Corp.

ORACLE

FUJITSU COMPUTERS SIEMENS

hp invent

COMPAQ


CONFERENCE
PROCEEDINGS 2000
www.eoug.com
EOUG Conference & Exhibition
EOUG Conference & Exhibition 2000
Madrid / Spain, June 21st - 23rd
European Oracle User Group
EOUG
In Cooperation with Oracle Corp.
ORACLE
FUJITSU COMPUTERS
SIEMENS
hp
invent
COMPAQ

EOUG CONFERENCE & EXHIBITION 2000
Madrid, Spain, June21st - 23rd

Welcome

"Global Internet Business" ...
Update on Best Practice ... User Experience ... Oracle Products & Futures

FUJITSU COMPUTERS SIEMENS

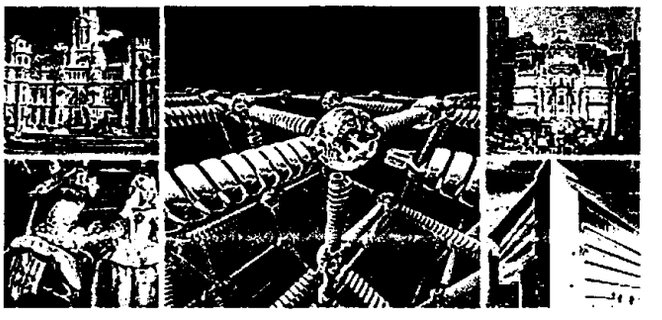
hp invent

COMPAQ

Conference Info

Programme

Registration



Exhibition

Hotels & Excursions

Tutorials

Programme Overview

How to contact the EOUG

ORACLE

Programme Streams

SERVER TECHNOLOGY & DBA TOPICS
Incorporating presentations of interest to all those who have the task of keeping Oracle databases 'Up & Running', efficiently & continuously. Includes how to maximise the options on the latest releases of Oracle8 & Oracle8i.
Abstracts: SERVER TECHNOLOGY & DBA TOPICS

eBUSINESS & WEB ENABLED COMPUTING
Targets the rapid developments in eBusiness and Web based computing. How to develop applications for it; how is it evolving.
Abstracts: eBUSINESS & WEB ENABLED COMPUTING

DEVELOPING APPLICATIONS
The stream for applications Designers and Developers. Focuses on the use of Designer, Developer, Reports, Forms & integration with other packages.
Abstracts: DEVELOPING APPLICATIONS

USING JAVA, OBJECTS & COMPONENTS
This stream focuses on the use of the Java language, Objects & Components for development. It covers the Applications Server, the JDeveloper suite and Oracle8i where appropriate.
Abstracts: USING JAVA, OBJECTS & COMPONENTS

BUSINESS INTELLIGENCE & DATA WAREHOUSING
This stream addresses all aspects of Decision Support. It includes the use of Data Warehouses, Repository, Data Marts, Discoverer etc.
Abstracts: BUSINESS INTELLIGENCE & DATA WAREHOUSING

IT MANAGEMENT & STRATEGIC PLANNING
Addresses general Information Technology management challenges.
Abstracts: IT MANAGEMENT & STRATEGIC PLANNING

ORACLE MINI LESSONS
Presented by Oracle Education, these highly concentrated educational sessions offer a combination of technical and training insights which often play to full houses.
Abstracts: ORACLE MINI LESSONS

ORACLE APPLICATIONS
Covers Oracle own package solutions including CRM, ERP Financials etc.
Abstracts: ORACLE APPLICATIONS

PARTNER SOLUTIONS
Focuses on solutions provided by Oracle's Business Alliance Partners. This includes how Oracle software performs on particular hardware platforms or how software designed by Business Partners provides solutions to specific problems.
Abstracts: PARTNER SOLUTIONS

ORACLE ON LINUX
Focuses on the developments in the fastest growing operating system in the Oracle environment.
Abstracts: ORACLE ON LINUX

TELECOMS & MEDIA
Targets developments in four specific industries, namely: Financial Services, Telecommunications and Media, Utilities and Public Services (i.e. services provided by Central and Local Government).
Abstracts: TELECOMS & MEDIA

ORACLE INDUSTRIES
Abstracts: ORACLE INDUSTRIES

Programme Guide

For the EOUG Conference 2000 we received more than 450 abstracts to review for the 250 slots planned. The quality of the abstracts submitted seems to improve each year, which makes the task of selection that much harder!
This year's abstracts again came not just from the EMEA region, but from the USA, Canada, Asia and Australia ... giving us presenters from 165 companies from over 30 countries.

Our thanks as always to our friends in Oracle for their excellent contribution to the programme.

The difficult task of deciding which papers to accept for a comprehensive but balanced Conference Programme falls to the EOUG Paper Selection Committee - an international team of Oracle Users with a wide range of experience. This year's Paper Selection Committee consisted of:

Ronnie Billen, Switzerland
Bert Spencer, UK
Carl Dudley, UK
Gerard Uiterwaal, The Netherlands
Dr. Alexei Golossov, Russia

Our thanks to them for assembling such a full and comprehensive programme.

To help identify those presentations of particular interest to you, we have published the full abstracts.

As in previous years, the presentations accepted for the EOUG Conference 2000 are sorted by Streams for your convenience.

Please note that the presentations listed on the following pages, their abstracts published on our web site, and their allocation to a certain stream **are subject to change**. Please refer to the full abstracts, which we will update if changes occur.

On our Web Site you will also find Session Preference Form for each Programme Stream. Please take a few minutes to complete it when you register, since that information is vital in helping us schedule the presentations, so that the size of the audience matches the size of the presentation rooms!

STREAM: SERVER TECHNOLOGY & DBA TOPICS

#003 Fifteen Key Tips for the Oracle DBA New to Windows NT
Robert Farrington, Oracle UK

Abstract: This paper describes several key areas the Oracle DBA on Microsoft Windows NT should be aware of, especially if moving to Windows NT from a UNIX background. The areas covered are based on questions received by Oracle Support Services, with the intention of helping the DBA utilise Oracle-related Windows NT features more effectively. Emphasis is placed on practical, easily to implement guidelines that are known to give proven benefits.
Topics:

- Windows NT Services
- Using the Control Panels
- Processes and Threads
- Filesystems
- Memory Resources
- Editing the Registry
- Install Accounts
- Relinking
- Number of Oracle Homes
- Automatic Startup and Shutdown
- Batch Jobs and Backups
- Raw Partitions
- Clustering and High Availability
- Built-in Diagnostic and Tuning Utilities
- Optional Diagnostic and Tuning Utilities

#004 Administrating SAP R/3 databases with Oracle 8.
Jan van Cruchten, Royal KPN (The Netherlands)

Target audience: Experienced or Advanced
In this presentation I want to tell about the things an oracle DBA should do to keep a SAP production system in good shape. SAP/R3 has a complex db structure. For example is SAP release 4.0B you have about 13.000 tables and 15.500 indexes to handle.
Subjects of the presentation:

- Introduction
  Speaker is sr. DBA and has 4 years experience with managing SAP R/3 Oracle environments.
  Together with 2 other dba's we maintain 22 production SAP systems, based on SLA agreements.
- House keeping
  What we do daily, weekly...
  What we have automated
  What and how we monitor proactively our environments
  Backup and recovery with SAP tools
- The role of the (oracle 8) cost-based optimiser in a SAP R/3 system
  How we feed the optimiser with actual data
- What to do if you want your system 7x24 hourly running
  Hardware failover- and software replication systems
- Big transports of data from production environments to project/ or acceptance systems without oracle utility's export vs. import
- What you should reorg and what not. What are the benefits ...

#005 Let's Tune Oracle8 For NT
Marlene Theriault, DBMentor, USA

When it comes to tuning an Oracle8 or 8i database, the Windows NT environment poses a whole new set of problem for the DBA. Since Oracle on Windows NT is a single executable with multiple, threaded processes, how does the DBA approach tuning this machine and the databases that reside on it?
This presentation will examine the approaches and tools available for tuning a database on Windows NT as well as some of the impacts of running more than one database on an NT server.

#009 Listen Here! (A Beginner's Guide to Networking)
Marlene Theriault, DBMentor, USA

Oracle Corporation provides the ability to communication with an Oracle database through the use of either SQL*Net2 and/or Net8 (used for both versions 8.0 and 8i).
How do you "talk" to your databases? What are the necessary components that are used to make Oracle networking, using SQL*Net2 or Net8, work and what do you, as a DBA, need to know to successfully configure the listener?
This presentation will provide a detailed look at Oracle networking is set up and will help you to understand the terms used as well as showing you how to work with the networking products.

#010 DBA 101: A Refresher Course
Marlene Theriault and Rachel Carmichael

There are many tasks that a Database Administrator should perform on a routine basis. Often, Oracle documentation is unclear or does not address these tasks well. New DBAs are either unaware that the tasks should be performed or do not know how to do them. Even senior DBAs sometimes are unaware of actions that they can take to make their job easier or more effective.
This paper is a collect of tasks and overview information that DBAs should know. The information was gathered by two DBAs over a period of several years and includes:

- Those things you do (actions which a DBA takes upon creating a database)
- Scripts to gather information to pro-actively monitor and document your database(s)
- Tips and traps which can cause a DBA problems and many wasted hours and much more.

#016 Dealing with Users, a Guide to the DBA_ Views
Rachel Carmichael DBAmentor.com, USA

As DBAs, we spend a great deal of our time worrying about the physical world of our database. Tablespace, table and index fragmentation are monitored daily, response times are carefully checked and performance tuning is done on a regular basis. The users of the system and their access and privileges tend to become the neglected stepchild of the overworked DBA. Oracle has provided a set of privileges and controls to help the DBA manage and maintain control of the users world and a set of data dictionary views to monitor the controls we impose. This paper will review the DBA_ views that relate to the users' world, explaining the new features and controls we have available in Oracle 8i to manage and monitor access to our databases.

#024 Oracle8: GIS (Geeky Internal Stuff: Data Block Internals)
Dan Hotka Quest Software, UK

This Segment of GIS will cover Oracle8 data block internals. The presenter offers an inside look at Oracle data blocks, how they work, how row data is physically stored, the high water mark, and an understanding of how space is allocated, managed and controlled. This presentation is a MUST for developers and DBAs.

#026 Now that I have the Statistics, Where Do I Go From Here?

Dawn Campbell, Quest Software, USA

Oracle's ability to validate the structure of an index or index partition provides the DBA with a set of statistics detailing the structure of the B-tree index. These raw statistics leave even the experienced DBA asking the question: Now that I have the statistics, where do I go from here? This paper focuses first on how to build a historical perspective of these index statistics. Second, the paper helps the DBA answer the question of whether an index should be rebuilt, coalesced, or left alone. The paper details the internal structure of the B-tree index, common index modification patterns, and how trending branch utilization, leaf utilization, optimal vs. actual height helps the DBA determine which indexes are candidates for rebuilding or coalescing. Two scenarios that utilize these concepts are shown.

#029 Oracle on UNIX and Oracle on Windows NT: Comparisons and Contrasts
Tim Foxon, Metron Technology Ltd, UK

Oracle is very well established as the leading database on UNIX systems. Oracle users and administrators are therefore familiar with the architecture of Oracle on such platforms, and also with the tools and techniques for deriving maximum benefit from their database. By contrast, Oracle on Windows NT is less widely understood.
This presentation examines the similarities and differences between Oracle on UNIX and Oracle on Windows NT, from a performance perspective.
It discusses, among other issues:

- The architecture of the product on the two platforms
- The range of tools that are available for database administration
- The sources and accuracy of the available performance and resource consumption data
- The scalability of the product on NT including a section on clustering technologies

A comparison will also be made between Oracle on Windows NT and Microsoft's own flagship RDBMS, SQL Server 7.0.

#042 Database Replication: Off-Load Production Systems to Enable Maximum Performance
Vince Lucey , VERITAS Software Corporation, UK

In today's global, 24x7, business environment, system downtime windows are rapidly becoming a thing of the past. Administrators managing critical Oracle database applications find it difficult to perform ongoing maintenance tasks, such as backups, reporting and data extraction. The larger the database, the more difficult it is to perform these operations without degrading production performance.
The solution is moving these tasks to another system with an up-to-date copy of the production data, with minimal impact on the primary system. Database Replication provides an easy, automated way to maintain a replicated Oracle database instance for off-host processing. This approach not only allows you to perform backups and reporting against the replicated system, you can also change the replicated data for testing or training. New advances in replication techniques support resynchronisation of the systems quickly and easily, with minimal impact on the production server. Today's hardware independent solutions offer more flexibility, a less disruptive installation and easier cost justification.

#057 Planning Your Move to the Cost-Based Optimizer in Applications Release 11i
Robert Farrington, Oracle UK

This paper gives a comprehensive introduction to the essential features of the Cost-Based Optimizer (CBO), use of which is required by Oracle Applications release 11i. The presentation is aimed at Applications DBAs, developers and managers planning to move to release 11i, who may not be familiar with the CBO. The intention is to provide a framework for the detailed planning and testing required before moving to Applications release 11i and hence the CBO. Particular emphasis is placed on the management strategies that are needed to utilise the CBO effectively.
Topics:

- Query Processing Overview
- Query Optimization Strategies
- CBO Terminology
- CBO Principles of Operation
- Initialization Parameters
- Introduction to Statistics
- Using Histograms
- Troubleshooting Checklist
- Coding Style
- Using Hints

#061 Oracle8i Parallel Server in Enterprise-Level Clusters
Paul Baumgartel, Adept Computer Associates, USA

Abstract: In order to minimize the risks of catastrophic business loss, critical databases must be deployed within computing environments that are both highly available and easily expandable. The computing paradigm that best solves this problem is the cluster.
Oracle has provided the Parallel Server option for use in clustered systems for several years. With the release of Oracle8i, Oracle has dramatically improved this technology with such innovations as Cache Fusion, integrated Distributed Lock Manager, and enhanced inter-process communication.
This presentation will explain how properly designed Parallel Server databases can meet the challenge of enterprise-level systems in e-commerce and critical internal systems. It will discuss how to configure Oracle8i Parallel Server on a clustered system to provide a fault-tolerant, high availability, high-performance computing environment. It will show how systems architects can increase RDBMS throughput, provide load balancing, and leverage clusters' inherent robustness and redundancy.

#067 Implementing Snapshot Deployment Templates in Oracle 8i
Zoran Jovanovic, IN2 d.o.o., Croatia

In Oracle 8i new feature Snapshot Deployment Templates is introduced which simplifies deployment of snapshot groups on multiple remote sites. Using this feature you can define a Snapshot Deployment Template on a master site, which contains definitions of all snapshots, and other objects (tables, views, indexes...), which you want to create on multiple remote sites. Besides that Snapshot Deployment Template can contain snapshot data, definitions of users who are authorized to instantiate template at remote sites and user parameters for snapshot subsetting.
Once that the template is defined you need to create a script to instantiate it. You store this script to a file, transfer it to a remote site and instantiate it there. Thereafter your snapshot group (one or more) is created and you can continue refreshing it using fast snapshot refresh if possible.
In my presentation I want to present how I have implemented Snapshot Deployment Templates in one insurance company with distributed subsidiaries.

#077 Creating High-Availability applications using Oracle Failsafe v3
Alan Ball, Oracle Corporation, UK

Previous releases of Oracle Failsafe concentrated on providing high availability for the Oracle database server running on a Windows NT cluster. With version 3, the product has been enhanced to enable the inclusion of Forms, Reports, Application servers and whole web sites into the Failsafe managed environment. This session will briefly review the new functionality, and then describe how it can be used in a practical application. Particular reference will be made to those features of the Oracle 8i server that enhance the failover mechanisms of Oracle Failsafe.

#078 More Tips and Techniques for DBAs and Developers
Ari Kaplan, Independent Consultant, USA

This presentation, brought on by demand from the original "Tips and Techniques" presentations, will demonstrate techniques (not covered in the original presentation series) used by expert DBAs and developers to become more productive in their environments. There are dozens of little-known methods that are frequently asked for but not documented well, if at all. This presentation will focus on some unique tricks and tips for getting more out of Oracle. Examples include "gotchas" of dropping and recreating tables, seeing hidden init.ora parameters, identifying and deleting duplicate records, changing the SQL prompt, preventing passwords from being seen in the operating system, and many, many more topics.

#079 A Bag of Tips and Techniques for DBAs and Developers
Ari Kaplan, Independent Consultant, USA

There are dozens of little-known facts and methods that most Oracle professionals frequently ask for to improve their environments. Many are not documented well, if at all. This presentation will focus on getting more out of Oracle by showing powerful techniques used by DBAs and developers to become more productive in their environments. Examples include using SQL to generate SQL, dumping the contents of a table to an ASCII file in a tab-delimited format, finding and deleting duplicate records in SQL, getting around some limitations of the LONG datatype, connecting to a database without knowing a user's password, and many, many more topics.

#094 Explain Plan explained.
Jonathan Lewis, JL Computer Consultancy, UK

The ultimate effectiveness of an Oracle-based application is inevitably down to the efficiency of the database design and subsequent data access. The key, practical, tool that Oracle offers to assist in this task is the EXPLAIN PLAN utility, which takes in SQL statement as an input and produces a description of the access path that the database engine will use to execute that SQL. Without a thorough understanding of what execution paths are available, and how they work, it is virtually impossible to make sensible decisions about the physical database design.
This presentation will start by describing the various sources of execution plans, the mechanism for invoking them, and the cost, risk, and accuracy of the different sources. The second part of the presentation will describe the different mechanisms available to Oracle for accessing data from single table, and for joining two tables. Finally we will walk through a complex plan applying the two simple principles of 'access path' and 'join mechanism' layer by layer to determine the execution path that Oracle would take.

#118 Oracle and high availability
Maarten de Vos, Origin/Nederland BV

Abstract: More and more organizations are getting heavily dependent of the availability of their databases. Examples of these are databases to support Internet and E-commerce applications and databases to support the workflow of large production plants. These databases require a 24-hrs operation, no loss of data and an availability of at least 99.5%. An extra complexity is the fact that these databases are getting larger every year and need to support hundredths of concurrent users.
Keyword in high availability environments is redundancy. This presentation provides a survey of hardware supplies and Oracle-tools and options to achieve this redundancy. Examples are RAID-configurations, Standby databases and Oracle Parallel Server.
Another point of attention will be the backup method.
Target audience: Beginner and Experienced

#133 Effectively using Oracle8i performance features
Herman de Boer, KPN Telecom BV, The Netherlands

Oracle8i offers a lot of interesting new performance features. Some of them seem obvious - but they are not. Other features seem difficult to understand and to use - but the effect of using them might lead to great benefits. In this article, some of these important new features are investigated and compared.
First, the area of database design for performance and tuning is briefly explored, so all features can be put into context.
Secondly, three interesting KPN Telecom performance problems will be investigated. Each problem has

been solved using 8i techniques. The process of testing alternatives and implementing the solution is described in depth. Used features are index-only tables, object tables, and materialized views.
Finally, TPC benchmark reports are discussed, for investigating how Oracle gets the most out of a configuration. Useful techniques are projected on situations at KPN Telecom.

#141 The Huge Effect of the Small FIRST_ROWS and ALL_ROWS Hints
Haim Kopans, Precise Software Solutions, USA

In order to select an access path to the data, Oracle's cost-based optimizer uses knowledge and assumptions about the amount of data that needs to be selected. One way to influence the assumptions is to use the FIRST_ROWS and ALL_ROWS hints and parameters. These hints instruct Oracle to select the best possible way to retrieve either the first rows, or all the rows, of the result table.
These hints have a major effect on the access path that is selected, on the creation of a temporary result table, and on the query response time
This paper will describe the FIRST_ROWS and ALL_ROWS hints, the places where they can be specified, and the effect they have on Oracle's access path selection.
The paper will describe the types of applications that will make the best use of each hint and will demonstrate the effects these hints have on response time.
Target Audience: Experienced or Advanced

#168 Tuning for Advanced DBAs; Others will require Oxygen
Rich Niemiec, TUSC, USA

Abstract: Tired of going to presentations only to get a single new concept... this talk is for you. Drilling into the heart of the Oracle database will be the focus of this fast-faced advanced tuning session. Exploring the information in the V$ views and X$ tables can lead to significant tuning insight that can help you accelerate even the most stubborn database performance. Forget the GUI tools and glitz, zoom into views and tables that store the impact areas of Memory, Disk I/O and Query information that will give you the advantage you need to accelerate performance. Not for weak heart!
Target Audience: Advanced

#199 The integration of the internet's LDAP standard with Oracle 8i
Danny Gielen, Cronos, Belgium

The integration of the internet's LDAP standard with Oracle 8i Release 2 provides high availability services and eases manageability for mission critical applications and environments. The IT market demands a common infrastructure for data and resources to reduce complexity and costs of configuration and management. Managing information about users, keeping user information current, and securing access to all the information in an enterprise can become easier using the Oracle Internet Directory, the LDAP-based directory service of Oracle 8i Release 2. This paper explains the advantages of using LDAP as the central repository for all databases, network components, servers, users and authorisation.
Target Audience: Beginner, Experienced

#218 Implementing Data Security using Fine Grained Access Control
Shankaran Iyer, Renaissance Worldwide Inc. USA

Oracle8i provides a flexible method of implementing security policies for tables and views using Fine Grained Access Control. This presentation is designed to provide information and share the author's project experiences in implementing security policies using fine grained access control and application context. The session will be divided into three sections. The first section will provide an in depth introduction to Fine Grained Access Control and situations where it can be effectively used. The discussion will cover setup issues, defining policy functions, predicates, defining context and context functions. The second section of this presentation will provide a glimpse into how fine grained access control works. It will also cover useful tips for DBAs including scripts to generate documentation on various security policies along with tips, recommendations and lessons learned when implementing Fine Grained Access Control methods. The third part will include an example of FGAC in a packaged environment and in a custom development environment.
Target Audience Experienced

#223 Oracle8i Indexing Choices: Best of Breed
John King, King Training Resources, USA

Oracle8i provides expanded options for indexing. Index choices include: traditional b-tree indexes, reverse key indexes, bit-map indexes, hash indexes, and index-organized tables. In addition, indexing is now available when using a function or expression on an index column. Choosing the proper index style is an important part of database and application design. Dramatic performance differences can be achieved by the careful creation of indexes and their subsequent use in applications. This session compares and contrasts the various options available and how to choose from among them. Specific topics include: b-tree indexes, reverse key indexes, bitmap indexes, hash indexes, index-organized tables, the ability to base indexes on functions or expressions, SQL to create/alter indexes, negative impacts of indexes, and hints involving indexes. Those present will be better equipped to build effective Oracle8i applications.

#224 Cube, Rollup, and Materialized Views: Mining Oracle Gold
John King, King Training Resources, USA

Oracle8i provides new features that reduce costs of summary queries and provide easier summarization. The CUBE and ROLLUP extensions to GROUP BY allow more-complete summaries that would otherwise require additional queries or coding. Taking advantage of Materialized Views allows organizations to pre-build tables of summary data used frequently. The results of common summaries in Materialized Views may be indexed helping speed typical queries. Materialized Views may be refreshed in the same manner used for Snapshots allowing summarization to occur with a frequency that meets the needs of the user community. By including CUBE and ROLLUP in a Materialized View along with other GROUP BY results, more complete information is available for reduced-cost queries. Those present will learn how to create and use Materialized Views and how to use CUBE and ROLLUP to create more complete summaries.

#230 Monitoring and Tuning of the Shared Pool
Michael R. Ault, DBMentors International, USA

In many of the tuning manuals the coverage for tuning of the shared pool consists of the simple statement "Make it bigger." In my experience this is not always the answer. In this paper I will discuss the various shared pool usage models and how they are each tuned. In addition I will discuss the methods which should be used to monitor not only how much of the shared pool is used, but whether the pool is being used effectively. After participating in this presentation the attendees should have a deeper, more practical understanding of the shared pool, its use, monitoring and tuning.
Target audience: Experienced

#232 Optimal Physical Database Design for Oracle8i
Dave Ensor, BMC Software, UK

The objective of this presentation is to help both database designers and DBAs to understand how the new indexing and storage management features in Oracle8i can be used to increase the performance and availability of Oracle-based applications, and when more traditional Oracle features are indicated.
ABSTRACT: Oracle8i adds a number of significant new features in the areas of indexing and space management along with major upgrades to the support for Index Organized Tables (IOT's).
The paper presents performance figures on index compression and IOT's, and analyzes the performance and operational impact of online table reorganization, which is now supported for IOT's. A strong case is made for migrating certain types of application to use IOT's though, as discussed, the change may not be transparent at the application level.
The paper also looks at temporary tables, and explains how they can both improve performance and reduce one specific type of application failure. The circumstances under which they should be used are detailed along with the application changes that should be considered. Performance improvements achieved from both locally managed and transportable tablespaces are also presented, and the potential implications are explained.

#248 Compaq and Oracle TestDrive
Nigel Bridgeman, Compaq EMEA

Launched at OOW'99, the Compaq TestDrive program lets you try out the latest internet development tools and solutions from Oracle on the hardware and operating systems of your choice.
This presentation introduces the TestDrive concept pioneered by Compaq and Oracle focusing on the tools and technologies that you can evaluate today from your desktop. The presentation also looks at the workings of the TestDrive site, including an in-depth look at the hardware and software infrastructure capable of supporting thousands of users per day. We will discuss how the systems are architected and deployed, particularly focusing on the integration of Unix , Linux and Windows NT.
So if you keen to evaluate WEBDB, or are sceptical that your Oracle java code is completely portable between Tru64 Unix, Windows NT, Linux and the Oracle Appliance - try it out for yourself at http://www.testdrive.compaq.com and attend the presentation to find out how this has all been made possible and how you can maximise this free service for your company.
Target Audience : BEA

#256 High Availability (HA) Solutions using Oracle and EMC Technology
Hardev Sanghera - EMC (UK) and Gerard Pillai - Enterprise Technology Center, Oracle Corporation (UK),

As organisations web-enable their business processes, the requirement for very high availability and speedy recoverability of their Oracle Databases becomes vital. With internet customer loyalty based on the speed at which a mouse button can be clicked, downtime of even minutes can result in significant financial loss.
While Oracle 8i offers a large number of high availability features, they can be further enhanced by using EMC software technology. This presentation includes an overview of Oracle 8i HA features and a detailed overview of EMC technology such as EMC Timefinder, EMC Symmetrix Remote Data Facility (SRDF) and EMC Data Manager (EDM) and how these technologies can be used to create non-disruptive backup strategies, very fast recovery strategies and Disaster Recovery(DR) solutions with minimal data loss.
Audience: IT Managers, DBA's, Anyone interested in HA

#257 APPLYING METRICS TO ORACLE DATABASES
Coral Calero, Mario Piattini, Marcella Genero, Grupo Alarcos, E.S.Informatica, Spain

Software measurement is an effective means to manage software development and maintenance projects. In the past decades a huge amount of software metrics has been proposed, but primarily focused on programs. Metrics for databases have been neglected, mainly because databases have developed a secondary role in Information Systems (IS) infrastructure until a few years ago. But nowadays, databases are the core of IS, influencing considerably their complexity. It is also necessary made a formal verification of the metrics in order to improve its usefulness. This paper proposes a metrics suite for measuring relational database complexity, presents different approaches for metrics formal verification, apply the proposed metrics to the ORACLE Dictionary and makes an overview of the metrics tool which supports the metrics.

#265 SERVER TECHNOLOGY & DBA Q&A PANEL

#266 Oracle 8i Log Miner - A heart of Gold!
Carl Dudley, Staffordshire University, UK

Oracle8i has a built in redo log analyser called LogMiner. Redo logs contain very useful information which can not be obtained from other sources. The LogMiner can reconstruct any Data Manipulation Language statement (and its inverse) from the information it finds in the redo log. This talk will illustrate how to set up the LogMiner and how it can be used to perform fine-grained recovery of user errors, track changes due to entire transactions, and assist in auditing and tuning. Its main features and limitations will be evaluated and examples of automatic construction of column mappings will be shown. The column mappings are then used to analyse changes to whole schemas during a specified period of database activity.

#306 Scaleability in Web-based solutions using the Oracle Application Server
Heimir Þór Sverrisson, Teymi hf, Iceland

There are several different architectures now supported and available to Web enable existing businesses. This presentation gives an overview of the different characteristics of solutions developed in old traditional CGI, PL/SQL using the PL/SQL cartridge of the OAS and the Distributed Object approach using JavaServer Pages and Enterprise Java Beans (EJB).
There are many new problems that arise when applications are deployed on the Web. Scaleability issues arise very quickly and are often very hard to diagnose. The problem has to be isolated to the correct pier or to the interconnection between them. This presentation describes methods that have been used with good results to isolate such problems. Some common technical pitfalls are discussed and security issues are mentioned.
Audience: Experienced to Advanced

#314 Managing your e-business
Jay Rossiter, Oracle Corporation, USA

Today's e-everything environments have set a standard of zero tolerance for system downtime, performance bottlenecks and inefficient administration tools for day-to-day tasks. Administrators need robust tools that work together to manage the complete stack -- from the applications, to the middle-tier Web servers, to the databases on the back-end. This session covers Oracle Enterprise Manager, Oracle's unified solution for managing Oracle products. Topics will cover the breadth of the Oracle solution, from the Enterprise Manager scalable framework, to pinpoint monitoring and event management, to database and application tuning to the integrated administration of entire Oracle software stack.

#315 8i Appliance Performs!
John O'Connor, Oracle Corporation, Ireland

Oracle8i Appliance offers a platform which provides an environment for speeding deployment, reducing costs and simplifying administration. It comes pre-packaged and pre-tuned. Oracle8i Appliance runs on a stripped down Unix Intel Operating System which immediately reduces a sizeable amount of operating system overhead, thus releasing more memory and CPU power for the dedicated Oracle8i Server on the Appliance box. As well as describing some of Oracle8i Appliance's pre configured features, this presentation will discuss the tuning that has gone into the Oracle8i Appliance by Oracle Corporation and show some comparisons against other operating systems such as Windows 2000/NT.

#316 iHost - Oracle's hosting initiative
Andy Cleverly, Oracle Corporation, UK
ABSTRACT AWAITED

#318 From Beta to Production - Eurostar and the database appliance
Simon Harrison, Oracle Corporation, UK

Eurostar UK Ltd, the International Passenger Train company were chosen by Oracle to be the only Beta Customer on the Oracle Raw Iron platform. This presentation will give an insight into the product, now renamed the Oracle 8i Appliance, from the early days of the Beta program, right up to the implementation of the 8i Appliance into Production. The presentation will detail what a 8i Appliance is, how it is implemented and the Benefits that Eurostar UK Ltd make from implementing the 8iAppliance. The Oracle 8i Appliance is a simple way of deploying Oracle databases and Applications straight out of the box, the presentation will explain this and what Eurostar UK Ltd would like to see in future releases of the 8I Appliance. The 8i Appliance Beta project utilised resources from Oracle UK and Eurostar UK, the presentation will explain what benefit can be gained from a beta program and more importantly what the customer is expected to provide.

#319 Enabling the Internet Enterprise with the 8i Appliance
Venkat Malla, Oracle Corporation

The Oracle8i Appliance provides a simple, yet complete e-business infrastructure for the Internet Enterprise. It is the fastest possible way to deploy Oracle with the added benefits of being reliable, low-cost and requiring no professional on-site management. The Oracle8i appliance is self-configuring, and self-managing yet provides all the power and performance associated with Oracle8i. This presentation will give an overview of the system architecture and components, show how the various benefits outlined above are achieved, and show how applications can be run against the appliance.

#320 Oracle 8i Intermedia - The Platform for Managing e-business Content
Joe Mauro, Oracle Corporation

e-Business lets companies exploit new business opportunities offered by the Internet. As a result, web-based internet and enterprise applications are rapidly emerging as the preferred means of conducting business. This trend has created a demand for richer web experiences utilizing a wider variety of media content -- images, audio, video, animation, text and documents. Web site designers, application developers, and IT administrators must now handle and manage rich media in addition to the more traditional structured content. A major component of internet platforms is content management services. These services consolidate internet content and support the development and deployment of media-rich, dynamic web sites and portals. This presentation highlights Oracle8i interMedia, management services that extend the benefits of the Oracle8i internet platform to rich content. This session will demonstrate the use of Oracle8i interMedia to create more powerful. flexible, and manageable media-rich web applications.

#336 Oracle and ESRI: Bringing Integrated Spatial Information Management to the enterprise
Frank Holsmuller, ESRI-Europe

Traditional GIS (geographic information systems) and MIS (management information systems) vendors now are developing integrated spatial information management technology that enables users to share spatial data across the enterprise. This paper describes how ESRI and Oracle are working together to deliver standards-compliant spatial information management capability. The paper illustrates the underlying architecture and customer benefits of deploying a combined ESRI/Oracle solution.
Users can now manage their spatial information resources using spatial technology from Oracle, the market-leader in data management technology. Likewise the capability to access, visualization and spatial analysis functionality is provided by ESRI.

#337 Delivering Business Benefit with Spatial Data Management
Mark Edgerley, Enghouse Systems, UK

Much is talked about Spatial Data Management, Open GIS and putting the G into the IS. This has been discussed to death over the last couple of years. Perhaps the GIS vendors have been neglecting the customer in a rush to have all the right ticks in the right technology boxes, rather than actually deliver what the customer wants and needs to provide benefit to their business. There are certainly important areas that vendors have to ensure they can match. Areas such as unified databases, web access, SQL standards etc. etc. This has basically been a technology driven "me too" approach by vendors to say "we can do that as well!" It is time for suppliers of GIS and Automated Mapping/Facility Management software to understand what their customers need to deliver to their customers and staff.
The technology within industry standards such as Oracle 8i with its advanced object model, its standard development tools, and web enablement allows software developers to concentrate on industry applications and utilise these standards. The use of Oracle's investment in spatial geometry, version and long transaction management etc. allows the application developer to concentrate on their areas of expertise. A telecommunications, Cable, Electricity, water or gas provider requires a user-friendly tool that allows them to input, maintain, analyse and disseminate data to deliver business benefit. This benefit may be to deliver information via ultra-thin clients, implement data standards and rule bases, share hereto-closed GIS data with other corporate systems, and have easy and fast access to the information they need to do their everyday job.

#363 Oracle8i Release 2: SQL functionality and performance enhancements
Lex De Haan, Oracle Corporation, The Netherlands

Oracle8i Release 2 comes with a lot of exciting SQL language enhancements, such as five new categories of powerful analytic functions for data mining, and the CASE operator as the powerful (and ANSI/ISO standard) alternative for the Oracle DECODE function. Oracle8i Release 2 also comes with a new feature which allows auto-substitution of literals to bind variables, thus reducing SQL parse time and optimizing shared SQL; this is particularly helpful in case of third-party applications that do not use bind variables. Subquery optimization (flattening) is significantly
improved, as well as query rewrites based on materialized views. This session will give participants an overview of all these SQL functionality and performance enhancements, and show the benefits.

#385 Explain Plan Part 2: Parallel, Partitioned & Problematic
Jonathan Lewis, JL Computer Consultancy, UK

As more and more features are added to the database, the scope for new and complex execution paths grows.
Building on the information in the first session on explain plan, the presentation covers interpretation of Parallel Execution plans, Plans relating to Partitioned Tables, and plans which are sufficiently ambiguous that you have to use alternative methods to find out what is really going on.

#398 XML for Real
Doug Burns, Independent, UK

XML is a fashionable technology with a bright future but hard information about how to get started with XML development is limited. This presentation will explain in detail where to find and how to install the various components you'll need to set up your first XML-enabled web database server that will accept user requests for XML pages containing SQL queries, pass the query on to an Oracle8i database and return the result to different user devices, using conditional e-formatting defined by your own stylesheets. Examples of XSQL pages, XML documents, DTDs and XSL Stylesheets will be discussed to provide a starting point for your own applications. The base operating system used is Windows 2000, so that you can set up your own development system on a PC, but the information will apply equally well to other Operating Systems. Stop putting XML off until you have the time, come and find out how to get started now.

#399 Dot.Ready
Alan Edwards, Sun Microsystems, UK

As the growth of the new e-conomy accelerates so the pressure to rapidly deliver a production ready environment optimised for your net application and database increases. The days of predictable loads and managable user bases have gone, to be replaced by realtime links with customers and suppliers. The clicks and mortar world has the scope for outstanding successes and catastrophic failures.
The bedrock for building a net business is the technology that will run it. Technology that once supported the business, now is the business. But firm foundations depend on more than just the technology platform itself, they depend on how the platform is designed, implemented, and managed in support of the business and end users.
Sun Microsystems, a long time partner with Oracle, and a leading technology provider to the new e-conomy recognises these issues. People, Processes and Products, have to be considered together when building an optimal platform for e-business. Issues like time to market, availability, rapid scalability, and consistent performance, become design criteria. This paper looks at these issues, the platform needs of the e-conomy and how to build a
server environment optimised for the business application.

Back

APPLYING METRICS TO ORACLE DATABASES[1]

Coral Calero, Mario Piattini, Manuel A. Serrano, Marcela Genero
Grupo ALARCOS - E.S.Informática.
Universidad de Castilla – La Mancha
Ronda Calatrava, 5 13071 Ciudad Real
Tfno: +34926295300 Ext:3715
Fax: +34926295354
SPAIN
E-mail:{mpiattin, ccalero, mserano,mgenero}@inf-cr.uclm.es

Abstract : - Software measurement is an effective means to manage software development and maintenance projects. In the past decades a huge amount of software metrics have been proposed, but primarily focused on programs. Metrics for databases have been neglected, mainly because databases have developed a secondary role in Information Systems (IS) infrastructure until a few years ago. But nowadays, databases are the core of IS, influencing considerably their complexity and to dispose on specific metrics for databases is very important. However, it is also necessary made a formal verification of the metrics in order to improve its usefulness. This paper presents a metrics suite proposed for measuring relational database complexity, presents different approaches for metrics formal verification and makes an overview of the metrics tool which supports the metrics.

1 Introduction

In the last years we have beheld an extraordinary diffusion of relational databases, which have been installed on macrocomputers but also on personal computers. Databases, especially the relational ones, have became the core of the most relevant Information Systems. However, their design is a long and hard task, although it is based on solid mathematical fundaments derived from the definition of the relational model (Codd, 1970).

Most of the methodologies in designing relational databases (Batini et al., 1992; Connolly et al., 1998; MacFadden et al., 1999) are divided into three steps: conceptual modelling, logical design and physical design; applying a set of rules that allow the transformation of a conceptual schema to a relational schema and that guarantee the "quality" of the relational schema applying the normalisation theory. Although this theory is very useful, it is necessary to develop new metrics to evaluate the quality of a relational database. In fact it is widely recognised that software measurement is an effective form to understand, monitor, control, predict and improve the development and maintenance of software projects (Briand et al., 1996). Also, they can help the practicioners and the investigators to make the best decisions (Pfleeger, 1997), helping the designer to choose among alternative database schemas.

Unfortunately, most of the metrics ever since McCabe proposed the cyclomatic complexity (McCabe, 1976) to the present time, have been centered in the program characteristics, disregarding databases (Sneed and Foshag, 1998). Nevertheless, design metrics for data aspects is important since their size and nature contribute to many aspects of a information system like the amount of effort to develop (MacDonnel et al., 1997).

One of the most important aspects relative to quality, is maintainability (ISO, 1994) which depends on three factors: understandability, modifiability and testeability, which are influenced by complexity (Li and Chen, 1987). But to take a general complexity measure is the impossible holy grail (Fenton, 1994). Henderson-Sellers(1996) distinguishes three types of complexity: representational, computational and psychological. The last one is composed by the problem complexity, the human cognitive factors and product complexity. The last one is our focus.

In this paper we propose a set of metrics for measuring the relational database schemas complexity: NA (Number of Attributes), RD (Referential Degree), DRT (Depth of the Referential Tree) and COS (COhesion of the Schema).

Proposing metrics is not difficult, the problem is that much published work on software metrics is theoretically flawed and it is necessary that software measurement adheres to the science of measurement to be valid and accepted (Fenton, 1994). Several frameworks for measure characterisation have been recently put forward (Briand et al., 1996; Morasca and Briand, 1997; Weyuker, 1998; Zuse, 1998).

[1] This research is part of the MANTICA project, partially supported by the CICYT and the European Union (CICYT-1FD97-0168) and by CALIDAT project carried out by Cronos Ibérica (supported by the Consejeria de Educación de la Comunidad de Madrid, Nr. 09/0013/1999)

Rigorous measurement of software attributes can provide substantial help in evaluation and improvement of software products and processes. In this paper we present two different approaches to verify formally software metrics. We also present an example of a metric formal verification in one of these frameworks. Section 2 presents the proposed metrics. A general presentation of formal verification with an example is carried out in section 3, section 4 presents the metrics tool which collects these metrics. Conclusions and future work come in the last section.

2 Metrics for relational databases

Since the proposal by Dr. Codd of the relational model in the late sixties (Codd, 1970), the database field has been extensively researched and relational database products have developed a very important industry.

Date defines a relational database management system as "a system in which, at minimum :

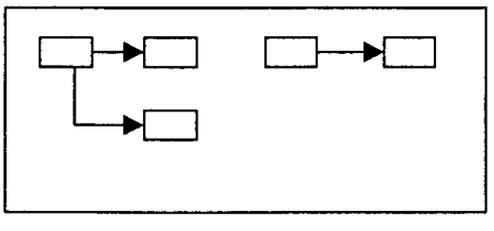
- The data is perceived by the user as tables (and nothing but tables); and
- The operators at the user's disposal are operators that generate new tables from old, and those operators include at least SELECT, PROJECT, JOIN" (Date, 1995).

The only indicator used to measure the quality of relational database has been the normalisation theory, upon which (Gray et al., 1991) propose to obtain a normalisation ratio.

We propose a set of metrics that can be divided in table-oriented metrics and schema-oriented metrics (two of them, NA and RD, can be used as table-oriented and schema-oriented metrics).

- As table-oriented metrics we propose:
  Number of attributes (NA(A)) . NA(A) is the number of attributes of the table A.
  Referential Degree (RD(A)). RD(A) is the number of foreign keys in the table A.
- As schema-oriented metrics we propose:
  Number of attributes (NA). NA is the number of attributes in the schema.
  Referential Degree (RD). The RD metric is defined as the number of foreign keys in the schema.
  Depth Referential Tree (DRT) . DRT is defined as the length of the longest referential path in the database schema. The cycles are only considered once.
  Cohesion of the schema (COS). COS is defined as the sum of the square of the number of tables per unrelated subgraph in the database:

$COS = \sum_{i=1}^{\lvert US \rvert} NTUSi^2$	|US| number of unrelated subgraphs NTUS$_i$ number of tables in the unrelated subgraph "i"



In this example there are two unrelated graphs and the value for the metric is

$COS = 3^2 + 2^2 = 9 + 4 = 13$

3 Metrics formal verification

There are a few proposals for obtaining accurate and unambiguous software attributes definition. All of them are based in axiomatic approaches or in measurement theory. Axiomatic approaches formally define desirable properties for measures for a given software attribute, so axioms must be used as guidelines for the definition of a measure. Measurement theory specifies the general framework in which measures should be defined.

Software metrics axioms sets have been developed without a consensus and sometimes without a common understanding of the data to which they apply. The main goal of axiomatisation in software metrics research is the clarification of concepts to ensure that new metrics are in some sense valid. However, if an axiom set cannot itself be shown to be fit for purpose, it cannot be used to validate metrics. We cannot tell whether a measure that does not satisfy the axioms has failed because it is not a measure of the class defined by the set of axioms (e..g. complexity, length...) or because the axiom set is inappropriate. Since the goal of axiomatisation in software metrics research is primarily definitional, with the aim of providing a standard against which to validate software metrics, it is not so obvious that the risks outweigh the benefits (Kitchenham and Stell, 1997). The strength of measurement theory is the

formulation of empirical conditions from which we can derive hypothesis of reality. Measurement theory gives conditions for the use of the metrics on certain scale which is important for the proper use of statistics.

3.1. Example

We will follow the Zuse's formal framework (Zuse, 1998) in order to make the formal verification of the NA metric. This framework is based on an extension of the classical measurement theory, which gives a sound basis of software measures, their validation and criteria for measurement scales.

Zuse(1998) describes measurement as a detour, "*necessary because humans mostly are not able to make clear and objective decisions or judgments*". Measurement is more than producing numbers, it is the combination of empirical entities with numerical entities. This process starts with the real world, which contains the objects that should be measured. People are interested in the establishment of "empirical relations" between objects, such as "higher than" or "equally high or higher than". These empirical relations will be denoted with the symbols "• >" and "• >=", respectively. An Empirical Relational System is a triple: **A** = (A, • >=, *o*), where A is a non-empty set of objects, • >= is an empirical relation on A, and *o* is a closed binary (concatenation) operation on A.

In many cases we are not able to produce directly relevant empirical results due to the difficulty of the question we deal with. There is an "intelligence barrier" that impedes to reduce information without help. With the aid of mathematics and statistics "the intelligence barrier" can be overcome: the empirical objects and relationships are mapped into proper numerical objects and relationships. A Numerical Relational System can be defined as B = ($\Re$, >=, +), where $\Re$ are the real numbers, >= a relation on $\Re$, and + a closed binary operation on $\Re$. A measure is then a mapping u: A -> $\Re$ such that: a • >= b $\Leftrightarrow$ u(a) > = u(b); $\forall$ a, b $\in$ A.

Once the mapping is established, mathematics and statistics can be used to process the information (e.g. working out means or variances). Measurement theory gives also conditions for the translation of numerical statements back to empirical statements. in order to check whether the measure satisfies the user needs, Zuse proposes an internal validation, based on the comparison between the empirical interpretation of numbers and the empirical statements in the real world.

The combination rule must be defined as: u(A1oA2) = f (u(A1) , u(A2)), where A1,A2, A1oA2 $\in$ A and f(A1,A2): A x A $\rightarrow$ A. This concatenation operation (*o*) can be contra-intuitive in the area of software engineering because it is not necessary to combine objects in reality. However, it provides a means to build up complex measurement structures, giving a more precise interpretation of numbers (Zuse, 1998).

On this framework, Zuse defines a set of axioms for measures which gives rise to distinct structures, in table 1 we present the most important ones:

MODIFIED EXTENSIVE STRUCTURE	INDEPENDENCE CONDITIONS	MODIFIED RELATION OF BELIEF
Axiom1: (A, • >=) (weak order) **Axiom2**: A1 o A2 • >= A1 (positivity) **Axiom3**: A1 o (A2 o A3) ≈ (A1 o A2) o A3 (weak associativity) **Axiom4**: A1 o A2 ≈ A2 o A1 (weak commutativity) **Axiom5**: A1 • >= A2 ⇒ A1 o A • >= A2 o A (weak monotonicity) **Axiom6**: If A3 • > A4 then for any A1, A2, then there exists a natural number n, such that A1o nA3 • >A2 o nA4 (Archimedean axiom)	**C1**: A1 ≈ A2 ⇒ A1 o A ≈ A2 o A and A1 ≈ A2 ⇒ A o A1 ≈ A o A2 **C2**: A1 ≈ A2 ⇔ A1 o A ≈ A2 o A and A1 ≈ A2 ⇔A o A1 ≈ A o A2 **C3**: A1 • >= A2 ⇒ A1 o A • >= A2 o A, and A1 • >= A2 ⇒ A o A1 • >= A o A2 **C4**: A1 • >= A2 ⇔ A1 o A • >= A2 o A, and A1 • >= A2 ⇔A o A1 • >= A o A2	**MRB1**: $\forall$ A, B $\in$ $\Im$: A • >= B or B • >= A (completeness) **MRB2**: $\forall$ A, B, C $\in$ $\Im$: A • >= B and B • >= C ⇒ A • >= C (transitivity) **MRB3**: $\forall$ A $\supseteq$ B ⇒ A • >= B (dominance axiom) **MRB4**: $\forall$ (A ⊃ B, A ∩ C = ϕ) ⇒ (A • >= B ⇒ A U C • > B U C) (partial monotonicity) **MRB5**: $\forall$ A $\in$ $\Im$: A • >= 0 (positivity)
As we know, binary relation • >= is called weak order if it is transitive and complete: A1 • >= A2, and A2 • >= A3 ⇒ A1 • >= A3 A1 • >= A2 or A2 • >= A1	Where A1 ≈ A2 if and only if A1 • >= A2 and A2 • >= A1, and A1 • > A2 if and only if A1 • >= A2 and not (A2 • >= A1).	

Table 1. Zuse's formal framework properties

There exist five scale types that are defined by admissible transformations (table 2). They are, in hierarchical order: nominal, ordinal, interval, ratio and absolute. Each scale type is defined by admissible transformations. Software measurement starts with the ordinal scale (Zuse, 1998).

Name of scale type	Admissible transformation g	Statistics	Statistics Tests
Nominal scale	Any one-to-one	Frec, mode	Non parametrics order independents
Ordinal scale	g: strictly increasing monotonic function	Median, quartil	Nonparametrics, Karrell coef, Spearman coef, Kendall
Interval scale	g(x)=ax+b, a>0	Mean, standard deviation, variance	Param, non param
Ratio scale	g(x)=ax, a>0	Geometric median, variation coef	Param.. non para
Absolute scale	g(x)=x	Equal than ratio	Equal than ratio

Table 2. Scale types and their transformations

Measures may be classified in a scale type, depending on whether they assume an extensive structure or not. When a measure accomplishes this structure, it also accomplishes the independence conditions and can be used on the ratio scale level.
If a measure does not satisfy the modified extensive structure, the combination rule will exist or not depending on the independence conditions. When a measure assumes the independence conditions but not the modified extensive structure, the scale type is the ordinal scale.
Object oriented measures assume more complicated properties related to concatenation operations than classic measure. Most of these metrics do not assume an extensive structure, so that we must use the modified relation of belief to characterise them "above" the level of the ordinal scale (the characterisation of measures above the ordinal scale level is very important because we cannot do very much with ordinal numbers).
We adapt this framework to the NA metric, verifying the fulfilment of each axiom. Seeing figure 1, we can define the combination rule for NA as: NA (Ri o Rj) = NA (Ri) + NA(Rj) – NA(Ri ∩ Rj), w here NA(Ri ∩ Rj) is the number of attributes which are common to (belong to the intersection[2] of) Ri and Rj.

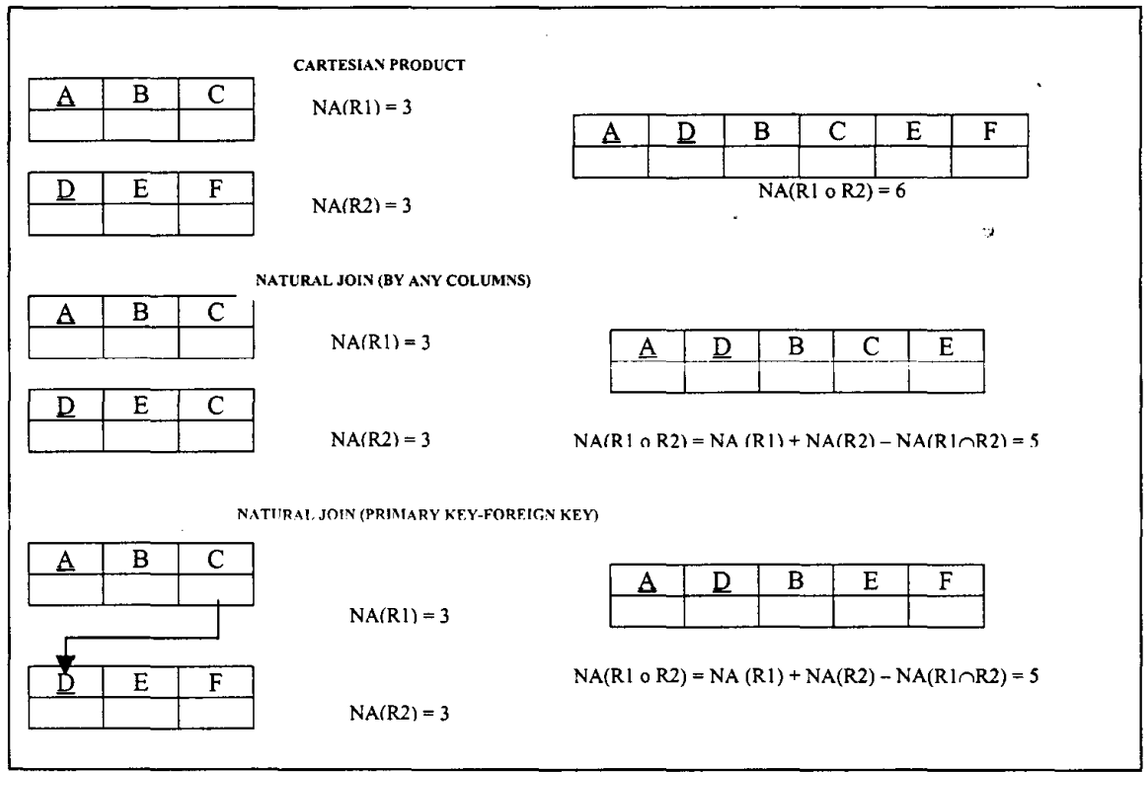


Figure 1. NA of combined tables

NA fulfils the first axiom of weak order, because if we have two relations R1 and R2, it is obvious that NA(R1) >= NA(R2) or NA(R2) >= NA(R1) (completeness) and let R1, R2 and R3 three relations, transitivity is always fulfilled: NA(R1) >= NA(R2) and NA(R2) >= NA(R3), then NA(R1) >= NA(R3). NA do not fulfil positivity, because if we combine a relation R1 with itself without cycles: NA(R1 o R1) is not greater than NA(R1). But it fulfils weak positivity , because it is always true that: NA(R1 o R2) >= NA(R1) for all R1, R2 ∈ R. NA fulfils axiom 3, because the natural join operation is associative and so NA(R1 o (R2 o R3)) ≈ NA((R1 o R2) o R3). NA also fulfils weak commutativity because natural join is commutative. NA but not fulfils weak monotonicity as we see in the figure 2.

[2] This intersection is different from the relational algebra operation of Intersection between two relations.

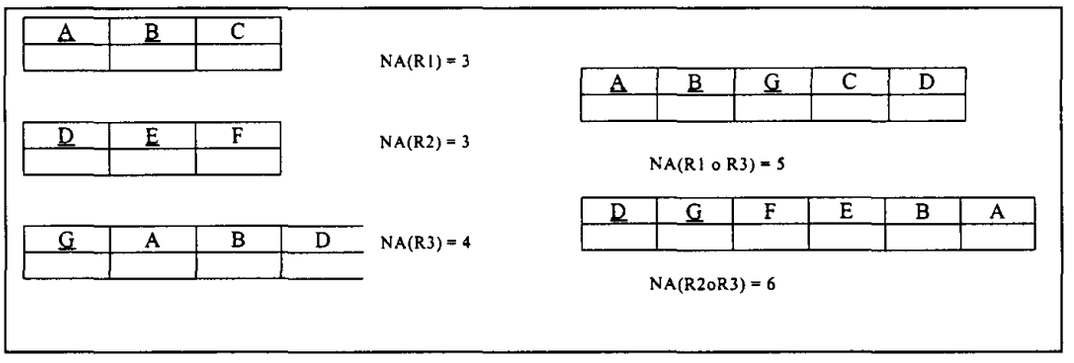


Figure 2. NA does not fulfil weak monotonicity.

Previously to verify if NA accomplishes the Archimedean axiom, we must verify, with the figure 3, if the metric is idempotent:

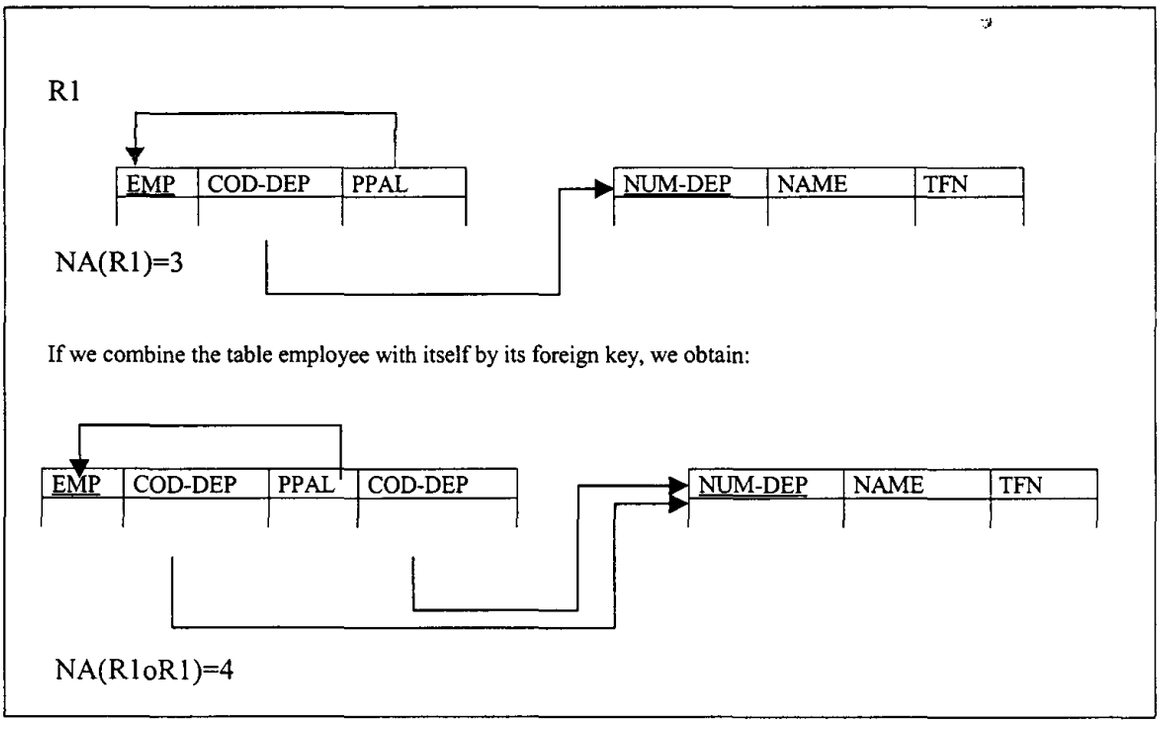


Figure 3. Idempotency of a table

It is easy to see that the number of attributes changes when we combine one table with itself, so it is not idempotent and is necessary to prove the Archimedean axiom. When two tables are combined by natural join successively, the number of attributes changes and may be possible to obtain two tables R1, R3 that do not accomplish the Archimedean axiom, as we can see in figure 4, where the result of the first concatenation is shown (the tables obtained in successive concatenations will be the same). We can conclude that NA does not assume an extensive structure.

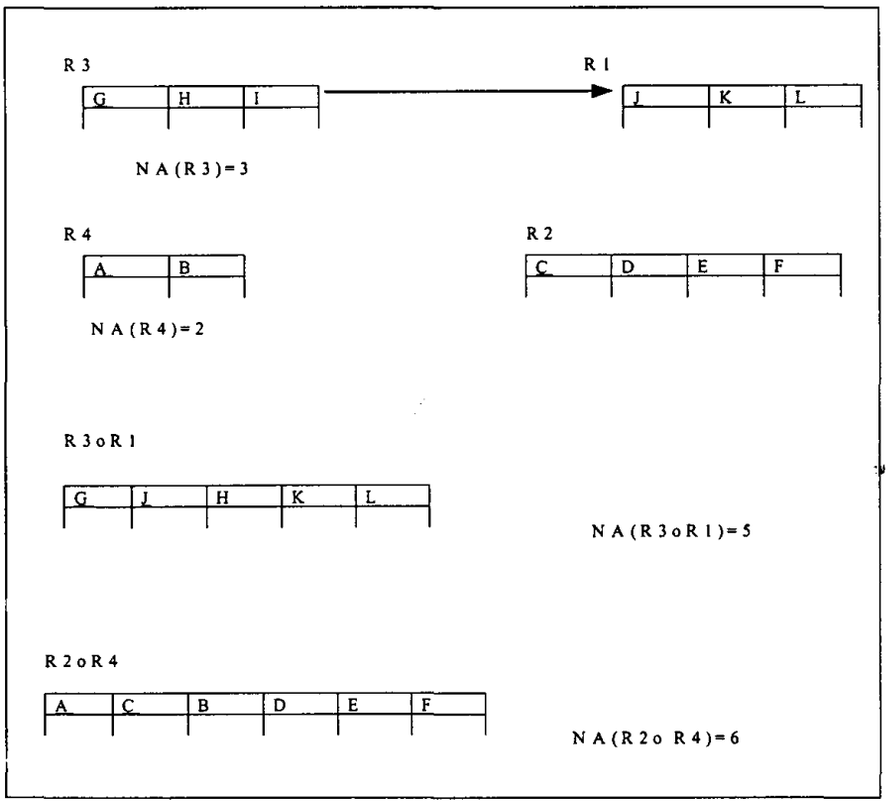


Figure 4. Archimedean axiom

Would NA verify the independence conditions?. With figure 3 we can ensure that NA does not accomplish the first condition, nor the second one. NA does not accomplish the weak monotonicity, so it cannot accomplish the third and the fourth independence conditions. In fact, this type of combination rules does not assume the independence conditions. The part $-NA(R_i \cap R_j)$ rejects the condition C1 that implies the rejection of the axiom of weak monotonicity, monotonicity and extensive structure.

Then we must study if NA fulfils some of the modified relations of belief. MRB1 is fulfilled, because giving two relations R1 and R2 $\in \Im$ ($\Im$ is the set of all the possible relations made with the attributes of the relational schema) NA(R1) >= NA(R2) or NA(R2) >= NA(R1). MRB2 is also fulfilled (transitivity of natural join). For MRB3 we must consider a relation R1 $\supseteq$ R2 where all the attributes of R2 are present in R1. In this case is evident that NA(R1) >= NA(R2), and MRB3 is fulfilled. MRB4 is fulfilled because if a relation R1 $\supset$ R2 then NA(R1) > NA(R2) and NA(R1 U R3) > NA(R2 U R3), being R1 $\cap$ R3 = ϕ. The relations R1 and R3 do not have any attribute in common. If we add the attributes of R3 to both R1 and R2, (if R1 subsumes R2), then the number of attributes of R1 and R3 is greater than the number of attributes of R2 and R3. MRB5 is fulfilled because a relation must always have zero or more attributes.

Then, we can conclude that NA accomplishes all of the axioms of the structure of belief, so we can characterise NA as a measure above the level of the ordinal scale, assuming the modified relation of belief.

3.2 Rest of metrics

In table 3 we present the results obtained from the formal verification of the rest of metrics, on the Zuse' s formal framework, presented in this paper. The complete verificatrion can be found in Calero et al (2000) for this framework and in Piattini et al (2000) for the Briand et al (1996) framework:

Properties	DRT	RD
Axiom 1	YES	YES
Axiom 2	NO	NO
Axiom 3	YES	YES
Axiom 4	YES	YES
Axiom 5	NO	NO
Axiom 6	NO	NO
Ind Cond. 1	NO	NO
Ind Cond. 2	NO	NO
Ind Cond. 3	NO	NO
Ind Cond. 4	NO	NO
MRB 1	YES	YES
MRB 2	YES	YES
MRB 3	YES	YES
MRB4	YES	YES
MRB5	YES	YES

Table 3. Characterization of the DRT and the RD metrics

4. The MANTICA tool

The objective of our tool is to acquire, manage, process and represent the values of a set of metrics that are applied to a database schema. The tool must be independent of the data model that we use to represent the database schema; in this way, we can measure a conceptual, logical or physic schema. The tool must be also platform independent, so we can measure databases that are implemented in different DBMS.
Due to this restrictions, our tool can indicate the database syntax that we wish to measure and the metrics that we are to use.
The tool have a user interface module which communicates with the others metrics modules. So all the metrics values are shown in the same way, making the tool easier to maintain, to extend and to use.
All the metrics modules need to access to a central database, which contains all the information needed to calculate the metrics values. This metadatabase store the information of the database schematas we want to measure.
By using this metadatabase, we get some advantages:

- We can simplify the measure process since the tool only need to process the information that is in its own metadatabase, without the need of have a lot of drivers to access different types of databases.
- It is possible to make statistic analysis with a set of databases schemas since we have all the information stored in the metadatabase.
- The tool can measure any type of database; we only need to make small conversion applications to store the database schema we want to measure in our metadatabase. (see figure 5)
- We can apply new metrics to a database schema without accessing to the actual database; we only need to capture the database schema once.

In the figure 6 we can observe the aspect of our tool.

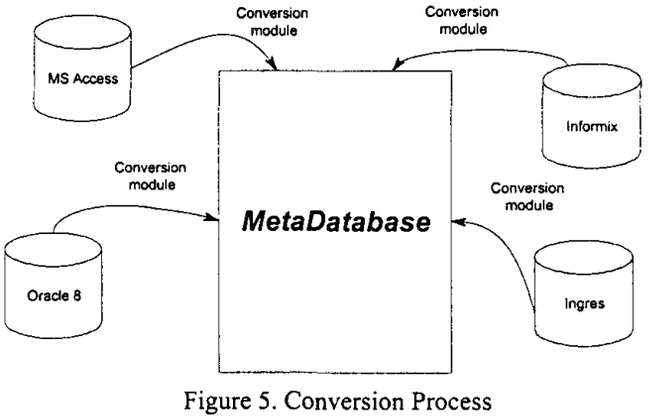


Figure 5. Conversion Process

Figure 6. View of our tool

4.1. Metadatabase

The design of our metadabase permit us to represent any database schema type. Any object of the schema is one "element" of a specific "element type" which has relationships with another elements.
Each element is associated with a set of "metrics" and obtains a set of values for this metrics (see figure 7). We can represent any database schema in this database design, and also we can store the values obtained for the metrics.
In order to calculate the values of the metrics of a given schema, we need to use some views of the metadatabase, converting the data stored in the metadatabase to a graph in memory, since most of the metrics need to use the graph theory to calculate the values.

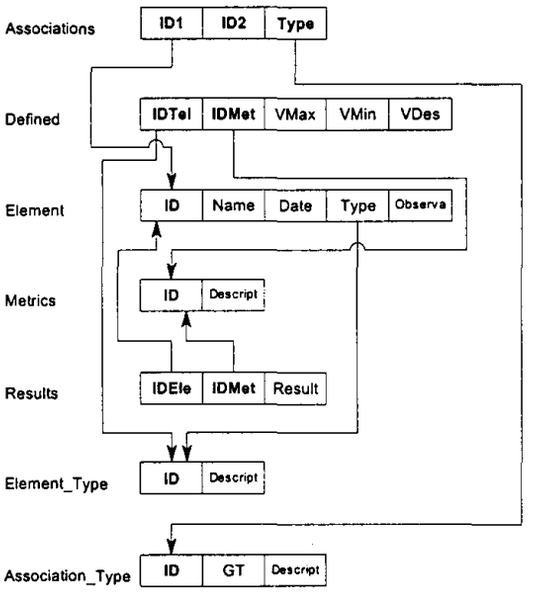


Figure 7. Database Relational schema

4.3. Representation of a relational schema in the metadatabase

In order to represent a relational schema in our metadatabase, we use different records. In table 4 we present the records generated:

Element Table	Associations Table
We add a record for each: • Schema RDB (RDBS) • Table (T) • Column (C) • Primary Key (PK) • Foreign Key (FK)	We add a record for each: • Table that belongs to a schema (TRDBS) • Column that belongs to a Table (CT) • Primary Key that belongs to a Table (PKT) • Foreign Key that belongs to a Table (FKT) • Column that belongs to a Primary Key (CPK) • Column that belongs to a Foreign Key (CFK) • Foreign Key that references to a Table (FKT) • Column that references to a Column (CCFK)

Table 4. Rules to store elements in the metadatabase

And we can get the views of our relational schema using the next queries:

Schemas. Shows all the relational schemas stored in the metadatabase

```
SELECT Element.ID AS ID, Element.Name_ele AS Name, Element.Date_ele AS Date
FROM Element
WHERE (((Element.Type)='RBDS'));
```

Tables. Shows all the tables of a relational schema.

```
SELECT Element.ID AS ID, Element.Name AS Name, Element.Date AS Date
FROM Element INNER JOIN (Associations INNER JOIN Element AS Element_1
        ON Associations.ID2 = Element_1.ID) ON Element.ID = Associations.ID1
WHERE (((Element.Type)='T') AND ((Associations.ID2)=[Num schema?]) AND
        ((Associations.Type)='TRDBS') AND ((Elemento_1.Tipo_ele)='RDBS'));
```

Columns by Table. Shows all the columns of all the tables within a relational schema

```
SELECT Element.ID AS ID, Element_1.Name AS Table, Element.Name AS Column, Element.Date AS
        Date
FROM Element INNER JOIN (((Associations INNER JOIN Element AS Element_1 ON Associations.ID2 =
        Element_1.ID) INNER JOIN Associations AS Associations_1 ON Element_1.ID =
        Associations_1.ID1) INNER JOIN Element AS Element_2 ON Associations_1.ID2 = Element_2.ID)
        ON Element.ID = Associations.ID1
WHERE (((Element.Type)='C') AND ((Associations.Type)='CT') AND ((Element_1.Type)='T')
        AND ((Associations_1.Type)='TRDBS') AND ((Associations_1.ID2)=[Num schema?])
        AND ((Element_2.Type_ele)='RDBS'))
ORDER BY Element_1.Name;
```

Primary Keys by Table. Shows the columns that belongs the primary key of each table of a relational schema

```
SELECT Element.ID AS ID, Element_1.Name AS Table, Element.Name AS PK, Element.Date AS
        Date
FROM Element INNER JOIN ((((Associations INNER JOIN Element AS Element_2 ON Associations.ID2 =
        Element_2.ID) INNER JOIN (Associations AS Associations_1 INNER JOIN Element AS
        Element_1 ON Associations_1.ID2 = Element_1.ID) ON Element_2.ID =
        Associations_1.ID1) INNER JOIN Associations ASAssociations_2 ON Element_1.ID =
        Associations_2.ID1) INNER JOIN Element AS Element_3 ON Associations_2.ID2_aso =
        Element_3.ID) ON Element.ID = Associations.ID1
WHERE (((Element.Type)='C') AND ((Associations.Type)='CPK') AND (Associations_1.Type)='PKT') AND
        ((Element_1.Typo)='T') AND (Element_2.Type)='PK') AND ((Associations_2.Type)='TRDBS') AND
        ((Element_3.Type)='RDBS') AND ((Element_3.ID)=[Num schema?]))
ORDER BY Elemento_1.Name;
```

Foreign Keys by Table. Shows all the columns that belongs to each of the foreign keys of each of the tables of a relational schema.

```
SELECT Element.ID AS ID, Element_1.Name AS Table, Element.Name AS FK, Element.Date AS Date
FROM Element INNER JOIN ((((Associations INNER JOIN Element AS Element_2 ON Associations.ID2 =
        Element_2.ID) INNER JOIN (Associations AS Asociations_1 INNER JOIN Element AS
        Element_1 ON Associations_1.ID2 = Element_1.ID) ON Element_2.ID =
        Associations_1.ID1) INNER JOIN Associations AS Associations_2 ON Element_1.ID =
        Associations_2.ID1) INNER JOIN Element AS Element_3 ON Associations_2.ID2 =
        Element_3.ID) ON Element.ID = Associations.ID1
WHERE (((Element.Type)='C') AND ((Associations.Type)='CFK') AND (Associations_1.Type)='FKT') AND
        ((Element_1.Type)='T') AND (Element_2.Type)='FK') AND ((Associations_2.Type)='TRDBS') AND
        (Element_3.Type)='RDBS') AND ((Element_3.ID)=[Num schema?]))
ORDER BY Element_1.Name;
```

Foreign Keys by Table by Target. Shows the correspondence between columns of each of the FK of a relational schema.

```
SELECT Element_5.ID AS ID_FK, Element_1.Name AS Table, Element.ID AS ID_Field, Element.Name AS
        FK, Element_3.Name AS TTarget, Element_2.ID AS ID_Target, Element_2.Name AS
        Target, Element.Date AS Date
```

```
FROM ((((Associations AS Associations_2 INNER JOIN Element AS Element_2 ON Associations_2.ID2 =
        Element_2.ID)  INNER JOIN Element ON Associations_2.ID1 = Element.ID_ele) INNER JOIN
        Associations AS Associations_3  ON Element_2.ID = Associations_3.ID1) INNER JOIN Element AS
        Element_3 ON Associations_3.ID2 = Element_3.ID)  INNER JOIN (Associations INNER JOIN
        (Element AS Element_5 INNER JOIN (((Associations AS Associations_1  INNER JOIN Element AS
        Element_1  ON Associations_1.ID2 = Element_1.ID) INNER JOIN Associations AS Associations_4
        ON Element_1.ID = Associations_4.ID1) INNER JOIN Element AS Element_4  ON
        Associations_4.ID2 =          Element_4.ID) ON Element_5.ID = Associations_1.ID1)  ON
        Associations.ID2 = Element_5.ID) ON Element.ID = Associations.ID1
WHERE (((Element.Type)='C') AND ((Associations.Type)='CFK') AND (Associations_1.Type)='FKT') AND
        ((Element_1.Type)='T') AND (Associations_2.Type)='CCFK') AND ((Associations_3.Type)='CT')
        AND ((Element_3.Type)='T') AND ((Associations_4.Type)='TRDBS')
        AND ((Element_4.Type)='RDBS') AND ((Element_4.ID)=[Num schema?])
        AND ((Elemento_5.Type)='FK'))
ORDER BY Element_1.Name;
```

For example, if wehave the relational schema showed in figure 8:

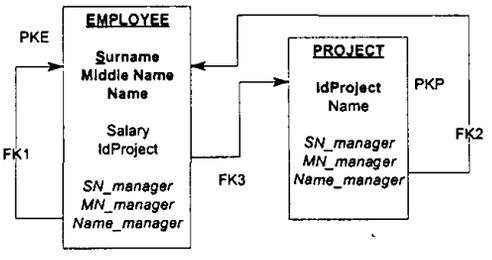


Figure 8. Example Relational Schema

We can represent it in the metadatabse as:

Element Type

ID	Description
C	Column
RDBS	RDB Schema
FK	Foreign Key
PK	Primary Key
T	Table

Association Type

ID	TC	Description
CCFK	S	Column referencing to a Column (FK)
CFK	A	Column belonging to a FK
CPK	A	Column belonging to a PK
CT	A	Column belonging to a Table
FK-T	S	FK referencing to a Table
FKT	A	FK belonging to a Table
PKT	A	PK belonging to a Table
TRDBS	A	Table belonging to a RDB Schema

Element

ID	Name	Date	Type
1	Employee	28/07/99	T
2	Project	28/07/99	T
3	Surname	28/07/99	C
4	Middle_Name	28/07/99	C
5	Name	28/07/99	C
6	Salary	28/07/99	C
7	SN_manager	28/07/99	C
8	MN_manager	28/07/99	C
9	Name_manager	28/07/99	C
10	Id_Project	28/07/99	C
11	Name	28/07/99	C
12	SN_manager	28/07/99	C
13	MN_manager	28/07/99	C
14	Name_manager	28/07/99	C
15	PKE	28/07/99	PK
16	PKP	28/07/99	PK
17	FK1	28/07/99	FK
18	FK2	28/07/99	FK
19	FK3	28/07/99	FK
20	Id_Project	28/07/99	C
21	Example 1: Relational Schema	28/07/99	RDBS

Associations		
ID1	ID2	Type
1		21TRDBS
2		21TRDBS
3		1CT
3		15CPK
4		1CT
4		15CPK
5		1CT
5		15CPK
6		1CT
7		1CT
7		3CCFK
7		17CFK
8		1CT
8		4CCFK
8		17CFK
9		1CT
9		5CCFK
9		17CFK
10		2CT
10		16CPK
11		2CT
12		2CT
12		3CCFK
12		18CFK
13		2CT
13		4CCFK
13		18CFK
14		2CT
14		5CCFK
14		18CFK
15		1PKT
16		2PKT
17		1FK-T
17		1FKT
18		1FK-T
18		2FKT
19		1FKT
19		2FK-T
20		1CT
20		10CCFK
20		19CFK

5. Conclusions and future work

More research is needed in software measurement (Neil, 1994), both from theoretical and from practical points of view. We think it is very interesting to dispose of metrics for relational databases. However, it is not enough with proposing metrics, it is also necessary to put them under formal verification and empirical validation, so we can assure its utility. We have presented a set of metrics to measure the database schema complexity in order to have a mechanism to improve and assure the quality of databases. Different basis to carry out formal validation have been presented with an example. We have carried out an empirical validation with the referential integrity related metrics (Calero et al., 1999) obtaining that the number of foreign keys (RD) is a solid indicator of the database schema understandability and the length of the referential path (DRT) can modulates the RD metric effects.

Metrics for others kinds of databases are being researched as active (Díaz and Piattini, 1999) and object-relational ones (Piattini et al., 1998). We are also improving the MANTICA tool to support more metrics and a richer user interface.

REFERENCES

- Batini, C., Ceri, S. and Navathe, S., (1992), *Conceptual database design. An entity relationship approach.* Benjamin Cummings Publishing Company.
- Briand, L., Morasca, S. and Basili, V. (1996). Property-Based Software Engineering Measurement. *IEEE Transactions on software Engineering,* Vol. 22, No 1, January, pp.68-86
- Calero, C., Piattini M., Polo, M and Ruiz, F. (1999).Validating referential integrity as a database quality metric. *ICEIS* 99, 27-31 March, pp.45-50
- Calero, C., Piattini, M., Genero, M., Serrano, M.A., and Caballero, I.(2000). Metrics for relational database maintainability. UKAIS 2000, March, Cardiff, UK.
- Churcher, N.J. and Shepperd, M.J. (1995), Comments on "A Metrics Suite for Object-Oriented Design. *IEEE Trans. on Software Engineering* Vol. 21, N° 3, 1995, pp. 263-265.
- Codd, E.F. (1970). A Relational Model of Data for Larged Shared Data Banks. *CACM,* 13 (6), pp.377-387.
- Connolly, T., Begg, C. and Strachan, A., (1998), *Database systems*, 2ª Ed. Addison-Wesley
- Date, C.J. (1995*). An Introduction to Database Systems.* 6th. Addison-Wesley, Reading, Massachusetts.

CAISE'99. Heidelberg, June, pp.16-18.

- Fenton, N. (1994). Software Measurement: A Necessary Scientific Basis. *IEEE Transactions on Software Engineering*, 20(3). Pp.199-206.
- Gray R.H.M., Carey B.N., McGlynn N.A. and Pengelly A.D., (1991), Design metrics for database systems. *BT Technology* J, Vol 9, 4 Oct, pp.69-79
- Henderson-Sellers, B. (1996). *Object-oriented Metrics - Measures of complexity*. Prentice-Hall, Upper Saddle River, New Jersey.
- ISO, (1994). Software Product Evaluation-Quality Characteristics and Guidelines for their Use. *ISO/IEC Standard 9126*, Geneva.
- Kitchenham, B. and Stell, J.G. (1997). The danger of using axioms in software metrics. *IEE Proc.-Softw. Eng.* Vol.144, No 5-6, October-December, pp 279-285
- Li, H.F. and Chen, W.K. (1987). An empirical study of software metrics. *IEEE Trans. on Software Engineering*, 13 (6), pp.679-708.
- McCabe, T.J. (1976). A complexity measure. *IEEE Trans. Software Engineering* 2(5), pp.308-320.
- MacDonell, S.G., Shepperd, M.J. and Sallis, P.J. (1997). Metrics for Database Systems: An Empirical Study. *Proc. Fourth International Software Metrics Symposium – Metrics'97*, Albuquerque. IEEE Computer Society, pp. 99-107.
- McFadden, F.R., Hoffer, J.A. and Prescott, M.B., (1999), *Modern database management*, 5ª Ed. Addison-Wesley
- Morasca, S. and Briand, L.C. (1997). Towards a Theoretical Framework fror Measuring Software Attributes. *Proc. of the Fourth International Software Metrics Symposium*, Albuquerque, IEEE Computer Society, pp.119-126.
- Neil, M. (1994) Measurement as an Alternative to Bureaucracy for the Achievement of Software Quality. *Software Quality Journal* 3 (2), pp.65-78.
- Piattini, M., Calero, C., Polo, M. and Ruiz, F. (1998). Maintainability in Object-Relational Databases. . *Proc of The European Software Measurement Conference* FESMA 98, Antwerp, May 6-8, Coombes, Van Huysduynen and Peeters (eds.),pp. 223-230.
- Piattini, M., Calero, C., Polo, F. and Ruiz, F.(2000). Towards a metric suite for relational databases complexity. IDPT2000, June, Texas, USA
- Pfleeger, S. L. (1997). Assessing Software Measurement. *IEEE Software*. March/April, pp. 25-26.
- Sneed, H.M. and Foshag, O. (1998). Measuring Legacy Database Structures. *Proc of The European Software Measurement Conference FESMA 98*, Antwerp, May 6-8, Coombes, Van Huysduynen and Peeters (eds.), pp.199-211.
- Weyuker, E.J. (1998). Evaluating Software Complexity Measures, *IEEE Transactions on Software Engineering*. 14(9), pp.1357-1365.
- Zuse, H. (1998). *A Framework of Software Measurement*. Berlin, Walter de Gruyter.