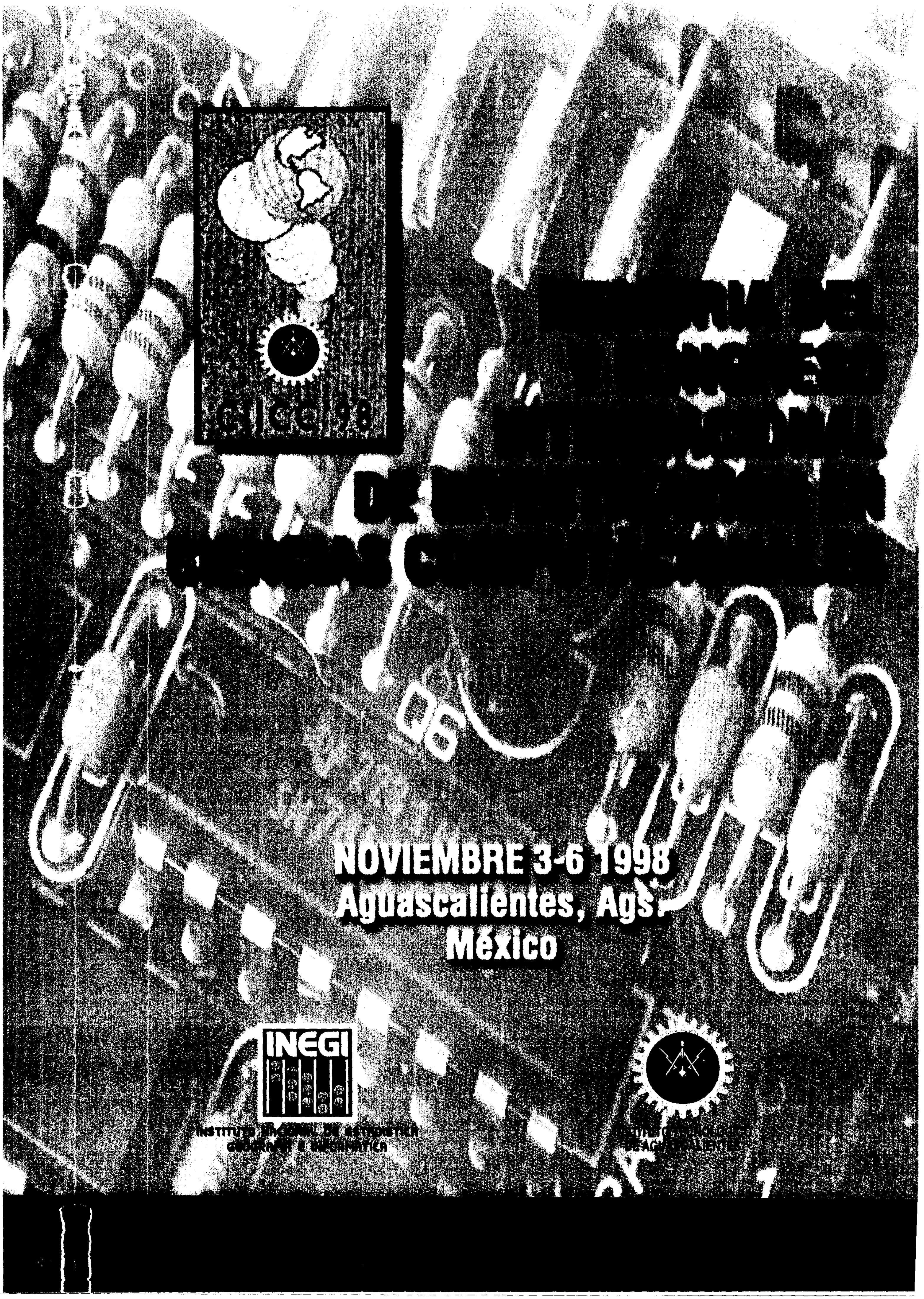
CICC'98
NOVIEMBRE 3-6 1998
Aguascalientes, Ags.
México
INEGI

Edificio Sede
Av. Héroe de Nacozari Núm. 2301 Sur
Fracc. Jardines del Parque, CP 20270
Aguascalientes, Ags.

http://www.inegi.gob.mx
usuario@cis.inegi.gob.mx

Memoria del V Congreso Internacional de Investigación en Ciencias Computacionales

Impreso en México
ISBN 970-13-2261-4

COMITÉ ORGANIZADOR / ORGANIZING COMMITTEE.

Director del Instituto Tecnológico de Aguascalientes
Ing. Angel Castro Cortés

Presidente del Congreso
Dr. Esteban Hernández Pérez
Director General del Sistema de Institutos Tecnológicos

Presidente Honorario de Congreso
Lic. Víctor Guerrero Mendieta
Secretario Técnico De la Comisión Nacional para la Conversión Informática Año 2000
Instituto Nacional de Estadística Geografía e Informática

Vicepresidente del Congreso
Dr. José de Jesús Grande Puigdomenech
Instituto Tecnológico de San Luis Potosí

Presidente del Comité Técnico
Dr. Rodolfo Pazos Rangel
Centro Nacional de Investigación y Desarrollo Tecnológico

Vicepresidente del Comité Técnico
Dra. Nieves Rodríguez Brisaboa
Universidad de la Coruña (España)

Coordinadores Generales
M.C. Marco Antonio Hernández Vargas
Ing. Joaquín Amador Macías
Instituto Tecnológico de Aguascalientes

Coordinación de Impresión
Lic. Héctor Salvador González León
Ing. J. Miguel Becerra de Luna

Coordinación de Conferencias Magistrales e Instructores
Ing. Héctor Jesús Macías Figueroa
M.C. Enrique de Luna Ramírez
M.C. Arturo Nerí Zuñiga

Coordinación de Eventos Sociales y Culturales
Lic. María Luisa Trujillo Avalos
Lic. Lourdes del R. Sánchez Delgado
Lic. Martha C. Lira Ortega
Lic. María Elena Esparza Arellano
Lic. Ma. De Lourdes Moreno Castanedo

Coordinación de Tiraje y Elaboración de Memoria del Evento
Ing. Rodolfo Reyes Solís
Lic. Hugo Hernández Ramos
Ing. Francisco Mora Medina

Coordinación de Organización de Ruedas de Prensa y Difusión
Arq. Luis Antonio Morales Padilla
Ing. Francisco Muro Cornejo

Coordinación de Logística
Ing. Rene Tristán Ávila
Lic. Cesar Dunay Acevedo Arreola
Lic. Arturo López Ponce
Lic. Ilda Díaz Ramos
Lic. Carlos López Pérez
Lic. Frida Eugenia Mora García

Coordinación de información e Inscripciones
Ing. Alejandra Espinosa Guzmán
Lic. J. Asunción Ortíz Martínez

Coordinación de invitación a Patrocinadores
Arq. Luis Antonio Morales Padilla
Ing. Francisco Muro Cornejo

Coordinación de Finanzas
Depto. Recursos Financieros
Del Instituto Tecnológico de Aguascalientes

Comité Técnico _ Technical Committee

M.C. Anastacio Antolino H.
Inst. Tecnológico de Morelia

M.C. Arturo Díaz P.
Centro de Invest. y de Est. Avanzados

Dr. Bárbaro J. Ferro C.*
Centro de Investigación en Computación

Dr. Carlos A. Reyes G.*
Inst. Tecnológico de Apizaco

Dr. David Romero*
Inst. de Matemáticas UNAM

Dr. Eduardo Morales M.*
ITESM Campus Morelos

Dr. Eduardo Oliva L.*
Esc. Sup. de Ingía. Mec. y Eléc. IPN

Dr. Ernest Teniente
Univ. Politécnica de Cataluña

Dr. Esperanza Marcos
Univ. Carlos III de Madrid

Dr. Guillermo Morales L.*
Centro de Invest. y de Est. Avanzados

Dr. Guillermo Rodríguez O.*
Inst. de Investigaciones Eléctricas

M.C. Héctor Fraire H.
Inst. Tecnológico de Cd. Madero

Dr. Héctor J. Hernández L.*
Univ. Autónoma de Cd. Juárez

Dr. Hugo C. Coyote E.*
Centro de Investigación en Computación

M.C. Hugo Estrada E.
Centro Nal. de Inves. y Des. Tecnológico

Dr. Jorge S. Benítez R.*
Inst. Nal. de Investigaciones Nucleares

M.C. José A. Martínez Flores
Inst. Tecnológico de Cd. Madero

Dr. José A. Segovia de Los R.*
Inst. Nal. de Investigaciones Nucleares

M.C. José F. Carpio T.
Inst. Tecnológico de Zacatepec

Dr. José Ruiz A.*
Centro Nal. de Inves. y Des. Tecnológico

Dr. José Torres J.*
ITESM Campus Morelos

M.C. Juan G. González S.*
Centro Nal. de Inves. y Des. Tecnológico

M.C. Laura Cruz R.
Inst. Tecnológico de Cd. Madero

M.C. Leopoldo Zepeda S.
Inst. Tecnológico de Culiacán

M.C. Manuel Mejía L.
Inst. de Investigaciones Eléctricas

Dr. Manuel Pérez C.
Universidad de Vigo

M.C. María L. Barrón E.
Inst. Tecnológico de Culiacán

M.C. Mario Guillén R.*
Inst. de Investigaciones Eléctricas

M.C. Máximo López S.*
Centro Nal. de Inves. y Des. Tecnológico

Dra. Nieves Rodríguez B.*
Universidad de La Coruña

M.C. Raúl Pinto E.*
Centro Nal. de Inves. y Des. Tecnológico

M.C. René Santaolaya S.*
Centro Nal. de Inves. y Des. Tecnológico

Dr. Rodolfo A. Pazos R.*
Centro Nal. de Inves. y Des. Tecnológico

M.C. Ruth Saez de N. R.
Inst. Tecnológico de León

M.C. Victor J. Sosa S.*
Centro Nal. de Inves. y Des. Tecnológico

* Miembro del Comité Técnico Permanente de la ANICC.

Contenido - Contents

SESIÓN 2A: COMPUTACIÓN DISTRIBUÍDA
SESSION 2A: DISTRIBUTED COMPUTING

SESIÓN 2B: REUSO DE SOFTWARE
SESSION 2B: SOFTWARE REUSE

SESIÓN 2C: SISTEMAS EXPERTOS
SESSION 2C: EXPERT SYSTEMS

Measures for Control Database Quality[1]

Mario Piattini Velthuis, Macario Polo Usaola, Coral Calero Muñoz, Francisco Ruiz González
Escuela Superior de Informática – Universidad de Castilla-La Mancha
Ronda de Calatrava, 7 – 13071-Ciudad Real (SPAIN)
E-mail: {mpiattin, mpolo, ccalero, fruiz}@inf-cr.uclm.es

Abstract

Although databases are nowadays one of the most important component of an Information System, measuring their quality have been a neglected area. Most of the existing works on metrics only deals with programs metrics. In this paper a set of metrics for control database quality during its life-cycle are presented.

Key words: database design, quality, software metrics.

1. Introduction

Metrics for databases have been neglected in the metric community [13]. Most all of the metrics proposed from the McCabe [8] famous cyclomatic number until today have been centered in measuring programs complexity, quality, maintenance, etc. However, in modern Information Systems (IS) the database has become a crucial component, so there is a need to propose and study some quality measures to assess the database quality.

In [1] several database quality attributes are proposed such as: completeness, correctness, minimality, expressivity, legibility, self-explanation, extensibility and normality, but none quantitative measure is proposed. Also in [12] guidelines to improve conceptual schema quality are given, but without any quantitative reference. Exceptions are [7], where a set of metrics for conceptual database schemas is proposed (based on the number of generalization hierarchies, relationships, entities, etc., but which requires long manual procedures), and [5] where a normalization index is developed. Regarding relational databases, there is also hardly any reference to database quality, excepting normal forms.

Recently, a more complete work on measuring databases has been published [13], proposing quantitative, complexity and qualitative factors for legacy databases. These authors outline five characteristics which can be associated with data structure: accessibility, storage efficiency, adaptability, convertibility and reusability.

In this paper we present a complete framework for measuring databases in both the conceptual and the logical stages, following the philosophy of ISO 9126 International Standard [6], which considers quality as an aggregation of several factors. The factors chosen in this work followed the attributes proposed by Batini et al. [1] complemented by others factors, focused mostly in database maintainability. Maintenance is considered the most important concern for modern IS department and requires greater attention by the software community [11].

[1]This work is partialy supported by MANTICA project (CICYT-1FD97-0168).

This article is organized as follows: in section 2, we expose a set of metrics for conceptual schemas. Metrics for logical (relational) schemas are presented in section 3. The application of these metrics to an example is shown in section 4. Finally, section 5 summarizes the paper and draw our conclusions.

2. Metrics for the conceptual schema

The objective of the conceptual design is to obtain a good representation of the organization information resources with user and specific applications independence. It is based on a conceptual model (usually a variation of the Entity-Relationship Model) which is used for the elaboration of a conceptual schema. The following properties and measures are proposed for the conceptual schema.

2.1 Completeness

Completeness is measured by a nominal metric that can be *high*, *low* or *very low*.

2.2. Correctness

The correctness metric can have also the *high*, *low* or *very low* values. Batra (1993) taxonomy can be used for giving one of these values.

2.3 Minimality

Minimality is measured as a function of the number of duplicated attributes present in some entity or relationship or attributes that can be derived The minimality ratio (MR) is defined as:

$$MR = 1 - \frac{NDA}{NAS}$$

...where: NDA is the Number of Derived or Duplicated Attributes

NAS is the Number of Attributes in the Schema

2.4 Expressivity

For measuring expressivity three ratios are proposed:

$$SFAR = \frac{SFA}{NAS}$$

...where SFA is the number of attributes with full semantic content (Semantic Full Attributes) and NAS is, as we said in the previous subsection, the Number of Attributes in the conceptual Schema.

$$SFER = \frac{SFE}{NES}$$

...where SFE is the number of entities with full semantic content (Semantic Full Entities) and NES is the Number of Entities in the Schema.

$$SFRR = \frac{SFR}{NRS}$$

...where SFR is the number of relationships whose names have full semantic content (Semantic Full Relationships) and NRS is the Number of

Relationships in the Schema.

An attribute, an entity or a relationship has full semantic content when its name is descriptive and accordant to the rules of the organization, in the sense given by Reingruber and Gregory [12].

2.5 Legibility and self-explanation

Although this metric is the most subjective of ours, we must not forgot Fenton and Pfleeger [4], who say: "it is important to recognize that subjective measurements can be useful, as long as we understand their imprecision". This metric is, as completeness and correctness, a nominal metric with values in *high*, *low* or *very low*.

2.6 Extensibility and normality

Since the Third Normal form (3NF) is the most common and natural to represent the universe of discourse, we propose to use the Gray et al. [5] metric:

$$N = \frac{A_e \times E_e + R_e \times A_e + R_e \times E_e}{A_3 \times E_3 + R_3 \times A_3 + R_3 \times E_3}$$

..where the numerator is calculated in function of the real schema, and the denominator in function of the same schema in 3NF.

Parameter "A" is the number of attributes, "E" is the number of entities and "R" is the number of relationships.

2.7 Cohesion of the schema

We propose the use of the following "Cohesion Of the Schema Ratio" to measure this property:

$$COSR = \frac{\sum_{i=1}^{|US|} NEUSi^2}{NES^2}$$

|US| is the number of Unrelated Subgraphs.

$NEUS_i$ is the Number of Entities in the Unrelated Subgraph number "i".

2.8 Intra-entity complexity

Intra-entity complexity is defined as the number of relationships of an entity with itself. Following Melton [9], we can provide a valid metric for the complete schema giving the mean and standard deviation of IEC, in this manner:

IEC_{SCHEMA}= (mean of IEC, standard deviation of IEC)

2.9 Schema complexity

The schema complexity is measured by the following ratio:

$$SCR = \frac{SC}{MNR}$$

...where: SC is the number of relationships in the schema and MNR is the Theoretical Maximum possible Number of Relationships in the schema. MNR is calculated through this expression (being NES the Number of Entities in the Schema):

$$MNR = \frac{NES \times (NES - 1)}{2}$$

2.10 Relationship's length

It is the length of the longest relationship's path in the schema, excluding reflexive ones

2.11 Size

We measure the schema size as a triple, formed by the number of entities, the number of relationships and the number of attributes in the schema. So:

S=(NES, NRS, NAS)

3. Metrics for the logical schema

The objective of the logical design is to transform the conceptual schema into a logical schema which organizes information according the logical model supported by the database management system. In this paper we focus, for obvious reasons, on relational databases, proposing the following metrics:

3.1. Completeness

Completeness, as for conceptual schemas, is a nominal metric and it will be ***high***, *low* or *very low*.

3.2. Correctness

Correctness can also be *high*, *low* or *very low*.

3.3 Minimality

A database has minimality when each aspect of the requirements appears only once in the relational scheme. We measure the minimality as a function of the number of derived or duplicated fields (NDF) in the tables of the database, similarly to conceptual schema minimality (see 2.3).

3.4 Expressivity

Expressivity is measuring by the following ratios:

$$SFFR = \frac{SFF}{NFS}$$

...where SFF is the number of fields with full semantic content (Semantic Full Fields) and NFS the Number of Fields in the relational Schema.

$$SFTR = \frac{SFT}{NTS}$$

...where SFT is the number of tables with full semantic content (Semantic Full Tables) and NTS is the Number of Tables in the Schema.

$$FPCR = \frac{FPC}{FKS}$$

...where FPC is the number of foreign keys whose names agree their respective primary keys (Foreign-Primary Concordance), excluding foreign keys to the proper table and FKS is the number of Foreign Keys in the Schema (excluding those whose respective primary keys are in the same table).

3.5 Legibility and self-explanation

We define legibility directly as a ratio proposed by Fenton and Pfleeger (1996) (named "Comments density"):

$$LR = \frac{CLOC}{LOC + CLO}$$

...where:

LR is the Legibility Ratio

CLOC is the number of Commented Lines Of Code.

LOC is the number of Lines Of Code.

According to Fenton and Pfleeger [4], the most extended definition of "Line Of Code" was proposed by Hewlett-Packard: any statement, excepting commentary lines and blank lines. We adopt this definition, also valid for databases context and, in particular, for SQL language.

3.6. Extensibility and normality

We define normality as the number of tables in third normal form (3NF) or higher divided by the number of tables in the schema. So:

$$NR = \frac{NT3NF}{NTS}$$

...where:

NT3NF is the Number of Tables in the 3NF

NTS is the Number of Tables in the relational Schema.

Every table in a normal form higher than 3NF is also in third normal form; so tables in BCNF, 4NF and 5NF are counted within this category.

3.7 Cohesion of the schema

Cohesion Of the Schema (COS) is defined as the number of unrelated subgraphs that we can find in the relational graph which represents the data base. Based on this metric we propose to use the "Cohesion Of the Schema Ratio":

$$COSR = \frac{\sum_{i=1}^{|US|} NTUSi^2}{NTS^2}$$

|US| is the number of unrelated subgraphs.

$NTUS_i$ is the number of tables in the unrelated subgraph number "i".

3.8 Intra-table complexity

For this property we propose to use the intra-complexity ratio of a table, defined as:

$$ITCR = \frac{ITC}{NFT}$$ ...where:

ITC is the number of FK whose respective PK is PK of its own table

NFT is the number of fields in the table we are considering.

ITC ant ITCR are metrics for each table in the database. For the schema we propose to use the mean and the standard deviation of ITC and ITCR:

ITC_{SCHEMA}= (Mean of ITC, Standard deviation of ITC)

$ITCR_{SCHEMA}$=(Mean of ITCR, Standard deviation of ITCR)

3.9 Schema complexity

The complexity of the schema is defined by the Schema Complexity Ratio:

SCR=SC/NFS ...where: SC (Schema Complexity) is defined as the number of Foreign Keys in the Schema and NFS is the Number of Fields in the Schema.

3.10. Referential length

It is the length of the longest referential path in the schema.

3.11 Size

We measure the schema size as the number of tables in the schema (NTS).

4. Example

Below, we apply our metrics to an example developed in De Miguel and Piattini [3] around a library. First we show in figure 1 the conceptual schema, and the SQL-92 Schema is resumed subsequently.

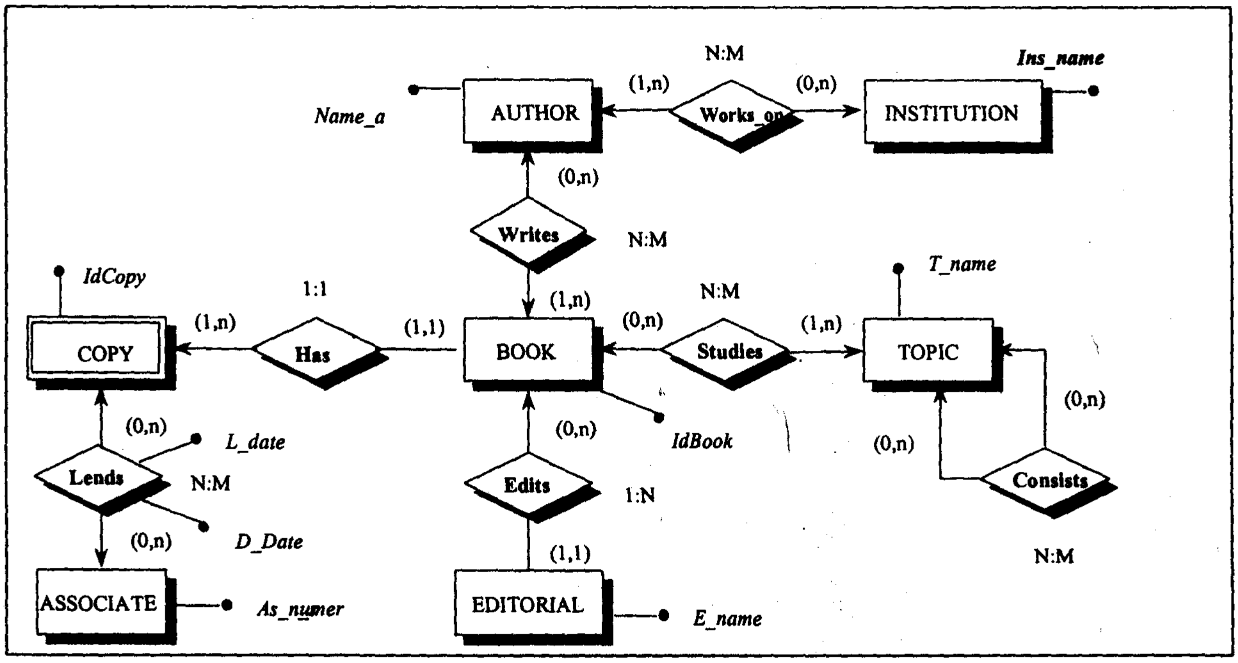

Figure 1. Conceptual schema for the example.

```
CREATE SCHEMA LIBRARY
          ********** DOMAINS *****************
          CREATE DOMAIN names CHAR(20)
          CREATE DOMAIN nacionalities CHAR(15)
          CREATE DOMAIN institutions CHAR(30)
          CREATE DOMAIN codes CHAR(4)
          CREATE DOMAIN titles CHAR(40)
          CREATE DOMAIN languages CHAR(1)
            CHECK (VALUE IN ('I','F','E','T','A','C') )
                    DEFAULT 'I'
          CREATE DOMAIN copies_no INTEGER(2)
                    CHECK (VALUE >= 0)
          CREATE DOMAIN as_number INTEGER(4)
                    CHECK (VALUE >=0)
          CREATE DOMAIN places CHAR(20)
          CREATE DOMAIN telephones INTEGER(9)
                    CHECK (VALUE > = 0)
          CREATE DOMAIN as_type CHAR(1)
                    CHECK (VALUE IN ('N','D','P') )
          CREATE DOMAIN dates DATE
          CREATE DOMAIN years INTEGER(4)
                    CHECK (VALUE > = 1960)
          CREATE DOMAIN topics CHAR(15)

          ************ TABLES *****************
          CREATE TABLE author
          (a_name            names,
           nationality       nacionalities,
           PRIMARY KEY (a_name) )
```

```
CREATE TABLE works_on
(a_name            names,
 ins_name          institutions,
 PRIMARY KEY (a_name, ins_name),
FOREIGN KEY (a_name) REFERENCES author
          ON UPDATE CASCADE)

CREATE TABLE institution
(ins_name          institutions,
 address           places,
 Tel               telephones,
 PRIMARY KEY (ins_name))

CREATE TABLE book
(Book_IdCodes,
 title             titles NOT NULL,
 languagelanguages NOT NULL,
 copies_no         copies_no NOT NULL,
 e_name            names NOT NULL,
 year              year,
PRIMARY KEY (Book_Id),
FOREIGN   KEY   (e_name)   REFERENCES
editorial
          ON UPDATE CASCADE)

CREATE TABLE writes
(a_name            names,
Book_Id            Codes,
PRIMARY KEY (a_name, Book_Id),
FOREIGN KEY (a_name) REFERENCES author
          ON UPDATE CASCADE,
FOREIGN KEY (Book_Id) REFERENCES book
          ON DELETE CASCADE
```

```
ON UPDATE CASCADE)

CREATE TABLE copy
(Book_Id codes,
identificative      copies_no,
PRIMARY KEY (Book_Id, identificative),
FOREIGN KEY (Book_Id) REFERENCES book
        ON DELETE CASCADE
        ON UPDATE CASCADE)

CREATE TABLE associate
(as_number      Associate_codes
name            names NOT NULL,
address         places NOT NULL,
tel             telephones,
as_type    types_of_associates NOT NULL,
PRIMARY KEY (as_number))

CREATE TABLE lends
(Book_Id codes,
identificative      copies_no,
as_number           Associate_codes,
L_date              dates,
D_date              dates,
CHECK (d_date > = l_date),
PRIMARY KEY (Book_Id, identificative,
as_number, l_date),
FOREIGN KEY (Book_Id, identificative)
REFERENCES copy,
FOREIGN KEY (as_number) REFERENCES
associate
        ON DELETE CASCADE
        ON UPDATE CASCADE)
```

*** Observation: the constraint verifies that the devolution date is higher than the lend date.

```
CREATE TABLE topic
(t_name         topics,
t_desc          descriptions,
PRIMARY KEY (t_name) )

CREATE TABLE consists
(t_name         topics,
s_topic         topics,
PRIMARY KEY (t_name, s_topic),
FOREIGN KEY (t_name) REFERENCES topic
        ON DELETE CASCADE
        ON UPDATE CASCADE,
FOREIGN KEY (s_topic) REFERENCES topic
        ON DELETE CASCADE
        ON UPDATE CASCADE)

CREATE TABLE editorial
(e_name         nombre,
address         places NOT NULL,
city            places NOT NULL,
country         places NOT NULL,
PRIMARY KEY (e_name) )

CREATE TABLE studies
(Book_Id Codes,
t_name          topics,
PRIMARY KEY (Book_Id, t_name),
FOREIGN KEY (t_name) REFERENCES topic,
FOREIGN KEY (Book_Id) REFERENCES book,
ON UPDATE CASCADE
                ON DELETE CASCADE)
```

4.1 Metrics for the conceptual schema

a) Completeness: *High,* supposing that all the requirements are present in the schema.

b) Correctness: it is, obviously, ***high***.

c) Minimality: we cannot find attributes in excess in the conceptual schema. So:

$$MR = 1 - \frac{0}{9} = 1$$

d) Expressivity: let us suppose that rules of the organization say that all the names used in designs must be the real names of every object and using, for example, Hungarian Notation, and that the only exception is the rule for the key attributes (which must be suffixed by "Id").

In the ER schema, there are only two attributes (*Copy_Id* and *Book_Id*) which fulfil this rule. So:

$$SFAR = \frac{2}{9}$$

$$SFER = \frac{7}{7} = 1$$

$$SFRR = \frac{6}{7} = 0.85$$ ...because *Works_on* is not in Hungarian Notation (it would be *WorksOn*)

e) Legibility and self-explanation: it's *high*.

f) Extensibility and normality: as the ER schema is in third normal form, N=1.

g) Cohesion of the schema:

$$COSR = \frac{7^2}{7^2} = 1$$ ...because there is only a unrelated subgraph, with the seven tables that there are in the schema.

h) Intra-entity complexity: as we have only one entity with reflexive relationships (*Consists*), the value of this metric is:

$$IEC_{SCHEMA} = (0.14, 0.38)$$

i) Schema complexity:

$$SCR = \frac{7}{\frac{7 \times 6}{2}} = \frac{7}{21} = 0.33$$ ...because there are 7 relationships (numerator) and 7 entities (denominator).

j) Relationship's length: its value is 4.

k) Size: S=(7, 7, 9)

4.2 Metrics for the logical schema

a) Completeness: *High*.

b) Correctness: we see that the relationship between *Institution* and *Works_on* expressed in the conceptual schema is not present in the logical one: although the designer has reserved a field (*ins_name*) to this foreign key, he has forgotten to write "FOREIGN KEY (ins_name) REFERENCES institution" in the *Works_on* relation. So, the value for this metric is *low*

c) Minimality:

$$MR = 1 - \frac{1}{37} = 0.97$$

...because the *copies_no* field (in table *Book*) can be derived from *Copy*.

d) Expressivity:

$$SFFR = \frac{1+0+1+4+0+1+2+1+0+0+3+0}{37} = \frac{13}{37} = 0.35$$

$$SFTR = \frac{11}{12} = 0.92$$

...because *Works_on* is not in Hungarian Notation.

$$FPCR = \frac{11}{11} = 1$$

e) Legibility and self-explanation:

$$LR = \frac{3}{110+3} = 0.026$$

f) Extensibility and normality:

N=1 ...because all the tables are in 3FN.

g) Cohesion of the schema:

$$COSR = \frac{1^2+11^2}{12^2} = \frac{122}{144} = 0.85$$

...because *Institution* is not related with any other table.

h) Intra-table complexity:

ICR=0 ...because there are not foreign keys with respective primary keys in the same table.

i) Schema complexity:

$$SCR = \frac{11}{37} = 0.29$$

j) Referential length: its value is 2.

k) Size: it is 12 because there are twelve tables.

5. Conclusions and future work

Database measures is a disregarded Software Engineering research area. We have proposed some metrics for control database quality for both the conceptual and the logical steps of database design. These metrics can help 1) designers in choosing between alternative design options, 2) software quality engineers and auditors to quantify the quality of database design, and 3) managers to estimate maintenance costs.

In table 1 we summarize the metrics and ratios proposed for the database life-cycle.

These metrics are being validated in several experiments both with students and in companies. We will try to obtain a set of values that could be used as limits for judging a conceptual or a logical design.

We are also working on object-relational databases, in which some of these metrics can be complemented with other based on object technology [10].

Attribute	Metrics		Possible values
	Conceptual	Logical	
Completeness	Nominal	Nominal	*High, low, very low*
Correctness	Nominal	Nominal	*High, low, very low*
Minimality	MR=1-(NDA/NES)	MR=1-(NDF/NFS)	[0, 1]
Expressivity	SFAR=SFA/NAS SFER=SFE/NES SFFR=SFR/NRS	SFFR=SFF/NFS SFTR=SFT/NTS FPCR=FPC/FKS	[0,1] [0,1] [0,1]
Legibility and self-explanation	Nominal	LR=CLOC/(LOC+CLOC)	[0,1]
Normality and extensibility	$N = \frac{A_4 \times E_4 + R_4 \times A_4 + R_4 \times E_4}{A_3 \times E_3 + R_3 \times A_3 + R_3 \times E_3}$	NR=NT3NF/NTS	[0,1]
Intra-element complexity	IECR=(mean of IEC, standard deviation of IEC)	ITCR=ITC/NFT	[0, infinite) for the conceptual schema [0,1] for the logical one
Schema complexity	SCR=SC/MNR	SCR=SC/NFS	[0, infinite) for the conceptual schema [0,1] for the logical one

Attribute	Metrics		Possible values
	Conceptual	Logical	
Relationship/Referential length	LR = length of the longest relationship's path	LR = length of the longest referential path	[0, infinite)
Size	S=(NES, NRS, NAS)	S= No. of tables	[0, infinite) for all the metrics
Cohesion of the schema	$COSR = \frac{\sum_{i=1}^{\lvert US \rvert} NEUSi^2}{NES^2}$	$COSR = \frac{\sum_{i=1}^{\lvert US \rvert} NTUSi^2}{NTS^2}$	(0,1] (zero is a limit value, unreachable in the practice)

Table 3. Summarized metrics and ratios.

References

[1] Batini, C., Ceri, S. and Navathe, S. (1992). *Conceptual database design. An entity relationship approach.* Benjamin Cummings Publishing Company.

[2] Batra, D. (1993). A framework for studying human error behavior in conceptual database modeling. *Information & Management*, No. 25, pp. 121-131.

[3] De Miguel, A. and Piattini, M. (1997). *Fundamentos y modelos de Bases de Datos. Madrid,* Ed. Ra-Ma.

[4] Fenton, N. E. and Pfleeger, S. L. (1997). *Software metrics. A rigorous & practical approach.* International Thomson Computer Press.

[5] Gray et al. (1991). *Design metrics for database systems.* BT Technology, Vol. 9.

[6] ISO (1994). *Software Product Evaluation-Quality Characteristics and Guidelines for their Use.* ISO/IEC Standard 9126, Geneva.

[7] Kesh, S. (1991). Evaluating the quality of entity relationship models. *Information and Software Technology*, vol. 37, no. 12, pp. 681-689

[8] McCabe, T. (1976). A complexity measure. *IEEE Trans. Software Engineering* Vol 2, no. 5, pp. 308-320.

[9] Melton, A. (ed.) (1995). *Software Measurement.* International Thompson Computer Press.

[10] Piattini, M., Calero, C., Polo, M. and Ruiz, F. (1998). Maintainability in object-relational databases. *Proc. of The European Software Measurement Conference FESMA'98*, Coombes, Hooft and Peeters (eds.), pp. 223-229.

[11] Pigoski, T.M. (1996). *Practical Software Maintenance. Best Practices for Managing Your Investment.* John Wiley & Sons. USA, 1996.

[12] Reingruber, M. C. and Gregory, W. W. (1994). *The Data Modeling Handbook. A best-practice approach to building quality data models.* John wiley & sons, Inc.

[13] Sneed, H.M. and Foshag, O. (1998). Measuring Legacy Database Structures. *Proc. of The European Software Measurement Conference FESMA'98*, Coombes, Hooft and Peeters (eds.), pp. 199-210.