EXPERIMENTAL VALIDATION OF SQL METRICS

ANTONIO MARTÍNEZ[+1], MARIO PIATTINI[+]
[+]GRUPO ALARCOS
DEPARTAMENTO DE INFORMÁTICA
UNIVERSITY OF CASTILLA-LA MANCHA
RONDA DE CALATRAVA, 5
13071, CIUDAD REAL (ESPAÑA)
E-MAIL: {amartinez, mpiattin} @inf-cr.uclm.es

[1]EXCMA. DIPUTACION DE CIUDAD REAL
CALLE TOLEDO, 17
13001, CIUDAD REAL (SPAIN)

Abstract[1]

Fourth Generation Language environments are substituting, more and more Third Generation Language, as a platform of system computer development. This is why, it is essential to control its complexity and maintenance. A way of carrying out this control is through the use of specific metrics for these environments, which is a field of software engineering where little research is done.
The Fourth Generation Language is classified as a sub-language because it is composed of sentences of a different nature (i.e. heterogeneous) to apply definite metrics.
In this article we describe three kinds of metrics of database manipulation sub-language, in particular, the SELECT sentence. An empirical study to demonstrate that these metrics affect the maintenance of SELECT sentence, and consequently, the maintenance of databases is presented. Eight cases to validate empirically the influence of definite metrics, in the maintenance of designed Fourth Generation environments. Considering the obtained results, we conclude that the number of tables (NT), the number of nesting (NA) and the grouping (A) affect the maintenance of Fourth Generation environments.

Keywords: Empirical software engineering, Measure, SQL Language.

[1] This work is part of MANTICA project; it is financed by CICYT and the European Union (1FD97-0168)

1. INTRODUCTION.

Many organisations which use management information systems are now aware that computer systems constructed using third generation languages such as COBOL can be more effectively produced and maintained using modern productivity-enhancing tools. These tools have been given various names, including fourth-generation languages (4GLs), application generators, or more recently fourth-generation systems (4GSs). (Holloway, 1990).
Metrics are useful mechanisms in improving the quality of software products, specially maintenance, which is the most important problem of software development, ranging between 60 and 90 percent of life-cycle costs (Card and Glass, 1990; Pigoski, 1997). Software measurement is widely recognised as an effective means to understand, monitor, control, predict and improve software development and maintenance projects (Briand et al., 1996). Measurement is used not only for understanding, controlling, and improving development, but also for determining the best ways to help practitioners and researchers.
Maintainability is achieved by means of three factors: understandability, modifiability and testability, which are in turn influenced by complexity (Li and Cheng, 1987). However, a general complexity is *"the impossible holy grail"* (Fenton, 1994). Henderson-Sellers (1996) distinguishes three types of complexity: computational, psychological and representational, and for psychological complexity he considers three components: problem complexity, human cognitive factors and product complexity. The last one is our focus.

In this paper three different types of metrics for the database manipulation sub-language, are proposed. They measure the fourth generation environment complexity. In section 2, we describe the metrics. In section 3 we show the empirical validation for the proposed metrics. Finally, in section 4 we summarise the paper and present the conclusions.

2. METRICS FOR DATA BASE MANIPULATION SUBLANGUAGE.

Different types of metrics have been defined for 4GL. So far, some projects have been developed to estimate the effort of development and the correlation of these with the size of a program (Dolado, 1997; Verner and Tate, 1988), but we think that projects to control the quality of 4GL programs are necessary too. In 4GL environments we have identified different database maintenance sub-languages (Martinez and Piattini, 1998), we propose three kinds of metrics for the sub-language and singularised it to SELECT sentence as follows:

Metrics NT
It expresses the number of tables that the SELECT sentence contains.

Metrics NA
Number of nesting in the SELECT sentence.

Metrics A
In the SELECT sentence, it indicates marks whether there is grouping (A=1) or not (A=0).

```
select f.p0_nom_nomi, p.num_ficha, p.fecha
      from prueba p, fper020 f, hor_personal h
      where p.nif not in (select h.nif
                           from prueba p, fper020 f, hor_personal h
                          where p.num_ficha=h.num_ficha
                            and f.p0_nif=h.nif
                            and p.nif=f.p0_nif
                            and p.fecha='171298'
                            and p.control='SM'
                            and p.estado='A'
                            and f.p0_sexo='V'
                            and p.hora in (select hora
                                             from prueba p, fper020 f,
                                                  hor_personal h
                                            where p.num_ficha=h.num_ficha
                                             and f.p0_nif=h.nif
                                             and p.nif=f.p0_nif
                                             and p.fecha='171298'
                                             and p.control='SM'
                                             and p.tipo='A0'
                                             and h.saldot=0
                                          )
                         )
   and p.fecha='151298'
   and p.num_ficha=h.num_ficha
   and p.nif=f.p0_nif
   and f.p0_nif=h.nif
   group by f.p0_nom_nomi, p.num_ficha, p.fecha
```

Figure 1. SELECT sentence example

The SELECT sentence we can caracterise in base to the values NT=3, NA=3 y A=1.

3. EMPIRICAL VALIDATION OF THE PROPOSED METRICS.

An empirical validation has been carried out following the experimental method applied to software engineering (Basili, 1998; Pleeger, 1997).
Our aim is to demonstrate that the proposed metrics can be used for measuring the complexity of the database manipulation sub-language (in particular for the SELECT sentence).

3.1. EXPERIMENT

We work with the NT, the NA and the A metrics in order to test if some of them are relevant for measuring the understandability of the SELECT sentence.

Hypotheses

- Null hypothesis: Different values of metrics do not affect the comprehension of the SELECT sentence.
- Alternative hypothesis 1: The value of the NT metrics affects the comprehension of the SELECT sentence.

- Alternative hypothesis 2: The value of the NA metrics affects the comprehension of the SELECT sentence.
- Alternative hypothesis 3: The value of the A metrics affects the comprehension of the SELECT sentence.
- Alternative hypothesis 4: The combination of NT and NA metrics affects the comprehension of the SELECT sentence.
- Alternative hypothesis 5: The combination of NT and A metrics affects the comprehension of the SELECT sentence.
- Alternative hypothesis 6: The combination of NA and A metrics affects the comprehension of the SELECT sentence.
- Alternative hypothesis 7: The combination of NT, NA and A metrics affects the comprehension of the SELECT sentence.

Subjects

The participants in the experiment are Computer Science students at the University of Castilla-La Mancha (Spain), who were enrolled in the databases course for the last two semesters.
Until the day of the experiment, the students did not Know that they were going to do it. The experiment was developed by 34 students, but only 19 were finally accepted.
We have tried to minimise variability among participants by choosing people of the same degree, in particular from the third and fourth year. Effects of irrelevant variables were minimised by making the same trials for all the subjects with the same duration (forty five minutes per test).

Experimental materials.

To test the hypotheses eight separate software designs were required. In each one the values of the two metrics were different. There were two possible values for NT metrics (one or three), for NA metrics (one or three) and for A metrics (zero or one).
The documentation accompanying each design was approximately twelve pages long including the tables and the queries (see appendix A).
The subjects were asked to perform eight tasks and they had to write down the initial and the final time, and the result for each query. Each subject wrote down the time manually; we knew that there was error in the measure. In addition, the queries were given to them in distinct order.

Experimental Design.

Each level of one factor appears with each level of the other one, so we have selected the crossing design. This crossing relationship is denoted as A x B x C. For us, A is the NT metrics, B is the NA metrics and C is the A metrics. See table 1.
To increase the power of the test, α has been set to 0.1 instead of 0.05 level which is more common (Briand et al., 1997).

		FACTOR A (NT)			
		LOW		HIGH	
		FACTOR B (NA)			
		LOW	HIGH	LOW	HIGH
FACTOR C (A)	LOW	1,1,0	1,3,0	3,1,0	3,3.0
	HIGH	1,1,1	1,3,1	3,1,1	3,3,1

TABLE 1.- Crossed Design for the experiment

Experimental Results

There are three major items to consider when choosing the analysis techniques: the nature of the data collected, the reason why the experiment is performed and the type of experimental design used (Pfleeger. 1995).
Due to the type of experiment used, we use the SPSS v. 7.5 software and F statistic because it is the technique to obtain the results (Rohatgi, 1976).
Table 2 shows the results for the F-statistic:

```
            TIEMPO
      by    NT
            NA
            A

            UNIQUE sums of squares
            All effects entered simultaneously


                                    Sum of                    Mean                  Sig
Source of Variation                 Squares      DF           Square        F       of F

Main Effects                       1392,132       3         464,044    382,878    ,000
   NT                               796,737       1         796,737    657,380    ,000
   NA                               553,289       1         553,289    456,514    ,000
   A                                 42,105       1          42,105     34,741    ,000

2-Way Interactions                   43,184       3          14,395     11,877    ,000
   NT        NA                      42,105       1          42,105     34,741    ,000
   NT        A                         ,658       1            ,658       ,543    ,462
   NA        A                         ,421       1            ,421       ,347    ,557

3-Way Interactions                    2,132       1           2,132      1,759    ,187
   NT        NA        A            2,132       1           2,132      1,759    ,187

Explained                          1437,447       7         205,350    169,432    ,000

Residual                            174,526     144           1,212

Total                              1611,974     151          10,675


   cases were processed.
   cases (,0 pct) were missing.
```

Comparing these values with $F_{1, 151} = 2.71$, we can ensure that:

- Alternative Hypothesis 1: " The value of the NT metrics affects comprehension of SELECT sentence."
  As 657.380>2.71, NT affects results of experiment, so that the alternative hypothesis 1 is valid.
- Alternative Hypothesis 2: The value of the NA metrics affects comprehension of SELECT sentence.
  As 456.514>2.71, NA affects results of experiment, so that the alternative hypothesis 2 is valid.
- Alternative Hypothesis 3: The value of the A metrics affects comprehension of SELECT sentence.
  As 34.741>2.71, A affect to results of experiment, so that the alternative hypothesis 3 is valid.
- Alternative Hypothesis 4: Combination of NT and NA metrics affects comprehension of SELECT sentence.
  As 34.741>2.71, the interaction of NT and NA affects results of experiment, so that, the alternative hypothesis 4 is valid.
- Alternative Hypothesis 5: Combination of NT and A affects comprehension of SELECT sentence.
  As 0.543<2.71, there is no significant effect of interaction between NT and A
- Alternative Hypothesis 6: Combination of NA and A affects comprehension of SELECT sentence.
  As 0.421<2.71, there is no significant effect of interaction between NA and A
- Alternative Hypothesis 7: Combination of NT, NA and A affects comprehension of SELECT sentence.
  As 1.759<2.71, there is no significant effect of interaction between NT,NA and A

We can conclude that the three kinds of metrics proposed have proved to be valid to evaluate the complexity of SELECT sentence. In addition, considering the level of confidence established, there are only exits meaningful differences in relation to the interaction between the number of nesting and the number of tables, but not in relation to the other interactions.

5. CONCLUSIONS OF EXPERIMENT.

There is a big necessity to measure the quality of the applications based on Fourth Generation Languages. Measurements can help to get some software quality attributes, which are used to build the best software products (Zuse,1998).
We have proposed and validated three types of metrics to measure the complexity of database manipulation sub-languages (singularised to the SELECT sentence). These metrics are not sufficient to evaluate the quality of products developed with Fourth-Generation Languages.
Presently, we are elaborating other metrics for the different sub-languages that we have identified in fourth generation environments (Martinez and Piattini, 1998).

REFERENCES.

Basili, V. (1998). Using Experiments to Build a Body of Knowledge". *Conferencia celebrada en la Facultad de Informática de la Universidad Politécnica de Madrid.* Madrid.

Briand, L., Bunse, D., Daly, J., Differding, C. An experimental comparison of the Manintainability of Object-Oriented and Structured Design Documents, Proc. Int. *Conf. On Software Maintenance, Harold, M.J. Visaggid, G. (eds), Bari, 1-3 Oct*, 130-138. 1997.

Briand, L.C., Morasca, S. y Basili, V. (1996). Property-based software engineering measurement. *IEEE Transactions on Software Engineering*, 22(1): 68-85.

Card, D.N. y Glass, R.L. (1990). *Measuring Software Design Quality*. Englewood Cliffs. USA.

Dolado, J.J. (1997). A Study of the Relationships among Albrecht and Mark II Function Points, Lines of Code 4GL and Effort. *J. systems software*, 37:161-173.

Fenton, N. (1994). Software Measurement: A Necessary Scientific Basis. *IEEE Transactions on Software Engineering*, 20 (3): 199-206.

Fenton, N. y Pfleeger, S. L. (1997). *Software Metrics: A Rigorous Approach* 2nd. edition. London, Chapman & Hall.

Henderson-Sellers, B. (1996). *Object-Oriented Metrics - Measures of complexity*. Prentice-Hall, Upper Saddle River, New Jersey.

Holloway, S. (ed.) (1990). *Fourth-Generation Systems, their scope application and methods of evaluation*. London: Chapman and Hall.

Kriz, Jürgen. "Facts and artefacts in social science: an ephistemological and methodological analysis of empirical social science". Research techniques, McGrawHill Research. 1988.

Li, H.F. y Cheng, W.K. (1987). An empirical study of software metrics. *IEEE Trans. on Software Engineering*, 13 (6): 679-708.

Martínez, A. & Piattini, M. (1998). Validation of 4GL metrics. *Proc. of the Software Measurement in Practice, 10th Anniversary Conference*. United Kingdom Software Metrics Association, Londres, octubre 1998, X 1-19.

Pigoski, T.M. (1997). *Practical Software Maintenance*. Wiley Computer Publishing. New York, USA.

Pfleeger, S. L. (1997). Assessing Software Measurement. *IEEE Software*. March/April, 25-26.

Pfleeger, S. L. (1995) Experimental Design and Analysis in Software Engeneering. *Anuals of Software Engineering 1*, p. 219-253. JC. Baltzer AG, Science Publishers.

Rohatgi, V.K., (1976). "An introduction to Probability Theory and Mathematical Statistics", *Wiley Series in Probability and Mathematical Statistics.*

Verner J. And Tate G.(1988) "Estimating Size and Effort in Fourth-Generation Development". *IEEE Trans. On Software Engineering*. July 1988.

Zuse, H. (1998). " A framework of software measurement" , *Ed. Walter De Gruyter.* 1998

APPENDIX A

TABLE DESCRIPTIONS.

DESCRIPTION TABLE fper020

```
Name:                   fper020
Owner:                  ingres
Created:                24/03/1999 14:08:38
Type:                   user table
Version:                OPING1.2

Column Information:
```

Column Name	Type	Length	Nulls	Defaults	Key Seq
p0_nif	varchar	9	no	no	1
p0_nom_haci	varchar	40	no	no	
p0_nom_nomi	varchar	30	no	no	
p0_direc	varchar	40	no	no	
p0_cpo	varchar	5	no	no	
p0_dom_pro	varchar	2	no	no	
p0_dom_pue	varchar	3	no	no	
p0_telefono	varchar	9	yes	null	
p0_sexo	varchar	1	no	no	
p0_estado	varchar	1	no	no	
p0_hijos	varchar	2	no	no	
p0_f_nac	date		no	no	
p0_nac_pro	varchar	2	no	no	
p0_nac_pue	varchar	3	no	no	
p0_nac_pais	varchar	3	no	no	
p0_nss	varchar	14	no	no	
p0_nmunpal	varchar	6	yes	null	

The fper020 table has 4 rows.

DESCRIPTION TABLE hor personal

```
Name:                   hor_personal
Owner:                  ingres
Created:                20/11/1998 18:05:17
Type:                   user table
Version:                OPING1.2
Column Information:
```

Column Name	Type	Length	Nulls	Defaults	Key Seq
cod_serv	varchar	3	no	no	
c_subs	varchar	2	no	no	
nif	varchar	9	no	no	
num_ficha	varchar	4	no	no	
saldo1	float	8	no	no	
saldo2	float	8	no	no	
saldo3	float	8	no	no	
saldo4	float	8	no	no	
saldot	float	8	no	no	
saldof	float	8	no	no	
periodo	integer	1	no	no	
anio	integer	1	no	no	
tipo_horario	integer	2	yes	null	
subtipo	integer	2	yes	null	
clave	varchar	15	yes	null	
f_alta	varchar	6	yes	null	
situacion	varchar	1	yes	null	
fsaldo	date		yes	null	

The hor_personal table has 3 rows.

```
DESCRIPTION TABLE prueba

Name:                   prueba
Owner:                  ingres
Created:                04/03/1999 13:46:31
Type:                   user table
Version:                OPING1.2

Column Information:

                                                                  Key
Column Name                     Type       Length Nulls Defaults Seq
nif                             varchar         9   no      no
num_ficha                       varchar         4   no      no
fecha                           varchar         6   no      no
hora                            varchar         4   no      no
id_incidencia                   varchar         2  yes    null
control                         varchar         2  yes    null
estado                          varchar         1  yes    null
cod_centro                      varchar         2  yes    null
tipo                            varchar         2   no      no
```

The prueba table has tiene 72 rows.

QUERIES (DATABASE MANIPULATION SUB-LENGUAGE)

1.- select hora from prueba where num_ficha='0959' and fecha='181298'

2.- select num_ficha, fecha, count(hora) as n_fichajes from prueba where num_ficha='0800' group by num_ficha, fecha

3.- select num_ficha, fecha from prueba where num_ficha not in (select num_ficha from prueba where fecha='171298' and control='SM' and estado='A' and hora in (select hora from prueba where fecha='171298' and control='SM' and tipo='A0')) and Fecha>'131298'

4.- select num_ficha from prueba where num_ficha not in (select Num_ficha from prueba where fecha='171298' and control='SM' and estado='A' and hora in (select hora from prueba where fecha='171298' and control='SM' and tipo='A0')) and fecha>'131298' group by num_ficha

5.- select f.p0_nom_nomi, h.hora, p.clave from fper020 f, prueba h, hor_personal p where f.p0_nif=h.nif and h.num_ficha=p.num_ficha and h.fecha='171298'

6.- select f.p0_nom_nomi, p.clave from fper020 f, prueba h, hor_personal p where f.p0_nif=h.nif and h.num_ficha=p.num_ficha and h.fecha='171298' group by p0_nom_nomi, clave

7.- select f.p0_nom_nomi, p.num_ficha, p.fecha, h.f_alta from prueba, fper020 f, hor_personal h where p.nif not in (select h.nif from prueba p, fper020 f, hor_personal h where p.num_ficha=h.num_ficha and f.p0_nif=h.nif and p.nif=f.p0_nif and p.fecha='171298' and p.control='SM' and p.estado='A' and f.p0_sexo='V' and p.hora in (select hora from prueba p,fper020 f, hor_personal h where p.num_ficha=h.num_ficha and f.p0_nif=h.nif and p.nif=f.p0_nif

```
and p.fecha='171298' and p.control='SM' and p.tipo='A0' and
h.saldot=0 ) ) and p.fecha='151298' and p.num_ficha=h.num_ficha
and p.nif=f.p0_nif and f.p0_nif=h.nif
```

8.-
```
select f.p0_nom_nomi, p.num_ficha, p.fecha from prueba p, fper020
f, hor_personal h where p.nif not in (select h.nif from prueba p,
fper020 f, hor_personal h where p.num_ficha=h.num_ficha and
f.p0_nif=h.nif and p.nif=f.p0_nif and p.fecha='171298' and
p.control='SM' and p.estado='A' and f.p0_sexo='V' and p.hora in
(select hora from prueba p, fper020 f, hor_personal h where
p.num_ficha=h.num_ficha and f.p0_nif=h.nif and p.nif=f.p0_nif
and p.fecha='171298' and p.control='SM' and p.tipo='A0' and
h.saldot=0 ) )and p.fecha='151298' and p.num_ficha=h.num_ficha
and p.nif=f.p0_nif and f.p0_nif=h.nif group by f.p0_nom_nomi,
p.num_ficha, p.fecha
```