8th Workshop on Quantitative Approaches in Object-Oriented Software Engineering (QAOOSE 2004)*

Coral Calero[1], Fernando Brito e Abreu[2], Geert Poels[3], and Houari A. Sahraoui[4]

[1] ALARCOS Research Group, Departamento de Informática, Universidad de Castilla-La Mancha, Ciudad Real, Spain
Coral.Calero@.uclm.es

[2] QUASAR Research Group, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Monte da Caparica, Portugal
fba@di.fct.unl.pt

[3] Faculty of Economics and Business Administration - Ghent University and Centre for Industrial Management - Katholieke Universiteit Leuven
geert.poels@rug.ac.be

[4] Département d'Informatique et Recherche Opérationnelle, Université de Montréal, Montréal, Quebec, Canada
sahraouh@IRO.UMontreal.CA

1 Summary

The workshop was a direct continuation of seven successful workshops, held at previous editions of ECOOP in Darmstadt (2003), Malaga (2002), Budapest (2001), Cannes (2000), Lisbon (1999), Brussels (1998) and Aarhus (1995). This time, as in previous editions, the workshop attracted participants from both academia and industry that are involved / interested in the application of quantitative methods in object oriented software engineering research and practice.

As a result of the previous edition and in order to open the workshop participation to a broader audience, the 2004 edition extended the scope of the workshop to quantitative approaches to other than object-oriented modelling, specification and programming methodologies and technologies. In particular component-based systems (CBS), web-based systems (WBS) and agent-based systems (ABS) will also fit into this new edition.

Like in previous years, submissions were invited, but not limited, to the areas of metrics collection, quality assessment, metrics validation, and process management.

This year we received 15 position papers. 9 authors were invited to present theirs positions as discussion topics and 3 others to present posters. Informal proceedings of QAOOSE'2004 were distributed to the participants.

* The title of this report should be referenced as "Report from the ECOOP 2004 8th Workshop on Quantitative Approaches in Object-Oriented Software Engineering (QAOOSE 2004)".

J. Malenfant and B.M. Østvold (Eds.): ECOOP 2004, LNCS 3344, pp. 23–35, 2004.

The workshop was divided in four sessions, which follow:

- Session A. Metric definition and validation
- Session B. Methodology and application of measurement
- Session C. Functional size and quality
- Session D. Short papers

At the end a "Discussion and closing session" took place.

The workshop had 17 participants (included de four organizers). Among them, we had people from several European countries and Canada and also from industry and from University.

As each year the workshop was very active and new topics and future directions for the workshop were addressed at the end. Among them, were selected for be included as a topics on next editions of the workshop the following: metrics visualization, components and services, aspect oriented software development, paradigm independent product metrics, metrics and reengineering, process aspects vs. product aspects, early phase metrics, relationships between cost and quality aspects, influence of context on quality.

2 Organizers

This year, the workshop was organized by:

Coral Calero
ALARCOS Research Group
Escuela Superior de Informática,
Universidad de Castilla-La Mancha. Ciudad Real, Spain

Fernando Brito e Abreu
QUASAR Research Group
Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa
Monte da Caparica, Portugal

Geert Poels
Faculty of Economics and Business Administration - Ghent University, and
Centre for Industrial Management – Katholieke Universiteit Leuven
Gent, Belgium

Houari A. Sahraoui
Département d'Informatique et Recherche Opérationnelle
Université de Montréal
Montréal, Canada

3 Workshop Attendants

The information about the workshop attendants is shown in this section. Name and affiliation is shown for each attendant:

- **Miguel Goulão.** Faculty of Sciences and Technology, Lisbon New University, Portugal

- **Manuel F. Bertoa** Dpto. Lenguajes y Ciencias de la Computación. Universidad de Málaga
- **Valerie Paulus.** CETIC. Charleroi, Belgium
- **Olivier Beaurepaire.** SNCF - Délégation aux Systèmes d'Information Voyageurs. Nantes, France
- **Benjamin Lecardeux.** SNCF - Délégation aux Systèmes d'Information Voyageurs. Nantes, France
- **Christine Havart.** SNCF - Délégation aux Systèmes d'Information Voyageurs. Nantes, France
- **Parastoo Mohagheghi.** Department of Computer and Information Science, Trondheim, Norway
- **Silvia Abrahão.** Department of Computer Science and Computation. Valencia University of Technology. Valencia, Spain
- **Houari Sahraoui.** Département d'Informatique et Recherche Opérationnelle. Université de Montréal. Montréal, Canada
- **Mario Piattini.** ALARCOS Research Group. Escuela Superior de Informática, Universidad de Castilla-La Mancha. Ciudad Real, Spain
- **Denis Kozlov.** Department of Computer Science and Information Systems, University of Jyväskylä. Jyväskylä, Finland
- **Jean-François Gelinas.** Département de Mathématiques et d'Informatique. Université du Québec à Trois- Rivières Canada
- **Sherif Gurguis.** Department of Computer Science. The American University in Cairo
- **Briand Henderson-Sellers** University of Technology, Sydney
- **Coral Calero** ALARCOS Research Group. Escuela Superior de Informática, Universidad de Castilla-La Mancha. Ciudad Real, Spain
- **Fernando Brito e Abreu** Faculty of Sciences and Technology, Lisbon New University, Portugal
- **Geert Poels** Faculty of Economics and Business Administration - Ghent University, and Centre for Industrial Management – Katholieke Universiteit Leuven. Gent, Belgium

4 The Call for Papers

The call for papers of the workshop was distributed by e-mail basically through distribution lists (among them the one created during the 7 years of previous workshops with all the workshop attendants) but also with direct e-mails to the workshop organizers contacts. The Call for Papers was structured as follows:

4.1 Introduction

A brief introduction of the workshop including an explanation about the number of editions, the main goal of the workshop, etc.

4.2 List of Topics

Topics for the workshop submissions. This year, submissions were invited, but not limited, to the areas of metrics collection, quality assessment, metrics validation, and process management.

Area C (Metrics Collection)

- Automatic support for sharing research hypotheses, data and results
- Standards for the collection, comparison and validation of metrics
- Embedding metrics in CASE and application development tools
- Evaluation of metrics collection tools
- Automating collection from formal metrics definition
- Metrics collection in the development process (measurement planning)
- Public repositories for measurement data
- Metrics visualization (*)
- Metrics for component-based systems (*)
- Metrics for web-based systems (*)

Area A (Quality Assessment)

- Measuring non-functional requirements of OO systems
- Quantitative OO and CB design heuristics
- Metric-based design refactoring
- OOD and CBD quality characteristics assessment
- Quantitative impact analysis in OO and CB architectures
- Quantitative assessment of OO analysis/design patterns and frameworks
- Quantitative assessment of behavioral modeling in OO models
- Quantitative assessment of OR and OO database/datawarehouse schemata
- Measurement and quality assessment of components (*)
- Measurement and quality assessment of agent-based systems(*)
- Agent-based Web service architecture as a means of providing QoS(*)
- Quality of Service models(*)
- Instrumentation of Web services for QoS (*)

Area V (Metrics Validation)

- Meta-level metrics
- Formal and empirical validation of metrics
- Metrics and Measurement Theory
- Validation techniques and their limits
- Standard data sets for metrics validation
- Limitations of quality estimation techniques

Area P (Process Management)

- Reliability and rework effort estimates based on design measures
- Reuse evaluation
- Resource estimation models for OO and CB software development

– Quantitative tracking of OO, web services, and CBS development activities
– Empirical studies on the use of measures for process management
– Measurement support in a CBD life cycle

We explicitly solicited position papers related to topics marked with an asterisk (*) as well as papers that document and/or motivate the use of quantitative methods in industrial software processes. These topics were identified as important open research issues in QAOOSE'2003 workshop.

4.3 Important Dates Information

Information about the important dates of the workshop were included on the call for papers

4.4 Web Page

The web address of the workshop was included on the call for papers. On this page the people could find all the information related to the workshop. The address of the workshop is: `http://alarcos.inf-cr.uclm.es/qaoose2004`.

5 The Workshop Sessions

The day started with the session on *Metric definition and validation*. In this session we had three presentations (one of the presenters failed and another from the short papers session asked for a change).

Independent Validation of a Component Metrics Suite. Miguel Goulão, Fernando Brito e Abreu.
The session started with the paper of Miguel Goulão titled "Independent Validation of a Component Metrics Suite". The paper describes an independent validation study for a suite of reusability metrics for component based design. The authors present a formalization that combines the UML 2.0 metamodel with OCL. By doing that they provide a formal, portable and executable definition of the metrics set that can be used by other researchers and practitioners to perform independent validations of the metrics suite. Also they present a prototype working environment to perform such independent validation experiments. A workshop attendant asked about the formalization of implementation metrics and Miguel asked that this approach cannot be used for this kind of metrics because only works with structural metrics. A question about the definition of one of the metrics (the RCO) used for the technique utilization example was done. Problems related to the difficulty of filling values in the metamodel and the validation of the metrics was discussed and the ideas for solving them were exposed by the presenter. A participant asked if the technique was limitative because the authors use reverse engineering and Miguel asked that no because it can be used in forward engineering as well, in fact with less effort since component diagrams will be available for obtaining the required information. The only thing to be in mind is that, sometimes, another meta-model would be needed,

for instance, when applying the work to relational database metrics (planned as future work). Also a question about the accuracy of the experimental results presented was done and the speaker explained that their objective in this paper was not to assess the accuracy of the results presented by Washizaki et al., but rather showing that a better formalization of their metrics allows the replication of their experiment in a far more reusable (because OCL interpreters are becoming widespread), efficient (because OCL clauses are executable and as such can help automate the collection process) and objective way (as long as the corresponding metamodel is made available). The next question was if the results mean that the analyzed components are of bad quality and Miguel said that they didn't get experts opinion on the components quality to perform such kind of conclusion. However their aim here was not to discuss the appropriateness of the thresholds proposed by Washizaki el al., but rather showing how independent cross can be easily performed without the typical problems of metrics definition interpretation, metrics collection (lack of tools), among others.

Usability Metrics for Software Components. Manuel F. Bertoa and Antonio Vallecillo

The session followed with the paper presented by Manuel F. Bertoa about usability metrics for software components. The work presented on this paper is justified on the need to select a component among a set of possible candidates that offer similar functionality and the fact that this selection can be done using metrics. In the paper they define, in a consistent way, usability metrics for software components based on the ISO 9126 Quality Model. They also define the basic concepts on software measurement used in the paper, what they understand for usability in a CBSD framework, and the component information available to be measured. One of the attendants asked about how the authors had considered the fact that, when a component has been used more than one time, it becomes easier and so, perhaps a more complex component become easier than other that using the proposed approach is considered as the easier one. Manuel answered telling that this was a good point to consider for future work because now they are centring their work only considering that the components to select are not known by the user. The difference between understandability and learnability was explained by the author after a question of one of the workshop attendees. Also someone asked about the metamodel the authors have used and Manuel explained that as a result of the research, the metamodel is always under revision. The importance of the 'point of view' concept was pointed to the speaker who answered that they mainly take the user point of view, and make it explicit. About the relationship between the metrics and the quality sub-characteristics Manuel said that they investigate this in future work when they have external metrics for the quality sub-characteristics. This investigation is done by means of empirical studies. The fact that the usability metrics cannot have an absolute value was remarked by an attendant and Manuel said that their point of view is to support the selection of components, meaning the choice between alternative components and so, absolute values are not needed.

Finally a question about how to evaluate non-functional aspects like QoS was raised and the speaker explained that this was not included yet, but they plan to extend the proposal in this respect.

Classification of Metrics Related to the Software Development Process as a Prerequisite for Its Improvement. Denis Kozlov
The session finished with the short paper of Denis Kozlov about Classification of metrics related to the software development process as a prerequisite for its improvement in which the author points out that estimation of object-oriented software quality remains as very urgent because there is a lack of generally accepted classifications of metrics related to final software and to the software development process, despite of numerous articles devoted to it. Then, the focus of the article is the classification of metrics of the software development process. The correlations between quality characteristics and metrics of the software development process have been revealed. An approach for using the presented classifications for improvement of the software development process was also proposed. At the end of the presentation some attendees remarked to Denis that the fact of mixing the ISO9126 (which is a product norm) with the process could be a bit dangerous and he would take this into account. Briand Henderson-Sellers remarked to Denis that his work links somehow to SPI&A framework and he suggested him look into the SPICE standard where process enactment is assessed, based upon the metamodel. Another attendant recommended the author to think about the existing relationship between his work and CMM and SPICE. A question about how where "correlations" presented in the paper determined was done and Denis answered that the "correlations" were just a personal view of the strength of. He plans to detail the criteria in a follow-up version of this paper.

After a break, the second session, on *Methodology and application of measurement*, was held. There were three presentations:

On the Application of Some Metrology Concepts to Internal Software Measurement. Miguel Lopez, Simon Alexandre, Valerie Paulus, Gregory Seront
This paper was presented by Valerie Paulus. The authors investigate the applicability of classical metrology concepts to software measurement. In particular they explore the concepts of systematic and random error, repeatability and reproducibility, uncertainty, calibration, and etalon when measuring internal attributes of software with metric tools. Using a laboratory experiment the uncertainty of McCabe's measure of the cyclomatic complexity of a Java class was examined. Although several factors can impact the measurement method, the experiment did not show the occurrence of systematic or random errors. Hence the authors conclude that uncertainty might not be a relevant concept for internal software measurement. On the other hand, it was shown that calibration is relevant when different measurement instruments for the same measure are available. This was shown using three tools that measure the cyclomatic complexity

of a Java class. Measurement results obtained with these tools can be different as they implement some counting rules differently. However, with calibration, the correctness of the measurement results can be ensured. One workshop participant questioned the relevancy of the research question as measurement results are deterministic when a same tool is used (unless there is some software bug). Valerie replied that this was not sure, and therefore they wished to examine this. In future work the authors intend to introduce themselves some errors to further investigate the usefulness of the metrology concepts. It was further remarked that concepts like calibration are surely relevant when there is ambiguity in the measurement methods, such as with functional size measurement. Valerie responded that new functional size measures such as COSMIC-FFP even incorporate calibration and other metrology concepts (e.g. conversion) as part of their method. Another question related to the use of McCabe's cyclomatic complexity as the object of study in the experiments, given that it is not an object-oriented measure. Valerie admitted, but McCabe's number was only an example and other measures will be investigated. At the end of the presentation, Valerie argued for the introduction of a common terminology for reference fields such as measurement theory and metrology. Currently there is a lot of confusion in software measurement with respect to validation, and the separation of a measure from its methods and instruments could resolve some of these semantic problems. As a final comment it was suggested that the application of metrology is probably very promising for dynamic metrics, where there is uncertainty by definition.

Industrialisation of Software-Quality-Led Project Management Process at the S.N.C.F. (French Railways). Olivier Bearupaire, Benjamin Lecardeux, Christine Havart.

This 'industrial' presentation was made by Olivier Beaurepaire. The industry report of these authors was especially welcomed by the workshop organizers, as in the past there was a lack of participants from industry in the QAOOSE workshops. The contribution by Beaurepaire and colleagues demonstrated that the topic of the QAOOSE workshops is not only of academic interest. In the presentation, Olivier demonstrated the metrics program at the French Railways company. The program was established to assist project managers in assuring the quality of the produced software systems. The core of the program is a decision support tool (called the Software Quality Portal) that helps both managers and developers to interpret the measurement results (e.g. occurrence of anti-patterns, quality over time trend analysis). As a result of the introduction of the metrics program, an improvement in the level of quality has been observed. Further, the metrics initiative establishes a contractual framework for sub-contractors, allowing the French Railways to assure the quality of externally developed software. One of the metrics used is the severity of the software problems that are observed. Workshop participants wondered how this severity could be quantified. Olivier responded that it was based on the end-user's point of view. A participant further remarked that the severity of software problems is not necessarily directly related to their costs. Another question related to the improvements

that were observed after the introduction of the metrics program. It was replied that there were improvements, although it was too soon to quantify them. Some results are promised for the near future, in a next paper. One of the workshop organizers further remarked that is could be worthwhile to move to the model level, and not focus metrics efforts exclusively on the code level. This would, amongst other benefits, make the metrics program more independent of the languages used (as for each language other metrics thresholds may apply). Such an approach, though valuable, might be difficult to implement at French Railways, because models are rarely used, or not kept up to date. Unfortunately, this is the rule rather than the exception in software engineering practice.

Exploring Industrial Data Repositories: Where Software Development Approaches Meet. Parastoo Mohagheghi, Reidar Conradi

The presentation was done by Parastoo Mohagheghi. The paper deals with methods and problems of exploring large industrial data repositories in empirical software engineering research, taking into account that software data are not always obtained as part of a measurement program. The presentation gave an overview of recent research results, presented by the authors at other conferences. These are a study of defect reports, a study of change requests, and a study of effort spent in some releases of a large-scale telecommunications system. It was concluded that the integration of the results of such studies with other studies and with theory is a challenge, especially since measurement programs and metrics are tied to particular development approaches. As part of a solution for this problem, the authors present a set of metrics for a combined incremental, reuse-, and component-based development appoach. One workshop participant asked about the granularity of the collected effort data. Parastoo answered that effort data is needed per component and not only per use case, and that this is currently a problem with the repository that was analyzed. It was commented by another participant that components might have considerably different sizes, so the use of 'component' as a normalizing factor for effort is questionable. According to Parastoo this is true, however, in a well-defined context such as the one presented, components have similar sizes. Another participant noticed that the distinction between 'reused' and 'not reused' components is too coarse-grained. There might be a wide range of values between these two extremes with respect to the extent of software reuse. There was also the question of the validity of the data in the repository. Parastoo admitted that all data was used, even without being sure that a validation procedure was used before storing the data. There was however a configuration management system in use, which provides some assurance against storing invalid data. Another question related to the analysis techniques that were used. These were hypothesis testing and correlational analysis, but not multivariate analysis yet.

After lunch we had a session on *Functional size and quality* where three papers were presented.

Validation Issues in Functional Size Measurement of Object-Oriented Conceptual Schemas: The Case of OOmFP. Silvia Abrahäo, Geert Poels, and Oscar Pastor.
The paper, presented by Silvia Abrahao, introduced a framework for evaluating Functional Size Measurement (FSM) methods, based on a process model for functional size measurement. The authors also show how to apply this framework to evaluate OO-Method Function Points (OOmFP) focusing on the role of theoretical validation within the evaluation framework.

Silvia was asked about the metamodel that during the talk she mentioned that they used for defining the concepts in FP but it cannot be found in the paper. She also was asked about the relationship between Object Points and their approach and she said that they haven't yet come across it but they will have a look. Also an attendant asked about what meta-model was used for the FSM method? and Silvia asked that they used UML to formalize the IFPUG-FPA meta-model that was first mapped into the OO-Method meta-model.

A Proposal of a Multidimensional Model for Web-Based Applications Quality. Ghazwa Malak, Linda Badri, Mourad Badri and H. Sahraoui.
In this paper, authors propose a three-dimensional model for web-based applications quality. Houari, who presented the paper, was asked about the correctness of considering the application domain as a dimension for the web quality model taking into account that it is a nominal scale and he answered that two dimensions is not enough. Also related to the domain dimension he was asked if there should not be a different quality model for each domain instead of having the domain dimension by itself and he said that they looked for orthogonal dimensions and they thought that a dimension, which refers to the profile of the web application, which has its impact on quality characteristics, was necessary. He was also asked about how the information content or volume assessment was considered in their model and he said that they believe assessment is only a problem if the site is not well structured.

Measuring the Effects of Patterns on Object Oriented Micro Architectural Design. Javier Garzás and Mario Piattini.
In OO Micro-Architectural Design Knowledge, design patterns are a key and important technique. "Using design patterns increments design quality" is a famous sentence, but this is an ambiguous and imprecise sentence: what does "design quality" mean? Patterns affects on difference way to the quality of the micro architectural design. In this paper authors propose metrics for answering this important question. An attendant asked Mario, the presenter of the work, that the optimal number of patterns seems to be a magic number and Mario answered that they were conscious of that, especially since there are so many kinds of OO patterns. He also he was asked about the definition of a metric included on the paper.

During the *short paper* session, we started by a demo on software metrics visualization using perception and simulation presented by Houari Sahraoui. After the demo, two participants presented briefly their positions.

Aspect Cohesion Measurement Based on Dependence Analysis. Jean-François Gélinas, Linda Badri and Mourad Badri
Jean-François argued that Aspect-Oriented Programming is a promising new paradigm. Although several metrics have been proposed in order to assess object-oriented software quality attributes, new metrics must be developed to hold account Aspect's characteristics. As cohesion is considered as one of the most important software quality attributes, he proposes a new approach for aspect cohesion assessment based on dependence analysis. To illustrate his proposal, he introduced several cohesion criteria and built a new cohesion metric using them. After this short presentation a discussion took a place. The first topic that was addressed concerns the alternatives of adapting existing OO cohesion metrics or developing aspect cohesion metrics from scratch. The position of Jean-François is that AOP introduced many new concepts which made any adaptation hard and risky. The second topic addressed the problem of choosing the cohesion rather than other attributes like coupling. Jean-François explained that cohesion is a fundamental principle of aspects. An aspect is supposed to implement a cohesive behavior. The final discussion topic was about the coverage of the proposed metric. Jean-François recognized that more metrics are probably needed to cover all the different factors that can influence the cohesion.

Towards A Minimal Performance Metrics Suite for Agent-Based Systems. Amir Zeid and Maha Abdel Kader.
Sherif Gurguis was the last participant who presented the position of his team. Like the previous position, he claimed that agent-oriented paradigm is gaining popularity. With this respect, each scientific development that claims to provide a "new way" for approaching existing problems needs proper (i.e. formal and quantifiable) evaluation methods and consensus-based criteria for measuring the validity of its claims. As agents present some unique features that should be verified and evaluated, it is important to define useful metrics to measure them. The particularity of this work is that the authors start by defining a large set of metrics and then reduce this set by evaluating the dependency between the metrics using correlation techniques.

The workshop finished with a *Discussion and closing session*. During this session, the first part was devoted to the identification of important issues that emerged from the different sessions of the day.

- Participants from industry and academia agreed the fact that it is difficult to define objective measures for evaluating the ROI of adopting product measurement and quality programs.

- ISO9126 was used during the past years as a starting point for a large part of software measurement research work. Today it is obvious that this model is not application to software applications built using emerging technologies such as component and web-based applications. The adaptation of this model and/or the definition of technology-specific models are crucial issues.
- The adaptation of structural programming-based metrics to OO paradigm have reveled many problems. The same problem is occurring today with the adaptation of OO metrics to emerging technologies (aspects, components, agents, etc...). What are the lessons and how can we make this transition more successful that the previous one?
- Although many contributions have been made in the study of the relationships between quality attributes and metrics, the area still suffers from rigorous and exhaustive results.
- The success of evaluating the majority of software quality factors is deeply related to the ability of understanding the phenomena behind them such as software evolution. Even if there is a consensus on this statement, very few work addressed the theory and model that can represent these phenomena.

During the second part of this session, additional issues were identified as future challenges for the community and consequently as important topics for the next year workshop edition. These topics are:

- Quantitative visualization of large sets of software artifacts
- Measurement and quality evaluation of component and service-based application
- Measurement and quality evaluation for aspect-oriented software development
- Rigorous empirical studies for software quality evaluation
- Paradigm independent product metrics
- Process vs product attributes
- Metric-based reengineering
- Early development phase metrics
- Relationship between cost/effort and quality factors
- Influence of context (product and/or organization) on quality

6 Other Information Related to the Workshop

The links to the web pages of the previous editions of the workshop are:

- QAOOSE'2003:
  `http://ctp.di.fct.unl.pt/QUASAR/QAOOSE2003`
- QAOOSE'2002:
  `http://alarcos.inf-cr.uclm.es/qaoose2002`
- QAOOSE'2001:
  `http://www.iro.umontreal.ca/~sahraouh/qaoose01`

- QAOOSE'2000:
  http://ecoop2000.unice.fr/Program/Technical/Workshops/w10.html
- QAOOSE'99:
  http://ecoop99.di.fc.ul.pt/techprogramme/w20.html
- OO Product Metrics for Software Quality Assessment:
  http://www.crim.ca/~hsahraou/oopm.html
- Quantitative Methods for OO Systems Development:
  http://ctp.di.fct.unl.pt/QUASAR/ECOOP95