The Evaluation Method for Software Product

Regina Colombo
 e-mail: regina.thienne@cenpra.gov.br
UNICAMP - Universidade Estadual de Campinas
Faculdade de Engenharia Mecânica
Caixa Postal 6122 CEP: 13.083-970
Campinas, São Paulo Brazil

Ana Guerra
 e-mail: ana.guerra@cenpra.gov.br
CenPRA – Centro de Pesquisas Renato Archer
Caixa Postal 6162 CEP 13084-971
Campinas, São Paulo Brazil

Abstract

The quality of a software product is the result of the activities carried out along its development process. To evaluate the quality of a software product is to determine, by applying static and dynamic analyses, the degree of fulfilment of its explicit and implicit requirements. Such requirements, in general, express user needs in quantitative or qualitative terms and define the quality trends of a software product [8]. The quality of software is a real concern and efforts that have been carried out to increase the quality of the software product developed. Standards and methods are applied to check for their quality. However, software quality is still a great concern. A gap is detected in the efforts that have been carried out so as to improve the software quality. The software product concentrates its efforts in development and maintenance, whereas the software product quality is focused on this paper through quality evaluation of the final software product developed. The Evaluation method helps the process to improve the Software Product Quality. The Process of Evaluation of software products is based on ISO/IEC 14598-5 [4], the scope of which is to provide requirements and recommendations for the practical implementation of evaluation of software products, developed or under development, as a series of activities defined under common agreement between the customer and the evaluator [7]. The main objective of this method is to provide to evaluators mechanisms to support the evaluation of software products from the point of view of the end user, according to ISO/IEC 9126 (NBR13596) [5] and ISO/IEC 12119 [3] in what concerns quality characteristics and software packages, respectively. The evaluation process is carried out by simulating a normal operational use of the product, beginning by analysing the documentation, installing the product as instructed in the documentation and proceeding by using the product in the most complete way that can be done. All along the process, evaluators assign rates to the product according to the questions of the Checklist. In addition to assigning rates, evaluator’s record the time spent in the evaluation, specify the major functions of the product, and write down comments on specific issues they consider relevant concerning the product. The final step in the evaluation process is the preparation of the Evaluation Report, which should address the major positive aspects of the evaluated product as well as suggestions for its improvement. The Process of Evaluation of Software Products-MEDE-PROS is a fairly mature method applied in the “Laboratórios de Avaliação de Produtos de Software” - Laboratories for Evaluation of Software Products, which are licensed to use it to evaluate software products by CenPRA – Renato Archer Research Centre.

Keywords:
1. Introduction
The software product quality needs to be checked in an intensive and formal way. This procedure minimises the lack of quality trends on the software product and, a deviation that results from this objective can be detected and improved from viewpoint of the end user, thus providing the necessary changes. The growing demand of software in our lives requires the development of effective methods and tools to support software evaluation. This work is aimed at developing methods and tools for evaluation and improvement of software products according to internationally accepted criteria and standards. It must be remarked that not only must software product be evaluated but also the process of software production should be evaluated to increase quality[1]. This paper is organised as follows: Section 2 describes the concepts on Software Products Quality and the Standards ISO/IEC 14598-5 [4], ISO/IEC 9126 [5], ISO/IEC 12119 [3]. Section 3 provides the method. Section 4 presents the MEDE-PROS method, Section 5 Evolution and Results and finally the conclusion.

2. Software Quality Concepts
The issuing of ISO/IEC 9126 [5], whose Brazilian equivalent has been given number NBR 13596, had the advantage of establishing a quality basic model of software product, changed into a reference recognised by most of community. However the quality characteristics as defined this Standard are not known as directly measurable [6]. The revision of Standard ISO/IEC 9126[5] and the series 14598 explain the quality “in use” concepts and “external” and “internal” measurements. The “in use quality” is measured in terms of the software qualities. In use quality is the combined effect of quality characteristics for the end user. It can be measured by the extent which specific users can attain their objectives effectively, and efficiency of task and satisfaction. External evaluations aim at given attributes of the measured product with product under operation. Internal evaluations also aim at specific attributes but require product requirement examination such as project documents and source code. The ISO 8402[2] Standard “Quality: The total characteristics of an entity that provides, with the capacity to satisfy its explicit and implied needs”. The ISO/IEC 9126[5] Standard defines a software product quality characteristics which makes up a product quality model which can be adopted as a basic reference in evaluation. Besides having the power of an international standard, it provides a reasonable coverage of major aspects for any software product. Bearing these in mind, the problem to evaluate the quality of a software is presented. This problem raises two others:
- How to ascertain explicit and implicit needs to the filled?
- How to check whether they are filled or not?
In the first one, it must be distinguished what are need or users desire. The stands for a requirement that must be filled so some activity can be undertaken. It is shown in the user’s desire, as changed into cultural and individual aspects. When a product is evaluated, the real needs must be accounted for and not only the desires as they may include changes. The user can consider as wishes only parts of the needs, by imparting great importance to secondary or accessory aspects and letting other important ones aside. The needs under consideration must define the diversity of possible users. This activity is covered by Product Quality Requirement Analyses. The second matter must be answered through methods and techniques for evaluation, so they can measure in what degree the requirements will be filled. Accordingly, an Evaluation of Software Product Cycle is displayed in Figure 1.

![Figure 1 - Evaluation of Software Product Cycle](image)

This article concerns three international standards in the area of Software Quality Evaluate, which are ISO 9126[5], ISO 12119[3] and ISO 14598-5[4]. According to ISO 9126 [5] the Quality Model in Figure 2 can be applicable in the phase of requirement definitions of the quality of a software to be developed as well as in the phase of product acceptance by evaluating intermediary and final product. The idea is that starting as of the quality characteristic it can be doubled up into Subcharacteristic and the latter into Quality Attributes in such a way as to come to quality measurement that can be measured and rated.
The assessing of a Software Product in this context bears three important concepts: measurement, rating and judgement as shown in Figure 3. **Measurement** is the action of applying measure software product to a specific software. **Rating** is the action of changing the secured measurement into a pre-established number system. **Judgement** is assuring your thoughts about software quality.

According to ISO 9126 [5] six quality characteristics are also established for a software product, they are: Functionality, Reliability, Usability, Efficiency, Maintainability and Portability. These six characteristics are also doubled in their respective subcharacteristics. According to ISO 12119[3], a Software Package is the “complete and document proof programs supplied to several users for application and generic function”. Examples of Software package are: Text process, electronic sheets, database, software charts, program for technical or scientific functions and utilitarian programs and the like. This specific kind of software products is also known internationally as COTS – Commercial off the Self. The ISO/IEC 12119 [3] Standard is applicable to appraising software packages as they are available and supplied for use in the market. This Standard establishes Figure 4 requirements, for software package, quality requirements, instructions on how to test a software package as regards established requirements. The quality requirements as established on this Standard can be adopted for software products which are not specifically COTS, this can be done by checking among the requirements established on the Standard for those that can be applied for any software products.

According to ISO14598-5[4] the fundamental characteristics expected in the Software Products Evaluation Process are: **Repeatability** – The repeated evaluation of a same product, with the same evaluation specification as done by the same evaluator must produce results that can be accepted is identical; **Reproducibility** – The evaluation of same product, with identical evaluation specification as executed by a different evaluator, must produce results that can be accepted as identical; **Impartiality** – The evaluation must not be influenced compared to any result in particular; **Objectivity** – The evaluation results must be factual, that is, not influenced by the evaluator’s feelings or opinions.

3. Proposing the Evaluation Method
Each of the Software Product components which, according to ISO 12119[3] are: Product Description, User’s Documentation, and Program and Data, herein mentioned as Software and Interface, is evaluated as per features defined by ISO 9126 Standard; those being used here are: Portability, Usability, Functionality, Efficiency and Reliability; and according to the Completeness requirement as defined by ISO/IEC 12119[3] Standard. The characteristic Maintainability does not apply in this context, as there is no access to source programs and intermediate products created during software product development. A more detailed insight on method structure is shown in Figure 5.

![Figure 5 – Scheduling Method for Evaluation Software Product](image)

Included in it are the quality subcharacteristics of ISO 9126 Standard, which is the unfolding of each characteristic plus the Completeness concept requirements as included in ISO/IEC 12119 Standard. Following measurements to be measured are specified, their respective rating and finally judging products quality. As per Quality Model as defined in ISO 9126, the next step is to unfold quality characteristics into attributes which can be measured and rated. The method herein adopted was the making of Checklist where each software component with is respective attributes were unfolded into questions and items which can be checked answered by the evaluator. Regarding questions included in the Checklist the evaluator must consider that questions are logical propositions about an attribute to be checked in an evaluation. Each proposition which poses an attribute should be the most objective possible involving a quality feature only. The main answers for the questions are: “Y” (Yes) for true propositions; “N” (No) for false propositions; “NA” (Not Applicable) for sentences which have reference to a feature which does not fit in the product being evaluated. The evaluator must show attention to the fact that the absence of whatever is being evaluated does not always mean the assignment will be “N” (No). The evaluator must find whether the proposition is applicable to the software product. “HE” (Hindered Evaluation) for sentences which the evaluator is in no condition to evaluate either by lack of means, lack of importance or even lack of specific knowledge on the subject being approached. For instance, for the Software component the Completeness quality requirement, the following attribute can be measured trough the following measurement.

The software functions:

1. Specified in documents, have all been implemented?
2. Implemented, will serve, in a complete way, the objections stated on the documents?
3. Will fill the need of the task the product proposes to achieve?

After selecting measures, the next step is to set up rating levels for the measures. In this methodology, the measures are questions that have type concept answers. A chart for each kind of answers is made within a numerical value. Each attribute must have as a numerical value. This value must be the Average and must be between zero (0) and one (1). This is important because all attributes must have similar values, even if they have a number of variable items and this usually happens as each attribute is subdivided into items so it can be checked in a more objective and clear way. Hence, in an attribute where these are 2,3 or 4 items, the numerical values of these items will take on proportional values and will be between zero and one depending on how many items the attributes are made of. It is necessary that the average of these items, that is, the numerical values have a value between zero and one. The items that will receive the “NA” or “AP” kind of answer, must be handled differently, because such answer cannot hinder product evaluation, is this case a value named “missing” is ascribed and in this way no numerical value is ascribed to them. Table 1 shows how to grant a numerical value, where optional answers are “Y” and “N” and that “Y” is a true proposition and “N” a FALSE one.

<table>
<thead>
<tr>
<th>Kind of question</th>
<th>Kind of answer</th>
<th>Numerical Value</th>
<th>Means kind of answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>HE</td>
<td>1</td>
<td>Hindered Evaluation</td>
</tr>
<tr>
<td>01</td>
<td>N</td>
<td>0,00</td>
<td>No</td>
</tr>
<tr>
<td>01</td>
<td>NA</td>
<td>0</td>
<td>Not Applicable</td>
</tr>
<tr>
<td>01</td>
<td>Y</td>
<td>0,01</td>
<td>True</td>
</tr>
</tbody>
</table>

Table 1 – Value of the Answer
“To judge quality” means, in essence, to interpret measurement results. The first step toward it has already been taken in the previous item, when rating levels for the measurements have been established. What would be desirable next, is to secure conclusions about quality, beginning with the series of values obtained from applying measurements.

Following is the suggestion of a method to secure evaluation synthetic results:

1) To chart all the metric results for a (0,1) scale, in which 0 stands for the worst possible result, whereas 1 stands for the best. The measurements suggested in this method has followed this pattern.

2) Establish weights for software quality characteristics and subcharacteristics. The weights must stand for each item’s relative importance in the global judgement of product quality. Values must be secured mainly upon the evaluation requester, in the end, must say, what is or is not important in his insight about product quality as well as specialized professionals in the area of software application.

3) To figure weighted average by using metric values and weights of respective characteristics and subcharacteristics. It is not recommended to assign weights to attributes since they are not rigorously defined in ISO 9126; the same attribute can be utilized to evaluate various subcharacteristics.

This it is possible to figure out averages by grouping the measurements of each quality trait. To figure out an average or weighted rate for usability, considering the measurements related to quality trends and weights ascribed to each subcharacteristic. The same can be done with each characteristic level separately. The process can be repeated in the quality characteristic level, hence allowing for computing a single index standing for software quality. The Evaluation Process where this method can be applied is the one established by ISO 14598-5[4] Standard. At the end of this process an evaluation qualitative report is secured where through the software product components and their respective quality characteristics the points which agree with the Standards and the ones which do not are attended. An evaluation quantitative report can also be produced, by showing through charts how much each software product has attained.

4. MEDE-PROS Method

The main purpose of the MEDE-PROS is to provide evaluators with means to support the evaluation of software products from the viewpoint of the end user, according to ISO/IEC 9126 [5] and ISO/IEC 12119 [3] in concerning quality characteristics and software packages, respectively. The Process of Evaluation of software products of the MEDE-PROS is based on ISO/IEC 14598-5 [4], aiming at providing requirements and recommendations for the practical implementation of software products evaluation, developed or under development, as a series of activities defined under common agreement between the customer and the evaluator [7]. The tools for application of the MEDE-PROS are stored in the Evaluation DataBase. The evaluation process is carried out by simulating a normal operational use of the product, beginning by analysing the documentation, installing the product as instructed in the documentation and proceeding by using the product in the most complete way that can be done. All along the process, evaluators assign rates to the product according to Checklist questions. In addition to assigning rates, evaluator’s record the time spent in the evaluation, specify the major functions of the product, and write down comments on specific issues they consider relevant concerning the product. The final step in the evaluation process is the preparation of the Evaluation Report, which should address the major positive aspects of the evaluated product as well as suggestions for its improvement. In order to evaluate Product Description and Package, the quality requirements specified in ISO/IEC 12119 Standard has been considered. In order to evaluate the User’s documentation, the quality requirement as specified in ISO 12119, and ANSI/ IEE 1063 and ISO 9126 were considered. In order to evaluate the Software, the quality requirement as specified in ISO 9126 Standard were taken into consideration. In order to evaluate Interface the quality requirements as specified in ISO 9126 and ISO 9241 parts 10,11 and 12 were taken into consideration (see Figure 6).

5. Results and Evolution for MEDE-PROS

Throughout the last ten years the MEDE-PROS methodology has obtained positive results for the Brazilian market of software production with projects carried out in a total of 360 products evaluated, through the evolution of the
method showed in the Table 2. This method is in constant development, in each version it is improved. The evaluation must be reproductive and repetitive with the new version of the Method.

<table>
<thead>
<tr>
<th>VERSION</th>
<th># questions</th>
<th>Standards ISO/IEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>V 93</td>
<td>43</td>
<td>9126</td>
</tr>
<tr>
<td>V 94</td>
<td>85</td>
<td>9126, 14598-5 + CenPRA experience *</td>
</tr>
<tr>
<td>V 95</td>
<td>90</td>
<td>9126, 14598-5, 12119 + CenPRA experience</td>
</tr>
<tr>
<td>V 96</td>
<td>129</td>
<td>9126, 14598-5, 12119 + CenPRA experience</td>
</tr>
<tr>
<td>V 01/97</td>
<td>129</td>
<td>9126, 14598-5, 12119 + CenPRA experience</td>
</tr>
<tr>
<td>V02/98</td>
<td>587 **</td>
<td>9126, 14598-5, 12119, 9241 + CenPRA experience</td>
</tr>
<tr>
<td>V02/2000</td>
<td>540 **</td>
<td>9126, 14598-5, 12119, 9241, IEEE 1063 + CenPRA experience</td>
</tr>
</tbody>
</table>

(*): CenPRA experience: includes meta Evaluation, Statistical Analysis and Historical Database.
(**): The numbers of questions increased because it was opened each attribute in attribute’s items.

Table 2 - Summary of the evolution of the method

This method framework can be used as a starting point for a product improvement and it can be tailored to the needs of company. It can be part of a program for the company’s quality improvement.

References

[1] Guerra, Ana; Sant’Ana, Mary - Quality of Software Process or Quality of Software Product? Submission has been accepted for the 12th International Conference for Software Quality, Ottawa-Ca, 11/2002.