The Use of Automated Software Tools in Evaluating an e-Learning Platform Quality

Traian-Lucian MILITARU, George SUCIU, Gyorgy TODORAN
Quality Dept, 3CPS, Quality Dept
Politehnica University of Bucharest, ROMANIA
gelmosro@yahoo.com, george@beia.ro, todoman.gyorgy@gmail.com

Abstract: This paper proposes an expert system which can be used to evaluate the quality of an e-learning platform. The proposed expert system is using the modified version of the SEEQUEL Core Quality Framework and it was built using CLIPS expert system generator. The SEEQUEL Core Quality Framework originated from the collaboration between the e-learning Industry Group (eLIG) with a number of European expert organizations and associations, coordinated by the MENON Network, is a framework used to build the quality tree by selecting the quality characteristics from a list of common characteristics applicable to the whole e-learning experience. CLIPS is a productive development and delivery expert system tool which provides a complete environment for the construction of rule based expert systems.

In the first part of this paper the SEEQUEL Core Quality Framework and CLIPS expert system generator are presented showing the advantage of using an expert system for this task. In the second part, a case study of evaluating an e-learning platform is presented. The final conclusion of the experiment was that an expert system can successfully replace a human expert for the proposed task.

Key Words: e-learning; evaluation; SEEQUEL Core Quality Framework; computer assisted

1. Introduction

E-learning covers all forms of electronically supported learning and teaching. This article only refers to computer-based e-learning which is essentially the computer and network-enabled transfer of knowledge. The content is considered to be delivered via the Internet. It can be or not instructor-led and includes media in the form of text, image, video and audio streaming. Due to the rise of the informational volume available on the Internet, more and more users choose to use virtual environments for the learning process. Like for any other product, the necessity to select those virtual environments which satisfy the demands of the users appears. The extent to which user expectations are met is called quality.

The proposed evaluation method is using a modified version of the SEEQUEL Core Quality Framework. The modifications provide more accurate results by choosing a finer scale and taking into account the interdependence of characteristics.

2. The General Presentation of the Evaluation Method

The proposed evaluation method includes five steps:
A. the definition and specification of the demands referring to quality;
B. the building of the quality model;
C. the measurement;
D. the aggregation of the scores;
E. the analysis of the scores, the formulation of the conclusions, recommendations for improvement.

A similar method is Web Quality Evaluation Method (WebQEM) developed between 1998-2000 by a group of researchers from the National University of La Pampa (Argentina) led by Luis Olsina which was used as an inspiration resource for formulating this generalized evaluation method.

2.1. Defining and Specifying the Demands Referring to Quality

The quality is linked to individuals’ visions, perspectives values, roles, contexts, and that a “one size fits all” model for quality does not exist. There is
no simple definition of quality in e-learning and any definition we might wish to consider runs the risk of constraining people vision of what quality means and it significance in their particular context [1]. In other words, quality is different for different people and even for the same person in a different context. The designer of the e-learning environment must perceive quality as the degree to which the learning objectives are achieved by the type of learner for which it was designed [1].

A conceptual framework which is trying to integrate in a single structure the very different perceptions of the various stakeholders about e-learning quality can be an usefully tool used to evaluate such of environments. A particular perspective on quality can be achieved by giving different weights to the conceptual framework structure’s members.

In this article will be used a modified version of the SEEQUEL Core Quality Framework (http://www.menon.org/) which is a part of the SEEQUEL project, supported in frames of the EU e-learning initiative, originated from the collaboration between the e-learning Industry Group (eLIG) with a number of European expert organisations and associations, co-ordinated by the MENON Network. Only three quality characteristics will be kept while the aggregation methodology will be replaced with Logic Scoring of Preferences (LSP) proposed by Dujmovic. The SEEQUEL project activities addressed the need for a common strategy to define and implement international e-learning quality standards. The SEEQUEL Core Quality Framework is based on a matrix where a list of common quality criteria applicable to the whole e-learning experience can be weighted by the various user (people or organisation), enabling any category of stakeholders to position their perception of quality with respect to the perceptions of another stakeholders’ category [1]. So the remaining challenge is choosing quality characteristics and determining their interdependence and their weights.

The SEEQUEL Core Quality Framework proposes a quality model with three main quality characteristics [1]:

- The Learning sources which contain all the involved sources: the technical infrastructure, the learning materials and the human resources in their functions of teaching and supporting learning.
- The Learning processes: Any learning experience consists of a series of processes. There are two main types of processes. Firstly there are those that occur during the actual learning experience – learning processes. Secondly, there are a series of processes set up around a given learning experience. These support processes underpin any learning experience, but are separate from the learning experience itself – for example recruitment of teachers, which does not take place as part of the learning experience, but still has a definite impact on the learning experience itself. The processes identified are typical of any training action such as: training needs analysis, guidance, recruitment, design (macro and micro), delivery and assessment/evaluation.
- The Learning context is the environment in which learning takes place. It is viewed in its double perspectives: intrinsic and relative. The quality of a learning experience depends both on the quality of the context itself and on the relationship of the designed and implemented experience to the context in which it occurs.

The SEEQUEL Quality conceptual framework proposes the following scale for weights.

- 2 = key/ core criteria for defining quality of the object
- 1 = important criteria for defining quality of the object
- 0 = non relevant for defining quality of the object

For getting more accurate results, this article propose a modified scale for weights, the [0..1] scale, where 0 means the characteristic is not important for the research and 1 means o maximum importance. The sum of
weights of all subcharacteristics of a characteristic must be 1.
In order to be able to define the quality demands, the users’ needs must be established. Based on this information, the relevant characteristics of the quality model defined in the SEEQUEL Core Quality Framework are selected and the weight of each characteristic is established.

2.2. The Measurement

This stage includes two main activities: the measurement projection where metrics and the measurement methodology of the measurable features (set up in the anterior step) are established and the measurement implementation where the measurement itself is done and the scores are calculated for each measured feature. The measurement methodology establishes how the measurements are done and also the order of the measurements. The measurement means to associate for each measurable feature identified as \( A_i \) a variable \( X_i \) whose numerical value is obtained by applying some direct or indirect metrics according to the measurement methodology. In view of obtaining scores for evaluating the quality level, a preference scale is set up by defining a preference criterion. A preference criterion can be the \( E: [0, X_i^{\text{max}}] \to [0,1] \) function, which generates an \( E(X_i) \) preference which represents the measure in which the \( A_i \) attribute satisfies the demands associated to him \([2]\). This preference represents the score. The 0 value indicates the fact that \( X_i \) does not satisfy the demands and the 1 value - which it carries out perfectly. The most common preference criterion and the one which was used in this study is the following:

\[
E(X_i) = \begin{cases} 0 & \text{if } X_i > X_i^{\text{max}} \\ \frac{X_i}{X_i^{\text{max}}} & \text{if } 0 \leq X_i \leq X_i^{\text{max}} \end{cases}
\]

### Table 1. Example of \( E(X_i) \) function

<table>
<thead>
<tr>
<th>The value of the ( X_i ) variable</th>
<th>The ( E(X_i) ) preference criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_i &gt; X_i^{\text{max}} )</td>
<td>( E(X_i) = 0 )</td>
</tr>
<tr>
<td>( 0 \leq X_i \leq X_i^{\text{max}} )</td>
<td>( E(X_i) = \frac{X_i}{X_i^{\text{max}}} )</td>
</tr>
</tbody>
</table>

2.3. The Aggregation of the Scores

In LSP, the \( Q \) quality index (or the total quality) is obtained by aggregation in conformity with the defined quality model. The aggregation is done ‘from down to up’, in several steps, in the tree-like structure of quality characteristics. The final result is a global plan which allows the calculation of the quality indicators for all the characteristics and also the total quality indicator. This last global preference represents the general satisfaction level referring to fulfilling the requirements concerning quality valued \([2]\).

The \( E \) aggregated preference of some \( E_i \) preferences from a subsequent inferior level of the quality tree is the following function:

\[
E(r) = (w_1 \cdot E_i^1 + w_2 \cdot E_i^2 + \ldots + w_k \cdot E_i^k)^{\frac{1}{r}} \quad (1)
\]

where

\[
-\infty \leq r \leq \infty \quad (2) \\
0 \leq E_i \leq 1 \quad (3) \\
0 \leq w_i \leq 1 \quad (4) \\
w_1 + w_2 + \ldots + w_k = 1 \quad (5)
\]

\[
E(-\infty) = \min(E_1, E_2, \ldots, E_k) \quad (6) \\
E(\infty) = \max(E_1, E_2, \ldots, E_k) \quad (7)
\]

The \( r \) parameter is given by the simultaneity level of preferences which are to be aggregated. Small values of \( r \) lead to small values of the \( E(r) \) function, which inside CPL (Continuous Preference Logic) means conjunction while high values of \( r \) lead to high values of the \( E(r) \), the equivalent of disjunction. So, low values of \( r \) are chosen when they wish to punish the lack of simultaneity of the preferences which are aggregated while high values appear in opposite cases\([3]\). This aggregation method also offers the opportunity to achieve more accurate results.

2.4. The Analysis of the Results, Formulating and Gathering Documentary Evidence for Conclusions
This step includes the analysis of results, the formulation of conclusions and, eventually, of some recommendations concerning the improvement of quality. All these are registered as documents to be presented to those who demanded the evaluation.

In view of an easy interpretation of preferences, the preference scale can be divided in levels of acceptability. Such an example could be:

- inadequate: from 0% to 39%
- medium: from 40% to 79%;
- satisfying: from 80% to 100%.

3. Expert Systems

An expert system is a computer system that emulates the decision-making ability of a human expert [4]. They are designed to solve complex problems by reasoning about knowledge, like an expert, not by following the procedure of a developer as is the case in conventional programming [5]. It can be concluded that if a human expert can specify the steps and reasoning by which a problem may be solved, then an expert system can be created to solve the same problem.

The architecture of an expert system is illustrated in the following figure:

![Figure 1. The structure of an expert system.](image)

The knowledge base of expert systems is a collection of rules encoded as metadata in a file system, or more often in a relational database and contains both factual and heuristic knowledge. Factual knowledge is that knowledge of the task domain that is widely shared and commonly agreed. Heuristic knowledge is the less rigorous, more experiential, more judgmental knowledge of performance. In contrast to factual knowledge, heuristic knowledge is rarely discussed and is largely individualistic. It is the knowledge of good practice, good judgment and plausible reasoning in the field. It is the knowledge that underlies the "art of good guessing" [5].

Knowledge representation formalizes and organizes the knowledge. The one used in CLIPS is the production rule, or simply rule. A rule consists of two parts: an IF part (also called a condition) and a THEN part (also called action). The IF part is a logical combination of some conditions. The piece of knowledge represented by a particular production rule is relevant to the line of reasoning being developed if the IF part of the rule is satisfied; consequently, the THEN part can be conclude or its problem-solving action taken. Expert systems whose knowledge is represented in a rule form are called rule-based systems [5]. The working memory initially contains the data that is received from the user as task input which is used to evaluate antecedents (the IF parts of the rules) in the knowledge base. Fired consequents from rules (actions whose conditions are true) in the knowledge base may create new values in working memory, update old values or remove existing values.

The inference engine is that part of an expert system which is designed to produce reasoning on rules. CLIPS uses forward chaining and rule prioritization for reasoning. The inference engine inspects the condition part of every rule from the knowledge base until it matches one for which the logical of condition is true based on the data available in the working memory. When the rule condition part is true, the inference engine executes the action part of that rule which is determining the change of the working memory. This process continues until the solution to the problem is found or no new rules can be fired (in this case, the user can be prompted for offering new information) or the necessary resources (time, memory, etc.) are over.

In an expert system (especially a large one with a large amount of rules), at any moment in time there may be a series of rules that are ready to fire which are known as the conflict set. A conflict resolution strategy is required to make the decision as to which rule should be fired first. CLIPS provides seven conflict resolution strategies:
depth, breadth, simplicity, complexity, LEX, MEA, and random. The default strategy is depth, which means the newly activated rules are placed above all rules of the same salience. For example, given that fact- a activates rule- 1 and rule- 2 and fact- b activates rule- 3 and rule- 4, then if fact- a is asserted before fact- b, rule- 3 and rule- 4 will be above rule- 1 and rule- 2 on the agenda. However, the position of rule- 1 relative to rule- 2 and rule- 3 relative to rule- 4 will be arbitrary [6].

The engine has two ways to run: batch or conversational. In batch, the expert system has all the necessary data to process from the beginning. For the user, the program works as a classical program: he provides data and receives results immediately. Reasoning is invisible. The conversational method becomes necessary when the developer knows he cannot ask the user for all the necessary data in the beginning, the problem being too complex. The software must "invent" the way to solve the problem, request the missing data from the user, gradually approaching the goal as quickly as possible. The result gives the impression of a dialogue led by an expert.

CLIPS is a productive development and delivery expert system tool which provides a complete environment for the construction of rule based expert systems. An expert system generator consists of an empty knowledge database which can be filled through an interface by a knowledge engineer, a customizable inference engine, user interface and explanatory component building tools and a working memory building tool. The inference engine, the filled database, the user interface and the explanatory component (both created by the system engineer) and the working memory will form the future generated expert system.

4. Case Study

The proposed expert system was used to design the quality evaluation process of the e-learning platform belonging to Open University from UK (http://www.open.ac.uk/). The SEEQUEL Core Quality Framework is permitting a quality evaluation from the perspective of any of the stakeholders, be it student, teacher or technical staff. In this study, the evaluation has been carried out from the student’s perspective, the person who wants to accumulate knowledge.

In order to have the right user perception about quality, the expert system must communicate with the user to find what is really important for him. Knowing the user expectation from the e-learning platform, the expert system can select and weigh the quality characteristics in the proper way to obtain his right perception about quality. First of all, the expert system must find the ‘stakeholders’ category where the user is included. For this study, the ‘students’ category was used. Of all proposed features of SEEQUEL Core Quality Framework it will selected and weighed those that were considered important for the present study. These are:

1. Learning sources (w=0.4)
1.1. Teaching staff (w=0.4)
1.1.1 Ability to support learning (based on past experience) (w=0.5)
1.1.2 Ability to support the motivation to learn (w=0.25)
1.1.3 Adaptability to change/flexibility (w=0.05)
1.1.4 Communication skills (w=0.10)
1.1.5 Ability to monitor learning progress (w=0.05)
1.1.6 Ability to assess learning progress (w=0.05)

1.2. Learning materials (w=0.4)
1.2.1 Content reliability and updated (w=0.35)
1.2.2 Technical quality (w=0.2)
1.2.3 Aesthetic quality (w=0.05)
1.2.4 Conformance to standards (w=0.2)
1.2.5 Ease of use (w=0.05)
1.2.6 Interactivity (w=0.05)
1.2.7 Materials are learner driven (w=0.05)
1.2.8 Low cost (w=0.05)

1.3. Learning infrastructure (w=0.2)
1.3.1 Accessibility (e.g. people with disabilities, organizational difficulties) (w=0.05)
1.3.2 Adaptability to users' need (w=0.3)
1.3.3 Interoperability with other systems (w=0.05)
1.3.4 Ease of use (w=0.2)
1.3.5 Reliability of the Technical infrastructure (w=0.2)
1.3.6 Availability of different communication tools (w=0.1)
1.3.7 Uniform interface (w=0.05)
1.3.8 Personalized interface (w=0.05)

2. Core learning processes (w=0.4)
2.1. Learning design (w=0.4)
2.1.1 The learning course is aimed at developing knowledge (w=0.1)
2.1.2 The learning course is aimed at developing competences (w=0.1)
2.1.3 Clear definition of the target groups (w=0.1)
2.1.4 Coherence of the didactic strategy with course objectives (w=0.2)

2.1.5 Granularity of the content (w=0.2)
2.1.6 Flexibility of the learning path is assured in the module or course (w=0.1)
2.1.7 Learning accreditation system is available and linked to the national/European accreditation system context (e.g. ECTS) (w=0.1)
2.1.8 Technical assistance in course development is assured (w=0.1)

2.2. Learning Delivery (w=0.6)
2.2.1 Accessibility (w=0.4)
2.2.2 Reliability and robustness of the online services (w=0.4)
2.2.3 Throughout the course students are provided with technical and pedagogical support for using the services available (individual and group level) (w=0.15)
2.2.4 Measures to monitor and rectify common technical problems (w=0.05)

3. Learning context (w=0.2)
3.1. Learning Environment (w=0.7)
3.1.1 Supporting personalization (w=0.2)
3.1.2 Supporting competence representation (mapping) and recognition (w=0.05)
3.1.3 Supporting interaction/communication (w=0.4)
3.1.4 Supporting team work (w=0.2)
3.1.5 Supporting individual behavior patterns rather than common behavior patterns (w=0.1)
3.1.6 Supporting divergent thinking (w=0.05)

3.2. Legislation (w=0.1)
3.2.1 IPR and Copyrights related to the materials' content are respected (w=0.5)
3.2.2 The accreditation system is recognized by professional associations (w=0.5)

3.3. Financial setting (w=0.2)
3.3.1 Specific financial aid measures are devoted to disadvantaged categories (e.g. grants, participation fees) (w=1).
Next, the proposed expert system will build the aggregation model by selecting the appropriate values of \( r \), according to the desired conjunction/disjunction degree (andness/orness).

The aggregation model has been built using information from the knowledge base. The human expert has established the simultaneity level of preferences which are to be aggregated.

To realize a continuous transition from disjunction to conjunction with 16 equidistant steps of andness/orness, the following special particular cases will be used:

- \( D ( r = +\infty ) \)
- \( D++ ( r = 20.63 ) \)
- \( D+ ( r = 9.52 ) \)
- \( D++ ( r = 5.8 ) \)
- \( D+ ( r = 3.93 ) \)
- \( D++ ( r = 2.79 ) \)
- \( D+ ( r = 2.018 ) \)
- \( D++ ( r = 1.449 ) \)
- \( A ( r = 1 ) \)
- \( C-- ( r = 0.619 ) \)
- \( C-- ( r = 0.261 ) \)
- \( C-- ( r = -0.148 ) \)
- \( CA ( r = -0.72 ) \)
- \( C-- ( r = -1.655 ) \)
- \( C+ ( r = -3.51 ) \)
- \( C++ ( r = -9.06 ) \)
- \( C ( r = -\infty ) \).

The small values of \( r \) (C++ aggregator for example) are chosen when desired to punish the lack of simultaneity of the preferences which are aggregated while high values appear in opposite cases.

For example, the aggregation model for "Learning materials" is shown in Figure 4:

![Figure 4. The aggregation model for “Learning materials”](image)

One can notice the expert is punishing the simultaneity of preferences associated to 1.2.2 (Technical quality) and 1.2.4 (Conformance to standards) by using D++ as aggregator. Also the lack of the simultaneity of preferences associated to 1.2.8 (Low cost) and the preferences associated to the rest of the...
quality characteristics are punished by using C++ as aggregator. The first step of the assessment process, the building of the tree-like quality model is ready according with the user’s view about the importance of the e-learning platform characteristics. Using this model, a questionnaire having the questions which covers the user interest can be generated. The user’s answers to these questions are defining his perception about quality characteristics of the e-learning platform and once they are provided, the aggregation can be started ending up with the calculation of the total quality indicator. Also, some conclusions and recommendations can be offered to the e-learning platform developers in order to help them to improve the users’ experience in their virtual environments.

<table>
<thead>
<tr>
<th>Main characteristics</th>
<th>Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning sources</td>
<td>94.45</td>
</tr>
<tr>
<td>Learning processes</td>
<td>88.75</td>
</tr>
<tr>
<td>Learning context</td>
<td>92.25</td>
</tr>
<tr>
<td>TOTAL</td>
<td>91.81</td>
</tr>
</tbody>
</table>

One can notice a greater attention of the creators of this e-learning environment for learning sources and for learning context. The Learning sources characteristic score could be bigger if the following aspects will be improved: Adaptability to change/ flexibility, Ability to support the motivation to learn, Ability to give encouragement, Aesthetic quality, Pre-requisites identification, Low cost and Motivating. The weak points of Learning processes are: Different payment facilities for the course fee, Coherence of the didactic strategy with course objectives and Granularity of the content. The following aspects need to be improved in order to have a bigger score for Learning context: Collaborative working procedures and tools, Encouragement for competition and Supporting contextualization rather than abstraction.

The final conclusion is that the analyzed e-learning environment has a satisfactory quality (from the student point of view), all the scores being higher than 80 on a scale from 0 to 100.

5. Conclusions

Applying this method of e-learning evaluation, the weak aspects of the analysed environment were highlighted, offering the possibility of a later improvement of quality. The main advantages of the method are its power given by the used mathematical support and also the ease of quality measurement and recording process, the ease of global quality calculation, the ease of quality comparison of similar products and observation of the weakness propagation.

Using an expert system to design the quality evaluation adapted to the evaluator's profile, the guarantee of an evaluation which will offer a complete image of the user’s perception on the
quality of the evaluated e-learning platform is ensured. The final conclusion is that an expert system can successfully replace a human expert for this task, the main advantages being the increased availability, low-cost and fast service.

References