Fuzzy Logic-Based System for the Estimation of the Usability Level in User Tests

G.E. Chanchí-Golondrino, L.M. Sierra-Martínez, W.Y. Campo-Muñoz

Gabriel Elías Chanchí-Golondrino*
Faculty of Engineering
University of Cartagena, Colombia
Ave. del Consulado, Cll. 30 # 39 B-192, Cartagena, Colombia
*Corresponding author: gchanchig@unicartagena.edu.co

Luz Marina Sierra-Martínez
Faculty of Electronic Engineering and Telecommunications
University of Cauca, Colombia
Cll. 5 # 4-70, Popayán, Colombia
lsierra@unicauca.edu.co

Wilmar Yesid Campo-Muñoz
Faculty of Engineering
University of Quindío, Colombia
Cra. 15 Cll. 12 norte, Armenia, Colombia
wycampo@uniqindio.edu.co

Abstract

Starting from the challenge of obtaining the usability level in numerical and linguistic terms within a user test based on the three attributes that define usability, the development of a system based on fuzzy logic is proposed for estimating the level of output usability in a test with users based on the ISO 9241-11 standard. The attributes are effectiveness, efficiency, and satisfaction, which have different metrics and numerical scales. For the development of the system, five methodological stages were defined: characterization of the structure of a usability test, definition of membership functions for the inputs and outputs of the system, design of the inference rules that relate the inputs and outputs, design and implementation of the fuzzy system, and development of the case study. The proposed system was implemented using the FCL (Fuzzy Control Language) and the jFuzzyLogic API that takes as inputs the values calculated for the attributes of effectiveness, efficiency, and satisfaction, and obtains the usability level as output considering the membership functions of the inputs and outputs, as well as a set of inference rules defined by a set of experts. As a case study, the proposed fuzzy system was validated from the results obtained in a usability test with 5 users which was developed on the Sigma Electrónica website. From the results obtained in the case study, it could be concluded that the implemented system is adequate in terms of obtaining a level of usability in numerical and linguistic terms in conventional usability tests developed in a usability laboratory considering the attributes of ISO 9241-11.

Keywords: ISO 9241-11, fuzzy logic, usability test, usability.
1 Introduction

From the increase in the number of applications available in the cloud and in software stores, user-centered design and usability has become key aspects in order to improve the quality of the software and promote competitiveness of companies in the market [2, 31, 33]. According to the ISO 9241-11 standard, usability can be defined as the extent to which a software product can be used by specific users to achieve specific objectives with effectiveness, efficiency, and satisfaction in a specific context of use [15, 16, 17, 30]. In this sense, to consider a software usable, it must comply with the attributes of effectiveness, efficiency, and satisfaction. In the context of usability, effectiveness refers to the fulfillment of tasks by the user in an adequate manner within the software; efficiency refers to the optimization of resources by the user in the interaction, where speeding the times for the development of tasks within the software is one of the key aspects; finally, satisfaction according to this standard refers to the existence of positive attitudes towards the use of the software [1, 6, 9].

On the other hand, according to the ISO 9126-1 standard, usability can be defined in terms of the ability of a software product to be understood, learned, operated, or used and that is attractive to the user under specific conditions [12, 19, 25, 28]. There is a relationship between the attributes defined in the ISO 9241-11 and ISO 9126-1 standards, in such a way that the ability to be understood and learned can be related to the efficiency attribute, while the ability to be operated or used is related to the effectiveness attribute. Likewise, the possibility that the software is attractive to the user can be associated with the satisfaction attribute.

There are different methods to evaluate the usability of a software product. Among them, the ones that stand out are the usability tests where a group of users develops a set of tasks within a given software; they are developed in a usability laboratory and supervised by a group of observers or coordinators of the test. They are in charge of remotely monitoring the tasks carried out by the user in order to apply the metrics that allow determining the effectiveness, efficiency, and satisfaction for the ISO 9241-11 standard [1, 5, 20, 22]. In this way, the user effectiveness percentage can be determined in terms of the number of sub-tasks developed by each task designed to be executed within the software that has to be evaluated. The percentage of efficiency can be determined by relating the time estimated by the coordinators for the development of tasks regarding the time employed by users to develop those tasks. Finally, the satisfaction calculation can be carried out from the processing of the survey done after the test [13].

One of the challenges from the calculation of the attributes of effectiveness, efficiency, and satisfaction of the different users who participate in a user test is the precise determination of an output usability level, taking into account the different characteristics and scales obtained in the 3 attributes, as well as the need to determine not only a result in numerical terms but also linguistic. In this sense, given the particularities of the problem, a system based on fuzzy logic can be designed and adapted to the context of the tests with users in order to estimate the level of output usability, taking as input the values calculated by the coordinators of the tests for effectiveness, efficiency, and satisfaction, as well as a set of inference rules that relate the inputs and the output level [10, 18]. Fuzzy logic can be understood as a mathematical instrument that facilitates the process of reaching a specific conclusion from input information that can be considered undefined, inaccurate, or vague. All of this considering the knowledge of experts as a basis [7, 14, 26, 32]. This way, the importance of fuzzy logic focuses on the fact of enabling the generation of intelligible results that relate numerical data with linguistic terms which are closer to natural language [3, 21, 29].

Different contributions have been developed in the field of usability in which fuzzy logic has been used. Thus in [7], the authors proposed a tool for the development of usability inspections based on fuzzy logic; it allows to obtain the level of usability from the qualification of usability criteria by expert users. In [4], the authors proposed a tool based on fuzzy logic to obtain the level of usability from the ratings made by users from the usability attributes of the ISO 9126-1 standard. In [23], the authors proposed a model based on fuzzy logic for evaluating software quality, taking as input the ratings made on the metrics proposed in the ISO 9126-1 standard. In [24], the authors developed a fuzzy system to obtain the level of usability from the qualification of software quality parameters for the evaluation of websites. Finally in [11], the authors proposed a fuzzy system for determining usability in virtual learning environments based on the qualification of criteria of the ISO 9126-1 standard adapted to that
context. Previous works have been focused on the application of fuzzy logic in usability evaluations based on the weighting of criteria by expert users, without focusing on the determination of usability from evaluation approaches such as tests with users developed in a usability lab.

In this article we propose the development of a software system based on fuzzy logic for estimating the level of usability output in tests with users developed in a usability laboratory, considering the attributes of effectiveness, efficiency, and satisfaction as defined in the ISO 9241-11 standard, as well as the metrics associated with these attributes. Thus, the proposed system takes as input the calculated values of effectiveness, efficiency, and satisfaction from the interaction of the users who participated in the test, in order to obtain a percentage level of output usability in linguistic and numerical terms, taking into account a set of inference rules defined by a group of experts in the development of the usability tests. The proposed system aims to guide the coordinators of a usability test to obtain clearer indicators regarding the level of usability of a software application evaluated in the context of a user test.

The rest of the article is organized as follows: section 2 presents the different stages of the methodology considered for the development of this research. Section 3 describes the results obtained from this research, which includes the design of the membership functions for the inputs and outputs of the fuzzy system, as well as the inference rules that allow to obtain the level of output usability from the system inputs. In the same way, the functionality of the fuzzy system developed to estimate the level of output usability in user tests is described, and a case study is presented where the functionality of the proposed system is verified with the results obtained from a test of usability developed with 5 users over the electronic accessories store website Sigma Electrónica [27]. Finally, section 4 shows the conclusions and future work derived from this research.

2 Material and Methods

For the development of this research, the following five methodological stages were defined: characterization of tests with users, definition of membership functions, design of inference rules, construction of the fuzzy system; and finally, development of the case study. The process was adapted from what is proposed in [8] (see Figure 1)

Figure 1: Methodology considered for this research (Own Elaboration)

In phase 1, the characterization of the structure of a test with conventional users developed in a usability laboratory was carried out. The test was made up of 4 stages: confidentiality agreement, pre-test questionnaire, list of tasks, and post-test questionnaire. The confidentiality agreement is a document that informs the user about the scope of the test and guarantees that the test information will be used for academic purposes only. The pre-test questionnaire is an instrument that seeks to obtain the user profile, as well as the previous experience in using applications similar to those that the user will evaluate in the test. The list of tasks corresponds to the tasks defined by the test coordinators and that the user will do within the application to be evaluated. Finally, the post-test questionnaire is a quantitative and qualitative instrument that seeks to inquire about the user’s experience with the evaluated application, which includes the perception of satisfaction in the use of the software. Based on the supervision done by the observers about the tasks performed by the user, the attributes of effectiveness, efficiency, and satisfaction are calculated according to the ISO 9241-11 standard. The percentage of effectiveness per task can be determined by the following equation, where each task in the test must have a set of sub-tasks.

\[
\text{%effectiveness} = \frac{(\text{number of subtasks performed} \times 100)}{\text{(number of defined subtasks)}}
\]
On the other hand, the percentage of efficiency per task can be determined from what is presented in the following equation, where the test coordinators have estimated a duration for each task.

\[
\text{\%efficiency} = \frac{(\text{estimated time per task} \times 100)}{\text{(time employed per task)}}
\]

Finally, regarding satisfaction, the evaluation given by the end users to the questions of the post-test questionnaire that involve the user’s perception of satisfaction is averaged, as it is presented in the following equation.

\[
\text{satisfaction} = \frac{(\text{sum of the questions ratings})}{(\text{number of questions})}
\]

In stage 2, the design of the membership functions for both the input (% effectiveness, % efficiency, and satisfaction) and output (level of output usability) variables of the system takes place. These functions were designed from the scale of the three attributes, and they were specified through the FCL (Fuzzy Control Language). In stage 3, a total of 27 inference rules were defined, which relate the inputs of the system to the output and were also specified in the FCL language. Starting from the membership functions and the inference rules, the development of the fuzzy system for determining the level of usability in user tests takes place in stage 4, making use of the Java language and the jFuzzyLogic library. Finally, in stage 5 of the methodology, a case study was developed in which the proposed fuzzy system was used to obtain the level of output usability from the data of a user test developed by five users over the Electronics store website: Sigma Electrónica [27].

3 Results

This section presents the description of the membership functions of the inputs and output of the system, as well as the ranges defined for the system. In the same way, the different views of the implemented system and a case study that demonstrates the functionality of the system are presented.

3.1 Membership functions

In Figure 2, the 3 membership functions defined are presented. On each one of them, three fuzzy sets (low, medium, and high) have been considered under the experts’ criteria, considering the equations associated with effectiveness, efficiency, and satisfaction presented in section 2, as well as the ranges belonging to these attributes.

Similarly, Figure 3 presents the membership function designed for the output usability level that according to the usability experts, is made up of four fuzzy sets: deficient, acceptable, good, and excellent. In this way, the system will not only obtain a level of usability as a result, but also the linguistic term associated with that value.

3.2 Inference Rules

Taking into account the membership functions defined for the input and output variables of the system, a total of 27 rules were defined that relate the fuzzy sets considered for effectiveness, efficiency, and satisfaction (low, medium and high), with the fuzzy sets considered for the level of output usability (deficient, acceptable, good, and excellent). Thus, Table 1 presents a fragment of the 27 inference rules defined, which are expressed in the FCL language.

3.3 Fuzzy system proposed and case study

From the membership functions of the input variables and the output variable, as well as the defined inference rules that relate the inputs to the output, a fuzzy system was developed in the Java programming language using the jFuzzyLogic library and the FCL fuzzy logic description language. The system allows to obtain the output usability level from the effectiveness, efficiency, and satisfaction attributes of the ISO 9241-11 standard, which are calculated from the metrics of each attribute. To
describe the operation and the different views or interfaces of the proposed fuzzy system, the results obtained in a usability test were used. The test was developed with 5 users in a virtual laboratory over the Sigma Electrónica website (see Figure 4). In this way, at the time the functionality of the implemented system is shown, the results obtained for the case study are shown.

In the first place, the proposed system allows to load the data captured by the coordinators of a user test from a plain text file for configuration. This way, it is possible to calculate the value of the usability attributes per task and at the test level. Thus, Figure 4 presents a plain text file that describes the results obtained in a test with users developed over the Sigma Electrónica website.

The file in Figure 5 shows how the test developed over the Sigma Electrónica website was carried out by 5 users who were assigned a total of 3 tasks (T1: register on the Sigma Electrónica website using a test email provided by the coordinators, T2: interact with the site’s shopping cart by searching for three products available in the store, T3: enter the blog section and comment on any of the articles in this section). In the same way, at the level of the effectiveness attribute, task 1 has 2 subtasks while task 2 has 3 subtasks associated, and task 3 includes 2 tasks. In this sense, it can be observed how user 1 (U1) developed the 2 subtasks corresponding to tasks 1 and 3. On task 2, U1 performed 2 of the 3 subtasks. At the efficiency level, it is possible to see how the coordinators of the user test established 1 minute for task 1, 3 minutes for task 2, and 2 minutes for task 3. It is also observed how user 4 (U4) developed task 1 in 2 minutes, task 2 in 2 minutes, and task 3 in 3 minutes. Finally, at the level of satisfaction, the 5 users rated 6 questions in the range from 1 to 5 related to the perception of the website after the test. Thus, in the case of user 5, the scores on the 6 perception questions were
Table 1: Fuzzy system inference rules

<table>
<thead>
<tr>
<th>Id</th>
<th>Inference rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IF effectiveness IS low AND efficiency IS low AND satisfaction IS low THEN usability_level IS deficient</td>
</tr>
<tr>
<td>4</td>
<td>IF effectiveness IS low AND efficiency IS low AND satisfaction IS low THEN usability_level IS deficient</td>
</tr>
<tr>
<td>8</td>
<td>IF effectiveness IS low AND efficiency IS high AND satisfaction IS medium THEN usability_level IS acceptable</td>
</tr>
<tr>
<td>12</td>
<td>IF effectiveness IS medium AND efficiency IS low AND satisfaction IS high THEN usability_level IS acceptable</td>
</tr>
<tr>
<td>15</td>
<td>IF effectiveness IS medium AND efficiency IS medium AND satisfaction IS high THEN usability_level IS good</td>
</tr>
<tr>
<td>17</td>
<td>IF effectiveness IS medium AND efficiency IS high AND satisfaction IS medium THEN usability_level IS good</td>
</tr>
<tr>
<td>21</td>
<td>IF effectiveness IS high AND efficiency IS low AND satisfaction IS high THEN usability_level IS good</td>
</tr>
<tr>
<td>24</td>
<td>IF effectiveness IS high AND efficiency IS medium AND satisfaction IS high THEN usability_level IS good</td>
</tr>
<tr>
<td>26</td>
<td>IF effectiveness IS high AND efficiency IS high AND satisfaction IS medium THEN usability_level IS excellent</td>
</tr>
<tr>
<td>27</td>
<td>IF effectiveness IS high AND efficiency IS high AND satisfaction IS high THEN usability_level IS excellent</td>
</tr>
</tbody>
</table>

Figure 4: Usability test developed (Own Elaboration)

3, 4, 4, 4, 2, and 4.

Figure 5: Test configuration file (Own Elaboration)

From a configuration file as the one presented in Figure 5, the proposed fuzzy system loads the test data and performs the calculation of the effectiveness and efficiency per task, as well as the satisfaction obtained from the evaluations made by the users about the 6 questions of the case study. Thus, Figure 6 shows the interface of the proposed fuzzy system, which has been loaded with the data from the configuration file shown in Figure 5. The graphical interface of the fuzzy system is composed of 5 tabs: Effectiveness, Efficiency, Satisfaction, Fuzzy Analysis, and Graphic Analysis.

In the "Effectiveness" tab, the user effectiveness per task is calculated from the configuration file of the test performed in the Sigma Electrónica website. In this way, it is observed in Figure 6 how the effectiveness of users 2, 3, and 5 is 100% in all tasks. Likewise, for the case of user 1 (U1), the
effectiveness for each task is 100%, 66.67%, and 100% respectively while for the case of user 4 (U4), the effectiveness for each task is 100%, 100%, and 50% respectively. Similarly, Figure 7 shows the 'Efficiency' tab, where the user efficiency per task is calculated from the configuration file of the case study test.

It can be seen how user 1 (U1) performed all the tasks of the user test in a time greater than that defined by the coordinators (the efficiency is less than 100%) while in the case of user 2 (U2), the time used to perform each of the tasks is less than what was defined by the test coordinators (the efficiency is greater than 100%). Regarding users 4 (U4) and 5 (U5), the time used is greater than what it was defined by the coordinators (the efficiency is less than 100%) in 2 of the 3 tasks while in the remaining task the time is less than what it was defined by the coordinators (the efficiency is greater than 100%). Concerning user 3, the time used is less than what it was defined by the test coordinators (the efficiency is greater than 100%) in 2 of the three tasks while in the remaining task, the time is greater than what it was defined by the coordinators (efficiency less than 100%). On the other hand, Figure 8 shows the 'Satisfaction' tab where the ratings assigned by the users of the case study test are loaded from the configuration file to each of the 6 questions related to the perception, which are done by the users after the development of the tasks.

In the 'Fuzzy Analysis' tab, the system calculates and presents the consolidated effectiveness,
efficiency, and satisfaction for the test of the case study taking into account the data loaded and presented in the first three tabs. Likewise, the system calculates and presents the level of output usability based on the degrees of belonging to the membership functions of the input variables (effectiveness, efficiency, and satisfaction) and the inference rules defined for each user and for the entire test. Similarly, the system graphically displays the membership role of the output usability level along with the resulting estimated level for each user and for the entire test. In addition, the system presents the membership levels of the three input variables, as well as the inference rules that are activated from the input values and the weighting provided to each of the rules that are activated in the lower text area. Thus, from the data of the case study shown in Figure 5, Figure 9 shows how the level of output usability for 4 of the 5 users (U2, U3, U4, and U5) is "excellent" while for the case of user 1 (U1) the output usability level is 61.92 and has been classified according to the membership functions as "acceptable". Likewise, it can be seen how for user 1 (U1), 2 inference rules are activated (rule 22 and 25), which correspond to the "acceptable" and "good" levels with a membership degree of 0.268 and 0.038 respectively. Therefore, the defined usability level of user 1 was classified as "acceptable". Finally, it can be seen how from the membership functions, the usability level obtained for the entire test of the case study is 93.04, which corresponds to the "excellent" classification with a degree of belonging of 0.6517.
Lastly, the ‘Graphic Analysis’ tab presents the membership functions of the inputs (effectiveness, efficiency, and satisfaction) and the output usability level, as well as the degree of membership estimated in each function (black vertical line on each membership function) from the values of these variables for each user of the usability test of the case study and for the whole test. As an example, based on the test data presented in Figure 5, Figure 10 shows the membership functions for the inputs and outputs of the fuzzy system corresponding to user 1. Thus, the level obtained for effectiveness is "high", the level obtained for efficiency is "medium", and the level obtained for satisfaction is "medium". Likewise, an "acceptable" output usability level is obtained through the defined inference rules; although, the rule associated with the "good" level is also activated as previously presented, but with a lower degree of membership. In the same way, in the "Graphical Analysis" tab, it is possible to obtain the level of output usability for the entire test in a numerical and graphical way from the average input values of effectiveness, efficiency, and satisfaction.

![Fuzzy logic-based system for usability level estimation in user tests](image)

**Figure 10:** Tab for the “Graphic Analysis” of the system (Own Elaboration)

## 4 Conclusions and future Work

In this study, the design and construction of a fuzzy system for estimating the output level of usability in tests with users was proposed based on the input values associated with the attributes of the ISO 9241-11 standard (effectiveness, efficiency, and satisfaction) and considering a set of inference rules defined by experts in the development of usability tests. The proposed fuzzy system aims to contribute to the problem of estimating the level of usability comprehending the different scales obtained in the three attributes, as well as the need to obtain not only a result in numerical but also linguistic terms.

The system proposed in this article allows, prior to the application of fuzzy logic, the calculation of effectiveness and efficiency per task, as well as the satisfaction evaluated by users in the post-test questionnaire. For the above, the data of a usability test is loaded into the system from a configuration file, in which the number of users, the number of tasks, the number of subtasks for each task, and the estimated time for the development of each task are specified. Likewise, the file contains the number of subtasks performed by the user regarding each task and the times used by each user to perform the tasks in the test. Finally, the file also includes the evaluations made by the users to the questions of the post-test questionnaire related to the perception of the users regarding the software evaluated in the usability test.

The FCL language and Java’s jFuzzyLogic API proved to be suitable for the definition of the membership functions of the inputs and outputs. It also proved right for the design of the inference rules and the implementation of the functional modules of the fuzzy system. In this sense, the data from the hypothetical test uploaded to the fuzzy system allowed to demonstrate its usefulness and
relevance in order to obtain the level of usability per user, as well as the level of usability throughout the test.

The application of the fuzzy system proposed to the data and/or that were obtained in the user test developed on the Sigma Electrónica website allowed to demonstrate the usefulness of the tool in terms of estimating the level of output usability for the five participating users and for the consolidation of the test. In this sense, it was obtained that for 4 of the 5 users the estimated output usability level is 'excellent' while the estimated output usability level is 'acceptable' for the remaining user. Based on the above, the usability level for the entire test of the case study is 93.04, which corresponds to the 'excellent' classification with a degree of membership of 0.6517 at that level.

Finally, the adaptation of the proposed system to tests with users supported by other quality standards, such as the ISO 250010 standard, is intended as future work derived from this research.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

References


[27] [Online]. Available: https://www.sigmaelectronica.net/


