UDC 004.055

Guchenko I.V., c.t.s.

APPLYING FUZZY MATHEMATICS TO SOFTWARE USABILITY EVALUATION MODEL

Institute of computer information technology
National Aviation University

The article is devoted to the issue of modification of earlier developed usability evaluation model with fuzzy multi-criteria weighted average approach. Usability is represented as a four-level hierarchical system of usability criteria. Usability subcharacteristics from ISO/IEC 25010:2011 were taken. The model supports usability management of software products based on the experts' judgments and evaluation of users' feedback. Fuzzy mathematics is applied to deal with the uncertainty and imprecision of the importance and rating of criteria on which usability depends

Introduction

Usability is an important characteristic of any product that is used by a customer. This applies to software. On the one hand, usability determines the consumption of resources (e.g. time and effort) against the accuracy and completeness of achieving goals. That affects the productivity of man and it is essential for software users. On the other hand, usability is one of the prominent factors in selecting software. It is related to its competitiveness and this is important for the software developer. Also usability is one of the characteristics of the software quality model and therefore is represented in the standards ISO / IEC 25010:2011, ISO 9241-11 and ISO / IEC 25060:2010.

Literature analysis

Many foreign scientists study usability: N. Bevan, B. Boehm, A. Cooper, Sh. Laskowski, J. McCall, J. Nielsen, A. Holzinger, J. Scholtz. Among ukrainian scientists it is explored indirectly through the concept of software quality by F. Andon, G. Koval', B. Konorev, T. Korotun, K. Lavryscheva, V. Suslov, I. Turkin, A. Kharchenko etc.

In many studies attempts to determine the usability are made, but often they are inconsistent [1]. Therefore, we will use the definition given in the standards ISO 9241-11 [2] and ISO / IEC 25010:2011 (updated ISO / IEC 9126-1:2001) [3]:

Usability – degree to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

In ISO / IEC 25010:2011 [3], which belongs to a series of standards SQuaRE (ISO / IEC 25000 - ISO / IEC 25099), usability is considered in two models: directly - in the product quality model; indirectly - in quality in use model.

According to the first model usability has six subcharacteristics: appropriateness recognisability, learnability, operability, user error protection, user interface aesthetics and accessibility.

Issues of usability subcharacteristics and properties definitions for specific software have been studied for over thirty years by various authors. Based on the standards, they are trying to expand the list and adapt it to real projects for usability measurement.

Usability, as noted above, is represented as a set of related subcharacteristics. They form the basis for the specification of usability requirements and its evaluation. Sets of software properties correspond to subcharacteristics. These properties match measures [3]. The value of a measure is obtained as the result of applying a measurement function to measure elements (Fig. 1).

For data derived from measurements in a variety of evaluation methods are used. Let us give the following definition:

Usability evaluation - the process of determining the actual state of the software regarding the desired state from the perspective of the possibility of usage by specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use; the result of such process in numerical form.

Under the *software state* we understand the set of quantitatively expressed properties that determine software behavior. Usability evaluation area is generalized as a set of all forms of methods that can help with understanding of ways to improve usability of software, problems with use in all contexts or even terms of use.

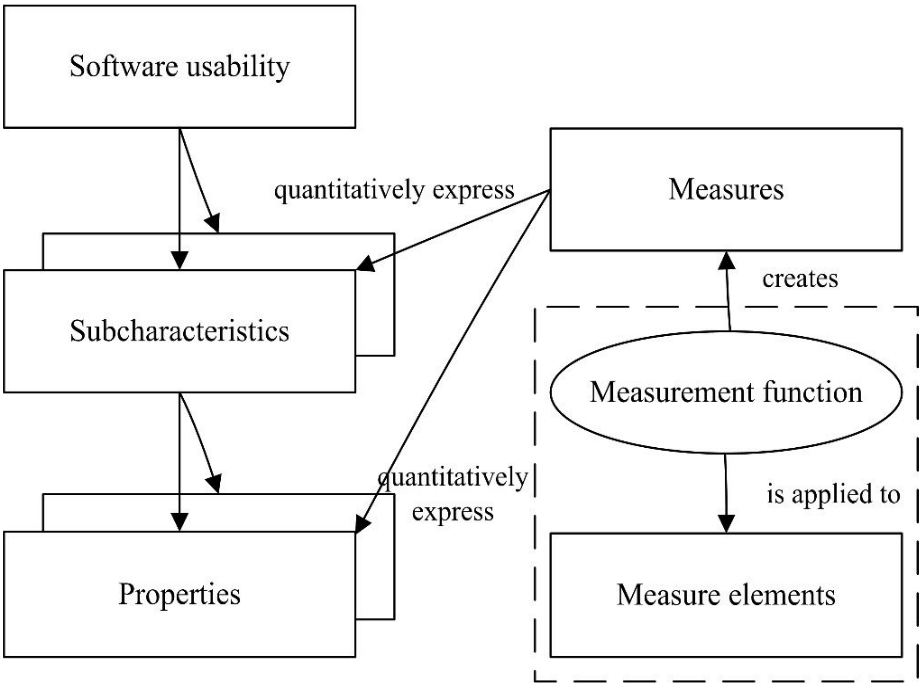

Figure 1. Usability measurement model

The history of approaches to usability assessment [4] includes about forty years. But most techniques are not well integrated and fail to cover all usability aspects. As it is shown in [5] only a few methods allow to obtain quantitative data. Moreover, those that do, cover not all usability subcharacteristics and depend on the expert’s or/and a user’s knowledge and experience. In such case, the issue of subjectivity of human judgments and vagueness of data arises. Last works on software quality (including usability) evaluation are devoted to fuzzy set theory usage [6-9]. Fuzzy logic helps to deal with the uncertainty and imprecision of the importance and rating of criteria on which usability depends.

Previous author’s works are devoted to the development of the method and the tool of software prod on the method of nested scalar convolutions and is represented by a scalar function of the additive convolution. It supports usability management of software products based on the automated evaluation of users’ feedback.

The aim of the present article is to apply fuzzy mathematics for multi-criteria usability evaluation that will make conducting of usability management according to the author’s method more precise.

The essentials of usability evaluation model

Usability (Q) is decomposed on subcharacteristics, each of which – on corresponding properties, and properties – on measures. Such decomposition is based on the relevant standards, guidelines, expert judgments etc. Each criteria of the higher level of the hierarchy contains the criteria of a lower level. It is allowed to introduce additional criteria at each level (Fig. 2). Criteria of each level of the hierarchy, except the first level, is characterized by two numerical parameters - numerical value and weight. The numerical values of usability properties are calculated on the basis of appropriate measures, the values of which are calculated as the average of user ratings.

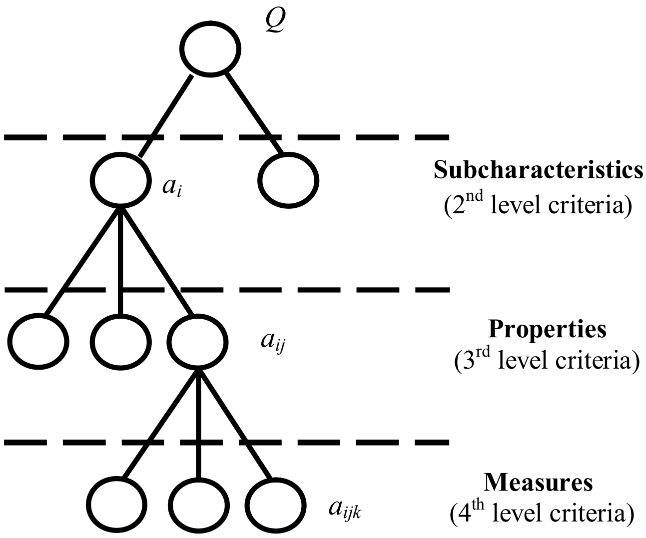

Figure 2. Graphic representation of the hierarchy of usability criteria

Usability evaluation model is represented in [10] by a scalar function of the additive convolution as follows:

$$\begin{aligned} Q(x) &= \sum_{i=1}^{n} p_i a_i(x) = \sum_{i=1}^{n} p_i \sum_{j=1}^{m} q_j a_{ij}(x) = \\ &= \sum_{i=1}^{n} p_i \sum_{j=1}^{m} q_j \sum_{k=1}^{l} r_k a_{ijk}(x), \end{aligned} \tag{1}$$

where $Q(x)$ – general criteria for user's ratings $x \in X$; $\{a_i(x)\}_1^n$, $\{a_{ij}(x)\}_1^m$, $\{a_{ijk}(x)\}_1^l$ – sets of criteria of corresponding hierarchical levels; *n, m, l* – number of criteria at levels; p_i, q_j, r_k - weights of a_i, a_{ij}, a_{ijk} criteria. For weights the normalization condition is true:

$$\sum_{i=1}^{n} p_i = \sum_{j=1}^{m} q_j = \sum_{k=1}^{l} r_k = 1. \tag{2}$$

A weight of each criteria is determined by the final experts' ranking using the Fishbern's scale. For usability criteria at all hierarchical levels a single scale from 0 to 1 is applied.

Sufficiency of the usability level is determined by comparing the general obtained score and calculated subcharacteristics' values with relevant analogues taken from a reference sample. Really existing software of the same functionality, with the same basic parameters, similar structure and operating conditions is chosen as analogue. Such an approach from a certain point of view is worse than modern neural network methods, but because of its simplicity, is more adapted to software development in terms of finance and time savings.

At the second level of the hierarchical structure, by which usability is represented (see Fig. 2), usability subcharacteristics from ISO / IEC 25010:2011 were applied [3]. List of properties was developed using QUIM model [11]. Interactions between subcharacteristics and properties are presented in Table 1. With regard to the measures that are calculated for each usability property, the corresponding list is presented in [10]. Measures are calculated using formulas or simple calculations.

Modification of usability evaluation model with fuzzy multi-criteria weighted average approach

In order to evaluate software usability triangular fuzzy sets are used to represent the corresponding criteria rating and weights. The algorithm of usability quantifying is as follows:

1. Specify fuzzification criteria for all usability criteria (subcharacteristics, properties and measures/measure elements).

2. Fuzzify ratings and weights obtained from users and experts into triangular fuzzy numbers (TFN). It is necessary to evaluate group TFN for expert's judgments (weights) of usability criteria (see fig. 1) and users' ratings of measure elements.

3. Evaluate the fuzzy moving average of the measures to evaluate the rating of the corresponding properties. Do similar calculations for obtaining rating of subcharacteristics and fuzzy rating for the usability.

4. Calculate a crisp usability value by the defuzzification process.

Table 1. Interactions between usability subcharacteristics and properties

Subcharacteristics	Properties					
	Operability	User error protection	User interface aesthetics	Learnability	Accessibility	Appropriateness recognisability
Time behavior	+					
Attractiveness			+			
Likeability			+			
Flexibility			+		+	
Minimal action			+	+	+	
Minimal memory load			+	+	+	
User guidance		+	+	+	+	
Consistency	+		+	+	+	
Self-descriptiveness				+	+	
Feedback	+					
Accuracy		+				
Fault-tolerance	+	+	+			
Readability					+	
Controllability	+		+		+	
Navigability					+	
Simplicity				+	+	
Familiarity				+		
Guide				+		+
Demonstrations				+		+
Help				+		+

Note that the values of all usability criteria and weights should be quantified in the range of 0 to 1 and thus normalized if it is necessary. The overall usability value will appear in the same range.

At the step 2 it is proposed to evaluate group TFN of the user's ratings $\tilde{x} = (U, M, L) \in X$ for measure elements, where X is a set of triangular fuzzy numbers, U, M and L denote the largest possible value, the most promising value and the smallest possible value that describes an event of fuzzy. There is only one measure element for each measure. Such group TFN defines a fuzzy rating for measure $a_{ijk}(\tilde{x})$ [7]:

$$a_{ijk}(\tilde{x}) = (U_{ijk}, M_{ijk}, L_{ijk}),\ U_{ijk} \geq M_{ijk} \geq L_{ijk},\ U_{ijk}, M_{ijk}, L_{ijk} \in [0,1]; \quad (3)$$

$U_{ijk} = \max\left(U_{(ijk)_s}\right)$; $L_{ijk} = \min\left(L_{(ijk)_s}\right)$;

$M_{ijk} = \sqrt[s]{\prod M_{(ijk)_s}}$, s – number of users.

Group TFN for expert's judgments (weights) of usability criteria can be obtained by applying the fuzzy weighted triangular averaging operator [8]:

$$\tilde{w} = \tilde{w}_1 \otimes e_1 \oplus \tilde{w}_2 \otimes e_2 \oplus \ldots \oplus \tilde{w}_t \otimes e_t, \quad (4)$$

where $\tilde{w}_1, \ldots, \tilde{w}_t$ – corresponding TFN criteria weights assigned by experts $E_1, \ldots, E_t$, respectively; $e_1, \ldots, e_t$ – weights allocated to expertsindicatesasis of their knowledge and/or experience, $e_1 + \ldots + e_t = 1$; **Ошибка! Закладка не определена.**and $\oplus$ indicates fuzzy multiplication and fuzzy addition operators, respectively, and defines as in [12, 13] for real fuzzy positive numbers.

Formula (1) according to step 3 of the algorithm can be modified:

$$\tilde{Q}(\tilde{x}) = \sum_{i=1}^{n} \tilde{w}_i a_i(\tilde{x}) = \sum_{i=1}^{n} \tilde{w}_i \sum_{j=1}^{m} \tilde{w}_j a_{ij}(\tilde{x}) = \\ = \sum_{i=1}^{n} \tilde{w}_i \sum_{j=1}^{m} \tilde{w}_j \sum_{k=1}^{l} \tilde{w}_k a_{ijk}(\tilde{x}), \quad (5)$$

where $\tilde{Q}(\tilde{x})$ – general criteria for user's fuzzy ratings $\tilde{x} = (U, M, L) \in X$, X is a set of triangular fuzzy numbers; $\{a_i(\tilde{x})\}_1^n$, $\{a_{ij}(\tilde{x})\}_1^m$, $\{a_{ijk}(\tilde{x})\}_1^l$ – sets of criteria of corresponding hierarchical levels; *n, m, l* – number of criteria at levels; $\tilde{w}_i$, $\tilde{w}_j$, $\tilde{w}_k$ - corresponding normal TFN for weights of a_i, a_{ij}, a_{ijk} criteria.

If fuzzy weights are not normal, than general formula cannot be applied in the form (5) and takes the form:

$$\tilde{Q}(\tilde{x}) = \frac{\sum_{i=1}^{n} \tilde{w}_i a_i(\tilde{x})}{\sum_{i=1}^{n} \tilde{w}_i}. \quad (6)$$

Many methods have been proposed in the literature for defuzzifying process mentioned in the step 4 of the algorithm. The most prevalent is the centroid method (also called center of area or center of gravity) [14]:

$$z^* = \frac{\int \mu(z) \cdot z dz}{\int \mu(z) dz}, \quad (7)$$

where z^* – the defuzzified crisp value of software usabofity, z is the value on the x axis, $\mu(z)$ – is the membership function.

As we use TFN the appropriate membership function is triangular:

$$\mu(z) = \begin{cases} 0, & z \leq a \text{ or } z \geq c \\ \dfrac{z-a}{b-a}, & a \leq z \leq b \\ \dfrac{c-z}{c-b}, & b \leq c \leq z \end{cases}, \quad (8)$$

where a, b, c – real numerical parameters and $a \leq b \leq c$.

Case study

For the fuzzy evaluation of represented usability model with the criteria given in table 1 and [10], web-browser Chromium 12 has been chosen. The survey inside groups of 10 users and 3 experts was conducted. Users were given a questionnaire with specified justification criteria for usability measure elements as it says in step 1 of the algorithm. Triangular fuzzy numbers (sets) for fuzzy ratings and weights were assigned (Table 2).

For example, there are 4 measures/measure elements for property "Minimal memory load": Visual coherence, Task visibility, Number of images, Number of unexplained acronyms [10]. The values of weights and ratings obtained from experts and users are given in tables 3 and 4.

Table 2. Values of fuzzification criteria and triangular fuzzy sets for fuzzy ratings and weights

Users' rating for measure elements (in points)	**Fuzzy linguistic therms for users' rating and experts' judgments**	**Fuzzy ratings**	**Fuzzy weights**
1-2	**VL** (very low)	(0.0, 0.1, 0.3)	(0.0, 0.0, 0.25)
3-4	**L** (low)	(0.1, 0.3, 0.5)	(0.0, 0.25, 0.5)
5-6	**M** (medium)	(0.5, 0.7, 0.9)	(0.25, 0.5, 0.75)
7-8	**H** (high)	(0.7, 0.9, 1.0)	(0.5, 0.75, 1.0)
9-10	**VH** (very high)	(0.9, 1.0, 1.0)	(0.75, 1.0, 1.0)

Table 3. Users' rating inputs for "Minimal memory load" measure elements

	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9	U_{10}	Group TFN (measure rating)
Visual coherence	H	H	H	M	M	L	H	H	H	H	(0.1, 0.76, 1.0)
Task visibility	M	L	M	M	H	L	M	M	L	M	(0.1, 0.56, 1.0)
Number of images	H	VH	H	H	M	M	H	M	VH	M	(0.5, 0.83, 1.0)
Number of unexplained acronyms	M	M	L	L	M	M	H	M	M	M	(0.1, 0.61, 1.0)

Table 4. Experts' weight inputs for "Minimal memory load" measures

	E_1 ($e_1 = 0.5$)	E_2 ($e_2 = 0.3$)	E_3 ($e_3 = 0.2$)	Group TFN (measure weight)
Visual coherence	L	L	L	(0.0, 0.25, 0.5)
Task visibility	M	M	L	(0.2, 0.45, 0.7)
Number of images	H	H	M	(0.45, 0.7, 0.95)
Number of unexplained acronyms	VH	H	H	(0.63, 0.88, 1.0)

Group TFN were calculated according to formulas (3) and (4). The fuzzy rating for the Minimal memory load property according to (5) and (6):

$$a_M = [(0.1, 0.76, 1.0)\times(0.0, 0.25, 0.5) + (0.1, 0.56, 1.0)\times(0.2, 0.45, 0.7) + (0.5, 0.83, 1.0)\times$$
$$\times(0.45, 0.7, 0.95) + (0.1, 0.61, 1.0)\times(0.63, 0.88, 1.0)] \,/\, [(0.0, 0.25, 0.5) + (0.2, 0.45, 0.7) +$$

$+ (0.45, 0.7, 0.95) + (0.63, 0.88, 1.0)] = (0.308, 1.56, 3.15) / (1.28, 2.28, 3.15) = = (0.241, 0.684, 1.0).$

Similarly, the ratings and the weights of other usability criteria were calculated. Finally fuzzy rating of usability in TFN form is:

$$\widetilde{Q}(\widetilde{x}) = (0.732, 0.85, 0.94).$$

Defuzzified usability value lies in the interval [0, 1] and was obtained with (7): $z^* = 0.84$. The respective membership function of usability is shown in figure 3, where $\mu_1(z) = 8.47z - 6.2$, $\mu_2(z) = 10.44 - 11.11z$ according to (8).

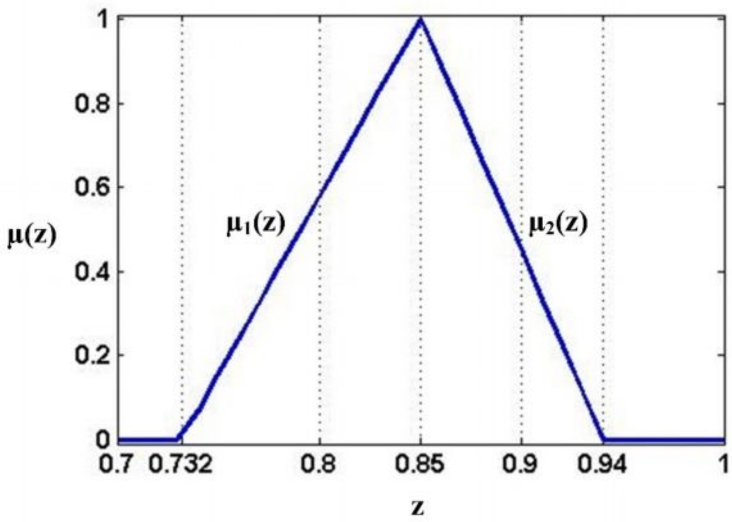

Figure 3. Fuzzy membership function for Usability

Conclusions

Usability is recognized as an important software quality attribute due to its social and technical aspects. Usability criteria are not easily measured and quantified. Most assessment techniques are not well integrated and fail to cover all usability aspects. Only a few methods allow to obtain quantitative data. Moreover, those that do, cover not all usability subcharacteristics. In the article, the four-level hierarchical usability model is represented using subcharacteristics stated in ISO/IEC 25010. Due to the unpredictable nature of usability criteria and subjectivity of experts' and users' judgments the fuzzy multi-criteria weighted average approach has been used to evaluate usability. Using modified usability evaluation model described in the article, usability of software with the same functionality, basic parameters, similar structure and operating conditions may be evaluated. It helps to identify more suitable software product for a defined set of users in certain environment.

References

1. Dubey S. K. Analytical Roadmap to Usability Definitions and Decompositions / Sanjay Kumar Dubey, Ajay Rana // International Journal of Engineering Science and Technology. – 2010. – № 2(9). – P. 4723-4729.
2. Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs), Part 11: Guidance on Usability: ISO 9241-11. – Geneva: International Organization for Standardization, 1998. – 22p.
3. Systems and software engineering, Systems and software Quality Requirements and Evaluation (SQuaRE), System and software quality models: ISO/IEC 25010:2011. – Geneva: International Organization for Standardization /International Electrotechnical Commission, 2011. – 34p.
4. Scholtz J. Usability evaluation / J. Scholtz // Encyclopedia of Human-Computer Interaction. – Great Barrington, MA: Berkshire Publishing Group, 2004.

5. Gulati A. Critical analysis on usability evaluation techniques / A. Gulati, S.K. Dubey // International Journal of Engineering Science and Technology. – 2012. – Vol. 4. – P. 990-997.

6. Dubey S.K. Usability evaluation of software systems using fuzzy multi-criteria approach / S.K. Dubey, A. Gulati, A. Rana // International Journal of Computer Science Issues. – 2012. – Vol. 9. – P. 404-409.

7. Dubey S.K. Measurement of object- oriented software usability using fuzzy AHP / S.K. Dubey, A. Mittal, A. Rana // International Journal of Computer Science and Telecommunications. – 2012. – Vol. 3. – P. 98-104.

8. Aydin S. Evaluation of E-commerce website quality using fuzzy multi-criteria decision making approach / S.Aydin, C. Kahraman // International Journal of Computer Science. – 2012. – Vol. 39.

9. Integrated software quality evaluation: a fuzzy multi-criteria approach / J.S. Challa, A. Paul, Y. Dada [et al.] // Journal of Information Processing Systems. – 2012. – Vol. 7. – P. 473-518.

10. Гученко І.В. Метод і засіб управління зручністю використання програмних продуктів: дис. ... кандидата технічних наук : 01.05.03 / Гученко Інна Володимирівна. – К., 2012. – 124 с.

11. Padda Harkirat K. QUIM: A Model for Usability/Quality in use Measurement / Harkirat K. Padda. – Colne: Lambert Academic Publishing, 2010. – 124 p.

12. Штовба С.Д. Проектирование нечетких систем средствами MATLAB / Штовба С.Д. – М.: Горячая линия – Телеком, 2007. – 288 с., ил.

13. Kaufmann A. Introduction to fuzzy arithmetic theory and applications / A. Kaufmann, M. Gupta. – N.Y.: Van Nostrand Reinhold, 1991. – 361p.

14. Ross T.J. Fuzzy logic with engineering applications / T.J. Ross. – [3rd ed/]. – Singapore: Fabulous Printers, 2010. – 607p.

Article to the Editor 29.09.2013