

# Formalization model of expert knowledge about a technical index level of engineering products

A A Zakharova<sup>1</sup>, V V Ostanin<sup>1</sup>

<sup>1</sup> Yurga Institute of Technology (affiliate) of Tomsk Polytechnic University, Leningradskaya Str., 26, Yurga, the Kemerovo region, Russia

E-mail: aaz@tpu.ru

**Abstract.** The authors set a timely problem that concerns development of decision making models, which allow formalizing expert subjective ideas about technical index level of engineering products. The authors proposed a formalization model of expert knowledge about technical index level of engineering products on the basis of fuzzy sets. The model has a method of membership-function construction for linguistic variable terms on the basis of exponential functions.

## 1. Introduction

The first paragraph after a heading is not indented (Bodytext style). Manufacture of innovative products is connected with many science-based engineering solutions, some of which could not be provided with a pure mathematical tool. That is why some of technological, constructive, organizational, and other decisions when people create and produce engineering products need experts in every branch. To support a decision making process it is necessary to form a method and model complex for decision support, which can process expert estimations and knowledge [1].

One of the problems, which experts have to solve when estimating a level and marketability of complex science-based units, is estimation of some technical indexes, quality merits and so on. During the process experts have to use qualitative (word) evaluations, for example, “high lifting capacity”, “weakness”, etc. [2]. That is why a timely problem is development of methods and models for decision making, which allow formalizing qualitative and quantitative evaluations, formalizing subjective expert ideas about any engineering product characteristic. It is proposed to use methods of a fuzzy set theory, which helps model smooth change of an object, and also unknown functional relationships presented as qualitative connections.

## 2. Formalization model of expert knowledge about a technical index level of engineering products

When scientists describe a decision making process in its hard formalized stages they take into account the following:

- a decision making process is characterized by some input parameters and one output parameter;
- some information given by experts about decision making strategies in standard situations is described by a set of conditional statements in terms of fuzzy and linguistic variables that connect input and output variables [3].



For example, if an expert arranges statements about a value of an output variable depending on values of two input variables, the system of standard fuzzy statements can be as follows (1):

$$\tilde{L}^{(1)} = \begin{cases} \tilde{L}_1^{(1)} : < \text{IF } E_{11} \text{ OR } E_{12} \text{ OR } E_{13} \text{ THEN } \beta_V \text{ is } a_{V_1} >; \\ \tilde{L}_2^{(1)} : < \text{IF } E_{21} \text{ OR } E_{22} \text{ OR } E_{23} \text{ THEN } \beta_V \text{ is } a_{V_2} >; \\ \tilde{L}_3^{(1)} : < \text{IF } E_{31} \text{ OR } E_{32} \text{ OR } E_{33} \text{ THEN } \beta_V \text{ is } a_{V_3} >. \end{cases} \quad (1)$$

Statements  $E_{ji}$  are statements of the type:  $< \beta_X \text{ is } a_{X_i} \text{ AND } \beta_Y \text{ is } a_{Y_k} >$ .

The statement  $E_{ji}$  represents itself an input fuzzy situation  $i$ , which can be if the linguistic variable  $\beta_V$  has a value of  $a_{V_j}$ .

For example,  $E_{11} : < \beta_X \text{ is } a_{X_1} \text{ AND } \beta_Y \text{ is } a_{Y_1} >$ .

$X, Y$  are the range of input parameters (technical data of products) that tangibly influences the output parameter  $V$ .

$< \beta_X, T_X, X >, < \beta_Y, T_Y, Y >, < \beta_V, T_V, V >$  are linguistic variables that are defined on the range  $X, Y$  and  $V$ .

A fuzzy variable is a variable of the type:

$< \alpha, X, C_\alpha >$ ,

where  $\alpha$  is a name for a fuzzy variable;

$X = \{x\}$  is a range of definition;

$C_\alpha = \{\mu_\alpha(x)/x\}$  is a fuzzy set on  $X$  that describes restriction to possible values of the fuzzy variable  $\alpha$  (its semantics). [4].

A linguistic variable is a variable of the type:  $< \beta, T, X >$  where:

$\beta$  is a name of a linguistic variable;

$T$  is its range (term-set) that represent fuzzy variable names;

$X$  is a range of definition for fuzzy variables [5].

Linguistic variables play important role in fuzzy model designing. With their help it is possible to formalize qualitative information about a decision making object, which is verbalized by specialists-experts [6].

Therefore the most important stage in a decision making process on the basis of fuzzy methods is the selection of methods for membership function construction of linguistic variables. Some indexes of engineering products should be presented as linguistic variables. It provides comparison of qualitative and quantitative expert evaluations. Depending on a factor type one applies different methods for membership function construction: for example, on the basis of pair-wise comparison, statistical data processing, standard functions, and so on [7, 8, 9]. This article considers the method based on standard functions.

During the process of expert estimation formalizing about the level of some technical indexes, in the majority of cases it is not necessary to have a model of high accuracy, rough description is enough. To construct membership functions of such notions it is possible to use direct methods based on direct expert assigning of a membership or function degree, which allows calculating a value.

To describe a membership function it is proposed to use a normal distribution frequency function of a continuous random quantity (Gaussian curve):

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/2\sigma^2}, \quad (2)$$

where  $a$  is an expected value of a random variable;

$\sigma$  is a mean-square deviation.

Membership function  $\mu_x$  takes on values in the interval  $[0, 1]$ , but conditions for terms description of linguistic variables suppose to use only normal fuzzy sets that describe basic values of linguistic variables. That is why it is necessary to normalize the frequency function of normal distribution. To do this it is necessary to divide it by a maximum value. The frequency function of normal distribution has its maximum value when  $x=a$ , here.

$$y = \frac{1}{\sigma\sqrt{2\pi}} \quad (3)$$

Consequently:

$$\mu_x = \exp\left[\frac{-(x-a)^2}{2\sigma^2}\right] \quad (3)$$

The parameter  $a$  is such  $x$  value as completely corresponds with a term described notion according to an expert point of view.

The parameter  $\sigma$  (i.e. the multiplier factor is  $\alpha = -\frac{1}{2\sigma^2}$ ) characterizes largeness of a definition range for a membership function or a fuzzy degree  $\mu_x$ .

Exponential functions are widely used for membership function description. Let us consider a possibility to specify function parameters (3) by an expert way.

The parameter  $a$  is a dominant element of a fuzzy set, the membership function is  $\mu_x = 1$ . An expert should select from a definition range of a linguistic variable the value that is “perfect” for describing a necessary notion. The parameter  $a$  is specified by an expert for every basis value of a linguistic variable.

Every value from a definition range of a linguistic variable should go with at least one notion (basis value of a linguistic variable), that is why membership functions of fuzzy variables, which describe neighbor basis values of a linguistic variable, are supposed to cross over. So the expert can set such a value  $x$ , when membership functions of neighbor term sets have the same values. It means that the expert set the value with which according to his opinion it is difficult to say what a neighbor value of a linguistic variable it refers to. The expert can also identify a membership degree of the value  $x$  to fuzzy sets of a neighbor term sets.

Now it is necessary to specify the parameter  $\sigma$ , that cannot be directly set by the expert, because a person can hardly imagine measure of characteristic dispersion related to its average value. The parameter  $\sigma$  can be expressed in terms of the formula (3):

$$2\sigma^2 = \frac{(x-a)^2}{-\ln \mu_x}; \quad (4)$$

where  $x$  is the value  $x \in X$ , when membership functions of neighbor term sets have the same values;

$\mu_x$  is a membership degree of the value  $x \in X$  to fuzzy sets of a neighbor term sets.

So, all function parameters (3) can be specified either by an expert or on the basis of his information. For terms construction of a linguistic variable the expert should set  $n$  values  $a_i$ ,  $(n-1)$  of the values  $x_{k_j}$  and  $\mu_{k_j}$ .

As membership functions have to have a final range of definition and exponential functions are infinite, we take up the range of  $\alpha$ -level when  $\alpha=0.05$  as a range of definition for fuzzy variables, which define basis values of a linguistic variable.

Membership functions of linguistic variable terms are set by the following functions (5):

$$\mu_{x_1} = \begin{cases} 1 & \text{when } x \leq a_1 \\ \exp\left[\frac{-(x-a_1)^2}{2\sigma_{11}^2}\right] & \text{when } x > a_1; \\ 0 & \text{when } x \geq x_{11} \end{cases}$$

$$\mu_{x_2} = \begin{cases} \exp\left[\frac{-(x-a_2)^2}{2\sigma_{21}^2}\right] & \text{when } x \leq a_2 \\ \exp\left[\frac{-(x-a_2)^2}{2\sigma_{22}^2}\right] & \text{when } x > a_2; \\ 0 & \text{when } x_{21} \geq x \geq x_{22}; \end{cases}$$

$$\mu_{x_n} = \begin{cases} 0 & \text{when } x \leq x_{nn-1}; \\ \exp\left[\frac{-(x-a_n)^2}{2\sigma_{nn-1}^2}\right] & \text{when } x \leq a_n; \\ 1 & \text{when } x > a_n \end{cases} \quad (5)$$

where:  $\mu_{x_i}$  is a membership function of  $i$ -term of a linguistic variable;

$i = \overline{1, n}$  is a number of a linguistic variable terms, term numeration is from left to right;

$a_i$  is a dominant element of  $i$ -term fuzzy set.

$$2\sigma_{ij}^2 = \frac{(x_{k_j} - a_i)^2}{-\ln \mu_{k_j}}; \quad (6)$$

where  $x_{k_j}$  is a value  $x \in X$ , when membership functions of neighbor terms have the same values;

$j = \overline{1, n-1}$  is a number of the value  $x_{k_j}$ , numeration is from left to right;

$\mu_{k_j}$  is a membership degree of the value  $x_{k_j}$  to fuzzy sets of neighbor terms (separation value);

$x_{nn-1}$  is  $x$  values that define a definition range of fuzzy variables.

This method simplifies the procedure of membership function construction, and function memorization by the computer, when it is provided meeting requirements for membership functions of linguistic variable terms [3]. The expert can also change membership functions, and a definition range of a linguistic variable. One more advantage is that a membership function is defined on a continuous carrier; it allows calculating its value at any variable value.

### 3. Examples of expert knowledge formalization about technological index level of engineering products on the basis of a fuzzy model

We construct a term sets for a linguistic variable  $\beta_Y$  – item technical complexity (unit of measurement is a ball) with a range of definition  $Y = [0;100]$  and basis range  $T_Y = \{\text{low, medium, high}\} = \{a_{Y_1}, a_{Y_2}, a_{Y_3}\}$ .

Expert estimations of parameters needed for membership function construction of  $\beta_Y$  variable are presented in table 1.

**Table 1.** Expert estimations of parameters for a linguistic variable  $\beta_Y$  – technical complexity of an item

Basis values	Dominant value $y$ of the fuzzy set that describes a term, $a_i$	Breakpoint values of neighbor terms $y_{k_j}$	Membership degree of neighbor values (separation value), $\mu_{k_j}$
$a_{Y_1}$ - low	0		
$a_{Y_2}$ - medium	50	$y_{k_1} = 25$	$\mu_{k_1} = 0.5$
$a_{Y_3}$ - high	100	$y_{k_2} = 75$	$\mu_{k_2} = 0.5$

We process the expert data.

1. We calculate the value  $2\sigma_{ij}^2 = \frac{(y_{k_j} - a_i)^2}{-\ln \mu_{k_j}}$  according to the formula (6).

$$2\sigma_{11}^2 = \frac{(y_{k_1} - a_1)^2}{-\ln \mu_{k_1}} = \frac{(25 - 0)^2}{-\ln 0.5} = 901.75; \quad 2\sigma_{21}^2 = \frac{(y_{k_1} - a_2)^2}{-\ln \mu_{k_1}} = \frac{(25 - 50)^2}{-\ln 0.5} = 901.75;$$

$$2\sigma_{22}^2 = \frac{(y_{k_2} - a_2)^2}{-\ln \mu_{k_2}} = \frac{(75 - 50)^2}{-\ln 0.5} = 901.75; \quad 2\sigma_{32}^2 = \frac{(y_{k_2} - a_3)^2}{-\ln \mu_{k_2}} = \frac{(75 - 100)^2}{-\ln 0.5} = 901.75.$$

2. We calculate the values  $y_{ik}$ , when  $\mu_{a_{y_i}} = 0.05$  according to the formula:

$$y_{11} = a_1 + \left| \sqrt{-2\sigma_{11}^2 \ln 0.05} \right|; \quad y_{21} = a_2 - \left| \sqrt{-2\sigma_{21}^2 \ln 0.05} \right|; \quad (7)$$

$$y_{22} = a_2 + \left| \sqrt{-2\sigma_{22}^2 \ln 0.05} \right|; \quad y_{32} = a_3 - \left| \sqrt{-2\sigma_{32}^2 \ln 0.05} \right|.$$

$$y_{11}=51.98; \quad y_{21}=-1.98; \quad y_{22}=101.98; \quad y_{32}=48.02.$$

As  $y_{21} = -1.98$ ;  $y_{22} = 101.98$  leave  $Y$  range of definition, so we take up  $y_{21} = 0$ ;  $y_{22} = 100$ .

3. We define membership functions according to the formulae (5) taking into account the limits (7):

$$\mu_{a_{y_1}} = \begin{cases} 1 & \text{when } y \leq 0; \\ \exp[-(y)^2 / 901.75] & \text{when } 0 < y < 51.98; \\ 0 & \text{when } y \geq 51.98; \end{cases}$$

$$\mu_{a_{y_2}} = \begin{cases} 0 & \text{when } 0 \geq y \geq 100; \\ \exp[-(y-50)^2 / 901.75] & \text{when } 0 < y < 100; \end{cases}; \tag{8}$$

$$\mu_{a_{y_3}} = \begin{cases} 0 & \text{when } y \leq 48.02; \\ \exp[-(y-100)^2 / 901.75] & \text{when } 48.02 < y < 100; \\ 1 & \text{when } y \geq 100. \end{cases}$$

Figure 1 presents graphs of linguistic variable terms “Item technical complexity”.

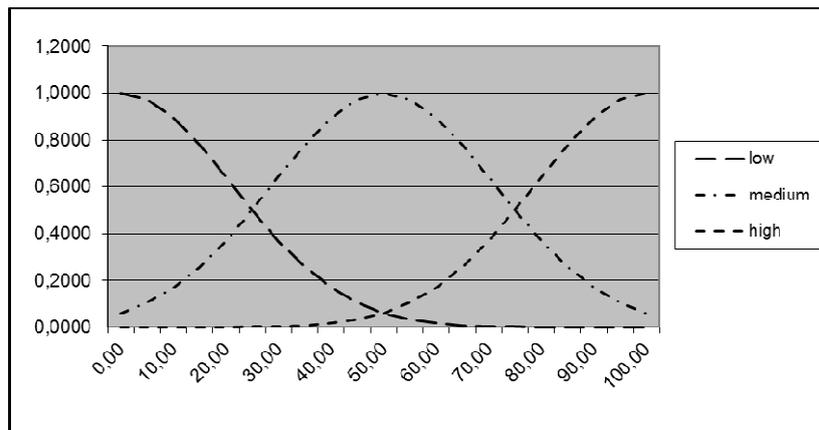


Figure 1. Graphs of linguistic variable terms “Item technical complexity”

In this example dominant values of linguistic variable terms are equally arranged in a defined definition range.

This method allows obtaining membership functions even when term intervals are unequally arranged.

For example, when formalizing expert knowledge for the linguistic variable  $\beta_Y$  – loading capacity (tons) with the range  $Y = [0;50]$  and the basis range  $T_Y = \{low, medium, high\} = \{a_{Y_1}, a_{Y_2}, a_{Y_3}\}$ , we got expert estimations of parameters, which are presented in Table 2.

**Table 2.** Expert estimations of parameters for the linguistic variable  $\beta_Y$  – loading capacity

Basis values	Dominant value $y$ of the fuzzy set that describes a term, $a_i$	Breakpoint values of neighbor terms $y_{k_j}$	Membership degree of neighbor values (separation value), $\mu_{k_j}$
$a_{Y_1}$ - low	0		
$a_{Y_2}$ - medium	20	$y_{k_1} = 10$	$\mu_{k_1} = 0.5$
$a_{Y_3}$ - high	50	$y_{k_2} = 35$	$\mu_{k_2} = 0.5$

We process the expert data.

1. We calculate the values  $2\sigma_{ij}^2 = \frac{(y_{k_j} - a_i)^2}{-\ln \mu_{k_j}}$ .

$$2\sigma_{11}^2 = \frac{(y_{k_1} - a_1)^2}{-\ln \mu_{k_1}} = \frac{(10-0)^2}{-\ln 0.5} = 144.16; \quad 2\sigma_{21}^2 = \frac{(y_{k_1} - a_2)^2}{-\ln \mu_{k_1}} = \frac{(10-20)^2}{-\ln 0.5} = 144.16;$$

$$2\sigma_{22}^2 = \frac{(y_{k_2} - a_2)^2}{-\ln \mu_{k_2}} = \frac{(35-20)^2}{-\ln 0.5} = 324.1; \quad 2\sigma_{32}^2 = \frac{(y_{k_2} - a_3)^2}{-\ln \mu_{k_2}} = \frac{(35-50)^2}{-\ln 0.5} = 324.1.$$

2. We calculate the values  $y_{ik}$ , when  $\mu_{a_{y_i}} = 0.05$  according to the formula (7):

$$y_{11}=20.78; \quad y_{21} = -0.78; \quad y_{22}=51.16; \quad y_{32}=18.84.$$

As  $y_{21} = -0.78$ ;  $y_{22}=51.16$  leave  $Y$  range of definition, so we take up  $y_{21} = 0$ ;  $y_{22}=50$ .

3. We define membership functions according to the formulae (5):

$$\mu_{a_{y_1}} = \begin{cases} 1 & \text{when } y \leq 0; \\ \exp[-(y)^2 / 144.16] & \text{when } 0 < y < 20.78; \\ 0 & \text{when } y \geq 20.78; \end{cases}$$

$$\mu_{a_{y_2}} = \begin{cases} 0 & \text{when } 0 \geq y \geq 50; \\ \exp[-(y-20)^2 / 144.16] & \text{when } 0 < y < 20; \\ \exp[-(y-20)^2 / 324.1] & \text{when } 20 \leq y < 50; \end{cases} \quad (9)$$

$$\mu_{a_{y_3}} = \begin{cases} 0 & \text{when } y \leq 18.84; \\ \exp[-(y-50)^2 / 324.1] & \text{when } 18.84 < y < 50; \\ 1 & \text{when } y \geq 50. \end{cases}$$

On the basis of the proposed model it is possible to determine values of input variables.

We can show a fuzzification process of three fuzzy statements: “low loading capacity”, “medium loading capacity”, and “high loading capacity” for the input linguistic variable “loading capacity” when we need to estimate the index “loading capacity” for products that have the index value 30 (tons) out of the possible value interval [0,50]. According to the formulae (9) we calculate values:  $\mu_{a_{y_1}}(30) = 0$ ;  $\mu_{a_{y_2}}(30) = 0,735$ ;  $\mu_{a_{y_3}}(30) = 0,291$ .

Therefore, the most membership degree of the input variable value  $y=30$  corresponds to the second statement: medium loading capacity. Obtained crisp values of the true input value for a linguistic variable terms can be used at the next stage of a decision making process, for example, in models of expert knowledge formalization about dependence of an output product parameter from several input parameters, with application of different logical derivation methods (1).

#### 4. Conclusion

The proposed fuzzy model for formalization of expert knowledge about technical index level of engineering products has an inherent value, because with its help it is possible to model a smooth intensive membership change of definite index values to an estimated level. As the result we have a possibility to transform qualitative expert estimations into quantitative values and vice versa, and make fuzzification of fuzzy statements. Expert confidence in a preferable index value is also taken into account in contrast to the general practice when they set threshold values.

In the model we realized the method of membership function construction of linguistic variable terms on the basis of exponential functions, which allow the expert to easily set membership functions

with the help of a few number of parameters. The membership function is determined on a continuous carrier, what allows calculating its value at any variable values.

Linguistic variables obtained on the basis of the proposed fuzzy model help characterize input and output factors in decision making models, which formalize expert knowledge about a decision making strategy in standard situations.

## References

- [1] Tceplit A, Grigoreva A, Osipov Y 2014 Developing the model for assessing the competitiveness of innovative engineering products *Applied Mechanics and Materials* **682** 623-630
- [2] Zakharova A A 2013 Fuzzy swot analysis for selection of bankruptcy risk factors *Applied Mechanics and Materials* **379** 207-213
- [3] Malyshev N G, Berstein L S, Bozhenuk A V 1998 Fuzzy model of decision making in CAD system Int. Conf. *Fuzzy Sets in Informatics* (Moscow) p.44
- [4] Zadeh L A 1965 Fuzzy sets, *Inf. Control.* **8(3)** 338-353.
- [5] Zadeh L A 1978 Fuzzy sets as a basis for a theory of possibility *Fuzzy Sets and Systems* **1** 3-28
- [6] Telipenko E V, Zakharova A A 2014 Bankruptcy risk management of a machine builder *Applied Mechanics and Materials* **682** 617-622
- [7] Frantti T, Mähönen P 2001 Fuzzy logic-based forecasting model *Engineering Applications of Artificial Intelligence* **14(2)** 189-201.
- [8] Ivezić D, Tanasijević M, Ignjatović D 2008 Fuzzy approach to dependability performance evaluation *Quality and Reliability Engineering International* **24(7)** 779-792.
- [9] Yadav O P, N. Singh Chinnam R B, Goel P S 2003 A fuzzy logic based approach to reliability improvement estimation during product development *Reliability Engineering and System Safety.* **80(1)** 63-74