# Using Fuzzy Logic in Knowledge Tests

Aleksandr Alekseev [1, *], Marika Aleksieieva [2], Kateryna Lozova [1], Tetiana Nahorna [1]

[1] Sumy State University, Faculty of Technical Systems and Energy Efficient Technologies, Sumy, Ukraine
[2] Harvard University, Graduate School of Arts and Sciences, Cambridge, Massachusetts, USA
alekseev_an@ukr.net, maleksieieva@fas.harvard.edu, katarina_lozovaya@ex.ua, nagorna-t@mail.ru

**Abstract.** An article describes the specialties of nonlinear scale formation of coincidence of standard answer with student's answer, basing on application of fuzzy logic during the test control of knowledge. It is given the detailed exposition of the mathematical apparatus that is used for substantiation the decision-making on the formation of the coincidence scale of answers. The author notes that using of the coincidence scale of answers gives the student an opportunity to express doubt and specify any degree of true answer ranging from "False" to "True". In this case test results are measured in the opposite terms from clear to fuzzy logic when the final mark is determined by the match of the answers. For example, if reference answer is equal to student's one it means that he/she knows the materials, and vice versa if the reference answer does not match, student does not know the topic. There are some types of the test tasks in the testing with using the coincidence scale of answers. The article describes the peculiarities off the parameter assignment of strictness the fuzzy-logic system. The results of experimental verification of the proposed innovations' effectiveness are given. These results allow stating the improvement of the measurement capabilities off the test with using the coincidence scale of answers basing on application of fuzzy logical calculations.

**Keywords:** Pedagogical measures, Knowledge test control, Measuring scale, Fuzzy logic, Strictness parameter, Test questions.

**Key Terms:** ICT Tool, Quality Assurance Process, Teaching Methodology, Teaching Process, Technology

## 1 Introduction

Fast and accurate evaluation of knowledge formation remains is a relevant task for long-existing forms of learning. Moreover, it has become an increasingly important for the recently emerged distance learning or blended learning (partial implementation

---

* Professor of Department of Manufacturing Engineering, Machines and Tools, Doctor of Pedagogy., Associate Professor

of distance learning technologies into classes that are conducted traditionally). The most important characteristics of the different forms of learning remains objective monitoring of students' academic achievements and construction of effective teaching methods based on it.

Further development of the theory and practice of the test control gives significant prospects for achievement of such goals. Using the information and telecommunication technologies, the test control successfully completes and improves existing traditional forms and methods of knowledge control. Computerized testing carries out a number of pedagogical functions assigned to knowledge test control, and becomes an effective means of summarizing the results of learning at all stages of education, from an entrance test to a comprehensive final exam.

## 2 Preconditions for using the fuzzy logic in the scaling of students' answers

Educational measurement technologies are developing in the close cooperation with the achievements of pedagogy, psychology, sociology and other empirical sciences, which are characterized by using the quantitative and quality indicators, differ by levels of manifestation of properties that are not measured directly. Due to this, there is no exception in the development of scaling tools applied to interpret the student's responses in higher education institutions with a computerized test control of knowledge.

The problem connected to the need in making available for the respondents involved in the questionnaire, or student who participates in the test control of knowledge, scale transfer of their judgments about the object of evaluation in the quantitative description of the level of assessed property has been known for a long time. Currently, there are several solutions proposed by different authors.

One of the first solutions was proposed by L. Thurstone [9]. The procedure of constructing the L. Thurstone scale is to provide an opinion about the level of assessed property in the frame of a set of evenly distributed judgments. Text description of each judgment is assigned a value of the level bar graph properties which corresponds to an interval scale. The scale constructed in such way is an interval one and its usage gives the possibility to apply a sufficiently wide range of statistical methods of processing the measurements results. However, a large amount of preparatory work related to the construction of the interval scale, relative equality of intervals, limit the possibilities of its application for the evaluation of students' knowledge.

The scale, developed by R. Likert [5], suggests the existence of the alternative judgments that reflect extreme levels of the assessed properties. These judgments can be formulated as "strongly agree" through "uncertainty" to "strongly disagree" for the test control. In addition, R. Likert scale bar graph set intermediate values associated with the specific levels of the assessed properties. In the text description no more than three of these intermediate values are commonly used, for example, "Somewhat agree", "neither yes nor no," "Somewhat agree".

The R. Likert scale is an ordinal scale, and despite the fact that, usually, its construction does not require the time-consuming preliminary work in the practice of educational measurement finds limited application. This is due to the fact that within the ordinal scales we can only arrange objects in ascending or descending order estimates of measurable properties at the lowest possible statistical treatment of the results of evaluations. Besides, the accuracy of pedagogical measurements with the use of scales that were developed due to approaches of R. Likert including numerical grading based scales limited by number of intermediate values, the number of which usually does not exceed 10-15.

Despite significant progress, reached for the last years in the different field of knowledge in the developing sphere, systematization and the field of analysis the results of practical application the methods of scaling the properties of qualitative and quantitative indicators, we have to realize that new approaches have limited application for the interpretation the student's responses on the knowledge test control.

To improve the accuracy of the pedagogical measurement, including the empowerment of statistical processing of the measurement results, it is necessary, in our view, to use such benefits in scaling the student responses using the ratio scale (name, by the definition of S.S. Stevens [8]). The mathematical apparatus of fuzzy logic will help to do things mentioned above.

## 3 Basic Provisions

During a traditionally organized examination the roles of a teacher and a student are allocated in accordance with the objectives of the oral control, when the teacher asks questions in order to identify student's generated knowledge. The student comprehends the questions and gives the answers, based on his/her idea of correctness of an answer. Then the teacher makes a judgment based on the results of such statements about the success of student's answers to particular questions, also evaluates the knowledge of all studied materials considering the total number of answers. Herewith, evaluating student's knowledge the teacher generally takes into consideration not only the formal correctness of the answers, but also how they were given, and whether the student was sure about the answers vs. showed a sign of insecurity, which may indicate instable knowledge. At the same time the student may use interpersonal contacts and consciously or unconsciously formulate the answer in such a way that it would let the teacher trace the causes of the seemingly unsuccessful answer. The doubt expressed in the answer provides an experienced teacher with another information channel that allows proper evaluating of the actual level of student's knowledge.

During classically organized test control, unlike oral examination, alienation of teacher's individuality occurs. Due to this fact, it is impossible to apply diagnostic capabilities of the teacher during the control process in order to identify the actual knowledge of the students.

Test control usually requires performing a task by selecting one of the possible answers or giving an unequivocal answer formulated from a limited set of words, letters, numbers, or graphics. In any case, the student should use his/her own experience and make such an answer, which would contain the conclusion of true judgment in terms of strict logic. However, it is not possible to express doubt or indicate how the answer may differ from the correct one.

Checking results of the written test control the teacher has only a report of a student containing no data about possible difficulties in formulating the response. Therefore, the final evaluation cannot indicate whether the student was sure in the answer, or just speculated it relying only on luck. Computerized test control is more formal and matches the reference answers with the student's ones.

Concerning that, the developed test control simulation model [1] proposes to perform computerized control of knowledge by using of an expert system (Fig. 1) based on fuzzy logic [4]. Application of this system gives to a student the opportunity to operate not only the classical values of logical variables like "false" and "true", but also to use their intermediate values fading from one extreme value ("false") to an opposite one ("true").

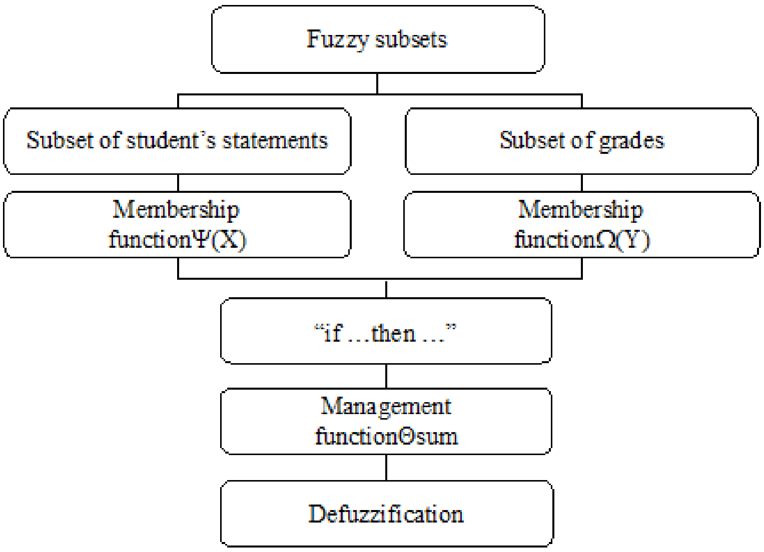


**Fig. 1.** Fuzzy logic expert system

The expert system uses piecewise continuous membership functions in order to define how evaluation of student's knowledge and his/her expressed statement relate to fuzzy logic subsets. These functions have transitional areas presented as segments a-b and b-c, connecting zero and one (maximum) levels of reliability (Fig. 2).

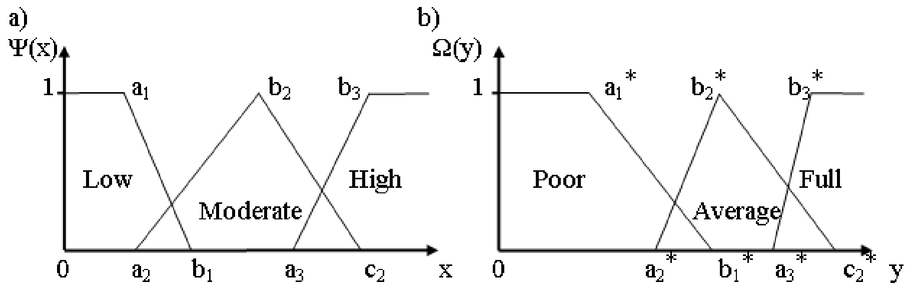


**Fig. 2.** Membership functions for subsets of student's statements (a) and evaluation of academic achievements (b)

A membership function for each term of the base term set of a logical variable "Level of matching answers" ($\Psi(x)$) is shown in Fig. 2,a. According to the mentioned above chart, all possible values of the function $\Psi(x)$ are characterized as low, moderate or high level of matching answers depending on how the student's answer is close to the reference one. At the same time, mismatch of the answers can be caused not only by the incomplete knowledge, but also by insufficient confidence in knowledge, excessive emotions, or any other reasons preventing the student from making an unequivocal judgment about trueness of his/her conclusions.

Similar situation is with a membership function of a logical variable "Evaluation of student's knowledge" ($\Omega(y)$), where the terms of a base term set are characterized by three gradations –"poor", "average" and "full" (Fig. 2,b) – depending on how the evaluation of the answer is close to one of evaluation scale criteria.

The presence of unrelated fuzzy logic sets allows to make such relevant fuzzy statements as "if ... then...". For example, clear logic accepts only two extreme statements: "if student's answer does not match the reference answer, then student's knowledge is unsatisfactory" and "if student's answer matches the reference answer, then the student has necessary knowledge". Fuzzy logic accepts both these extreme values, as well as any other intermediate statement linking the certain degree of answer accuracy and the corresponding answer evaluation.

The matching of subset items of the postulating and stating parts of a statement may apply a control function that is based on either "correlation –product encoding" method or "correlation – min encoding" method [4]. Currently there are no evidences confirming the preference of using one of these methods in computerized control of knowledge. However, "correlation – product encoding" method is used in the simulation model due to a number of reasons.

Sum combination method is used for getting a generalized logical statement. Herewith, superposition of membership functions of fuzzy sets is defined as

$$\Theta_{sum}(Z)=\Theta_i(Z) \quad \forall Z, \quad i\epsilon[1,3] \tag{1}$$

Transformation of a fuzzy set into a single decision taken on the basis of fuzzy logic statements requires using the gravity center of the fuzzy set membership function – centroid defuzzification method.

## 4 Strictness Parameter

The application of fuzzy logic relieves the student from necessity to speculate if he/she is not sure in the answer. Clearly indicating the degree of trueness in the answer, the student thereby provides the data giving possibility of mathematically differentiation of his/her academic achievements with high accuracy, and to perform unambiguous evaluation.

Mathematical application of fuzzy logic to the test control can also enter a "strictness" parameter. At the oral examination the teacher can somehow "forgive" a controversial answer deviating from his/her idea of trueness. But a stricter teacher will punish this controversial answer by a worse grade. Similarly to a traditional examinations conducted by teachers with different ideas of perfect knowledge of materials, the tests based on fuzzy logic may also be evaluated in different ways.

Fig. 3 shows an example of control which lets the student to give answers in relatively simple way in terms of fuzzy logic, if it is added to the test software interface. Indicating the degree of student's answer deviation from the reference one requires moving a slider to any position between the leftmost ("False") and the rightmost ("True), and clicking "OK". The slider location is determined and converted into relative coordinates, which are used for further calculations in a fuzzy logic expert system.

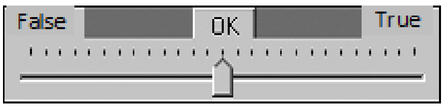


**Fig. 3.** Control of a fuzzy logic system

Rating answers and number of points accrued will depend on how the student indicated the degree of his/her answer matching the reference one. The number of points accrued depends also on the "strictness" parameter, which is indicated by sections in transition areas of membership functions –student's answer matches / does not match the reference answer, and the student learned / did not learn the controlled material. Despite the fact that the coordinating of these segments have quantitative indication, the level of strictness to student's knowledge is measured qualitatively, in such terms as "strict" (S), "stricter" (SS), "less strict" (LS), and "not strict" (NS), filling these concepts with quantitative measurement each time. Thus, if there is a necessity to compare the results of control, then introduction of the "strictness" parameter requires specified adopted coordinate values for transition sections.

Fig. 4 shows the charts illustrating changes in application rate of control for entering answers when implementing different "strictness" strategies of the fuzzy logic expert system. In the diagrams was considered such data – we applied the information about the students who used the element of fuzzy-logic system in the computerized tests.

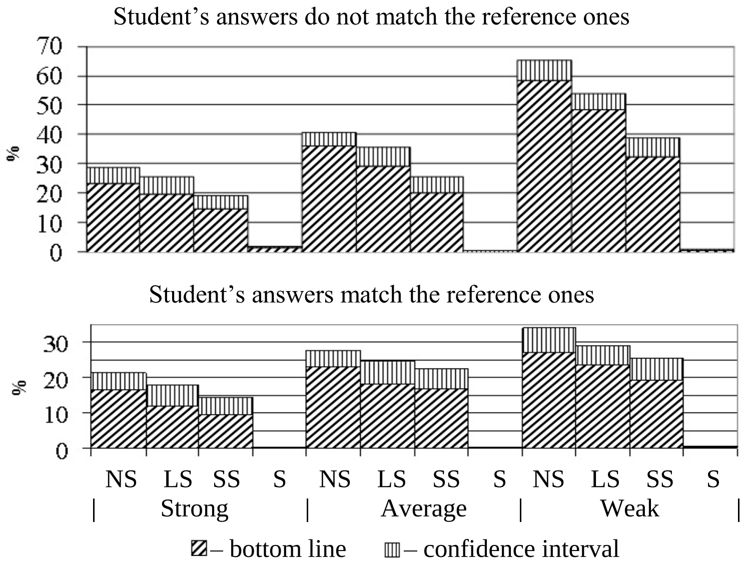


**Fig. 4.** Application rate of the control

The chart in Fig. 4 shows that in case of specified extreme level "Strict" the vast majority of students (over 98%), regardless of degree the preparedness (relatively strong, average and weak students) and degree of their answers matching with the reference ones, rarely use the opportunity to make a statement in terms of fuzzy logic. Absence of effective incentives upon the almost confident answer, and extremely punishable little doubt lead to the fact that students prefer to answer in terms of "False"-"True"." Thus, the fuzzy logic expert system capacities are practically not used. Therefore, it is not recommended to use a "strict" test system for practical purposes.

Other manifestations of "strictness" are popular enough to use fuzzy logic. It should be noted that the level "Not strict" is often demanded by weak students, as in case of student's answers matching the reference answers (30,5%±3,4%), and to even greater extent in case of answer mismatch (62,5%±4,1%). Therefore, this approach is not recommended to be a priority in order to ensure that all students are in equal conditions and none of them has any preference.

Table 1 shows the coordinates of the membership functions corresponding to the level of "Stricter". According to the chart in Fig. 4, it is often demanded and can be recommended as the core level in the absence of any other preferences. This recommendation can be confirmed by positive experience of use as the sole strictness parameter in the fuzzy logic system of test software SSUquestionnaire [7].

**Table 1.** Coordinates of membership function transition lines

| | $\Psi(X)$ | | | | | | | $\Omega(X)$ | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Low | | Moderate | | | High | | Poor | | Average | | | Full | |
| | $a_1$ | $b_1$ | $a_2$ | $b_2$ | $c_2$ | $a_3$ | $b_3$ | $a_1$* | $b_1$* | $a_2$* | $b_2$* | $c_2$* | $a_3$* | $b_3$* |
| Y | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| X | 0 | 0,25 | 0,15 | 0,5 | 0,85 | 0,75 | 1 | 0 | 25 | 15 | 50 | 85 | 75 | 100 |

Discussing the data shown in Fig. 4, it is necessary to underline that they do not directly recommend any of strictness degrees in the test system. So there may be different approaches to setting the "strictness" parameter. However, it is necessary to mention that regardless of the adopted approaches to setting the expert system strictness level, it must be set up prior to the test control. Any changes in the conditions of control through adjusting the strictness parameter for specific students, groups of students or disciplines are unacceptable. Like the oral examination, on the one hand there is contradiction between the desire to set up individual approach to each student and evaluate his/her achievements with the strictness degree that would enhance learning, vs. on the other hand, the requirement of compliance with the general approach to all students. Therefore, differentiating the "strictness" parameter in a fuzzy- logic test system can be justified for some special cases, but the general approach requires this parameter to be standardized, and academic achievement of any student should be equally evaluated, regardless of any subjective or objective circumstances.

## 5 Types of Tests

Despite the considerable variety of standardized test questions ([3], etc.), fuzzy logic expert system accepts only two types of tests.

The first type of tests covers the tasks containing the questions that can be answered using the full range of logical variables from "False" to "True". These are the questions that require to confirm or deny any statement, such as "Fish live in a water", "2 +2 = 4" or "The sun shines at night", "2 +2 = 5", etc.

When performing the test of the first type a student can move the control slider of a fuzzy logic expert system (Fig. 3) to any of the positions, which , in his/her opinion, corresponds to the degree of answer trueness. If one of the extreme positions is selected and student's answer matches the reference answer, the highest possible score will be awarded. If the selected extreme position of the slider does not match the reference answer, there will be 0 points. Intermediate position of the slider will allow giving intermediate (between zero and a maximum) number of points.

Another type of tasks includes the questions that can be answered within a half of the range of logical variables from "Not true" ("Not false") to "Truth". These tasks include questions along with two or more options of possible answers. At least one of them is correct and at least one is wrong. For example, if the task has a question "2 +2 =?" along with three answer options "3", "4" and "5", and it is offered to determine which one is correct, then examinee cannot select the wrong answer “3” or “5” stating that it is false.

When performing such task, the control slider of the fuzzy logic expert system can be moved within a range from the middle position “Not true” (or “False”) to the rightmost position “True”. In this case, the maximum possible score will be given if student’s answer matches the reference one and the rightmost slider position is selected. In all other cases, the amount of points accrued will be determined by how student’s answer matches the reference one (depending on the slider position).

Table 2 shows different scoring options for the two considered types of tests (maximum score for correctly completed task is 100 points).

**Table 2.** Points accrued for a completed task depending on student’s answer matching the reference answer

| Type of task | Student’s answer matches the reference one | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Type 1 | False | | ¼ | | 1/2 | | ¾ | | True | |
| Type 2 | ½ | | 5/8 | | 3/4 | | 7/8 | | 1 | |
| Strictness parameter value | Strict | Not strict | Strict | Not strict | Strict | Not strict | Strict | Not strict | Strict | Not strict |
| Answer evaluated, points | 0 | 0 | 1 | 30 | 3 | 70 | 6 | 90 | 100 | 100 |

## 6 Measurement Capabilities

For the evaluation of the impact of a fuzzy logic expert system on the measurement capabilities of test knowledge control was made an experimental research.

The experiment engaged 228 students divided between the experimental and control groups. The groups were formed on the basis of current students’ progress. Mann-Whitney [6] checks showed that the groups are homogeneous.

The students in the experimental group were given a fully functional test program SSUquestionnaire, also they had an opportunity to give fuzzy logical answers. Strictness parameter of the fuzzy logic expert system was set up as “Stricter” and did not change throughout the experiment.

The test program used in the control group differed from the fully functional one, since its fuzzy logic module was disabled. The students could not move the control slider of the fuzzy logic expert system to any intermediate position; they had been forewarned as well.

Test results of the experimental and control groups were processed mathematically. They helped to estimate the strength of links between successful execution of individual test items and the final estimates the students received for all of the test questions. Pearson correlation coefficient [2] was calculated for the test results of each group independently. It was believed that the closer the absolute value of Pearson correlation coefficient is to one, the tighter are links and measurement capabilities of the relevant test.

Comparison of the received data showed that the experimental group revealed closer linear dependence between the results of individual tasks and the general test results than the control group. Pearson correlation coefficient in the experimental group increased from 0,52 to 0,65 compared to the control group, that indicates better measurement properties of the test.

## 7 Conclusion

Elimination the identity of the person from the process of control enables using the diagnostic capabilities of the examiner during the test. This disadvantage of test control can be mitigated by use of an expert system developed on the basis of mathematical fuzzy logic.

The advantage of the fuzzy logic expert system hides in the fact that its introduction into a test program provides students with the opportunity not only to give the answers based on strict logic, but also to indicate any degree of answer trueness ranging from “False” to “True”. A student does not have to give a definite answer, even if he/she is required to go beyond the scope of their own knowledge. He/she can express doubt indicating how an idea of the true answer matches or does not match the reference answer. In this case, the test results are not measured in terms of clear logic (if the reference answer matches student’s answer, then the student knows the material, and vice versa), but in terms of fuzzy logic, when the final evaluation is determined by how these answers match.

The proposed justification of the decisions made by the examiner on the basis of the fuzzy logic expert system mitigates disadvantages of computerized testing as a tool for educational measurements, but does not eliminate these disadvantages entirely. Further efforts in the improving the theory and methods of test control, including methods directed on the fundraising the computer equipment for modeling diagnostic functions of the teacher in the control process will enhance the reliability of results of the evaluation of student’s knowledge.

## References

1. Alexeyev, A. N., Alexeyeva, G. V.: A simulation model of the test control of knowledge and skills. *Computer-oriented educational system*, Kyiv, NPU. M. Dragomanov, No 7 (14), 65 - 71. (in Ukrainian) (2009)
2. Glass, G. V., Stanley, J.C.: Statistical methods in education and psychology – Englewood Cliffs, N.J., Prentice-Hall (1970)
3. IMS Global Learning Consortium. Accessible at http://www.imsglobal.org/question
4. Korneev V.V., Gareev A.F., Vasyutin S.V., Reich V.V.: Databases. Intelligent Processing of Information. – Moscow: Knowledge (in Russian) (2000)
5. Likert R., Roslow S., Murphy G. A.: Simple and Reliable Method of Scoring the Thurstone Attitude Scales. Journal of Social Psychology, 1934. Vol. 5, 228 – 238 (1934)

6. Mann, H. B., Whitney, D. R.: On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 50–60 (1947)
7. New Opportunities of Knowledge Testing using software SSUquestionnaire Version 4.10 Accessible at: http://test.sumdu.edu.ua (in Russian)
8. Stevens S. S.: Experimental Psychology. Moscow: Foreign Literature, Vol. 1 (in Russian) (1960)
9. Thurstone L. L.: The Measurement of Social Attitudes. Journal of Abnormal and Social Psychology, Vol. 26, 249 – 269 (1931)