Igor D. Rudinskiy

Emin M. Askerov

Maksim A. Emelin

MULTIPLE CRITERIA VECTOR TESTING RESULTS EVALUATION MODEL

Abstract

In the paper one of the most important problems of modern testology which consists in quality definition of test task is discussed. Special attention is devoted to basic characteristic – the truth degree of test tasks answers. This characteristic is considered in pedagogical knowledge testing. A hierarchy of knowledge evaluation models is offered. The hierarchy allows to track features and conditions of the applications of various models. The examples of test tasks evaluation using one-parametrical knowledge evaluation model (binary, algebraic, or fuzzy) are presented. An idea and an example of using the fundamentally new vector (multiparametric) knowledge evaluation model are discussed. This model allows to evaluate the quality of answers to the test task with respect to several criteria simultaneously.

Keywords

Pedagogical testing, multiple criteria, knowledge evaluation model, knowledge control, truth degree, test tasks, automated knowledge testing.

INTRODUCTION

One of the most actual problems of modern pedagogy is the maintenance of quality of control materials (CM) [1]. The most important characteristic of CM that is considered in pedagogical knowledge testing is the truth degree of answers to test tasks.

The extension of the truth degree evaluation scale of test tasks answers allows to extend the analysis of a trainee's knowledge and to increase the reliability of pedagogical control. A hierarchy of knowledge evaluation

models is presented in this article. A diagram of this is shown in Figure 1. The hierarchy allows to track features and conditions of the applications of various models and, based on that, to synthesize an essentially new vector knowledge evaluation model. Vector models allows to estimate the quality of answers to test tasks by several criteria simultaneously.

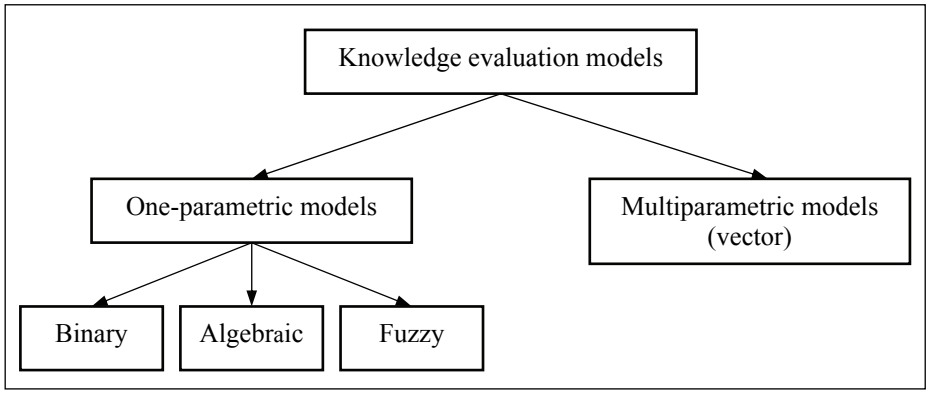

Fig. 1. Hierarchy of knowledge evaluation models

One of the most elementary one-parametrical model of truth degree evaluation (we shall name it ***binary***) is the popular approach to the truth degree evaluation of test tasks answers in terms of "true-false" [4]. Points 0 and 1 on a line (describing the truth degree of test task answers) correspond to this approach from the mathematical point of view (Figure 2).

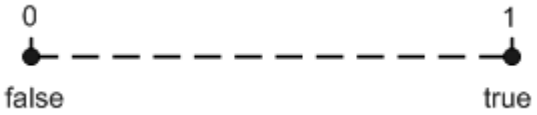

Fig. 2. Graphic representation of the truth degree for absolutely correct or absolutely wrong variant of the answer

The limitation of the acceptable truth degree of answer set to only two values allows to organize effective knowledge testing only in the case of knowledge revealing the concrete facts by the presentation to trainee of such types of questions, as: "What?", "Where?", "When?", "Who?", etc.

An example of answer evaluation for the elementary test task: "How much is 2×2 in decimal numeration?" using binary model is shown in Table 1.

Table 1

Variant of the answer	Truth degree
5	0
3	0
1	0
4	1

The second model with respect to its complexity, further named ***algebraic***, allows to evaluate the truth degree of variants of the test tasks answers by number whose values are taken from a range, for example, from the interval [0; 1] (Figure 3). The algebraic model gives an opportunity for a more flexible truth degree evaluation of variants of test tasks answers in comparison with the binary model, due to the wide use of incomplete, inaccurate, uncertain, etc. answers in pedagogical practice [4]. In the algebraic model the truth degree evaluation is represented by one of several points in the allowable values range. An example of the evaluation of the test task: "Choose the basic attributes of a parallelogram", using the algebraic model, is shown in Table 2.

Table 2

Variant of the answer	Truth degree
Opposite sides are equal	0,5
Opposite angles are equal	0,5
Opposite sides and angles are equal	1
Diagonals are equal	0

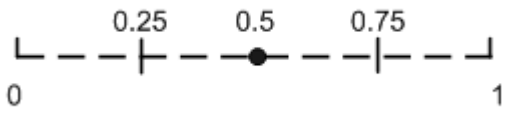

Fig. 3. Graphic representation of truth degree of the answer: "Opposite angles are equal"

One of the most complex one-parametrical models is the **fuzzy** model. The fuzzy model uses mathematical methods of fuzzy algebra for truth degree evaluation of answers to test tasks [2]. The quality of the answer is characterized by the membership function which is defined on the set of linguistic variable values "truth degree of answer" variants. The grade of membership of a variant of the answer to each category being estimated is defined by the number which values are taken from a range, for example, from the interval [0; 1]. In the fuzzy model the truth degree of an answer is characterized not by a unique point, but also by an aggregate of points which make up the membership function. The fuzzy model allows to estimate the truth degree of answers for test tasks in a more flexible manner as compared with the previous approaches to evaluation, in particular, in weakly-formalizable subjects (such as history, literature, etc.). An example of fuzzy evaluation of the truth degree of answers to the test task: "Where did Carlson live?" is shown in Table 3. In this example the symbol I designates a linguistic variable "the truth degree of the answer", containing the values: I = ["true", "not absolutely true", „false"].

Table 3

Variant of the answer	Truth degree		
	true	not absolutely true	false
In the house	0,5	0,7	0,4
On the roof of a house	0,7	0,4	0,2
In a house in the attic	1	0	0
In a cellar	0	0	1

The membership function for the answer "On the roof of a house" is represented on Figure 4.

All models considered above allow to estimate the quality of CM by a unique parameter (in our case this parameter is the „truth degree of the answer to the test task"). From the position of mathematical analysis, such approach can be named the **scalar** approach.

However, it is clear that the quality of answers to test tasks can be characterized not only by the unique parameter „truth degree", but can also be represented by the whole spectrum of parameters which can vary depending on the purpose of the pedagogical control, the character of discipline, the algorithm of testing, etc. An attempt to generalize scalar evaluation methods

used in multivariate evaluation of qualities of answers to test tasks is discussed in this article. A formal model of evaluation of the qualities of answers to the test task by several criteria will be called the ***vector model***.

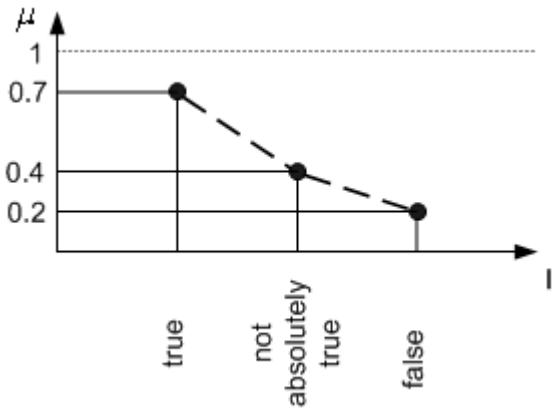

Fig. 4. Graphic representation of truth degree for the answer "On the roof of a house"

In the vector model the parameters playing the role of quality evaluation criteria are used as components of the vector basis which forms n-dimensional vector space (where *n* is the number of criteria used to evaluate the quality of answers to test tasks). Evaluation with respect to each criterion can be made using one of the abovelisted models (binary, algebraic, or fuzzy). The final estimation of the test task performed using the vector model will be determined by a point in n-dimensional space (for example, the point **A** in Figure 5 in the case of three-dimensional space of criteria). From the mathematical point of view, the final estimation can be exposed by each criterion separately or by the decision in a multiple criteria problem with various methods of deescalation, concession, method of the main criterion, method of an ideal point, etc. [3].

We shall consider the test task in the field "History of Russia", for example, "What role did the Communist Party of the Soviet Union play in the history of the development of Russia?" using the vector model. We shall assume that experts have estimated possible variants of answers, for example, by three criteria: truth (using binary model), completeness (using algebraic model), and originality (using fuzzy model, I = ["originally", "not very originally", "unoriginally"]). An example of the evaluation is shown in Table 4.

Table 4

Variant of the answer	Evaluation criteria				
	Truth	Completeness	Originality		
			Originally	Not very originally	Unoriginally
There are both positive and negative moments	1	0,5	0,5	0,3	0,9
Extremely positive role, democratic foundations put obstacles for the country	0	0,2	0,8	0,5	0,1
Extremely negative role, years of communism were the period of stagnation for the country	0	0,2	1	0	0
Domination of the Communist Party has not affected the development of Russia	0	0	0,3	0,1	0,5

A simplified illustration of the vector model by evaluation can be seen from Table 4 is shown in Figure 5. It is presented in a simplified manner as a coordinate on the axis “originality” where the fuzzy evaluation model is used and can be represented by a set of points which form a hyperplane. In general, a more complex case of three-dimensional space (which can be displayed as a hypercube) is considered.

As compared with the scalar quality evaluation models of answers to test tasks, the vector model presented above is more complex and full. It allows to expand the evaluation possibilities and to estimate the truth degree of test tasks more authentically. Thus, certainly, the labor input of expert evaluation of test tasks is increasing.

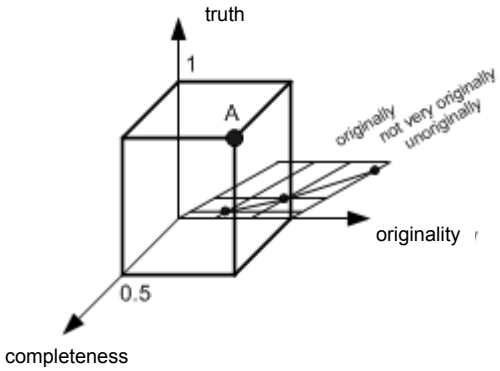

Fig. 5. A simplified illustration of the vector evaluation model

Nevertheless, the simplicity of test tasks duplication and their repeated use in knowledge control, and also the basic opportunity of application of a vector estimation of answers to test tasks in the scalar form (which can be demanded for knowledge evaluation with respect to one of the criteria) allows to consider this model to be a method for future use for realization in systems of automated knowledge testing.

REFERENCES

1. Mihalychew E.A.: Didactic testology. National Education, Moscow 2001.
2. Rudinskiy I.D.: Basics of Formally Structural Modeling of Educational Systems and Automation of Knowledge Testing. Hot Line Telecom, Moscow 2004, pp. 1-204.
3. Kini R.L., Rajfa H.: Decision Making at Many Criteria. Radio and Communication, Moscow 1981.
4. Questions of Testing in Education. "The All-Russia Guidance Magazine" 2001, No. 1.