

New approach to development of the expert system of control and knowledge assessment

Golovachyova Viktoriya, Tomilova Nadezhda, Tomilov Alexander

Karaganda State Technical University

Kazakhstan

Abstract - The expert systems of control and knowledge estimation are usually developed on the basis of different approaches to constructing a question and answer on test. The analysis of existing testing methods made it possible to conclude that questions with selective and selectively constructed types of answer do not always allow objectively assessing the knowledge of trainees, which reduces the stimulating effect of pedagogical evaluation on cognitive activity of trainees, and the quality of the learning process as a whole. The article proposes a method of knowledge estimation, which is based on a new approach to constructing a question and answer that allows a freely designed form of a test response. The proposed approach gives an opportunity to expand the functions of tests, thereby increasing the approximation degree of the estimation for the test to the level of the real knowledge of students

Keywords - The expert systems of control and knowledge estimation, testing methods, approaches to constructing a question and answer on test, quality of the learning process

1 Introduction

The pedagogical science is intensively searching for ways and means to improve the monitoring and estimation of knowledge in order to improve the quality of education. According to the opinion of the scientist K.Ingenkamp "If modern scientifically grounded didactics does not rely on the rich tools of the maximum objective methods of pedagogical diagnostics, then it is doomed to failure" [1].

Discussion, of course, is about objective control, that is, about such methods of testing knowledge and wider, pedagogical diagnostics, when a teacher or researcher uses a means that gives accurate and complete information about the level of knowledge, the quality of the educational process as a whole. At the new stage of development, the means of pedagogical science considers testing.

The advantage of such form of control in its objectivity, that is, independent testing and estimation of knowledge from an expert - a teacher, who has knowledge in this subject area.

Practice shows that among the rationally composed test tasks are those that require the presentation of the answer in one of the following forms [2,7]:

- 1) Selection of the correct and complete answer from a series of the proposed ones (correct or incorrect, complete or incomplete, accurate or inaccurate);
- 2) Selection comprising of two parts (the first part requires any selection, and the second one requires justification of the selection);
- 3) Alternative selection (yes / no, 0/1, true / false);
- 4) Arrangement of elements from the proposed list in the correct succession;
- 5) Matching elements from two lists;

6) Statement completion, indication or selection of missed words;

7) One word (or number) answers;

8) Answer in many word limited in the order or inter-word connection.

Existing possible forms of presenting a response to a test task are appropriate. The expediency of using them in the educational process is justified. Didactic tests, undoubtedly, by many features are the most advanced method of control and firmly hold the leading position among the traditional forms of knowledge control.

The practitioners do not consider it necessary to change something in forms, to learn form-making and the methodology for developing new test tasks. Because of this position, testing practice begins to degrade.

The analysis showed that modern computer tools allow building tests with selective, numerical, constructed answers [3]. In practice, tests with selective answers are most often used. They are easier to prepare (there is no need to create many standards of correct answers, to ensure the completeness of which is extremely difficult) and, most importantly, easier to use. In tests with selective answers, students spend most of their time on the task, rather than on a set of answers.

World practice shows that the existing methods of constructing questions and answers based on didactic tests do not always allow objectively revealing the real level of knowledge of students, especially in the subjects of the social and humanitarian cycle, and have negative aspects: serious distortions of perception by students of the integrity of material; development of stereotyped thinking; mechanical memorization of answers and others [4].

Such a situation has obvious negative consequences: reducing the stimulating effect of the pedagogical evaluation on the cognitive activity of students, on the quality of the educational process as a whole.

Test questions provide information for reflection. The fact is that not every discipline lends itself to formalization. For such subjects as physics, mathematics, mechanics and others, formalization is obvious. In terms of the same socio-humanitarian cycle, the knowledge used cannot always be clearly formalized. The lag in the development of methodological problems and the creation of new methods for estimating knowledge is the main reason for the gap between the potential and real capabilities of the computer.

In order for knowledge control with the use of testing to be really, effective for this cycle of subjects, at each stage of the training it is necessary to know the degree of understanding the knowledge. At the same time, it is necessary to cover all the useful characteristics of understanding knowledge by means of testing. For example, indicators, such as the ability to concretize your

answer with examples, knowledge of facts, the ability to logically connect, intelligently express your thoughts, etc.

In order for knowledge control with the use of testing to be really effective for this cycle of subjects, at each stage of the training it is necessary to know the degree of mastering the knowledge. At the same time, it is necessary to cover all the necessary characteristics of mastering knowledge by means of testing. Such, for example, indicators, such as the ability to concretize your answer with examples, knowledge of facts, the ability to logically connect, intelligently express your thoughts, etc. Only when using such forms of knowledge estimation that are not inferior to the content of verbal control can you achieve great advantages from using such an effective means of obtaining an accurate and objective estimation of knowledge like testing [5].

However, there is an important problem - the difficulty of recognizing the meaning of a textual response. The usual methods of analyzing systems and computer modeling, as is known, based on the accurate processing of numerical data, are essentially incapable of covering a huge complexity process of human thinking and decision-making. Hence, the conclusion suggests that to obtain significant conclusions about the behavior of humanistic systems, which includes education. In this case, we will have to abandon high standards of precision and severity, which we generally expect in the mathematical analysis of well-defined mechanical systems and to be more tolerant of other approaches that are approximate in nature.

The search for pedagogically effective ways and the development of a method to improve the process of monitoring and evaluating knowledge in order to improve the quality of instruction is a problem of study.

2 Method of Knowledge

Estimation

The main idea of this method is that student's knowledge is estimated by test questions, answers to which are given by them in a freely-constructed form [5].

For assessment of the analysis of such answer we use the following complex of criteria [4]:

2.1 Objectivity

It reflects the basic level of knowledge of the subject and is determined by comparing the correspondence of the descriptors used to the descriptors of the thesaurus or their synonyms. The quality of the criterion "objectivity" is characterized by a coefficient δ . The criterion for estimating δ is the ratio of the number of correctly used descriptors to the total number of descriptors corresponding to each question of the test in the thesaurus. The criterion is calculated by the following formula:

$$\delta = \frac{N}{M} \quad (1)$$

where N is the number of descriptors corresponding to the thesaurus for each test question;

M is the total number of descriptors corresponding to the thesaurus for each test question.

2.2 Literacy

The criterion is defined by the rules for building text documents. In this research, we are restricted with the conditions for the application of particular indicators, e.g. spelling check, which is determined by comparing the correspondence of every word in the answer text to the words in the spelling dictionary. The quality of the criterion "literacy" is characterized by the coefficient γ . The criterion for estimating γ is the ratio of the number of correctly written words to the total number of words in the text of the answer. The criterion is calculated by the following formula:

$$\gamma = \frac{K}{R} \quad (2)$$

Where K is the number of correctly written words in the text of the answer;

R is the total number of words in the text of the answer.

2.3 Presence of examples

The presence of an example in the text answer clearly gives an explanation of the main question in the test. The criterion is determined by comparing the correspondence of the examples used in the text to words in the sample database or their synonyms. The quality of the criterion "the presence of examples" is characterized by the coefficient φ . The criterion for estimating φ is the ratio of the number of correctly-mentioned examples in the text to the total number of examples corresponding to their each question of the test. The criterion is evaluated by the following formula:

$$\varphi = \frac{D}{F} \quad (3)$$

where D is the number of examples corresponding to the database or their synonyms for every test question;

F is the total number of examples from the sample database or their synonyms for each test question.

2.4 Logical connections between sentences

These connections are determined by the rules of language for communication between sentences. The quality of the criterion "logical connections between sentences" is characterized by the coefficient μ . The criterion for estimating μ is the ratio of the number of logical connections in the answer, which is determined by the answer incidence matrix (Figure 1) and the maximum possible number of logical connections between the sentences used in the answer text.

	1	2	...	N
1	1	1	1	1
2		1		1
...			1	1
n				1

Figure 1. Answer incidence matrix

The diagonal elements of the matrix determine the number of sentences in the text of the answer. Super diagonal elements of the matrix determine connections between the concerned sentences and the subsequent sentences of the text answer.

The criterion is evaluated by the following formula:

$$\mu = \frac{E}{L} \quad (4)$$

where E is the number of logical connections determined by the answer incidence matrix;

L is the maximum number of logical connections between the used sentences.

It evaluated by the following formula:

$$L = n \cdot (n - 1) - \sum_{i=1}^{n-1} i \quad (5)$$

where n is the number of sentences in the text of answer.

2.5 Complexity

It characterizes the quality of the test answer in general and is determined by the presence of connections between the concerned criteria: objectivity, literacy, presence of examples, and logical connections between sentences. The quality of the criterion "complexity" is determined by the coefficient η . The criterion is evaluated by the following formula:

$$\eta = \frac{(\delta + \gamma + \varphi + \mu)}{z} \quad (6)$$

where δ is the coefficient obtained by the criterion of "objectivity";

γ is coefficient obtained by the criterion "literacy";

φ is coefficient obtained by the criterion "the presence of examples";

μ is the coefficient obtained by the criterion "logical links between sentences";

z is the number of criteria considered.

The complexity analysis is determined by expert means. The general score for every question of the test is assigned by the following formula:

$$Ball = \left[\frac{\bar{\delta} + \bar{\gamma} + \bar{\mu} + \bar{\eta}}{\omega \cdot k} \right] \quad (7)$$

where $\bar{\delta}$ is a coefficient obtained by the criterion of "objectivity";

$\bar{\gamma}$ is a coefficient obtained by the criterion of "literacy";

$\bar{\varphi}$ is a coefficient obtained by the criterion "presence of examples";

$\bar{\mu}$ is a coefficient obtained by the criterion "logical connections between sentences"

$\bar{\eta}$ is a coefficient obtained by the criterion of "complexity";

ω is the total number of criteria of the answer analysis;

k is the complexity coefficient of every test question.

3 Specifics of construction the expert system of control and knowledge assessment

3.1 "Verbally – linguistic" contents assessment of the answer

We will review an example "verbally – linguistic" the answer, for example, on the question "What is an array of information?"

Can be one of answers to such question following:

"An array of information is the information structure formed of one or more records in such a way that records describe objects and the array - a class of objects. Several arrays make system or a set of arrays. Record is called precisely established data set, characterizing some object or process. Examples of records can be cash-memos, orders on performance of work, consignment notes, questionnaires, and statistical reports".

Array of information – the information structure formed of one or more records in such a way that records describe objects, and the massif – a class of objects. The array is a subject. Structure is a predicate, formed of one or more records, thus – the definition expressed by a participial phrase; that records describe objects, and the array – a class of objects – the additional structure consisting of two simple sentences connected by a coordinative conjunction /and/, and in the second sentence – incomplete – there is no predicate therefore the dash is put. All additional structure joins a subordinating conjunction / that/.

Second sentence: Several arrays make system, or a set of arrays. Several arrays – a subject; make system – a predicate; a set of arrays – the application.

Third sentence: *Record is called precisely established data set, characterizing some object or process.*

Set – a subject; *called* – a predicate; *established* – definition; *precisely* – a circumstance; *data* – definition; *characterizing some object or process* – the definition expressed by a participial phrase.

The fourth sentence: *Examples of record can serve commodity checks, work orders, invoices, questionnaires, statistical reports. Examples of recording are definition (inconsistent); can serve as a predicate; Checks, orders, invoices, questionnaires, reports - subjects; commodity - definition (agreed); on - an preposition; implementation - definition (uncoordinated); works - addition.*

The criterion for estimating γ can be a set of different rules that determine the grammatical quality of the text answer in general.

The presence of examples in the text of the answer, as is known, visually illustrates the main question. The quality of the "presence of examples" criterion characterized by the coefficient ϕ and determined by comparing the correspondence of the examples used to words from the sample database. In the above answer to the question "What is an array of information?" as indicated, five correct examples are shown: goods checks, work orders, invoices, questionnaires, statistical reports.

Based on the experience of teachers, the criterion for evaluating ϕ can be the presence of one example; the presence of two or more examples; absence of an example; wrong example.

Logical connections between sentences are determined by the rules of language communication between sentences. As you know, a text consisting of two or more sentences related to each other in meaning or structurally, and performing a composition-stylistic function in a coherent text, is a complex syntactic whole.

From the point of view of the structural-semantic, complex syntactic integers divided into complex syntactic integers with a chain connection and complex syntactic integers with a parallel connection between sentences.

In the reviewed example, the first sentence of a complex syntactic integer has a general meaning. All other sentences disclose the contents of the first. Of the four sentences presented in the response, the first two connected by a parallel connection and the last two connected by a parallel connection.

The evaluation criteria μ can be:

- All answers of sentences should have logical connections with the main sentences of the answer;
- All sentences, except for the main one, should have a subordinate character to the main sentence;
- Supplement its content, disclose in more detail the essence of the main sentence, explain the main sentence, classify the subject matter and so on.

In the answer to the question «What is an array of information?» the diagonal elements of the matrix consist of four elements, because in the text of the submitted answer - four sentences. Between the first and second sentences of the answer there is a logical connection, which determined by the keyword "array". Between the first and third, the first and fourth, as well as between the third and fourth sentences of the answer, there is a logical connection, determined by the keyword "record". The

calculation of the criterion "logical connections between sentences" determined by the number of the logical connections in the answer text.

The criterion "Complexity" characterizes the quality of each textual answer as a whole and determined by the presence of connections between the criteria: objectivity, literacy, the existence of an example, and logical connections between sentences. If examinee while answering to the basically question of the test highly defines "objectivity" criteria. Grammatically correctly constructs the text of the answer, demonstrates the existence of examples explaining the essence of the main question, there is a gradual development of the thought in the text of the answer, from the point of view of its general characterization it is complex. Analysis of complexity is calculated by the number of connections in answer and determined by expert means.

The general mark for each question of the test is put down on the formula (7).

However, from a pedagogical point of view, when assessing knowledge, it is necessary to take into account the structure of knowledge of each trainee. For this, we developed the following heuristic algorithm: each variant of the test consists of 10 proposed test questions, the numbers of which ranked as their complexity increases. Of these, the first 3 (three) test questions are simple, the next 5 (five) questions are average and the last 2 (two) questions are complex.

Using the heuristic approach to estimation, we have developed an algorithm that is a combinatorial problem and is formulated as follows: it is necessary to define all possible values of α , β and λ such that they satisfy the conditions for obtaining an estimate on a five-point system:

$$\begin{aligned} a &\leq \alpha_k + \beta_m + \lambda_n \leq d \quad (\text{excellent}) \\ B &\leq \alpha_k + \beta_m + \lambda_n < a \quad (\text{good}) \\ c &\leq \alpha_k + \beta_m + \lambda_n < B \quad (\text{satisfactorily}) \\ 0 &\leq \alpha_k + \beta_m + \lambda_n < c \quad (\text{unsatisfactory}) \end{aligned} \quad (9)$$

At limiting values k, m, n .

Where α, β, λ – all sorts of combinations of complex, medium and simple questions, respectively;

k – total number of complex issues in the test;

m – total number of average questions in the test;

n – total number of simple questions in the test;

d – total number questions in the test;

a, B, c – the number of correct answers to the test that satisfy the conditions for obtaining an assessment on a five-point system.

3.2 Features of development of the expert system of control and knowledge assessment

The method of knowledge assessment developed lies at the basis of construction an expert information system for monitoring and assessing of knowledge.

The core of such a system is the basic vocabulary of the subject area, which created and accumulated in the process

of construction and teaching students and is the main supporting didactic material. It serves for the analysis of each text answer and consists of the following components:

Theaurus. The main way to form a thesaurus is to sample descriptors, usually expressed by nouns, from a summary of lectures, a work program, their synonyms and other linguistic sources specialized in the subject.

The base of examples. The main way to form such a base is to sample the examples given by the teacher at the lectures, examples from students' answers in practical, seminar classes, etc.

The orthographic dictionary. The input information of the tested person fixed and entered into the lexical analyzer of the system. The lexical analyzer takes the original response text directly from the elements of the input interface and converts it into an array of lexical units. For each lexical unit, the parser performs the search procedure using the basic vocabulary of the subject domain. In the case of an exact match of the analyzed word with the descriptor of the basic dictionary of the domain, the search option considered successful and the corresponding information transmitted to the program.

The lexical. The program breaking the entrance text into the sequence of words or lexemes. The identifier is allocated, and then check on accessory to its set of keywords to be done.

The syntax analyzer. The program making the parse of the offers' text.

About Authors:



Golovachyova Viktoriya received her PhD in Computer Science and Education from ENU of L.N. Gumilyov, Kazakhstan in 2011 and has passed a scientific training in England, in the Center of the international education, Plymouth, in 2009. Currently, she is a professor at Department of Computer and Information Science of Karaganda state technical university, Kazakhstan. Her research interests include development and deployment of computer learning tools in educational process of higher



Tomilova Nadezhda is the manager of Department of Computer and Information Science of Karaganda state technical university, Academician of the international academy of informatization, Kazakhstan. The field of scientific research and achievements – digitalization of processes of education, development of the automated control systems for production of power supply



Tomilov Alexander is a teacher in Department of Computer and Information Science of Karaganda state technical university, Kazakhstan and PhD candidate in the specialty «Mining». The field of scientific research and achievements – digitalization of processes of education, development of the automated systems of design of fastening of excavations.

4 Conclusions

Cont

5 Conclusions

Contribution to the world pedagogical science of the conducted research is that its results can used, both in general pedagogy, and in theoretical and practical dentistry. The work theoretically substantiates the necessity of applying a new approach to constructing a response to a test question that allows a freely designed form of a test question and answer, as well as the need to develop criteria for such a response and a scientifically grounded approach to their evaluation. The full information contained in the test answer of the examinee taken into account using the developed algorithms for calculating test response analysis criteria and computer data processing tools, which allows obtaining a comprehensive objective assessment of the level of mastering knowledge. The procedure for conducting pedagogical testing strictly formalized, but the result unambiguously follows from the answers of the subjects.

The practical significance of the research is expressed in the fact that the expert information system for monitoring and evaluating knowledge has been introduced into the practice of the training process, which can be recommended to improve the monitoring and evaluation of knowledge in the subjects of the social and humanitarian cycle in order to improve the quality of instruction. This makes it possible to apply the results obtained to solve the scientific problem of objective and reliable assessment of knowledge using the test method of monitoring and evaluating knowledge.

References

- [1] Ingenkamp, K. (1991). Pedagogicheskaya diagnostika. M.: Pedagogika.
- [2] Golovacheva V, Abaeva N (2015). Theory and practice of designing pedagogical tests (pp. 5-210). Monograph (Germany: LAP LAMBERT Academic Publishing, 216 p. IL.).
- [3] Golovachyova V, Abaeva N, Kokkoz M, Mustaphina L, Mustaphina V (2015) (pp. 284-295). Specifics of Development of the Integral Method of Knowledge Estimation. Review of European Studies, Published by Canadian Center of Science and Education, Vol. 7.
- [4] Golovachyova V, Menlibekova G, Abayeva N, Ten T, Kogaya G. (2016) Construction of Expert Knowledge Monitoring and Assessment System Based on Integral Method of Knowledge Evaluation (pp. 2539-2552). International journal of Environmental & Science Education, Vol 11(9).
- [5] Golovachyova, V.N. (2001) Osnovi primeneniya testjvih tehnologiy v pedagogike. Almati: Isdatelstvo "Gilim".
- [6] Golovacheva, V. N., Abaeva, N. F. & Kokkoz, M. M. (2015). To the Issue of University Tests for Students. Higher School of Kazakhstan, 3, 35–39.
- [7] Golovacheva, V. N. (2000). Analiz sistem avtomatizirovanogo kontrolya znani. Karaganda: "Avtomatika i informatika" (No.1-2, C.54-57)