Computer Testing as a Form of Students' Knowledge Control

Dmitry Romanenko[0000-0002-1628-770X]

Belarusian State Technological University, Sverdlova str., 13a, 220006 Minsk, Belarus
rdm@belstu.by

Abstract. The article describes the experience of using testing to control students' knowledge in technical disciplines, as well as integrating the obtained results into the calculation of the final student's rating in the discipline. To organize the control of students' knowledge, a testing system was developed and characterized by a high level of security, a flexible system for assessing test results that takes into account a partially correct answer, flexible opportunities for the teacher in the context of the question complexity coefficients choice, a flexible testing organization system that allows the formation of tests of the same complexity with any number of difficulty levels of questions. The developed testing system allows using the following types of questions: one-line answer, multi-line answer, univariate question, multivariate question, logical sequences, pairwise matching. The experience of using the presented types of questions within the discipline "computer networks" is described. It is suggested the principles of the tests formation, taking into account the complexity of the questions, the method of calculating the results for individual questions and the entire test, which allows taking into account the partially correct answer of the student, reduce the final result on the question when the student chooses incorrect answers, take into account any (selected by the teacher) question difficulty factors. A method is proposed for calculating the final assessment of a student's knowledge in a discipline, which implies an integrated accounting of his achievements in the study of both theoretical and practical components.

Keywords: testing, knowledge control, rating, significance coefficient, question, logical sequences, pairwise matching, univariate question, multivariate question.

1 Introduction

In modern education, computer testing can perform three main interrelated functions: diagnostic, teaching and educational [1]. The diagnostic function is to identify the level of student knowledge, abilities, skills. In terms of objectivity, breadth and speed of diagnosis, testing surpasses all other forms of knowledge control. The teaching function of testing is to motivate the student to intensify the work on the assimilation of educational material. The educational function is manifested in the frequency and inevitability of test control. This disciplines, organizes and directs the activities of

Proceedings of the 4th International Conference on Informatization of Education and E-learning Methodology: Digital Technologies in Education (IEELM-DTE 2020), Krasnoyarsk, Russia, October 6-9, 2020.

students, helps to identify and eliminate gaps in knowledge, forms the desire to develop their abilities [1-3]. Testing is the fairest method. It puts all students on an equal condition, both in the control process and in the evaluation process, practically eliminating the subjectivity of the teacher.

2 Testing system for knowledge control

The main aim of the work is to develop a testing system, methods for conducting and calculating the results of knowledge control. Another important task is to develop a methodology for using the results of knowledge control in the form of testing for a complex assessment of knowledge in a discipline in the form of calculating the final (rating) assessment.

A testing system was developed in the form of a web-portal, which makes it possible to remotely control knowledge, for example, through the Internet [1]. In this regard, only a web browser is sufficient to organize testing on computers. From the point of implementation, the following main features can be distinguished: PHP programming language, Symfony 2 framework, Percona Server 5.6, Redis 2.8, web-server Nginx, PHP-FPM, support for the multi-node architecture is implemented. Note that the system contains both the roles of the teacher and the student, and the role of the operator, which allows you to plan tests. The administrator role is required to manage all processes.

2.1 Using types of questions for knowledge control

The most important stage that allows you to objectively assess knowledge will be the creation of a high-quality and comprehensively thought-out base of questions, which will allow students to demonstrate all their knowledge and abilities in the subject area. Questions in the developed testing system can be of six types: single-line answer, multi-line answer, choice of several correct answers, choice of one correct answer, logical sequence, pairwise matching [4].

One-line answer. In this type of questions, the teacher must enter the correct answer in text, for example, the result of solving the problem. The result of the student's answer is calculated using the Damerau-Levenshtein algorithm. If the answer is more than 70% similar to the correct one (can be changed in the system settings), then he gets 100% for this question, otherwise - 0% [5].

Multi-line answer. It is assumed that the student must give a detailed answer to questions of this type, which he introduces himself. The student also enters the answer to this type of question independently (example of a question: write the definition of the computer network topology). The result is determined by the teacher manually in the range from 0% to 100%, depending on the completeness of the answer [5].

Multivariate question. This type of question includes a list of several options, in which there may be several correct, all correct, or even no correct answer. The teacher needs to add possible answer options and choose the correct ones. The result is calculated as the difference between the percentage of selected correct and incorrect answers (will be discussed in more detail in this article). If a negative value is obtained, for example, the student noted only incorrect answers, the result is taken equal to zero (0%) [5].

Univariate question. In this type of question, there can be only one correct answer. Therefore, the student receives for this question 100% if he chose the correct answer, and 0% in any other case [5].

Logical sequences. Here the teacher needs to create a logical chain. The teacher should immediately add a chain with the correct arrangement of elements. The student will display it chaotically. The result is the percentage of the elements in the correct positions to the total number of elements. This type of question can be very useful for assessing knowledge of algorithms [5].

Pairwise matching. This type of question involves making logical pairs. The teacher should immediately create knowingly correct pairs. For a student, during testing, they will be mixed. The result for this test is the percentage of correctly formed pairs in relation to the total number of pairs [5].

2.2 Method of calculating test results

Consider further the system for evaluating the answers to questions. Each question in the test is assessed separately in a percentage from 0 to 100. The system checks the answers automatically if possible. The formula for calculating the result (*R_t*) of the *i* question is as follows.

$$R\_t_i = \left(\frac{(k_{correct}^{marked})_i}{(k_{correct})_i} - \frac{(k_{incorrect}^{marked})_i}{(k_{incorrect})_i} \right) \cdot 100\%, \tag{1}$$

where $(k_{correct}^{marked})_i$ and $(k_{incorrect}^{marked})_i$ – respectively, the number of correct and incorrect answers marked by the student within the *i* question, $(k_{correct})_i$ and $(k_{incorrect})_i$ – respectively, the total number of correct and incorrect answers within the *i* question.

Note that this form of assessing answers to a question allows you to take into account partially correct answers, but at the same time encourages the student to be attentive and judicious when choosing answers.

The total test result (*R_test*) is determined by the following formula:

$$R\_test = \frac{\sum_{i=1}^{n} ks_i \cdot R\_t_i}{100 \cdot n} \cdot 100\%, \tag{2}$$

where ks_i – difficulty coefficient of the *i* question, *n* – number of questions in the test.

Such a formula for calculating the final test result gives the teacher complete freedom in choosing the coefficients of the difficulty of the questions.

2.3 Test organization methodology

The process of creating a test involves choosing a test type, subject, list of topics, maximum number of questions and difficulty. Complexity is relevant only for the type of random test with a complexity factor. There are three types of tests in the system: test by variants; test with random questions of the same difficulty; test with random questions taking into account difficulty coefficients.

The first type assumes that in addition to creating a test, the teacher will also need to create variants and add questions to them. When scheduling a test, the teacher selects a variant to be uploaded to students. A test with random questions of the same difficulty, although it seems possible, in practice, in technical disciplines, including disciplines of the IT profile, is actually not realizable, because in any discipline there is always both harder and easier material, both more important for the formation of the relevant competencies in a future specialist, and less important. Therefore, for an objective assessment of students' knowledge, the use of questions belonging to different categories of difficulty is required (third option). The number of difficulty coefficients and their values are not limited by the system. The contribution of each question to the final result should be calculated taking into account the difficulty coefficients assigned to the questions. For the correct formation of a test with a random set of questions and different difficulty coefficients, it is necessary to introduce the concept of test difficulty ($\underline{D_{test}}$), which will be defined as the sum of the question difficulty coefficients (ks_i) that make up the test (N is the number of questions in the test).

$$D_{test} = \sum_{i=1}^{N} Ks_i. \tag{3}$$

The main task of using a difficulty map is to form tests for students with the same total difficulty. In the proposed system, this is implemented as follows: the teacher creates a map of the difficulty of the test − how many questions and what category of difficulty the student should select at random by the system. For example, it is necessary to form a test consisting of 30 questions with three categories of difficulty and a total difficulty equal to 60. Several variants of the test difficulty map are presented in the following Fig. 1.

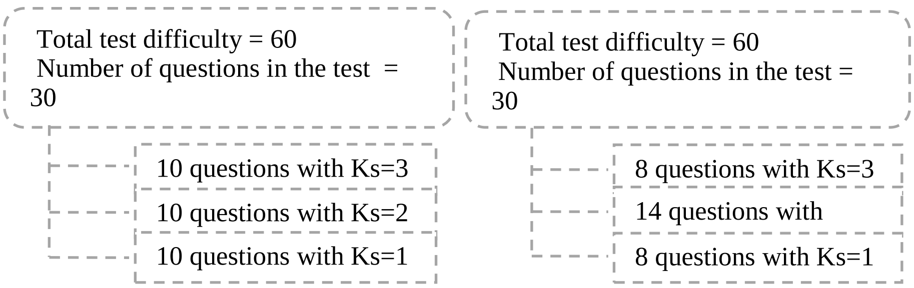

Fig. 1. Examples of test difficulty map

This logic of test planning can be complicated by adding restrictions on the number of questions from certain parts of the discipline, which will be useful for uniform control of knowledge throughout the educational material.

The teacher can also schedule tests for other teachers. This mechanism is also available in the operator role. Based on this, it should be noted that the teacher fully includes the operator's mechanism.

After completing the test, it is necessary to check students' answers. With automatic verification (depending on the test settings and the types of questions used), the results will immediately go to the verified ones and will appear in the students' personal account, otherwise the teacher checks them manually.

2.4 Experience of using testing for knowledge control

Generally, the test results with a competent organization of the knowledge control process, as well as a high-quality preparation of test questions, can give an objective assessment of knowledge. In the course of the research, the following scale of test results and the corresponding assessments were accepted: the range of test results 60% - 64.99% corresponds to the mark “three (3)”, the range of results 65% - 69.99% − the mark “four (4)”, the range of 70% -74.99% − the mark “five (5)”, in the range 75% - 79.99% − the mark “six (6)”, in the range 80% - 84.99% − the mark “seven (7)”, in the range 85% -89.99% − the mark “eight (8)”, the range of 90% - 94.99% − the mark “nine (9)”, and finally the mark “ten (10)” is given to the results in the range of 95% -100%.

The experience of using testing on the disciplines "Computer networks", "Administration of information systems" and other technical disciplines has shown that the following types of questions are the most effective for assessing students' knowledge:

1. Multivariate question - used to create complex questions, to answer which the student needs to be able to logically compare (combine) various educational information on the discipline.
2. Logical sequences - used to test students' knowledge of algorithms, sequences of actions that must be performed to solve a problem or within processes. An example is the algorithm of the CSMA / CD access method, divided into logically completed actions (blocks), which are showed for student in the test in a random order.
3. To check the results of solving problems (for example, to determine the broadcast address for a given network or to determine the network or host identifier by a known IP address/mask), you can use either a one-line answer with the student's input of the obtained result, or a single-variant answer with the choice of one single correct answer.

In general, when developing a data analysis subsystem, the main attention was paid to the issues of correct interpretation of the level of students' knowledge in the form of test results, as well as the perception of the information received. To conduct a meaningful analysis of test results, each subject area of educational activity was considered as an information system, consisting of a certain amount of educational elements. At the same time, it is supposed that the test tasks are developed with thw aim to cover the content of each block and the most important educational elements. Therefore, according to the result of solving each test task, it is possible to determine the quality of mastering not only the course as a whole, but also each block, and in it - those elements of knowledge that are presented in this test task. In the developed subsystem for analyzing test data, it is supposed that the teacher pre-selects a certain scale of results gradation, as well as the minimum threshold at which the success of the test is established by students. Obviously, such a formation of test results allows teachers to set different requirements for the level of knowledge of students, depending on the specific test or subject. It is necessary also to note that the teacher has access to data analysis for the following categories: subject, test, questions and students. The teacher simply selects the desired statistics display option, performs parameter settings depending on the analyzed criterion. Further, in accordance with the set criteria, the analyzed information will be displayed in a graphical form.

It should also be noted that the teacher has access to data analysis for the following categories: subject, test, questions and students. The teacher simply selects the desired statistics display option, performs parameter settings depending on the analyzed criterion. Further, in accordance with the set criteria, the analyzed information will be displayed in a graphical form.

3 Methodology for the formation of a student's overall rating by discipline

Now let's consider the methodology for taking into account test results in the final assessment of knowledge in the discipline. Suppose that the discipline is logically divided into two components: theoretical material, knowledge of which will be assessed by testing, and practical skills (results of laboratory or practical work). Note that the second component can be assessed both by the teacher through the process of defending the work, and through testing with prevailing practical issues and tasks.

To begin with, two maps of the significance of the results are drawn up: one for tests on theoretical material, the second for laboratory work. As a rule, all theoretical material is divided into a several tests (in this example there will be four), each of which is assigned weight coefficients, the sum of which is 100 (Table 1).

Table 1. Coefficients of significance of tests within the theoretical material of the discipline

№ test	1	2	3	4
Test significance coefficients	15	30	35	20

The resulting test score can be calculated using the following formula:

$$R\_theory = \frac{\sum_{i=1}^{n}(k\_test_i \cdot R\_test_i)}{100 \cdot R\_\max}, \quad (4)$$

where $R\_test_i -$ result of *i* test (mark), k_test_i − significance coefficients *i* tests, $R\_\max -$ maximum test mark.

Similarly, it is necessary to calculate the final grade for the practical component of the discipline (for example, laboratory work), introducing also the significance coefficients of laboratory work, which actually evaluates their contribution to the general practical training within the discipline. Suppose that within the framework of discipline, 9 laboratory works are performed with the following significance coefficients, the sum of which is also equal to 100 (Table 2).

Table 2. Coefficients of significance of laboratory work in the framework of the practical material of the discipline

№ laboratory work	1	2	3	4	5	6	7	8	9
Significance coefficient of laboratory work	8	15	15	12	5	24	8	8	5

Now you can calculate the resulting grade for the practical component of the discipline using the following formula:

$$R\_practice = \frac{\sum_{i=1}^{n}(k\_practice_i \cdot R\_practice_i)}{100 \cdot R\_\max}, \quad (5)$$

where $R\_practice_i -$ result of *i* laboratory work (rating), $k\_practice_i$ − significance coefficient of *i* laboratory work, $R\_\max$ − maximum rating for laboratory work.

Next, it is necessary to determine the influence of all components (theoretical material, laboratory work) on the final assessment, i.e. it is necessary, again, to assign coefficients that determine the contribution of each of the discipline components to the overall result (Table 3).

Table 3. Coefficients that determine the contribution of theoretical and practical (laboratory) material to the final assessment

Significance coefficient of theoretical material	4
Significance coefficient of practical (laboratory) material	2
Significance coefficient of individual task	2

Note that there may be more components of the final grade, and each of them should have its own coefficient.

The final rating (mark) *R* is defined as the arithmetic mean, taking into account the established weight coefficients according to the following formula:

$$R = \frac{\sum_{i=1}^{n}(k_i \cdot R_i)}{\sum_{i=1}^{n} k_i}, \quad (6)$$

where k_i − the coefficient of the contribution of theoretical (practical) material to the final assessment, R_i − result (rating) in the theoretical (practical) component of the discipline.

In general, the presented methodology for determining the final rating (mark) of a student in a discipline allows the teacher to freely determine the contribution of various components of the discipline, and therefore will give a correct representation of both the theoretical and practical skills of the student, and will show his level of competence.

4 Conclusions

Thus, a method for calculating the final rating of a student's knowledge is proposed, which is based on combining, taking into account the significance of test results (rating) for various components of the discipline. The number of components that influenced the final results is not limited (teacher can take into account control works, colloquia, seminars, results of individual tasks, etc.), and is determined by the teacher depending on the specifics and characteristics of the discipline materials (in the examples considered, the priority was given knowledge on theoretical material, which, for example, is typical for fundamental disciplines). It is important to correctly and reasonably determine all the necessary significance coefficients (tests, laboratory work, etc.). It can be proposed that this technique, with a competent and thoughtful teacher approach to the formation of the questions base, to the determination of the necessary coefficients, will increase the objectivity of the knowledge rating and practical skills in the discipline, including distance learning.

References

1. Mayorov, A. N.: Theory and practice of creating tests for the education system. National education, Moscow (2010).
2. Noskov, M.V., Dyachuk, P.P., Dobronets, B.S., etc.: Evolution of education in the conditions of information. Krasnoyarsk (2019).
3. Skorin, Y. Shcherbakov, O.: Introduction of electronic test tools with the aim of increasing the effectiveness of the knowledge test. Development Management, 4(4), 14-25 (2018).
4. Avanesov, V. S.: The form of test tasks. Test Center, Moscow (2005).
5. Romanenko, D. M., Kudlatskaya, M. F.: Features of the organization and conduct of computer testing for knowledge control in the disciplines of IT specialties. Information technologies in education, science and production: V International scientific-technical Internet conference, November 18-19, 2017 Section Information technologies in production and scientific research [Electronic resource] (2017).