Development of knowledge testing systems based on discrete optimization models

L A Zaozerskaya[1], **V A Plankova**[1]

[1] Sobolev Institute of Mathematics, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

E-mail: zaozer@ofim.oscsbras.ru, plankova@ofim.oscsbras.ru

Abstract.
We present an approach to creating computer testing systems and an overview of our results obtained based on it. Special attention is paid to the problem of forming the optimal test content. This problem consists of finding a subset of the test tasks allowing us to get an objective conclusion about the degree of assimilation by a student the training course. Several formulations are described for solving this problem and the problem of forming an optimal test kit. Corresponding mathematical models are constructed on the base of the set covering problem and its generalizations. A specialized computer testing system, developed by us based on this approach, is described. The experience of this system use and the prospects of the proposed approach for creating knowledge testing systems are discussed.

1. Introduction

Various testing problems arise in manufacturing, economics, transportation, education, and other areas. Testing is widely used to verify product quality, identify customer preferences, determine the level of training of specialists. For example, in [1], this method is applied to study knowledge of junior medical employees in the field of modern research related to their professional field.

Many decision-making problems arising in real applications can be solved using discrete optimization models. A set of the integer programming problems that model practical problems arising from scheduling, routing, facility location, etc is presented in [7]. Using of knapsack problems for resource allocation can find in [12]. One from supply management problems is considered in [22].

One of the problems encountered during testing is to minimize the number of tests necessary to analyze all the properties of an investigated object. This problem can be solved using the set covering problem (SCP) [6]. The SCP is a well-known problem of discrete optimization and has numerous applications. In the field of education, discrete optimization models are used, for example, for the timetabling problem of university courses [3], and for distributing the academic load between teachers [23]. In [9] the knapsack problem models the way students are graded after the test.

Preparation of high-quality knowledge testing is a difficult, creative and very time-consuming process included the development of test tasks creating a database and test task variants. In the literature, there are a large number of works devoted to educational testing, for example, the choice of the type of test task, its validity and reliability as well as the development of tests in individual subjects. For example, strategies using a research-validated multiple-choice question sequence for leaning one of the topics of quantum physics to study are is explored in [11]. In [16], the application of a diagnostic tool which is a survey to investigate students' understanding of introductory quantum physics concepts, is analyzed.

In this paper, we will be interested in the problem of forming the test content, i.e., selection of test tasks from those available for inclusion in the test. The test content must accord to the given goals, the level of training of students and some other criteria. We suggest using Boolean programming models to solve this problem.

The modern information technologies provide extensive opportunities for automating the learning process, including the stage of knowledge control and testing as one of its forms. In universities, computer testing is widely used at the stages of intermediate, thematic and final control. To create computer tests on various academic disciplines of higher professional education, universal integrated development environments of tests or tools of e-learning systems are used most often. In such systems, usually, there is no stage of automation of the process of forming the optimal test content. Tests are created based on a scenario developed by the teacher within the capabilities of the shell used.

In this paper, we describe an approach to the development of computer systems of knowledge testing in which the models and methods of discrete optimization applied to solve the problems of forming the optimal content of the test. For this, we introduced the notions of the knowledge element of the studied course, test task template, and test structure. Using these notions, the problem of finding the optimal structure test [18] and the problem of the optimal test kit [19] are formulated. The models of Boolean programming are constructed for indicated problems, and algorithms for solving them are discussed.

Based on this approach, we have created a computer testing system EmmTest on Linear Programming of the university discipline Economic-mathematical Methods [20]. In this system, the individual test generation occurs in an on-line mode following the found optimal test structure. Variants of most test tasks are formed at the time of testing using specially designed procedures. The results of using system EmmTest for training and exams are discussed.

2. Mathematical models for knowledge testing problems

We will consider the several problems of forming the test content when it is required to choose such a subset of tasks that meets the specified testing goals, for example, for intermediate, thematic or final control of knowledge.

2.1. Forming the test for separate theme

As we have already noted, the set covering problem serves as a suitable mathematical model of the problem of testing knowledge for the small training course. Let the training course (or its separate topic) is considered as a set of knowledge elements, that is, definitions, properties, statements, algorithms, etc. Each test task checks a specific subset of knowledge elements. The test size is determined as the number of tasks included in the test. It is required to construct a test that guarantees the quality control of student assimilation of the entire training course and has a minimum size.

So we assume that a set from m knowledge elements of the training course has been formed and there are n test tasks, each of which is intended to test a certain subset of knowledge elements. A boolean matrix $A = (a_{ij})$ of size $m \times n$ is given. Column j of this matrix is the characteristic vector of test task j, i.e., $a_{ij} = 1$ if task j checks knowledge element i otherwise $a_{ij} = 0$ for $i = 1, ..., m$, $j = 1, ..., n$.

We introduce the Boolean variables as follow $x_j = 1$ if task j is included in the test and $x_j = 0$ otherwise for $j = 1, ..., n$. Now the problem of forming the test of the minimum size can be written as unweighted SCP

$$\sum_{j=1}^{n} x_j \to \min \tag{1}$$

subject to

$$\sum_{j=1}^{n} a_{ij}x_j \geq 1,\ i = 1, .., m, \tag{2}$$

$$x_j \in \{0, 1\},\ j = 1, ..., n. \tag{3}$$

The objective function (1) is equal to the size of the test, and it is minimized. Constraints (2) express the conditions that each element of knowledge must be tested in at least one test task. Many results on the structure and complexity of this problem, its solving algorithms, and the experimental research can find in the literature. To search for the optimal solution, hybrid methods based on branch and bound algorithms [4, 5] and L-class enumeration algorithm [8] can be used. Algorithms from [10, 15] can be used to find approximate solutions.

We describe simple modifications to the consider testing problem. Let c_j be the average time to solve test task j. It is required to find the test checking all knowledge elements and having a minimum average execution time. In this case, we obtain the weighted SCP with the objective function $f(x) = \sum_{j=1}^{n} c_j x_j$.

To increase the reliability of the test, we require that the knowledge element i be checked at least in b_i test tasks for $i = 1, ..., m$. Then the conditions (2) take the form

$$\sum_{j=1}^{n} a_{ij}x_j \geq b_i,\ i = 1, .., m$$

and the model of such a testing problem is a generalized SCP.

The use of model (1)–(3) was successful for small topics. For testing throughout the discipline, the size of the test can be quite large, which makes it impossible to complete the test in the time allotted for the exam. Therefore, we propose other formulations of the problem of forming the optimal test content for various types of knowledge control including the final one. The corresponding mathematical models are given.

2.2. Forming the test for training course

To formalize the problems of knowledge control for the training course, we introduce the following notions. Recall that the training course is considered as a set of knowledge elements, i.e., most significant definitions, properties, statements, algorithms, etc. The typical test task or the test task template is a general formulation of the task for checking a certain subset of knowledge elements. A general formulation of the task includes the statement of the task, the suggested answers, and the rule or algorithm of generating the source data for a variant of the typical task. The source data is numeric or graphic data. Note a change of the source data does not affect the checked knowledge elements subset.

The test structure is the set of typical test tasks in the test and the test size is the number of tasks in such set. Optimal test structure (OTS) is a structure of test that allows us to make an objective conclusion about the degree of student's assimilation of discipline and satisfies a certain optimization criterion. Thus, an individual test consists of variants of typical test tasks determined by its optimal structure.

As noted above, if the test structure for the final testing obtained as the decision of set covering problem then it can include a large number of tasks. In this regard, we proposed several formulations of the optimal test structure forming problem [18]. They are based on the partition of the knowledge elements set on a basic set (it includes key elements of discipline) and an additional set (the rest elements).

We consider the following formulation of the OTS problem forming. Suppose the range of the total difficulty of typical tasks of the test structure and the multiplicity of checking the base set elements are set. It is necessary to find the test structure that the total importance of checked knowledge elements from the additional set is maximized under indicated conditions.

To construct a mathematical model, we introduce the following notations. Let the sets of indices $B = \{1, \ldots, m_1\}$ and $C = \{1, \ldots, m_2\}$ correspond to elements of the base and additional sets. Let n be the number of developed typical test tasks, and k the number of tasks in the structure of the test, i.e., its size. We denote by A a Boolean matrix of size $m_1 \times n$. Here $a_{ij} = 1$ if j-th task is intended to check i-th element of B otherwise $a_{ij} = 0$ for $i = 1, ..., m_1$, $j = 1, ..., n$. The Boolean matrix $\tilde{A}$ of size $m_2 \times n$ is defined similarly for knowledge elements from C and the set of test tasks. Let all test tasks be separated on l groups according to a certain attribute (for example, by themes or type of task: analytical, graphical, algorithmic, etc.). Denote the set of task numbers in the group t by J_t for $t = 1, \ldots, l$.

We introduce variables $x_j = 1$ if j-th task is included in the test structure and $x_j = 0$ otherwise for $j = 1, ..., n$. Further introduce auxiliary Boolean variables v_i for $i = 1, ..., m_2$ as follow. If $v_i = 1$ in some feasible solution then i-th knowledge element from C is checked by the selected test structure.

Let c_i be the importance of checking i-th knowledge element from C for $i = 1, \ldots, m_2$; p_j be the complexity (or the execution time) of j-th task for $j = 1, \ldots, n$; h_0, h_1 be the lower and upper bounds on difficulty of the test structure; b_i be the smallest number of tasks in the test structure for checking i-th knowledge element from B for $i = 1, \ldots, m_1$; d_t^0, d_t^1 be the minimum and maximum number of typical tasks included in the test structurefrom the group t for $t = 1, \ldots, l$. All parameters c_i, h_0, h_1, b_i, p_j, d_t^0 and d_t^1 take positive values.

The corresponding mathematical model is as follows:

$$\sum_{i=1}^{m_2} c_i v_i \to \max \tag{4}$$

subject to

$$h_0 \le \sum_{j=1}^{n} p_j x_j \le h_1, \tag{5}$$

$$\sum_{j=1}^{n} a_{ij} x_j \ge b_i, \quad i = 1, ..., m_1, \tag{6}$$

$$\sum_{j=1}^{n} \tilde{a}_{ij} x_j - v_i \ge 0, \quad i = 1, ..., m_2, \tag{7}$$

$$d_t^0 \le \sum_{j \in J_t} x_j \le d_t^1, \quad t = 1, ..., l, \tag{8}$$

$$x_j \in \{0, 1\}, \quad j = 1, ..., n, \tag{9}$$

$$v_i \in \{0, 1\}, \quad i = 1, ..., m_2. \tag{10}$$

Here $(x_1, ..., x_n)$ and $(v_1, ..., v_{m_2})$ are the characteristic vectors of the structure and of the subset of checked knowledge elements from the additional set. Inequalities (5) indicate the bounds on the total complexity of typical tasks included in the desired structure. The constraints (6) reflect the requirement of multiple verification of the knowledge elements of the basic set. Any variable v_i is included in only one inequality from (7). If element i from C is not checked by x then $v_i = 0$. Otherwise, the value of v_i is limited only by condition (10). Therefore, for any vector x satisfying (5), (6), (8), (9), there is admissible solution (x, v) that $v_i = 1$ if and only if element i from C checked by x. Since the objective function (4) must be maximized then for the optimal solution, it will be equal to the total importance of the elements being checked from the additional set. For each group, conditions (8) specify the minimum and the maximum number of typical tasks for including to structure. That ensures the variety of the formed structure.

Problem (4)–(10) is NP-hard [2]. However, if this problem has a relatively small dimension then it can be solved by one of the known integer linear programming algorithms. The variant of model (4)–(10) was successfully used by us when developing a computer testing system Emmtest. To solve the corresponding problem, we used the modification [17] of the L-class enumeration algorithm [8, 13]. This algorithm is based on the method of regular partitions [14].

2.3. Optimal kit of test structures

With a significant volume of the training course and limited testing time, using only one test structure may cause to the exclusion of a part of the studied material from the testing process. Therefore, when testing a group of students, we suggest using not one but several test structures (i.e., kit of structures). In this case, checking the knowledge of each student in the group, at the same time, we will get results on the assimilation of any element of knowledge of the studied course. This is provided the teacher with feedback and, therefore, allows the reasonable adjustment of the educational process.

Clearly, if the number of test structures in the resulting kit will be large, then for obtaining an objective conclusion, it may be no provided the necessary frequency of checking each knowledge element from the additional set. To increase the information content of the testing results, it is necessary to minimize the cardinality of the kit of the test structures.

We offer one of the formulations of the forming the optimal kit of test structures problem. It is required to find the kit of minimum power from fixed-size test structures. Moreover, all elements from the base set must be checked in each kit structure and any element from the additional set is checked at least in a given number of structures of this kit.

We assume that for the number of test structures in the formed kit, the upper bound T is given. Let k be the size of any structure being formed. We require that each element of the complementary set be checked at least d times. We introduce the variables z_t, x_{tj}, y_{ti} for $t = 1, ..., T$, $j = 1, ..., n$, $i = 1, ..., m_2$. Here $z_t = 1$ if the structure t is formed, and $z_t = 0$, otherwise; $x_{tj} = 1$, if test task j is included in structure t, and $x_{tj} = 0$ otherwise; $y_{ti} = 1$ if element i of C is checked in t-th structure, and $y_{ti} = 0$ otherwise.

The mathematical model of the considered problem has the form:

$$\sum_{t=1}^{T} z_t \to \min \tag{11}$$

subject to

$$\sum_{j=1}^{n} x_{tj} = kz_t,\ t = 1, ..., T, \tag{12}$$

$$\sum_{j=1}^{n} a_{ij} x_{tj} \geq z_t,\ t = 1, ..., T,\ i \in B, \tag{13}$$

$$y_{ti} \leq \sum_{j=1}^{n} \tilde{a}_{ij} x_{tj} \leq k y_{ti},\ t = 1, ..., T,\ i \in C, \tag{14}$$

$$\sum_{t=1}^{T} y_{ti} \geq d, \ i \in C, \tag{15}$$

$$z_t, x_{tj}, y_{ti} \in \{0,1\}, \ t = 1, ..., T, \ j = 1, ..., n, \ i \in C. \tag{16}$$

Here, the objective function (11) minimizes the power of the formed test structures kit. Conditions (12) describe the requirements for including exactly k tasks in the structure t if it is formed; otherwise, the structure does not contain tasks. Constraints (13) provide verification of all elements of the base set in each formed test structure. The inequalities (14) guarantee that variable y_{ti} is equal to one if and only if in t-th structure there is at least one task checking i-th element from set C. Conditions (15) guarantee the check of each element from the additional set at least by d structures from any kit.

3. Computer testing system EmmTest

The specialized computer testing system EmmTest for knowledge verification in the discipline of Economic-mathematical Methods [21] was created on base the presented approach.

The constructed models and methods of Boolean programming are used in the system when forming tests on the Linear Programming. This topic is an important part of the disciplines Game Theory and Operations Research, System Analysis, Optimization Methods, etc. Therefore, the EmmTest system can be widely used under teaching students of mathematical, economic, technical and other specialties.

The model (4)–(10) was used under the assumption that checking all elements of the additional set is equally important ($c_i = 1$ for all i), and test tasks are considered the equal in complexity or in the time allotted for their execution ($p_j = 1$ for all j). In addition, we assume $h_0 = h_1 = k$. The restriction (5) in this case takes the form

$$\sum_{j=1}^{n} x_j = k.$$

Here k is the size of the test. Any b_i equals 1, i.e., for any element from the additional set, there is at least one test that checks this element. From each group, at least one task is included in the test ($d_t^0 = 1$, $d_t^1 = |J_t|$ for all t).

When developing specialized computer testing systems, one of the most time-consuming steps is the preparation of typical test tasks and the definition for each test task of knowledge elements subset that are tested by it. This difficult task requires a generalization of the teachers' experience in this discipline.

In system EmmTest, 31 typical test tasks was realizated on Linear Programming. Each test task is an exercise or task to check knowledge elements at the level of reproduction of information from memory and its application in typical situations.

The system uses two ways to implement typical test tasks. In the first method, the test task numerical data is formed by a specially developed algorithm for it using a random number sensor. This guarantees a variety of individual variants for this test task. In this case, the full automation of the test variant forming process is provided. We note some positive aspects of this method. The participation of the subject teacher in creating the system is required only when developing typical test tasks and algorithms for generating their variants. Also, no test task database storage required. The second method is designed for multi-stage tasks. They are used to test the knowledge and skills of applying iterative algorithms for solving optimization problems. To form variants of such tasks, we created special procedures that allow the teacher to select data suitable for testing in a dialogue mode and save them in a database without the participation of a programmer.

The system EmmTest consists of two independent applications: a teacher's module and a testing module.

The teachers module is designed to tuning the system for the testing session, depending on the goals, preparation of the initial data of multi-stage tasks and processing of test results. The most important part of the described module is the Control Unit. The teacher uses this block to create a test or the tests kit on Linear programming. He sets the size of the test and determines in the list all the basic elements of knowledge, that is, elements subject to mandatory check.

For example, for the task of forming the OTS in the system for given data, the corresponding Boolean programming model is built. Further, it is solved by the L-class enumeration algorithm [17]. The result is the task numbers included in the optimal test structure. When students work with the testing module the individual test variants will be created in the on-line mode in accordance with the formed optimal structure.

To form the optimal kit of test structures, we used an iterative process [19] of a greedy type based on applying the describe above variant of the problem of forming the OTS. First, problem (4)–(10) is solved for a given set C of additional knowledge elements. The resulting optimal test structure is included in the required kit. All elements that are checked by this optimal structure are removed from C. For lessened C, the indicated variant of problem (4)–(10) is again solved. The procedure ends with the current set C empty.

4. Conclusion

We presented a new approach to solving the problems of knowledge testing. It is based on the use of discrete optimization models. Substantial statements of the problems of forming the optimal test structure and constructing optimal kit of test structures are proposed. Corresponding mathematical models constructed in the form of Boolean programming problems and algorithms for solving them developed.

Based on these models, the authors created a specialized computer system EmmTest to knowledge control in the discipline Economic-mathematical Methods. In it, the optimal structure of the test on the topic Linear Programming is formed in accordance with the current data set by the teacher. An individual variant of the test for the student is generated in accordance with obtained optimal structure.

The features of the tasks in this discipline allowed us to generate variants of the majority of test tasks using specially developed algorithms in on-line mode. In combination with the ability to control the test content depending on current goals, this allowed us to automate the process of creating tests and the testing itself.

In our opinion, the approach using discrete optimization models to solve problems arising in the field of knowledge testing is promising. It can be used both in creating specialized computer testing systems on various educational disciplines (in particular, mathematical ones) and in developing universal integrated environments for test creation.

Acknowledgments

For the first author, the work was supported by the program of fundamental scientific researches of the SB RAS I.5.1, project 0314-2019-0019.

For the second author, the work was supported by the program of fundamental scientific researches of the SB RAS I.1.3., project 0314-2019-0009.

References

[1] Abuelhia E 2016 Awareness of ionizing radiation exposure among junior doctors and senior medical students in radiological investigations *J Radiol Prot* **37** 1 pp 59-67

[2] Ageev A A and Sviridenko M I 1999 Approximation Algorithms for Maximum Coverage and Max Cut with Given Sizes of Parts *Integer Programming and Combinatorial Optimization. Lecture Notes in Computer Science*, ed Cornuejols G, Burkard R E and Woeginger G J **1610** (Berlin, Heidelberg: Springer) pp 17-30

[3] Avella P, Vasilev I 2005 A computational study of a cutting plane algorithm for university course time-tabling *Journal of Scheduling* **8** 6 pp 497514

[4] Balas E and Carrera M C 1996 A Dynamic subgradient-based branch-and-bound procedure for set covering *Operations Research* **44** 6 pp 875-890

[5] Caprara A, Toth P and Fischetti M 2000 Algorithms for the set covering problem *Annals of Operations Research* **98** pp 353371

[6] Christofides N 1975 Graph theory: An Algorithmic Approach (New York: Academic Press Inc.)

[7] Conforti M, Cornuejols G, Zambelli G 2014 Integer Programming Models *Integer Programming. Graduate Texts in Mathematics* (Cham: Springer) **271** pp 45-84

[8] Eremeev A V, Kolokolov A A and Zaozerskaya L A 2001 A Hybrid Algorithm for Set Covering Problem *International Journal of Mathematical Algorithms. Special issue: International workshop on discrete optimization methods in scheduling and computer-aided design*, ed Dolgui A, Gordon V and Proth Jean-Marie (Minsk, Belarus) 6 p

[9] Feuerman M and Weiss H 1973 A Mathematical Programming Model for Test Construction and Scoring *Management Science* **19** 8 pp 961-966

[10] Hochbaum D S 1997 Approximating Covering and Packing Problems: Set Cover, Vertex Cover, Independent Set, and Related Problems *Approximation algorithms for NP-hard problems*, ed D S Hochbaum (New York: PWS Publishing Company) pp 94-143

[11] Justice P, Marshman E and Singh C 2019 Improving student understanding of quantum mechanics underlying the SternGerlach experiment using a research-validated multiple-choice question sequence *European Journal of Physics* **40** 055702

[12] Kellerer H, Pferschy U and Pisinger D 2003 *Knapsack Problems* (Berlin, Heidelberg, New York: Springer)

[13] Kolokolov A A 1995 *Proc. 3rd IFIP WG-7.6 Working Conference on Optimization - Based Computer - Aided Modelling and Design* (Prague: IITA) pp 256–260

[14] Kolokolov A A 1996 Regular partitions and cuts in integer programming *Discrete Analysis and Operations Research. Mathematics and Its Applications*, ed Korshunov A D **355** (Dordrecht: Springer) pp 59–79

[15] Rubio A G et al 2016 Solving the Set Covering Problem with a Binary Black Hole Inspired Algorithm *Computational Science and Its Applications. Lecture Notes in Computer Science*, ed Gervasi O et al. (Cham: Springer) vol **9786** pp 207-219

[16] Sharma S and Ahluwalia P K 2012 Diagnosing alternative conceptions of Fermi energy among undergraduate students *European Journal of Physics* **33** 4 pp 883-895

[17] Zaozerskaya L A, Kolokolov A A and Plankova V A 2008 *Omsk Scientific Bulletin* **1** (64), pp 10-12

[18] Zaozerskaya L A and Plankova V A 2008 *Vestnik NSU. Series: Information Technologies* **6** 1 pp 47-52

[19] Zaozerskaya L A and Plankova V A 2012 *Omsk Scientific Bulletin* **3** 113 pp 33-36

[20] Zaozerskaya L A and Plankova V A 2013 The development of specialized computer knowledge control system using discrete optimization models *Proceedings include extended abstracts of reports presented at the VI International Conference on Optimization Methods and Applications "Optimization and Applications"* (Moscow: Dorodnicyn Computing Centre of RAS) p 168

[21] Zaozerskaya L A and Plankova V A 2013 On the development of specialized computer systems of knowledge control *Modern problems of science and education* 1

[22] Zaozerskaya L A and Plankova V A 2019 *J. Phys.: Conf. Ser.* **1210** 012164

[23] Zaozerskaya L A, Plankova V A and Devyaterikova M V 2018 Modeling and Solving Academic Load Distribution Problem *CEUR-WS (Proc. Int. Conf. OPTA-SCL, Omsk, Russia)* **2098** pp 438-445