THE SEMANTIC COMPLEXITY ESTIMATION OF THE LEARNING TASK EXPLANATION

R.V. Mayer

The Glazov Korolenko State Pedagogical Institute (RUSSIAN FEDERATION)

Abstract

The didactic process can be presented as the pupil's solution of the sequence of the learning tasks (LTs). The condition for successful learning is the correct sequence of the LTs presentation to the learner, in which an important didactic principle "from simple to complex" is realized. Therefore, the development of methods for rating the complexity of explaining the task is an actual problem of didactics. Its solution will allow to determine the didactic complexity of the LTs for the tests, exams and correctly assess the students' knowledge. LT should be understood as any educational task that requires intellectual action (tasks in mathematics, physics, chemistry, a specific formula derivation, proof of the theorem, etc.). As components of LT complexity various scientists in didactics call the numbers and complexity of elements, relations, closed circuits or loops, logical operations, formulas; the abstraction degree of the concepts and models; the presence of implicit factors that influence the study process; the redundancy of the LT conditions; the LT belonging to several types of tasks or to various subjects; the need for complex or cumbersome mathematical transformations, etc. At present, there is no effective method of "measuring" the complexity of the LT solution.

The purpose of the study is to develop an effective method for determining the didactic complexity of the LT solving, based on an analysis of its explanation and taking into account the used terms complexity. It is assumed that the didactic complexity of LT is proportional to the amount of semantic information in explaining its solution. To determine it, we propose calculating the total complexity of all terms. The methodological basis of this study is the works by Je.G. Gel'fman, M.A. Holodnaja, V.P. Bespal'ko, Ya.A. Mikk (the textbook theory), O.V. Zerkal', N.M. Solomatin (the semantic information), B. Davis, D. Sumara (the didactic objects complexity), A.I. Uemov, S.I. Shapiro (the knowledge folding), A.M. Sokhor (the educational texts informational capacity), Yu.A. Schrader (the thesaurus approach), V.I. Shalak (the content analysis), A.V. Gidlevsky, A.N. Kolmogorov, I.L. Lerner (the learning tasks complexity), N.K. Krioni, A.D. Nikin, A.V. Phillipova (the automated assessment of the text complexity), O.N. Podol'skaja, V.V. Vjazankova, K.V. Horoshun (methods of infometry in education).

The novelty of the work lies in the fact that the method for assessing the complexity of LT solving has been developed. It consists in the following: 1) to code the task condition and its solution (formulas and explanations) in text file F1; 2) to create text file F2, containing the list of terms used in the task solving; 3) to assess the terms complexity by calculating the number of words in the definitions; 4) with help of special program to analyze file F1 containing the task solving and determine the total complexity of the text, its volume and the average information folding coefficient. The complexity of 10 physical tasks from different sections of the physics course is estimated, the general informativeness of explanations and the average folding coefficient of information are determined. It has been established that the overall information content of the standard task explanation varies quite significantly (from 70 to 760 conventional units of information), and the average coefficient of folding is in the range of 1.5 – 7.7. The method of automated terms counting allows to evaluate the didactic complexity of any information block containing an explanation of the LT solution, the formula derivation, the theorem proof, etc.

Keywords: informativeness, information folding, semantic information, complexity, thesaurus, educational task.

1 INTRODUCTION

One of the modern didactics urgent problems is the development of the qualitative assessment methods of the educational texts complexity. Its decision will allow to estimate information volume and complexity of various didactic objects (manuals, test tasks, pictures, formulas) and processes (the subject studying, the task solution). Important characteristics of educational texts and other didactic objects (figures, tables, formulas) are the volume, complexity and general informativeness. The

ISBN: 978-84-09-14755-7

development of their measurement methods can help to implement the well-known didactic principle "from simple to complex" in the selection of the studied issues, to correctly assess the complexity of the test tasks, and thus more accurately determine the student's knowledge. The complexity of the studied issues (or used texts) should correspond to the students' intellectual level. According to the principle of complexity, to achieve the result we should choose a system with minimal complexity at a given level of quality [1]. In relation to the educational process this means that it is necessary to use those textbooks and methods for which the complexity–quality ratio is optimal. The textbook quality can also be characterized by the number of types of educational tasks that students are able to solve after studying the educational material presented in it.

An important component of the educational process is the solution by pupil the sequence of the learning tasks (LTs). The condition for successful learning is the correct sequence of the LT presentation to the learner, in which an important didactic principle "from simple to complex" is realized. Therefore, the development of methods for assessing the complexity of explaining the task is an urgent problem of didactics. Its solution allows to determine the didactic complexity of the various test items, to choose the LTs for textbooks and task-books correctly. In the broad sense of the word LT should be understood as any educational task that requires intellectual action (tasks in mathematics, physics, chemistry, a specific formula derivation, proof of the theorem, etc.).

Various scientists in didactics developed the problem of determining the LT complexity. For example, I.L. Lerner wrote that the complexity of the LT is due to: 1) the composition of the condition and the quantity of data; 2) the number of logical links connecting the task condition with the solution result; 3) the number of conclusions that can be made in the process of the task solving. According to the A.N. Kolmogorov's algorithmic concept, the complexity of the LT is proportional to the number of operations necessary to solve it. Therefore, it is determined by the length of the most rational algorithm for obtaining the correct answer, which includes the shortest way to understand the condition. O.B. Episheva and V.I. Krupich argues that complexity is an objective characteristic of the task which depends on the number of elementary actions (reasoning), links and types of links which should be taken into account by the pupil [2, p. 57]. V.M. Krotov proposes a table of the physical tasks complexity taking into account the solution structure, the number of phenomena, processes and objects, the number of unknown quantities, the explicit or implicit assignment of the task requirements, the mathematical apparatus complexity, the method of setting the condition [3]. At this he allocates drawing tasks, text, graphic, and experimental tasks. A.V. Gidlevskij uses a subject-predicate (i.e. logical) approach to estimating the difficulty of solving LT, based on the creation of appropriate graphological models [4] and taking into account the number of vertices and edges of the graph, which reflects the structure of the task solution. As components of the LT complexity, other researchers name the number and complexities of elements, relations, closed circuits, logical actions, formulas; abstractness degree of used concepts and models; the presence of implicitly defined factors affecting the process under study; the redundancy of LT conditions; the belonging of LT to several types of tasks or different topics; the need for complex or cumbersome mathematical transformations, etc. [5, 6]. At present, there is no effective method of "measuring" the complexity of LT solving.

The purpose of the research is to develop an effective method for determining the didactic complexity of LT solving, based on an analysis of its explanation and taking into account the complexity of the terms used. It is assumed that the didactic complexity of LT is proportional to the amount of semantic information in explaining its solution. To determine the task explanation informativeness, a thesaurus approach is used which demands the selection the elementary semantic units (words, elementary statements) in the LT explanation and their counting [7, 8].

2 METHODOLOGY

The methodological basis of the present study are the works by Je.G. Gel'fman, M.A. Holodnaja [9], V.P. Bespal'ko, Ja.A. Mikk [10] (theory textbook), O.V. Zerkal' [11], N.M. Solomatin [12] (semantic information), B. Davis, D. Sumara [13] (the complexity of the didactic objects), A.I. Uemov, S.I. Shapiro (folding of knowledge), A.M. Sokhor (information capacity of educational texts), Yu.A. Schrader (thesaurus approach), V.I. Shalak [14] (content analysis) N.K. Krioni, A.D. Nikina, A.V. Fillipova [15] (automated evaluation of the texts complexity), O.N. Podol'skaja, V.V. Vjazankova, K.V. Horoshun [16] (methods of infometry in education).

As basic we use the system analysis method, according to which the text containing the LT solution is considered as a system with the multilevel hierarchical structure. The text consists of information units (paragraphs, sections expressing a single idea), which contain one or more sentences. There are the following levels of the text understanding: words, sentences, information blocks, the text as a whole. To understand the meaning of the text (sentence), the student must establish links between the contained sentences (words).

Evaluation of the semantic complexity of the text containing the LT solution requires taking into account the abstraction degree of concepts used and the logical links between the individual statements. The content analysis method is used for this purpose. It consists in the application of a special computer program that calculates the terms used. To take into account their information capacity, the program should turn (address) to a text file which contains a list of terms with their complexities or abstraction levels.

3 DISCUSSION

Any learning task corresponds to a certain academic discipline, to one or several specific topics or themes. There are tasks on the Archimedes law, on the Pythagorean theorem, on the solution of a quadratic equation, etc. There are combined tasks that require knowledge of various topics, for example, on mechanics and electrodynamics. LT can be characterized by the subject, topic, quantity, keywords, etc. An important characteristic of the LT is the didactic complexity which is equal to the informativeness of its solution explanation.

When the pupil solves the task, he/she must guess or choose the solution method. This often requires additional assumptions, constructions, arguments, etc. The complexity assessment of these actions is practically impossible. Therefore, scientists usually talk about the difficulty of explaining the LT solution by the teacher. Proposed by A.V. Gidlevskij, subject-predicate approach to the assessment of the LT solving difficulty involves the analysis of problem solving by expert, building a graph structure of task solving, counting the number of vertices, links and duplicates [4]. This allows to create a scale of tasks difficulty. Trying to estimate the formula complexity, A.V. Gidlewski considers it as a system of four interconnected objects (1/2). He estimates the complexity of objects m and υ^2 as 2 points, and the coefficient 1/2 as 1 point, at this noting that he can not justify the criteria for such estimation. As a result, the complexity of the task on calculating the kinetic energy is equal to 5. According to this logic, the formulas $W = Li^2/2$ or $W = Cu^2/2$ must have same complexity. But it is not so! Pupils in 10–11th grades understand what mass and speed are, but they often find it difficult to explain what is called inductance, capacitance, etc. A.V. Gidlevskij's method does not take into account the abstractness degree of the used concepts and physical quantities.

The proposed method is based on the following ideas. To assess the LT solution complexity, we can write it on a sheet of paper (or create a text file), and then determine the didactic complexity of the resulting text. An important component of it is the text semantic complexity, which depends on the average abstractness of concepts, the complexity of ideas and reasonings. If the student has well-developed reading skills, then not the structure complexity (the average length of words and sentences), but the semantic complexity of the text has the basic influence on understanding the educational text in mathematics, physics, chemistry, biology, etc.

In general, the educational text is a simplified information model of its author's knowledge, the system of statements, according to which the student can recreate the knowledge itself. The student perceives the text not literally (as a computer program), but comprehends it, recreating the knowledge which the author of the text wanted to transfer. The same scientific knowledge can be transferred by different educational texts. The student extracts something from the information encoded in the text, which leads to the formation of new concepts in his consciousness and the links between them. Any sentence can be put in accordance with the semantic network; any text expressing a certain set of ideas corresponds to a finite set of semantic networks. The same set of ideas can be expressed in different ways, creating different sentences from the same set of terms. The resulting texts will have approximately the same difficulty.

Understanding the task explanation (proof of the theorem, logic derivation of the formula) leads to the pupil's penetration into the essence of the perceived material, the formation in his mind of ideas and meaningful generalizations that reflect objects, their properties and relationships, expressing relationships with other objects. At this the inclusion of new material into the system of the pupil's existing knowledge, the establishment of links between them occurs. In accordance with the main provisions of cognitive linguistics, a person thinks in concepts but expresses his thoughts verbally, creating sentences from words. Any thought is expressed in the form of the corresponding statement, and while it does not exist – it is not a thought, but a feeling. In accordance with the principle of economy of thinking, a person seeks the way to express his thoughts as shortly as possible. The most capacious concepts are used, the volume and content of which are optimized as a result of numerous applications by scientists.

Words with common root carry an approximately equal amount of semantic information, since they are not stored separately in the mind of a person, but are included in the single psycholinguistic union – a concept. For example, the terms “diffraction”, “diffractive” and “diffract” constitute the concept of “diffraction”, therefore, they have approximately equal information saturation. In the Russian language the informativeness of sentence does not depend from the words transposition practically. The word combinations “body moves quickly”, “quick movement of body”, “quick moving body”, “quickness of body movement” used in the text correspond to very similar semantic networks and therefore also carry approximately the same amount of information. But not everything is so simple. If the phrase “the brightness of the lamp” is in the physical text, then it should be noted that the scientific term “brightness” means a certain physical quantity, and its complexity (or informativeness) is significantly greater than complexity of the everyday concept “brightness”.

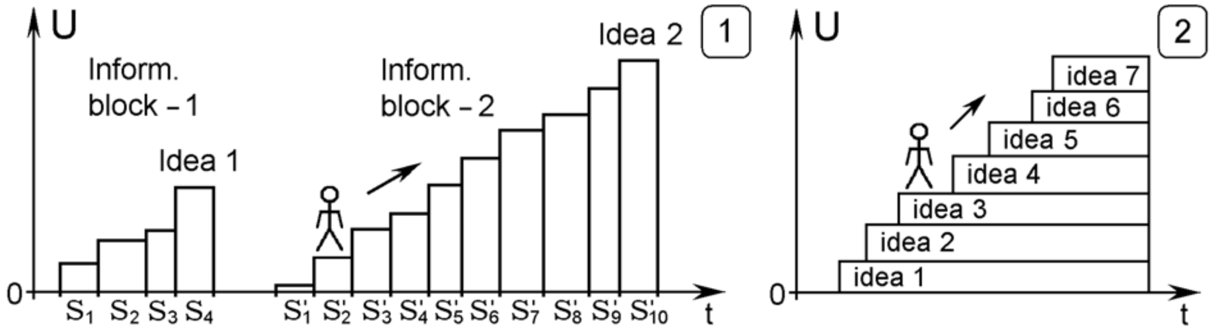

Figure 1. Understanding of the task explanation, like climbing stairs.

Let the teacher, explaining the LT solution or deriving a formula, consistently pronounce one sentence after another, demonstrate drawings, write down equations. All statements (verbal and mathematical) form an information block which consists of the logical reasoning, proving some idea. For the student, each statement is a logical task, a puzzle, which he solves in his head. If in the some information block IB-1 the concrete LT is solved (the formula is derived, the theorem is proved), the student, comprehending sentence by sentence (S_1, S_2, S_3, S_4), goes up the stairs to a new level of idea 1 understanding U (Fig. 1.1). After studying one information block, the student proceeds to the next. The stairs have different heights and lengths, and the student moves with a variable speed. The time of understanding of this or that statement (that is passing corresponding step) depends on its length and idea complexity. If the explanation of LT contains the simple statements, then steps are low, the student, without effort, appears at the top of the stairs. The more complex is (more difficult is) the sentence, the more the height of stair, on which the schoolchild should rise. If the stair is too high, the student can not understand the appropriate sentence, therefore it should be formulated easier or broken into two simple sentences, etc. Similarly, if the educational material contains several ideas, its understanding and assimilation is similar to going up a staircase, represented in a fig. 1.2.

Any explanation consists of sentences linked logically. Each sentence corresponds to a semantic network – the graph where vertices mean concepts, and edges – connections between them. The complexity of the LT explanation consists of the complexities of its statements components (verbal and mathematical). If, moving from one statement to another, the student has to realize (guess) some complex facts, which are not mentioned in the explanation, then this text is called tense or stressful. The LT explanation should have the sentences containing the most important information; at this the student does not feel any difficulty in understanding. Correctly composed LT explanation does not require much tension from the student, since all the important information is contained in the sentences explicitly (i.e. obviously, clearly). If the LT explanation is a stressful text, then it is necessary to bring together the signifier (statements) and the signified (transmitted thoughts), fill in the sense gaps (voids) so that the student with a thesaurus Z_0 can understand the explanation without stress (tension). The amount of information presented implicitly or latently should be close to zero, and can be neglected.

In computer science, the complexity of the system is determined by the minimum length of the binary message, which gives its full description. The human brain is not a digital computer, but a neural network operating with concepts; it creates and perceives information in the form of messages containing “simply” or “ordinary” words and scientific terms. Therefore, the amount of semantic information and semantic complexity of LT should be measured not in bits, but in phrases or words [7, 8]. A phrase is an elementary statement consisting of a subject, a predicate, an object and a circumstance that expresses a simple thought. This can be the fact that there is a phenomenon, the fact that the phenomenon has a certain property, or the fact that it has a relationship with other phenomena.

The text of the LT solution not only contains facts, but also includes logical considerations that are the result of speech-thinking actions, or such intellectual operations on the theoretical model of the research object as the conclusion justification, the hypothesis, the proof, etc. These operations are often clearly marked by the corresponding discursive words and expressions which serve the organizers of a scientific thought: “hence”, “therefore”, etc. These logical connections can be expressed implicitly. For example, consider the fragment of the text: “The circuit consists of a battery, a key and a lamp. When the key is closed, current flows through the wires. It heats up the spiral of the lamp, and it glows.” In this text, cause-and-consequence relationships are hidden; the student must identify them independently. Paraphrase the text so, that the connections are expressed explicitly: “{The circuit} contains {a battery, a key and a lamp}. IF {the key is closed}, THEN {current flows through the wires}. {The current heats the lamp coil}, SO {the lamp coil glows}”. In this case the following facts are obtained: F1: “the circuit contains a battery”; F2: “the circuit contains a key”; F3: “the circuit contains a lamp”; F4: “the key is closed”; F5: “the current flows through the wires”; F6: F4 – cause F5; F7: “the current heats the lamp spiral”; F8: “the lamp spiral glows”; F9: F7 are cause F8.

An elementary fact, i.e. a statement containing a subject and a relation, is often taken as a unit of semantic information [12]. The more facts, the more informativeness (meaningfulness) of the message is. To decompose the text into elementary phrases and to estimate the complexity of each fact “manually” is a laborious task. In our case, it is convenient to take a simple concept expressed by a word as a measurement unit of semantic information. If to apply the computer, this allows to automate the calculation of scientific terms and take into account their complexity. For the conditional unit of information (CUI) we take the amount of information contained in one significant word that does not require an explanation (for example: “a table”, “to run”, “green”). This is convenient from a practical point of view, since: 1) ordinary words and scientific terms reflect the objective features of human perception of the surrounding world and the laws of the science development; 2) the knowledge acquired by a person is a system of concepts and links between them; 3) the number of concepts learned by the student can be estimated by the testing method; 4) to determine the folding information coefficient K_{FI} in the text it is enough to count the number of significant words; 5) the text volume (that is, the number of words) is approximately proportional to the time of its reading and retelling on the lesson.

4 RESULTS OF RESEARCH

To assess the complexity of the LT solution, it is proposed to use the thesaurus approach which, taking into account the meaning and complexity of the scientific terms used, allows to determine the amount of semantic information in the explanation. A thesaurus is a systematic set of terms related to a particular subject area. The pupil understands the learning material, if his thesaurus corresponds to the teacher’s explanation. The semantic information amount received by a pupil who solves an LT is estimated by the degree of change in his knowledge. It is relative, as change of the pupil’s knowledge depends on his thesaurus Z_0. This corresponds to the pragmatic approach, according to which information is considered from the point of view of its usefulness for the LT solution.

Let us take as a basis that the explanation of the LT solution is characterized by: 1) the volume V which is equal to the number of symbols, words, elementary judgments used (volume approach); 2) the general informativeness (or semantic complexity) Inf. The abstractness degree the explanation of the task can be characterized by the average coefficient of information folding $K_{FI} = Inf / V$.

Since physics is studied in the 7–11th grades, to assess the complexity of physical tasks, for zero level Z_0 we should choose the knowledge of the 5th–6th grades pupil, who has not yet begun to study physics. Such student has well-acquired arithmetic operations, but his thesaurus does not contain the concepts of “derivative”, “integral”, etc. Within physical theory he well understand the concepts “time”, “distance”, “speed”, but the concepts of "electric field", "refractive index", "mass defect" are not included in his thesaurus and need to be defined and explained. To determine the information capacity of the scientific term, it is necessary to count the number of words in the term explanation, containing the concepts from the thesaurus Z_0, but not to explain this issue to a real fifth-grader. For example, “acceleration is the ratio of changes in velocity to the corresponding time interval.” If you do not decipher the concept of “velocity”, the complexity of the concept “acceleration” is equal 6. A text containing a brief LT explanation can be created; its integral informativeness with respect to the thesaurus Z_0 will be the complexity indicator of the LT explanation.

The complexity of the picture (or formula) regarding to the thesaurus Z_0 is equal to the average terminological folding (or informational density) of the short and full text description of these didactic objects, and informativeness (integral complexity) of didactic objects is equal to its volume. An important indicator is the average difficulty of the sentence understanding. It does not depend on the text length and is proportional to the average amount of information per sentence or the average complexity of the sentence typical for the text. It can be defined as the product of the average word complexity in a given text by the average number of words in sentence L. The complexity of the formula is equal to the product of the average complexity of all physics quantities (concepts) included in the formula by the number of characters in it. Since the short-term memory of a person holds about seven blocks of information, you can use the formula $S = K_{FI} L/7.$ If the sentence of seven words (it is easily retained in memory) does not contain scientific terms ($K_{FI} = 1$), then its complexity is $S = 1$. If the average K_{FI} for the text is 2.4 and the average sentence length is $L_S = 9.7,$ the complexity of the text is $S = 2.4 \cdot 9.7/7 \approx 3.3.$

In order to assess the complexity of an explanation with respect to a level Z_0, it is necessary to supplement the text with the missing statements so that it is compatible with level Z_0 and then add up the complexities of all sentences. There is no need to take into account its logical structure in detail, but it is enough to identify and calculate the number of logical connections. To do this, the words “because” or “therefore” should be added to the text of the LT solution. The computer calculates the number of these words and takes into account their complexity (we think it is 3).

Let us suppose, it is necessary to estimate the complexity of N tasks $Z_1, Z_2, \ldots, Z_N$. The condition and solution of the LT is a system consisting of the text, formulas and a pictures. The essence of the proposed method is to encode the LT solution as a text file, assess the complexity of separate terms and then automatically analyze it using a special computer program. The method for evaluating the general informativeness of the explanation of the LT solution consists in the following: 1) the expert reads the LT condition and solves it by drawing a picture, writing down formulas and explanations; at the same time the shortest LT solution is created; 2) to create a text file task.txt, containing the coded LT condition, pictures and its solution with formulas and explanations corresponding to the selected thesaurus Z_0; 3) to create a text file words.txt, containing a list of terms used in the LT solving; 4) to assess the complexity s_i of the terms, corresponding to the thesaurus Z_0, and record their meanings in words.txt; 5) with the help of a special program (referring to the file words.txt) to analyze the file task.txt and determine the total complexity of the text, its volume and the average coefficient of information folding. To determine the complexity s_i of the $i-$th term, it is necessary to count the number of words in its definition, which is understandable to the student with the thesaurus Z_0.

We have evaluated the following 10 physical problems, in which experiments from various branches of physics are considered:

1 The ball is thrown upwards at a speed 12 m/s. How high will it go?

2 A body weighing 300 g oscillates on a spring of rigidity 40 N/m. Determine the oscillation period.

3 Find the velocity of the silver vapor molecules, if their angular displacement in the Stern experiment is 5.8 degrees at the instrument rotation frequency 150 Hz. The distance between the inner and outer cylinders is 2 cm.

4 In a vessel, containing 2 kg of water at 20 degrees Celsius, a piece of iron weighing 0.6 kg, having a temperature of 90 degrees Celsius is put. Determine the temperature of the water after the establishment of thermal equilibrium.

5 A voltmeter is connected to the winding of 150 turns. Coil area is 3 sm ^{2}. Find the rate of the magnetic field change inside the winding, if the voltmeter indicates 12 V.

6 The galvanic cell, resistor and ammeter are connected consecutively. A voltmeter is connected in parallel with the resistor. The electromotive force of a galvanic cell is 2 V, internal resistance is 1 Ohm. The resistance of the resistor is 8 ohms. Determine the indications of the devices.

7 At a distance of 12 cm from the collecting lens with an optical power of 10 diopter is an object 1 cm high. Where should I place the screen to get a sharp image of the object? What is the height of the image?

8 A red light with a wavelength of 750 nm falls on a diffraction grating with a period of 0.01 mm. The distance from the diffraction grating to the screen is 1.3 m. Find the distance between the diffraction peaks of the second order.

9 To what minimum potential the zinc plate will be charged if it is irradiated with monochromatic light with wavelength 324 nm?

10 Radon is placed in the vial, the activity of which is $4.5\cdot 10^{10}$ decays per second. In what time the activity of radon will be equal to $2.3\cdot 10^{10}$ decays per second? Radon half-life is 3.82 days.

Let us consider the solution of task 9. The figure is: photons fall on a plate and beat electrons out. Explanation: When the zinc plate is irradiated, a photo effect occurs. The plate potential decreases until it becomes equal to the holding voltage. The formulas are:

$$h\nu = A + m_e\upsilon_m^2/2,\quad m_e\upsilon_m^2/2 = |A_{EF}|,\ -m_e\upsilon_m^2/2 = A_{EF},\quad A_{EF} = -eU_H,$$
$$h\nu = A + eU_H,\quad U_H = \varphi - \varphi_\infty,\quad \varphi_\infty = 0,\quad \varphi = (h\nu - A)/e.$$

The task solution is encoded in a text file. For example, the formula is coded like this: "Plank_constant multiply frequency equal exit_work plus charge electron multiply hold_voltage" ("The Plank's constant multiplied by frequency is equal to the exit work plus charge electron multiplied by holding voltage"). The text file gets added words "because" and "therefore" and the following rules, used in mathematical transformations: 1) if the same values are added to the left and right sides of the equality, then the equality remains true; 2) if the left and right sides of the equality are multiplied by the same values, then the equality remains true. The resulting text file is analyzed using a special program written on Pascal.

The results of evaluating the information content of the above tasks are presented in Table. 1. It consists of columns: 1) the number of the LT and the section of physics (mechanics, thermodynamics and molecular physics, electrodynamics, optics, quantum physics); 2) the amount of the text in words V (double terms are counted as two words); 3) the number of words in the text N_W, which are not scientific terms; 4) the total textural complexity (informativeness) of the explanation Q_T, which is found with the help of the computer program; 5) the total formula complexity Q_F, which is found using the program; 6) the average length of the formulas L_F (average number of mathematical symbols); 7) the overall complexity of the explanation S; 8) the general informativeness of explanations Inf; 9) the complexity coefficient K_C; 10) the coefficient of folding information K_{FI}. The following formulas are used:

$$Q = n_1 s_1 + n_2 s_2 + ... + n_N s_N,\quad S = N_W + Q_T + Q_F L_F/7,$$

$$Inf = N_W + Q_T + Q_F,\quad K_C = S/V,\quad K_{FI} = Inf/V.$$

Here n_i is the number of uses in the explanation of i–th term with complexity s_i.

Table 1. The results of the complexity estimation of the physical tasks explanation.

Task	V	N_W	Q_T	Q_F	L_F	S	Inf	K_C	K_{FI}
1-М	69	29	43	85	10,0	193,4	157	2,80	2,28
2-М	45	28	26	14	8,0	70,0	68	1,56	1,51
3-Т	121	76	46	90	5,7	194,9	212	1,61	1,75
4-Т	187	59	55	411	11,3	779,2	525	4,17	2,81
5-Э	58	14	279	152	5,5	412,4	445	7,11	7,67
6-Э	117	20	246	496	6,3	714,5	762	6,11	6,51
7-О	147	45	155	178	7,8	399,1	378	2,72	2,57
8-О	140	40	164	244	8,6	503,8	448	3,60	3,20
9-КФ	134	36	216	273	8,2	571,8	525	4,27	3,92
10-КФ	120	30	194	440	7,0	664,0	664	5,53	5,53

When calculating the complexity S we take into account the average length of the formulas L_F. The division by seven is explained by the fact that a person in a short-term memory easily holds about seven blocks of information. If the formulas are complex ($L_F > 7$), then $S > Inf$ and $K_W > K_{FI}$. From the table it can be seen that the general information content of the explanation of the standard task from the school physics course varies 10 – 11 times (from 70 to 760 CUI), and the average folding coefficient is in the range 1.5 – 7.7. The proposed method is also used to determine the complexity of the Pythagorean theorem proof; it has turned out that the total information content of the proof is equal $Inf \approx 250$ CUI, the average folding coefficient is 2.17.

5 CONCLUSIONS

The article considers the actual problem the complexity estimation of the learning tasks explaining to the student. The proposed method requires the application of the formal linguistic and information methods, which consist in creating a file that encodes the explanation, and analyzing this file using a special computer program. The program refers to the thesaurus containing a list of terms and their complexity, which is equal to the number of words in the term definition. Using this method, the complexity of 10 physics tasks for 10 – 11 grades is estimated. It has been established that the general informativeness of the explanation of the standard task varies quite significantly (from 70 to 760 CUI), and the average folding coefficient is in the range of 1.5 – 7.7. The method of automated counting of terms allows to estimate the didactic complexity of any information block containing an explanation of the LT solution, formula derivation, proof of a theorem, etc.

REFERENCES

[1] R. R. Gajsin “Princip slozhnosti v estestvennonauchnom poznanii: Metodologicheskij analiz” [The principle of complexity in natural knowledge: Methodological analysis]: *Diss. ... kand. filosofsk. nauk.* – Ufa, 2002. – 169 p.

[2] O. B. Episheva, V. I. Krupich “Uchit' shkol'nikov uchit'sja matematike: formirovanie priemov uchebnoj dejatel'nosti: Kn. dlja uchitelja” [To teach schoolchildren to learn mathematics: the formation of teaching methods: Book for the teacher]. – M.: Prosveshhenie, 1990. 128 p.

[3] V. M. Krotov “K voprosu o slozhnosti (trudnosti) fizicheskih zadach” [To the question of the complexity (difficulty) of physical problems] // *Fizika: prablemy vykladannja.* 1999. №3. pp. 69 – 74.

[4] A. V. Gidlevskij “Ischislenie trudnosti didakticheskoj zadachi” [Calculation of the difficulty of the didactic task] // *Vestnik Omskogo un-ta.* 2010. № 4. pp. 241 – 246.

[5] I. S. Naumov, V. S. Vyhovanec “Ocenka trudnosti i slozhnosti uchebnyh zadach na osnove sintaksicheskogo analiza tekstov” [Evaluation of the difficulty and complexity of educational tasks on the basis of the syntactic analysis of texts] // *Upravlenie bol'shimi sistemami: Sbornik trudov.* 2014. Vyp. 48. pp. 97 – 131.

[6] A. L. Sakovich “Slozhnost' fizicheskih zadach i ih urovni” [The complexity of physical problems and their levels] // *Fizika. Prablemy vykladannja.* 2004. №1. pp. 33 – 40.

[7] R. V. Mayer “Kontent-analiz shkol'nyh uchebnikov po estestvenno-nauchnym disciplinam: monografija” [Content analysis of school textbooks in the natural sciences: monograph]. – Glazov: Glazov. gos. ped. in-t, 2016. 137 p.

[8] R. V. Mayer “Kolichestvennaja ocenka slozhnosti izuchenija jelektricheskogo i magnitnyh polej v razlichnyh uchebnikah fiziki” [Quantitative assessment of the complexity of studying electric and magnetic fields in various physics textbooks] // *NIR. Social'no-gumanitarnye issledovanija i tehnologii*. - № 1(26), 2019. pp. 11-16.

[9] Je. G. Gel'fman, M. A. Holodnaja “Psihodidaktika shkol'nogo uchebnika. Intellektual'noe vospitanie uchashhihsja” [Psychodidactics of a school textbook. Intellectual education of students]. – SPb.: Piter, 2006. 384 p.

[10] Ja. A. Mikk “Optimizacija slozhnosti uchebnogo teksta: V pomoshh' avtoram i redaktoram” [Optimization of the complexity of the educational text: To help authors and editors]. – M.: Prosveshhenie, 1981. 119 p.

[11] O. V. Zerkal' “Semanticheskaja informacija i podhody k ee ocenke. Chast' 1. Semantiko–pragmaticheskaja informacija i logiko–semanticheskaja koncepcija” [Semantic information and approaches to its assessment. Part 1. Semantic-pragmatic information and logical-semantic concept] // *Filosofija nauki.* 2014. № 1. pp. 53 – 69.

[12] N. M. Solomatin “Informacionnye semanticheskie sistemy” [Information semantic systems] // *Perspektivy razvitija vychislitel'noj tehniki: V 11 kn.: Sprav. Posobie. Kn. 1.* – M.: Vyssh. shk., 1989. 127 p.

[13] B. Davis, D. Sumara Complexity and Education: Inquiries Into Learning, Teaching, and Research [Complexity and Education: Inquiries Into Learning, Teaching, and Research]. – Mahwah, New Jersey, London, 2006. 201 p.

[14] V. I. Shalak “Sovremennyj kontent–analiz. Prilozhenija v oblasti: politologii, psihologii, sociologii, kul'turologii, jekonomiki, reklamy” [Modern content analysis. Applications in the field of: political science, psychology, sociology, cultural studies, economics, advertising]. – M.: Omega–L, 2004. 272 p.

[15] N. K. Krioni, A. D. Nikin, A. V. Fillipova “Avtomatizirovannaja sistema analiza slozhnosti uchebnyh tekstov” [Automated system for the analysis of the complexity of educational texts] // *Vestnik UGATU (Ufa).* 2008, T. 11. № 1 (28). pp. 101 – 107.

[16] O. N. Podol'skaja, V. V. Vjazankova, K. V. Horoshun “Diagnostika obrazovatel'nogo processa na osnove infometrii” [Diagnostics of the educational process based on infometry] // *Azimuth of Scientific Research: Pedagogy and Psychology*. 2019. pp. 183-185.