ISSN 1512-1801

UDC 37.02

THE SEMANTIC COMPLEXITY OF FORMULAS AND METHODS FOR SOLVING TASKS IN MECHANICS: EVALUATION RESULTS

Robert Valerievich Mayer

Doctor of Pedagogical Sciences, Professor of Physics and Didactic of Physics Chair
FSEI of HE «Glazov Korolenko State Engineering and Pedagogical University», Glazov, Russia

Abstract: *A content analysis of Russian school physics textbook and task book has been conducted. All the formulas on mechanics have been identified and 9 topics have been defined, each corresponding to its own method of problem solving. The formulas have been encoded with a verbal code and are presented in a text file Formula.txt. The complexity of the concepts that denote physical quantities is assessed separately and their list is placed in the file Slovar.txt. With the help of the special computer program, the semantic complexity of the formulas is determined and the information folding coefficient is calculated for each formula. The difference between definition formulas and those expressing functional dependencies has been taken into account. The distributions of formulas and methods in the space of their characteristics have been analyzed. The uncertainty and complexity of choosing formulas and solving tasks independently are estimated.*

Keywords: *pedagogy, methodology, educational task, mechanics, semantics, complexity, formula.*

Introduction

Determining the complexity of various didactic objects (educational texts, their fragments, formulas, drawings, etc.) is one of the urgent problems of the learning theory. The assessment results of the educational material complexity depend on when, for how long, and by which method a specific topic is studied. Measuring the complexity of solving physics tasks (PTs) would allow arranging them in order of increasing difficulty and accurately evaluating pupils' work on control papers, exams, Olympiads, etc.

The method for assessing the complexity of educational tasks has been repeatedly discussed in the scientific and methodological literature [1 – 4]. To achieve this, scientists usually use methods based on: 1) expert assessments; 2) subjective assessments of the task difficulty by pupils; 3) measuring the time and quality of task solving; 4) registration of psychophysiological parameters of thought processes; 5) counting of scientific terms used by pupils. Obtaining expert assessments of the complexity of various tasks and conducting pedagogical or psychophysiological experiments requires a significant amount of time and the participation of a large number of experts or pupils. Therefore, methods for assessing the complexity of PTs, based on a structural, semantic, and logical analysis of their conditions and solutions, are of great practical interest [1; 3; 5].

Some researchers (for example, [3; 4]), focusing their attention on the logical structure of solving an educational tasks, do not take into account that: 1) the task complexity depends on the semantic complexity of the terms used to solve it; 2) the complexity or difficulty of the pupils' individual solution of PT is significantly higher than the complexity of understanding the finished task solution presented in the textbook, as the pupil must select the appropriate formulas (theoretical models, laws, etc.) regardless of the teacher.

ISSN 1512-1801

The monograph [6] shows how the semantic complexity (SC) of the considered didactic object can be estimated. To do this, all its components should be encoded with a single code, and then the complexity of the resulting message should be determined. The verbal code is the most convenient option. The resulting text file can then be analyzed by a computer program, which identifies scientific terms and calculates their semantic complexity, as described in [6; 7].

The aim of the study is to: 1) assess the semantic complexity of formulas and their corresponding theoretical propositions; 2) determine the complexity of methods for solving tasks in mechanics; 3) 3) take into account the uncertainty of choosing the right method and formula. **The methodological basis** for this research is the work of the following scientists: B.C. Babaev, M.V. Kulagina and Ju.Ju. Shkitina [1], G.A. Ball [2], A.V. Gidlevsky [3], V.M. Krotov [4], I.S. Naumov and V.S. Vykhovanets [5] (complexity of educational tasks); R.V. Mayer [6; 7], E.Ja. Tarshis [8] (content analysis method).

The results of research

The semantic complexity of a didactic object can be defined as the amount of information contained in its description. It is convenient to take the complexity of words well known to a fifth grader and included in his thesaurus Z_5 (man, water, air, table, etc.) as a unit of SC measurement. The semantic complexity of the abstract concept C relative to the thesaurus Z_5 can be calculated as the smallest number of such words needed to explain C.

When studying physics, pupils learn a large number of formulas. There are two main types of formulas: definitions of new physical quantities and formulas expressing the relationship between physical quantities. Each formula has a specific meaning and carries a certain amount of information. For example, the formula $A = \Delta E_K$ means that the change in the kinetic energy of a mechanical system is equal to the work of all forces acting on it. The more semantic information there is in a statement, the more simple words from the thesaurus Z_5 must be uttered to explain its essence, and the more difficult it is.

In physics lessons pupils get acquainted with various methods of solving PTs [9; 10]. Each method represents a set of techniques used in solving PTs on a certain topic. The tasks solved by the same method are similar conceptually, they use approximately the same physical models, concepts, theoretical ideas and formulas. PTs on the topics of "Uniform motion", "Work, energy, power", "Mechanical oscillations" are fundamentally different from each other and are solved in different ways: each method uses its own physical ideas, models, laws and formulas expressing them. Limiting ourselves to mechanics, we can highlight the following topics (and their corresponding task solving methods): 1. Uniform motion. 2. Uniformly accelerated motion. 3. Circular motion. 4. Newton's laws. Forces in mechanics. 5. Momentum and its change. 6. Work, energy, power. 7. Balance of bodies. 8. Mechanical oscillations. 9. Mechanical waves.

From the standpoint of General System Theory, a physical formula is a system consisting of separate elements: letters denoting physical quantities and mathematical symbols. The complexity of the formula depends on: 1) the number of elements (quantities) and the connections between them (i.e. mathematical symbols); 2) the complexity of mathematical operations and physical quantities (scientific concepts) included in the formula.

The following method is used to determine the semantic complexity of the formula [6; 7]: 1) to create a file *Formula.txt* in which the formulas are verbally encoded, that is, they are presented in the form of sentences: $W = CU^2/2$ – "the energy of a capacitor is equal to the capacitance multiplied by the voltage and divided by a number"; 2) to evaluate the complexity of the terms included in the formulas and write the results to a file called *Slovar.txt;* 3) to use the special computer program

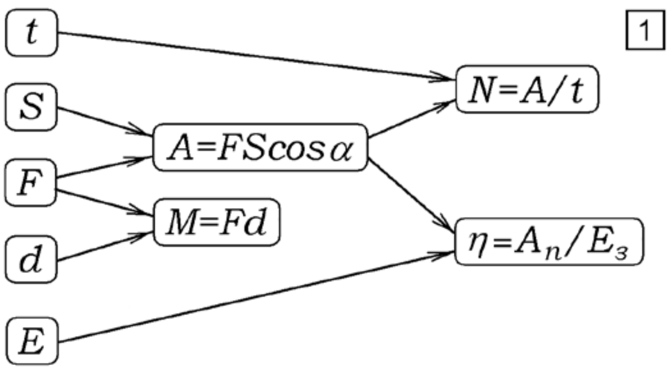
ISSN 1512-1801

written in *ABCPascal*, which accesses the file *Slovar.txt*, to analyze the file *Formula.txt* and to determine the total complexity of all the terms that make up the descriptions of the formulas.

To determine the semantic complexity of the formula, expressing functional dependence (for example, $E_\kappa + E_n = const$), we must replace this formula with a factual statement ("if only conservative forces operate in the system, then kinetic energy plus potential energy is equal to a constant") and summarize the semantic complexities of all its constituent terms. When assessing the semantic complexity of definition the physical quantity X, the complexity of the term denoting the quantity X is considered equal to 1. For example, in the sentence "capacitance – the ratio of the capacitor charge to the voltage on the plates", the semantic complexity of the word "capacitance" is assumed to be 1. This is logical, since this sentence contains a repeat: before and after the dash, we are talking about the capacitor capacitance. A different approach (when we consider that S(capacity)$\approx$15) will lead to the fact that the semantic complexity of the definition (and hence the formula $C = q/U$) will be overestimated by almost 2 times. When determining the number of words in a sentence (that is, the information volume) corresponding to the formula, all words, including the term being defined, are counted.

For similar reasons, the complexity of the proportionality coefficient (G, k, ρ etc.) is equal to 5. Let's consider the formula $F = Gm_1m_2/r^2$. If we express the gravitational constant ($G = Fr^2/m_1m_2$) and determine its complexity by calculating the total complexity of the terms, then the formula complexity will be overestimated by 1.5 – 2 times. The complexity of the numbers π, e etc. is assumed to be equal to two.

How to assess the complexity of physical concepts denoting physical objects, phenomena, quantities? Note that the definition of each new abstract concept of the $(k+1)-$th level contains previously introduced abstractions of the $k-$th level. For example, using such basic concepts as time t, distance d, displacement S, force F, energy E, it is possible to formulate definitions of the force moment M, work A, power N and efficiency η (Fig. 1.1). At the same time, it is obvious that the concept "force" is simpler than the concept "work", and "work" is simpler than the concept "power": SC(force) < SC(work) < SC(power). Since "power is equal to the ratio of work to time interval" (6 words), then we can write: SC(power) $\approx$ SC(work) + 5 (the word "power" should not be taken into account). Counting the number of words in the definitions and comparing the concepts with each other, we can conclude that their difficulties are equal: time – 2, distance – 2, movement – 3, force – 6, moment of forces – 9, work – 14, power – 19.



2

№	FORMULAS	SC	V	КСИ
	Method 4			
14	$\vec{F} = \vec{F}_1 + \vec{F}_2 + ... + \vec{F}_n$	31	11	2,8
15	$m\vec{a} = \vec{F}_1 + \vec{F}_2 + ... + \vec{F}_n$	45	14	3,2
...	...	...	...	...
22	$F_c = k_2 \upsilon^2$	36	8	4,5
	Method 5			

Fig. 1. The connections of physical concepts. The fragment of a table with formulas.

The method discussed above allowed to estimate the complexity of 46 school formulas in mechanics studied in 10-th grade [9], and for each formula to calculate the information folding coefficient IFC = SC/V, where V is the volume of information in the sentence replacing the formula; it is equal to the number of words. IFC shows the density of information concentration in terms; the higher IFC is, the more difficult to understand the corresponding text, sentence or formula. The results of the formula evaluation were summarized in an Excel spreadsheet, the fragment of which is shown in Fig. 1.2. Below they are presented in the format: "formula (definition (D), law or dependencies (L); SC ; V ; IFC)":

1. Uniform motion: 1) velocity $\vec{\upsilon} = d\,\vec{r}\,/\,dt$ (D; 22; 11; 2,0); 2) for uniform movement $\vec{r} = \vec{r}_0 + \vec{\upsilon}\,t$ (L; 36; 13; 2,8); 3) for uniform movement $x = x_0 + \upsilon_{0x} t$ (L; 32; 12; 2,7); 4) "the velocity vector of the point relative to the earth = the velocity vector of the point relative to the reference frame + the velocity vector of the reference frame relative to the earth": $\vec{\upsilon} = \vec{\upsilon}_1 + \vec{\upsilon}_2$ (L; 65; 21; 3,1).

2. Uniformly accelerated motion: 5) acceleration $\vec{a} = \Delta\vec{\upsilon}\,/\,\Delta t$ (D; 24; 11; 2,2); 6) for uniformly accelerated motion $\vec{\upsilon} = \vec{\upsilon} + \vec{a}t$ (L; 50; 15; 3,3); 7) for uniformly accelerated motion $\upsilon_x = \upsilon_{0x} + a_x t$ (L; 54; 15; 3,6); 8) for uniformly accelerated motion $x = x_0 + \upsilon_{0x}t + a_x t^2 / 2$ (L; 66; 22; 3.0).

3. Circular motion: 9) for uniform motion around the circle $a = \upsilon^2 / R$ (L; 37; 12; 3,1); 10) angular velocity $\omega = \Delta\varphi / \Delta t$ (D; 13; 9; 1,4); 11) angular velocity $\omega = 2\pi / T$ (L; 20; 9; 2,2); 12) angular velocity $\omega = 2\pi\nu$ (L; 20; 9; 2.2); 13) velocity $\upsilon = \omega R$ (L; 23; 10; 2,3).

4. Newton's laws. Forces in mechanics: 14) resultant force $\vec{F} = \vec{F}_1 + \vec{F}_2 + ... + \vec{F}_n$ (L; 31; 11; 2,8); 15) Newton's second law: $m\vec{a} = \vec{F}_1 + \vec{F}_2 + ... + \vec{F}_n$ (L; 45; 14; 3,2); 16) for two interacting bodies $\vec{F}_1 = -\vec{F}_2$ (L; 25; 9; 2,8); 17) gravitational interaction force $F = Gm_1 m_2 / r^2$ (L; 58; 16; 3.6); 18) gravity force $\vec{F} = m\vec{g}$ (L; 26; 9; 2,9); 19) elastic force $F_{ynp,x} = -k\,x$ (L; 27; 8; 3,4); 20) friction force $F_{mp} = \mu N$ (L; 25; 10; 2,5); 21) resistance force $F_c = k_1 \upsilon$ (L; 40; 10; 4.0); 22) for high velocities the resistance force $F_c = k_2 \upsilon^2$ (L; 42, 11, 3.8).

5. Momentum and its change: 23) body momentum $\vec{p} = m\vec{\upsilon}$ (D; 21; 8; 2,6); 24) change in body momentum $\Delta\,\vec{p} = \vec{F}\,\Delta\,t$ (L; 37; 10; 3,7); 25) for a closed system $m_1\vec{\upsilon}_1 + m_2\vec{\upsilon}_2 + \; + m_3\vec{\upsilon}_3 + ... = const$ (L; 47; 14; 3,4).

6. Work, energy, power: 26) work of force $A = FS\cos\alpha$ (D; 32; 13; 2,5); 27) engine power $N = \Delta\,A / \Delta\,t$ (D; 21; 7; 3,0); 28) kinetic energy of the body $E_\kappa = m\upsilon^2 / 2$ (D; 22; 10; 2,2); 29) change in kinetic energy $\Delta E_\kappa = A$ (L; 41; 8; 5.1); 30) the work of gravity $A = -\Delta E_n = m\,gh_1 - m\,gh_2$ (L; 77; 26; 3,0); 31) potential energy of the raised body $E_n = m\,gh$ (L; 37; 11; 3,4); 32) potential energy of the deformed body $E_n = k\,\Delta l^2 / 2$ (L; 39; 10; 3,9); 33) if only conservative forces operate in the system, then $E = E_\kappa + E_n = const$ (L; 79; 15; 5,3).

7. Equilibrium of bodies: 34) mass center is in equilibrium if $\vec{F}_1 + \vec{F}_2 + \vec{F}_3 + ... = 0$ (L; 12; 6; 2,0); 35) moment of force $M = \pm Fd$ (D; 38; 12; 3,2); 36) body is in equilibrium if $M_1 + M_2 + M_3 + ... = 0$ (L; 36; 12; 3,0).

8. Mechanical oscillations: 37) for spring pendulum $x''=-(k/m)x$ (L; 105; 14; 7,5); 38) for harmonic oscillations $x=x_m\cos(\omega t)$ (L; 35; 12; 2,9); 39) own oscillation frequency $\omega_0=\sqrt{k/m}$ (L; 38; 11; 3,5); 40) oscillation period of the spring pendulum $T=2\pi\sqrt{m/k}$ (L; 44; 14; 3.1); 41) oscillation period of the thread pendulum $T=2\pi\sqrt{l/g}$ (L; 39; 16; 2,4); 42) the oscillation energy of the spring pendulum $W=k\,x_m^2/2=m\upsilon_m^2/2$ (L; 64; 20; 3.2).

9. Mechanical waves: 43) wavelength $\lambda=\upsilon T=\upsilon/\nu$ (L; 48; 12; 4,0); 44) wave equation $s=s_m\sin(\omega(t-x/\upsilon))$ (L; 45; 16; 2,8); 45) intensity of the wave $I=\Delta W/S\Delta t$ (D; 30; 10; 3,0); 46) wave intensity $I=wc$ (L; 75; 8; 9.4).

Let's consider the distribution of formulas in the feature space "semantic complexity SC – volume V", for this we put points on the coordinate plane «SC – V» (Fig. 2.1). It can be seen that the absolute majority of points (43 out of 46) are located close to an increasing straight line passing near the origin. This fact and a rather high correlation coefficient (≈ 0.67) mean that there is a stochastic relationship between SC and V. It is explained by the fact that the more words in a sentence, the higher its semantic complexity. The slope angle α cotangent is approximately equal to the average value of IFC for these 43 formulas. Three points falling out of this pattern correspond to formulas containing concepts with high IFC: 33) if conservative forces are acting in the system, then $E=E_\kappa+E_n=const$ ($IFC\approx 5.3$); 37) for a spring pendulum $x''=-(k/m)x$ ($IFC\approx$ 7.5); 46) wave intensity $I=wc$ ($IFC\approx 9.4$). The high information density means that these formulas are more difficult to understand and explain, because for this we will have to use concepts with a high abstraction degree.

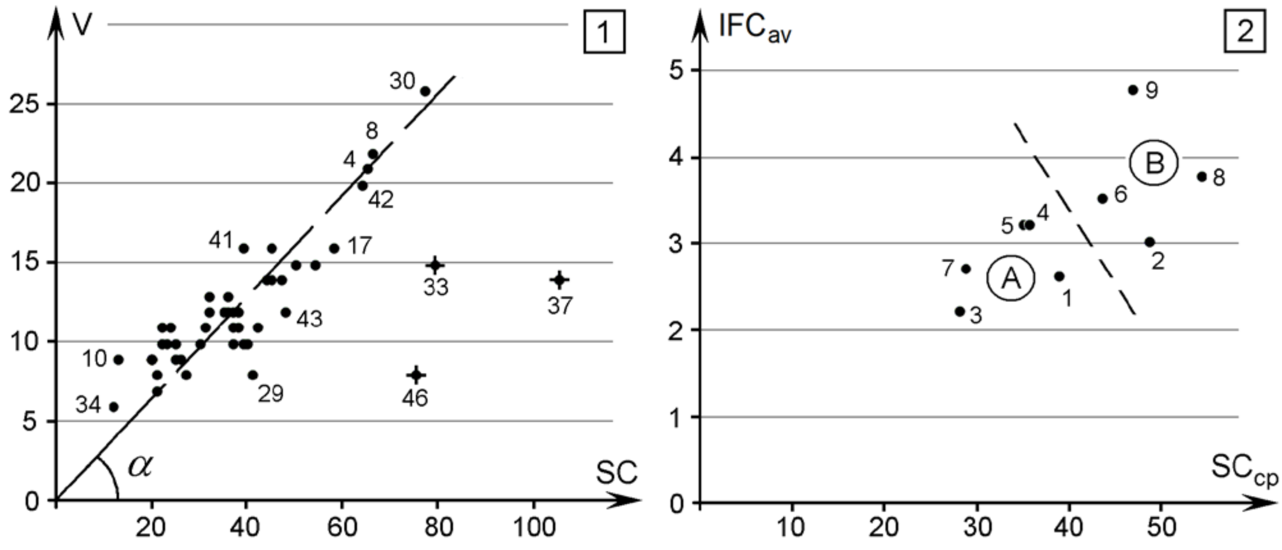


Fig. 1. Clouds of points corresponding to formulas (1) and methods (2).

It is possible to determine the average semantic complexity SC_{av} of each method and the average IFC_{av} for the constituent formulas. Fig. 1.2 shows the distribution of methods for solving tasks in mechanics in the feature space "average semantic complexity SC_{av} – average IFC_{av}". High SC_{av} means that a significant information amount is contained in the formulas of this method, and high IFC_{av} indicates a great difficulty in understanding the formulas. It turns out that the farther the points are from the origin (Fig. 1.2), the more difficult the corresponding methods are. They can be divided into two categories according to their complexity:

ISSN 1512-1801

A) Simple methods and topics: 1. Uniform motion. 3. Circular motion. 4. Newton's laws. Forces in mechanics. 5. Momentum and its change. 7. Balance of bodies.

B) Complex methods and topics: 2. Uniformly accelerated motion; 6. Work, energy, power; 8. Mechanical oscillations; 9. Mechanical waves.

The solutions of PTs, as a rule, has physical, mathematical and computational components. The physical complexity of task solving individually (without a teacher) depends on the complexity of choosing the initial formulas expressing the connections between known and desired physical quantities.

Let's consider the pupil solving the single-formula task (i.e. the PT that requires the use of a single physical formula). In practice, all possible options are implemented, concluded between the two extremes: 1) the pupil knows well how the PT is solved, so he/she remembers the solution; 2) the pupil faced this problem for the first time, and he will really to solve it. In the first case, the complexity of choosing a method and formula is minimal and equal $CCFM_{min} = 1$. In the second case, it is maximal and depends on: 1) the uncertainty of choosing the method or topic to which the PT relates; 2) the uncertainty of choosing a formula from the whole set of formulas related to this method (topic). If we consider the intermediate situation and take the arithmetic mean of $CCFM_{max}$ and $CCFM_{min} = 1$, we get $CCFM_{av} = (1 + CCFM_{max})/2$.

The complexity of choosing a suitable method from the 1st section of physics (mechanics) is characterized by the value $CCM_1 = \ln(M_1)$ where $M_1 = 9$ is the number of methods in the first section of physics – "Mechanics". The complexity of choosing the correct formula from $j-$method is equal $CCF_{1,j} = \ln(F_{1,j})$. It turns out that the total complexity of choosing a formula for solving a single-formula task is equal: $CCFM_{max}(1,j) = CCM_1 + CCF_{1,j} = \ln(M_1) + \ln(F_{1,j})$. So, the complexity of the original formula should be increased by $CCFM_{av}(1,j)$ times.

In our case, there are 9 methods, therefore $CCM_1 = \ln(9) \approx 2,2$; method 1 contains 4 formulas, and method 6 contains 8 formulas, hence $CCF_{1,1} = \ln(4) \approx 1.39$ and $CCF_{1,6} = \ln(8) \approx 2.08$. From here: $CCFM_{av}(1,1) \approx (2,2+1,39+1)/2 \approx 2,3$ and $CCFM_{av}(1,6) \approx (2,2+2,08+1)/2 \approx 2,6$. If several initial physical formulas are required to solve the PT, then the semantic complexity of each formula should be increased $CCFM_{av}(1,j)$ times and then added together.

Conclusions

The estimating complexity problem of formulas and methods used in solving school tasks in mechanics is considered. It is shown, that methods of assessing the complexity of a physical task, based on the analysis of its conditions, solutions and considering the semantic complexity of the terms used, are of practical importance. As a result of the analysis of school textbooks and task books on physics, all formulas related to mechanics have been written out and 9 topics have been identified, each of which corresponds to its own method of solving problems. The formulas are encoded with a verbal code in a text file. The article discusses how to assess the complexity of concepts denoting physical quantities. With the help of a special computer program accessing the dictionary, the semantic complexities of the formulas are determined and the coefficient of information folding is calculated for each one. At the same time, the difference between the definition formulas and formulas expressing functional dependencies between physical quantities was taken into account. This made it possible to identify the most complex formulas and methods for solving tasks in mechanics. The distribution of formulas and methods in the space of their features ("complexity – information volume" and "average complexity – average information folding coefficient") is analyzed. The number of formulas in each method is also taken into account, and the uncertainty of the formula choice for the solution of physics task is estimated.

ISSN 1512-1801

References

1. Babaev B.C., Kulagina M.V., Shkitina Ju.Ju. Opredelenie trudnosti i slozhnosti fizicheskih zadach // Fizicheskoe obrazovanie v vuzah. T. 11, № 4, 2005. pp. 93 – 101.
2. Ball G.A. Teorija uchebnyh zadach: psihologo-pedagogicheskij aspekt. Moscow: Pedagogika, 1990. 184 p.
3. Gidlevskij A.V. Ischislenie trudnosti didakticheskoj zadachi // Vestnik Omskogo univer-siteta. 2010. № 4. pp. 241 – 246.
4. Krotov V.M. K voprosu o slozhnosti (trudnosti) fizicheskih zadach // Fizika: prablemy vykladannja. 1999. № 3. p. 69 – 74.
5. Naumov I.S., Vyhovanec V.S. Ocenka trudnosti i slozhnosti uchebnyh zadach na osnove sintaksicheskogo analiza tekstov // Upravlenie bol'shimi sistemami: sb. tr. 2014. Vyp. 48. pp. 97 – 131.
6. Mayer R.V. Didakticheskaja slozhnost' uchebnyh tekstov i ee ocenka: monografija. Glazov: GGPI, 2020. 149 p.
7. Mayer R.V. Metod ocenki slozhnosti logicheskih rassuzhdenij // NIR. Social'no-gumanitarnye issledovanija i tehnologii. № 3(32), 2020. pp. 35 – 40.
8. Tarshis E.Ja. Kontent-analiz: Principy metodologii (Postroenie teoreticheskoj bazy. Ontologija, analitika i fenomenologija teksta. Programmy issledovanija). Moscow: LIBRO-KOM, 2013. 176 p.
9. Mjakishev G.Ja. Fizika: ucheb. dlja 10-h klassov obshheobrazovat. uchrezhdenij: bazovyj i profil. urovni / G.Ja. Mjakishev, B.B. Buhovcev, N.N. Sotskij. Moscow: Prosveshhenie, 2004. 336 p.
10. Rymkevich A.P. Zadachnik. 10 – 11 kl.: posobie dlja obshheobrazovat. uchrezhdenij. Moscow: Drofa, 2013. 192 p.

Article received 2024-07-20