Quasi-Cryptographic Methods in the Management of Educational Tasks

Oleg I. Krivosheev  
*Institute of Control Sciences RAS*Moscow, Russia  
o-kriosheev@ya.ru

F.F. Pashchenko  
*Institute of Control Sciences RAS*Moscow, Russia  
pif-70@yandex.ru

E.S. Duvanov  
*Lipetsk State Technical University*Lipetsk, Russia  
evgenyduvanov@yandex.ruY.I. Kudinov  
*Lipetsk State Technical University*Lipetsk, Russia  
kudinov.yi@mail.ru

*Abstract*—A method of organizing the educational process and knowledge control in large groups of students by issuing assignments from a large (almost unlimited) sample of options is proposed. The method of generating variants based on the last name, first name, patronymic of the student is shown. It is shown that a suitable sample of variants should be at least the square of the size of the stream/group. The possibility of accelerated verification of tasks with the help of functional connections in the space of answers, i.e., peculiar keys known to the teacher, is demonstrated. Separate examples of tasks important both in the methodological aspect and in terms of application in the spectrum of disciplines of the departments of applied mathematics are given.

Keywords—parametric tasks, knowledge control, automation of knowledge control, educational process management

# Introduction

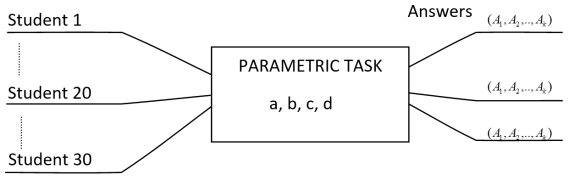
The system of parametric tasks provides for the generation of conditions for an arbitrarily wide variety of options (usually 1,000-10,000 options are enough), allows you to partially algorithmize the verification and, in the end, organize work with large flows. A related aspect was the fact that the author had the need to give lectures and simultaneously conduct seminars in large groups of 70-150 students (as a result, approximately 6,000 - 8,000 students were trained under the system at different stages of its development over 5 years).

The authors do not exclude that attempts to standardize the mass control of students' knowledge in Russia may make it relevant to actively use a fundamentally similar system, as well as individual private findings on a state scale.

When checking the assignment, it is enough for the teacher to control: the beginning of the solution with its own numbers, control ratios, which will be discussed further..

# The Main Idea of the Method

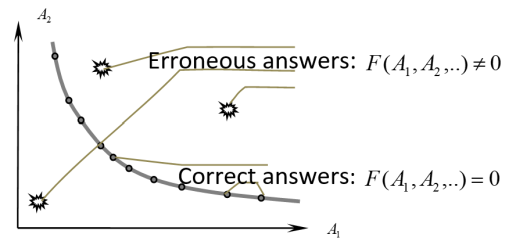
A method is proposed when a student receives a home, seminar or control task as a function of (approximately) 4 numerical parameters - the key for generating his personal condition (Fig. 1).



1. Each student receives a unique condition and and their own vector of answers.

The numerical parameters of each student are unique (for this, with the typical size of the group/stream at the university, the space of options should be from 1000 to 10,000 - which we will discuss further).

In the process of solving, the student receives a vector of answers (denote), which in practice in most cases turn out to be not independent, i.e. connected by one or more fairly simple symmetry relations (Fig. 2).



1. The correct answers belong to the surface of measure zero.

To be specific, there is one symmetry relation, for example, when solving direct and dual linear programming problems – this may be the equality of the values of objective functions in the optimum, and extremely multiple relations, which are good and convenient to check selectively in this variant, arise when solving graph problems (such as calculating the Bellman function in each in dynamic programming tasks) or other tasks related to processing arrays of data.

In addition, it is known that a teacher who regularly checks such tasks also sooner or later develops an idea of the range in which solutions for specific keys should be located (often with an accuracy of tens of percent). This need to calculate the answer corresponding to the invariant in conjunction with the optional knowledge of the range of finding the answer, allows, by setting the task of calculating a difficult one–sided function - it means solving the problem «head-on» -, the result of which satisfies a simple condition, to speed up the process of monitoring progress (speeding up learning in general).

A student cannot just invent an answer out of his head – he must get into the set of measure zero (whose co-dimension is equal to the number of test ratios):

,

and even if the key is known to students (and sometimes (and often) it makes sense for the teacher to disclose it to speed up (verification) procedures), an excessively «resourceful» student is obliged to get into a mess by grossly mistaken with the range in which the answer should be given to the student individual parameters

,

after that, this student will be brought to light with entry into the so-called «black list».

Using someone else's numbers, without radically simplifying, life is easily revealed. For example, if a student with a number 6, 8, 2, 4 uses the number of a permanent neighbor (best friend) 5, 8, 1, 7, then this can be seen from a bird's-eye view in the thickest notebook simply by a set of numbers, which are most often substituted into the conditions (for convenience) without changing when the number of numerical parameters of the problem is small, and in graphs or in a matrix, where very a large array of data, it is always enough to control only some of the substituted parameters – two or three are enough – at the same time, it is not necessary for students to know which parameters you control most often.

# Cryptographic Analogies

The methods used in this procedure are very similar to several basic ideas of encryption and authentication algorithms used in cryptography and included in all textbooks.

A one-way function in cryptography is called a function that is difficult to calculate (or difficult to reverse) by computer standards. Of course, we have ideas about complexity associated with people.

First, we implement a one–way function - the calculation of the response vector. By complexity, we understand that it makes no sense for a student (within the framework of his qualifications) to get this result except by directly solving the problem.

We also solve the tasks of assigning keys to generate conditions and verification-authentication tasks.

Probabilistic methods typical in cryptography and the control ratios described above are also used for verification.

When solving the problem of assigning keys to generate conditions and choosing the power of their space, the role is played by the task of preventing mass coincidence – a collision of variants (keys), equivalent to the task of preventing packet collisions during block encryption.

* 1. *Example 1 (simple). Linear programming*

TASK: A linear programming problem is given.

Solve it by a direct and dual method, compare the answers.

Verification is performed by comparing the results (which should be equal)

## Example 2 (relatively more complex). Cournot Duopoly

TASK: The demand function is given, the unit production costs are given - for the first firm, - for the second.  - total amount of demand, - the general market price.

It is required to find the equilibrium of the Cournot duopoly (profit, price and output volume) for these two firms.

 - variables,  - parameters.

This is the «canonical» problem of game theory

# Evaluation of the Required Number of Options

The key collision problem is well known in cryptography. It sounds like this. There are *n* randomly selected keys (in another formulation, *n* small blocks of data, network packets, etc.) from a much larger number of *N* evenly distributed objects of this kind. What is the probability that two packages will be the same.

In cryptography, this makes it possible to crack, and in our case, the emergence of a «sticky student». We will consider it permissible to have no more than one shirking «stickler» per stream. The number of variants should be large enough so that with a sufficiently randomized assignment to a stream of 100 people, the probability of coincidences is minimal. Estimates show that more than 1000 options are needed, the task is well solved with 10,000 options.

# Calculation of the Variant and Generation of Personal Data

We used (*a*) - the number of letters of the surname («Фамилия»), (*b*) - the name («Имя »), (*c*) - the patronymic («Отчество ») (*abc*) and the number of the month of birth *d*(«Месяц рождения »)*.*

Example of generating PERSONAL DATA for severalpeople (TableI).

1. Example of Generating Unique Personal Data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Фамилия | Имя | Отчество | Месяц рождения | abcd |
| Ломоносов(9) | Михаил(6) | Васильевич(10) | Ноябрь(11) | 9-6-10-11 |
| Пушкин  (6) | Александр  (9) | Сергеевич  (9) | Май  (5) | 6995 |
| Менделеев  (9) | Дмитрий  (7) | Иванович  (8) | Январь  (1) | 9781 |
| Можайский  (9) | Александр  (9) | Фёдорович  (9) | Март  (3) | 9993 |
| Мечников  (8) | Илья  (4) | Ильич  (5) | Май  (5) | 8455 |
| Кулибин  (7) | Иван  (4) | Петрович  (8) | Апрель  (4) | 7484 |
| Пирогов  (7) | Николай  (7) | Иванович  (8) | Ноябрь  (11/1) | 7781 |
| Попов  (5) | Александр  (9) | Степанович  (10) | Март  (3) | 5913 |

Let's briefly talk about other options for assigning numbers. In the author's experience, there was a university where long numbers of credit books and student cards were fairly evenly distributed. In this case, they could be used as a legal number. An unexpected circumstance in this case is that the group of students who lost student cards and credits, as well as those who were just enrolled and transferred to the group that are in the process of registration, is quite large, which is why the described FULL NAME («ФИО») + *d* system was also constantly used either as the main or as a backup.

Among other things, this experience shows what problems a teacher who has tried to distribute numbers by direct appointment will have to face – there will be unacceptably many "hung up" students who, under a plausible pretext, did not receive a number or did not find out, and the lists of groups in any university are in motion.

To assign options and numbers, there is experience in using small texts in cryptography tasks – surnames(«Фамилия»), first names(«Имя »), patronymics(«Отчество »), digitized (Table II).

1. Digitization of the Letters of the Alphabet

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **а** | **б** | **в** | **г** | **д** | **е** | **ё** | **ж** | **з** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| **и** | **й** | **к** | **л** | **м** | **н** | **о** |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| **р** | **с** | **т** | **у** | **ф** | **х** | **ц** | **ч** | **ш** |
| 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 |
| **щ** | **ъ** | **ы** | **ь** | **э** | **ю** | **я** |
| 27 | 28 | 29 | 30 | 31 | 32 | 33 |

Table II was used only when encoding texts for programmers engaged in information security tasks. Its disadvantage compared to Table I is the inability to calculate the option without referring to the table. However, in special cases, this table can also be useful - when it is necessary to distribute options from a small sample very tightly and evenly (it is important when the complexity of the task increases too sharply with the growth of the sample and you need to use very dense samples of options). In the author's practice, the codes of the first letters of the surnames («Фамилия»), first names («Имя »), patronymics («Отчество »)or having almost ideal properties for the uniformity of distribution of their residues from division by 10 (i.e. combinations of the last digits of these 33-letter codes) were most often used.

Since the number of letters with the standard full name encoding (used in Table I) is not evenly distributed over the interval from 1 to 9 – this can be seen from Table I – by the first three digits of *abc*, it actually varies in a much narrower interval – the standard deviation of the order of 2-3, then the effective number of variants obtained in this way is less than 10,000. If 5 independent variants are considered effectively for each *abc* digit, then the effective number of combinations of variants is

.

What corresponds to the flow of students

 students.

At the same time, it is clear that the month of birth in this system is ABSOLUTELY necessary. If we also excluded the month of birth, we would have an acceptable collision-free sample of only 10 people:

students.

# Verbal Description

In general, the student achieves the vector of three goals: understanding, design, explanation to a friend. The latter leads both to the consolidation of understanding and to the formation of an important PEDAGOGICAL skill and the SKILL of WORKING WITH PEOPLE. The student receives a vector of skills – in addition to the educational (*z*), also pedagogical (*u*).

During the course, the student solves problems from a certain set of sizes offered by the teacher in the classroom.

Each of them he can:

* Decide for yourself.
* After that, decide for a friend.
* Explain to a friend.
* Decide after the explanation given by a friend.
* Write off OFFICIALLY or «HONESTLY» write off. Ask a friend to solve them with their numbers (buy for communication bonuses). The author, as a teacher, accepts this option.
* To write off MEANLY («not honestly») – that is, to write off NOT your own option, but a friend's option (risking getting «in full» into the blacklist).

The first 4 options are welcome, the fifth – «honest cheating» due to the features of the system, which is associated with high costs (for a loafer) – it is allowed that capable students should not mess around (another option of honest cheating works as a kind of excess pressure relief), the sixth option is punishable by a complete recheck of all, including previously passed control papers.

The 5th option is equivalent to teaching a friend, only a friend is an excellent student who is somehow motivated by failing students, and his «motivation» is a demotivating penalty (in the usual sense) any capable student from the hack: excellent students usually do not want to solve a lot (a lot are options for the whole group, which, as a rule, is unacceptably much), which makes the resource of cheating extremely limited (rare in the economic sense) and it is used only by the most hopelessly «needy» quitters.

At the same time, the 5th option reduces or completely frees up the teacher's time, which he would have spent additionally explaining and organizing the nth iterations of consultations and retakes during extracurricular time or at the expense of normal students' time.

It is significant that many students are trying to deliberately torment the teacher, and if this strategy fails, then they solve it in much less time, with the wording «why didn't you say right away that you wouldn't accept fewer tasks». We can assume that this is the 7th strategy, it is usually applied in a session.

For completeness, we will also mention the «strategies»:

* To look for holes in the dam (to torment others - the head of the department and other teachers)
* To look for dishonest workarounds (somehow to bribe the dean's office, which is not everywhere, but it occurs).
* (Almost) criminal activity (forgery of signatures in the statements).

# The Task of Managing Resources when Checking Tasks

The task can be considered as a population game of students and a teacher.

The student's task:

* To minimize their costs depending on
* Personal abilities
* Time for full-time and correspondence classes
* Personal interest or ambition (which can be formalized as a lower limit on the average score or the mathematical expectation of the assessment)

The costs of a student consist in spending time and effort on:

* Cheating in class.
* The solution in the lesson.
* Cheating outside the audience.
* Solution outside the audience.
* Motivation of a friend to explain in the audience and outside the audience.
* Motivation and, even, hiring a friend to solve your option.

The student's goal is to set the required number of points for the target assessment.

The costs of the teacher consist in:

* Explanation of the material to students who were not in class
* Conducting additional tests for students who were not in the classroom / lecture
* Verification of works, explanation of errors.
* Conducting lectures and tests and classes.

Conducting control tests on the system when communication is impossible (one or three options) is more expensive than on the system when everyone has an individual option (the time and effort of the teacher is less).

The objective functions of the teacher are to maximize the total load on students and the amount of material covered.

# Mathematical Model

Depending on the specific circumstances, all this can be described by some functions (based on Fig. 3).

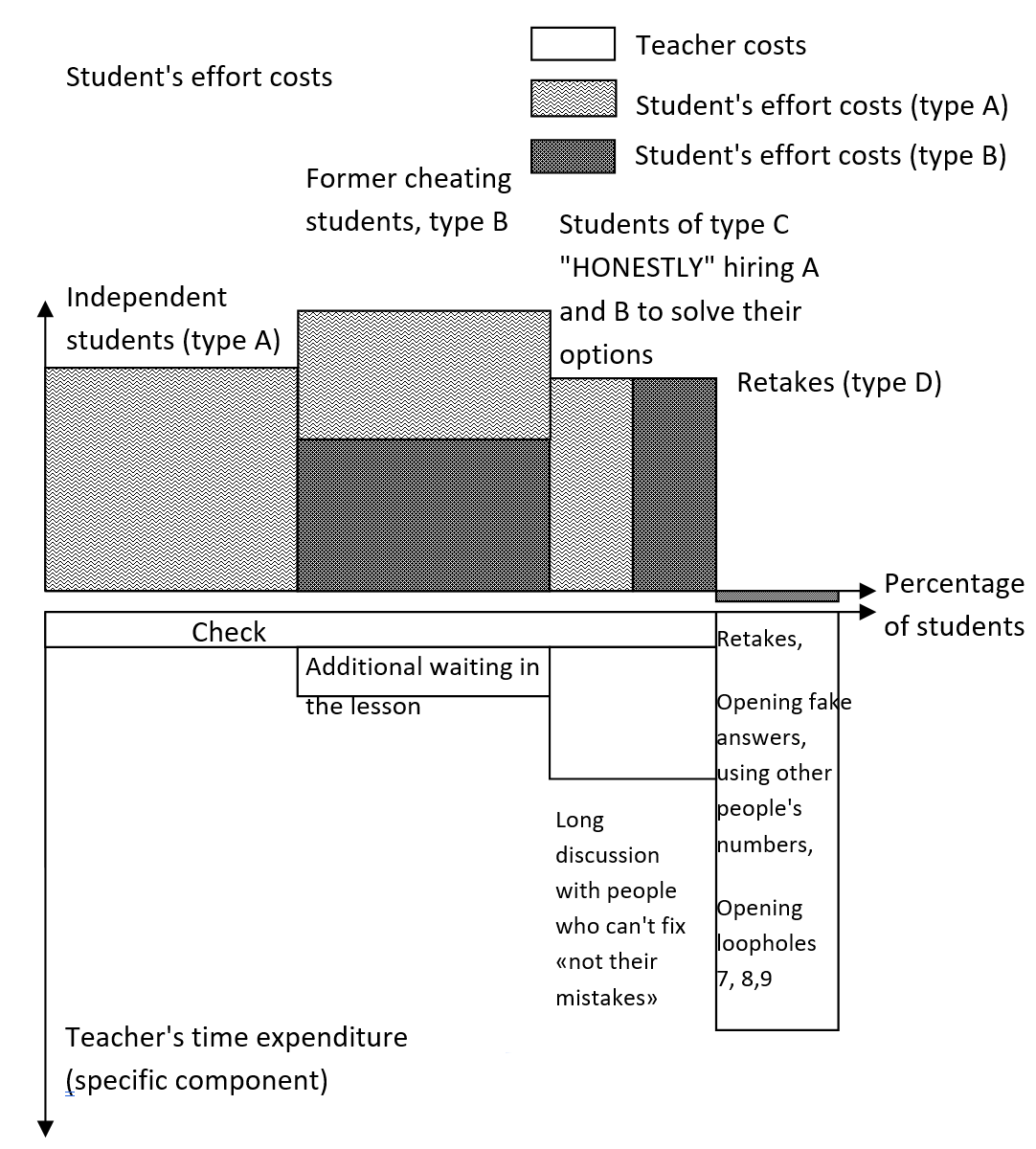
Denote - the proportion of students who,

- they do it themselves,

- resort to the help of a friend for additional instruction,

- honestly hire a friend to solve with their numbers,

- they try to solve the issue not during (retake or session) or dishonestly (from corruption to forgery).



1. The costs of the student's forces in the semester and the costs of the teacher's forces, including during retakes.

With the growth of the time T, which the teacher allocates for the performance of work and explanation in the distribution.

.

There is a «flow of probability» from right to left.

Suppose that each student is characterized by an ability, which we will describe as

 - the speed of problem solving. We assume that students are ordered by abilities, and is the inverse function of the probability distribution, more precisely its complement to 1:

.

During T, each student can (potentially) decide

tasks.

Approximate share will be found from the equation.

Thus, the share is determined by the last student who can barely decide for himself in the allotted time T.

 we will find, more precisely, approximately estimate from above, from the equation

.

The meaning of this equality is that some excellent students, especially the most capable, have time to help others.

In fact, the process is more complicated - for one student, one minor hint can lead to a breakthrough and increase his productivity at times (the hint falls on fertile soil - this is the very case when the «RAM» limit is simply being expanded), and the other does not understand anything «dead swan» excellent student, sorry, you need to pull out from under the rubble «on yourself», for the time for which you can briefly – at the level of ideas to suggest ten other students.

Students are those who were able (perhaps by some effort) to interest students from the two previous groups to decide for them (let's assume that events occur outside of school hours). Individually , this strategy is simplified for representatives of this class with the growth of the number of groups .

Group D students are those who have neglected previous opportunities. The more of them, the less time the teacher devotes to earning authority, i.e. the general idea of the inevitability of the penalty, in the sense of solving the minimum required number of tasks according to the standard and punishment for non-compliance with such a standard.

Most of the TEACHER'S PERSONAL TIME is spent on group D directly or indirectly in specific terms - payment for the need to earn authority (to maintain a system of punishments so that there are not too many violators).

# Optimization Problem

With the coefficients of the previous example



The load on the teacher at the same time (in order of magnitude) will increase by 10-100 times.

This means that in the case of traditional nonparametric control within the same time limit for total hours, the parametric form allows us to consider a larger number of tasks.

# Conclusion

A system is proposed for calculating individual conditions from PARAMETRIC conditions by substituting randomly distributed personal data of students in them – GENERATION KEYS, usually consisting of the number of letters in the surnames («Фамилия»), first names («Имя »), patronymics («Отчество ») (+ the number of the month of birth).

We have shown that the minimum cardinality of a set of variants determined by the size of the GENERATION key space should be 1000-10,000, and more generally the square of the number of the group. This is necessary to avoid a COLLISION of options when they are independently assigned. The uneven distribution of the number of letters in the surname, first name, patronymic in the compartment with the need to get 1000-10, 000 variants, led to the need to connect another fourth numeric parameter - the number of the month of birth.

The system has unexpected advantages:

* The task can be printed on a small piece of paper, as well as demonstrated on one slide and even in the corner of the slide.
* There is an opportunity for collective knowledge exchange. At the same time, strong students who decide quickly are engaged with weak ones, which ensures a more efficient use of the teacher's resources.
* There is an opportunity to combine a lecture, a control, a seminar with the involvement of three-year students who have got into the deepest wilds of the class, and, most importantly, with a threefold increase in the speed of mastering; at the same time, the system of Physical Technology is partially reproduced - to give tasks together with the development of theory, and sometimes even instead of and as the only means of mastering it.
* An experienced teacher's idea of what range the answers should be in, combined with the verification key, allows you to quickly weed out wrong decisions and prevent «resourceful» attacks of various types.
* Direct cheating is impossible, and solving all the flow options is very expensive in intellectual terms for such capable students who could do it (as a result, it is expensive in one sense or another and for those students who can resort to their services).

##### References

1. KrivosheevO.I.. «abcd-параметрический задачник» (abcd-parametric problem book), free web access: <http://yadi.sk/d/oGqU2cIYDYyid>
2. Knut D.. «Искусство программирования в 4 томах» (The Art of Programming in 4 volumes), MIR Press, 1977 (in Russian)
3. Automated Problem Generation for Education. Free web access: <http://research.microsoft.com/en-us/projects/problem-generation/default.aspx>
4. Generation of didactic materials in mathematics. Free web access: <http://comp-science.narod.ru/matem/matem.html>
5. Automatically Generating Algebra Problems. Free web access: <http://research.microsoft.com/en-us/um/people/sumitg/pubs/aaai12-full.pdf>