Kybernetika

Jozef Brody

Error frequency - regulator of the teaching algorithm

Kybernetika, Vol. 4 (1968), No. 6, (548)--569

Persistent URL: http://dml.cz/dmlcz/125461

KYBERNETIKA ČÍSLO 6, ROČNÍK 4/1968

Error Frequency - Regulator of the Teaching Algorithm

JOZEF BRODY

The article brings a description of the mathematical formulation of the effectivity problem in the teaching process. As an example of the problem formulated, the construction of a mathematical model of learning, applying the error frequency as an only regulator of this process, is herewith quoted. The resulting construction of the non-Markov algorithm is performed by means of an automatic computer.

INTRODUCTION

Reduction of the non-Markov algorithms

According to L. N. Landa [1] the algorithm is to be understood as "a strictly obligatory sequence of operations, solving all the tasks of some class given". On the basis of so presented definition of the algorithm, the algorithmization of a certain matter is then a determination of the obligatory sequence of operations solving all the tasks implied in this matter. Should thus the algorithm represented by the sequence of operations

$$(1) \qquad \{o_1, o_2, \ldots, o_n\}$$

teach the given matter, then the sequence must teach all the tasks of this matter given.

In this article I shall deal with the teaching algorithms of the following type:

$$(2) \qquad o_j \equiv i_j \ \& \ r_j\,, \quad j = 1, 2, \ldots, n-1\,,$$
$$o_n \equiv i_n,$$

where i_k is the k-th instruction and r_j is the student's reaction to the instruction i_j, while the index $k = 1, 2, \ldots, n$; and the index $j = 1, \ldots, n-1$. The instruction i_k is to be understood here as a direction (theory) for solving the tasks of a certain type, or directly the task that should be solved. Thus the reaction r_j to the instruction i_j is then directly the solution of the given task. In case the task was in the instruction i_j (as far as it is contained in this instruction) solved correctly, the reaction r_j will be substituted by the sign $+$, in an opposite case $r_j = -$. Provided that the instruction i_j does not contain any task to be solved, then $r_j = 0$.

Definition 1. The function f_j, which by means of the sequence of reactions

(3) $$\{r_1, r_2, \ldots, r_j\}$$

assigns the instruction i_{j+1}, is called "the decision function".

Definition 2. If the algorithm (1) is given, in which the relations (2) hold and the system of decision functions

(4) $$\{f_1, f_2, \ldots, f_{n-1}\}$$

forms those functions for which

(5) $$f_j(r_j)\,,\, i_{j+1}$$

holds, then we may say that the algorithm (1) is a Markov algorithm.

Note. The decision function f_j is here the function of an only variable r_j and its value is therefore independent on the reactions $r_1, \ldots, r_{j-1}$ .In other words: the instruction i_{j+1} depends on the reaction r_j only.

Definition 3. If the algorithm (1) is not a Markov one, it is called "non-Markov algorithm".

Note. As in the non-Markov algorithm the decision functions (4) are generally (for at least one j) also functions of the reactions r_k, where $k < j$, then the instruction i_{j+1} is generally dependent on the whole sequence of reactions (3), i.e.

(6) $$i_{j+1} = f_j(r_1, r_2, \ldots, r_j)\,.$$

Thus it may be said that the instruction i_{j+1} depends on the whole history of the reaction preceding it.

From the pedagogical point of view the non-Markov algorithms have a great importance just because they do not neglect the history of the whole process. That is the reason why a great number of teaching machines, solving this situation at least partly, is arosen. Partial solution consists in the fact that the decision in these machines is not conditional on the last reaction, but on a certain number of the preceding ones. The solution is performed in such a way that these machines have the s-step storage, where s is mostly smaller than 7, which limits the history length of the whole process This situation is given by the technical possibilities which, in conformity with the economical aspect, do not permit a more complicated solution of the whole problem. At the 5- (or 6-) step storage this solution appears practically as convenient. In spite of this it happens sometimes that the given task would require a storage of more steps. But due to a small number of these tasks, solution of this situation seems not to be necessary.

Nevertheless, I would like to seek a certain solution of this problem. Let us presume that the given algorithm contains the binary reactions $(+, -)$ only. This presumption is not to the detriment to the generality of the problem.

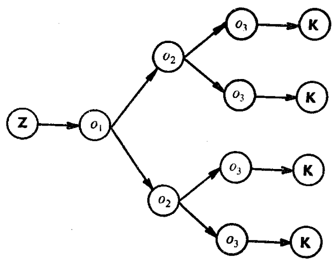


Fig. 1.

Whole the situation may be graphically plotted by the multigraph (Fig. 1), where **Z** is the beginning of each algorithm; o_i is the i-th operation in the respective algorithm; $\xrightarrow{+}$ – the respective reaction is $+$; **K** is the end in the resp. algorithm. Fig. 1 shows the multigraph containing 4 possible algorithms, because I have mentioned here the situation for $n = 3$ as an example. The sum m of all the possible algorithms is given, in general, by the relation

$$m = 2^{n-1}\,, \tag{7}$$

Each trajectory $\mathbf{Z} \to \ldots \to \mathbf{K}$ represents one possible algorithm.

As it is immediately evident, one of the greatest problems will be the extensive growing of the number m of all the possible algorithms in the multigraph plotted. Thus my endeavour consisted in the reduction of this great quantum of trajectories.

The construction of the concrete multigraph is always preceded by a certain criterion **H**, no the basis of which the number of all the possible algorithms involved in this multigraph may be already partly reduced. In the further stage we may then pass on to the testing, where on the basis of its results those trajectories, used only with a very small frequency, may be determined. Further reduction may be then performed by omitting these trajectories.

The construction of the multigraph by means of the criterion **H**, as well as the employment of a certain effectivity measure $\boldsymbol{F}$, will be described in the further chapters. The criterion **H**, as well as the effectivity measure F is namely coherent with the pedagogically-psychological structure of the teaching process. A very important question is here the solution of so called effectivity problem, the formalization of which is important just to enable to solve thus arosen problem by mathematical means.

Due to a great numerical difficulty, I have solve the concrete example by the help of a computer. On the basis of thus achieved results, I have set up a multigraph (Fig. 2) the construction of which I wish to clear up more closely in further chapters.

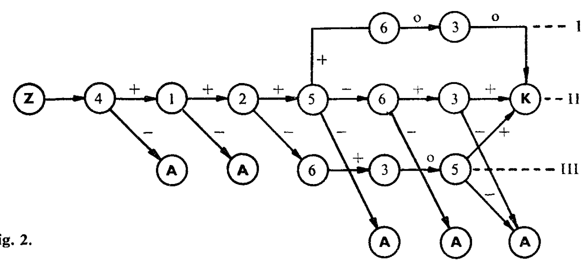


Fig. 2.

Points **A** are here the nodes in which the teaching process cannot be regulated without affecting the criterion **H**; *i*'s are the tasks solved in the respective step; **Z**, **K**, $\xrightarrow{\pm}$ have here the same signification as in Fig. 1.

This multigraph shows that the position of individual steps is divided into three strata. Then it is sufficient for the teaching device to store only the respective stratum in which the student is just situated. In this way the device simultaneously takes into consideration the whole history of reactions that later decides the further procedure. Meanwhile I have described only one multigraph in this way and it would be therefore very premature to declare a categorical judgment on the solution of non-Markov processes accordingly. A whole range of experiments is just performed at the Faculty of mechanical engineering that will surely suggest much more on the mentioned problem.

In further chapters I wish to clear up the pedagogically-psychological aspects which led me to the choice of the criterion **H** and the measure F, more closely. But first of all I would like to present some basic definitions on the basis of which I would elaborate a concrete example.

I. EFFECTIVITY PROBLEM IN THE TEACHING PROCESS

Likewise any activity, also the teaching is affected by a range of determining parameters (time, tiredness, feedback, economical stimuli, etc.). Therefore if we wish to speak about the effectivity of a given activity, it is necessary to consider this activity as a system of all its determinants.

Should the process **P** (an activity passing in time) and a system of its determinants $\mathscr{S}$ be given, then we must still construct the effectivity measure of the process **P** that will give a true picture of a certain aspect on the parameters' effectivity from the system $\mathscr{S}$. The effectivity measure will then be the real function $F(x)$, where x are the real values of parameters from the system. Because $\mathscr{S}$ may be generally a system of the infinite bulk, it is advantageous to choose a sub-system $\overline{\mathscr{S}} \subset \mathscr{S}$, that would be a final one, containing just all the parameters from the system $\mathscr{S}$, which are important for the process **P** (due to criterial function F). Importance of the respective parameter may be judged e.g. by means of the methods of the mathematical statistics [2] (e.g. the Student's test of importance); or their importance is given by the conception.

Thus we have arrived at the fact that the process **P** is determined by the system of $\mathscr{S}$ parameters, from which we have chosen the $\overline{\mathscr{S}}$ sub-system, the bulk of which is finite and which contains just all the important parameters of the process **P**. It is evident that the creation of the system $\overline{\mathscr{S}}$ is very complicated and largely laborious. It is a very important work for the creation of a mathematical model of teaching, where the teaching process is considered as a system. As those parameters, chosen into the system $\overline{\mathscr{S}}$ cannot acquire all values, it is further necessary to create a limitation system for values of those parameters.

For example the time – its values in hours must fulfill the prescribed inequalities; the expenses – counted in money units – must not exceed the certain limit prescribed; etc.

If X_i is the parameter of the system $\mathscr{S}$, I shall denote its (real) value by x_i. This value is quoted by a real number indicating it in an appurtenant degree, i.e. the time in hours, expenses in money units, etc.

Let us further presume that the system $\overline{\mathscr{S}}$ has the final bulk k, i.e. that the set of the parameters' values from $\overline{\mathscr{S}}$ is a sub-set of the room $\mathbf{R}_k$, the elements of which are arranged k-tuples $x(x_1, \ldots, x_k)$. This sub-set will be denoted by $\mathbf{M}$. Each point of the set $\mathbf{M}$ is then the k-tuple of values determining a certain state of the process **P**. The set $\mathbf{M}$ may be formed by means of a r-tuple of inequalities

$$(8) \qquad \begin{aligned} &g_1(x_1, \ldots, x_k) \geqq 0\,, \\ &g_2(x_1, \ldots, x_k) \geqq 0\,, \\ &\ldots\ldots\ldots\ldots\ldots \\ &g_r(x_1, \ldots, x_k) \geqq 0\,. \end{aligned}$$

The inequalities (8) determining the set $\mathbf{M}$ are a mathematical expression of the criterion **H**. Functions $g_i(x)$ are the real functions of real variables $x_1, \ldots, x_k$.

Now, when for the process **P**, the system $\mathscr{S}$ and its sub-system $\overline{\mathscr{S}}$, as well as the criterial function F and the set $\mathbf{M}$ are determined, we may already define the further terms:

Definition 3. We shall call the set $\boldsymbol{M}$ the set of all admissible values of parameters of the system $\overline{\mathscr{S}}$.

Definition 4. Let the process $\mathbf{P}$ be affected by a k-tuple of admissible values $x^1(x_1^1, \ldots, x_k^1)$ and by another k-tuple of admissible values $x^2(x_1^2, \ldots, x_k^2)$. The function F, which is the function of *parameters' values*, will then in x^1 acquire the value $F(x^1)$ and in x^2 the value $F(x^2)$. Let us say that the process $\mathbf{P}$, being affected by parameters the value of which are x^1, is more effective than if being affected by parameters with values x^2, if

$$F(x_1^1, \ldots, x_k^1) > F(x_1^2, \ldots, x_k^2) . \tag{9}$$

It may be analogously defined when the process $\mathbf{P}$, being affected by two different k-tuples, is likewise effective (instead of the inequality (9) there is an equality). Classifying the set $\boldsymbol{M}$ into several classes, we obtain the total arrangement defined by the inequality (9). Namely for two arbitrary admissible k-tuples of values, one of the following three relations must always hold:

1. the first is more effective than the second,
2. the effectivity of both is equal,
3. the second is more effective than the first.

Note. The set $\boldsymbol{M}$ which is limited by the criterion $\mathbf{H}$, is sometimes called "the variety reduction" [3]. In the theories of mathematical programming [4] and [5], the reader may be more closely acquainted with the concepts I have defined in definitions 3 and 4. The continuity with the theory of games and with the statistical decisions [8] will surely not pass unnoticed by an attentive reader.

Provided that it is possible to form the mathematical model of teaching up to this state, the further endeavour will be to choose a point with the optimum effectivity on the set $\boldsymbol{M}$. This problem often called "the planning of the teaching process" is in fact the problem of the optimum regulation. Thus our endeavour will now consist in the choice of such a k-tuple of values $x \in \boldsymbol{M}$ of parameters $X \in \overline{\mathscr{S}}$, pertaining to the class with the maximum value $F(x)$. At is will be further evident, in the concrete example the optimum arrangement of tasks will be considered and just the function F decides it.

II. CREATION OF THE SYSTEM $\overline{\mathscr{S}}$. THE ERROR FREQUENCY-REGULATOR OF THE TEACHING ALGORITHM

Algorithm, the construction of which I shall here perform, is set up of an operation sequence (1), where each the operation is composed of the instruction represented here by a task, and of the reaction as described by the relations (2). The reactions will be here of two kinds and I shall denote them by the signs + and −, as I have already described in the introduction.

Beside these n operations of the constructed algorithm (1), an aim task is given. This task has to enable the investigation on the fulfilment of the aim of the operations' sequence, the solution of which is subject of the algorithm (1).

It is possible to attain this aim theoretically, even without using the algorithm (1) in case the student was already somewhere acquainted with the matter before having used this algorithm. However, the fulfilment of the aim may be achieved also by some sub-sequence of operations

$$\{o_{i_1}, o_{i_2}, \ldots, o_{i_m}\} . \tag{10}$$

Strictly speaking; there is given an aim of the sequence in question, to which a sequence of tasks is determined by

$$i_1, \ldots, i_n . \tag{11}$$

Passing of this sequence (11) is thus possible by N means, where

$$N = \sum_{j=0}^{n} j! = 1 + 1 + 2 + 6 + \ldots + n! . \tag{12}$$

All these strategies in a number N form an area of admissible strategies $\boldsymbol{M}$. To each the strategy $\boldsymbol{Y}$ its value $\boldsymbol{F}(\boldsymbol{Y})$ is assigned and this may be achieved by means of the effectivity degree $\boldsymbol{F}$. This value will be set up of certain probabilities and I shall here described its construction in details later.

The model I am describing here is thus a stochastic model. At the construction of the criterial function $\boldsymbol{F}$, I shall use the error frequency parameter, i.e. the number of errors on the sequence, arosen at the solution of tasks (11). I have chosen the error frequency parameter from several reasons:

a) it is relatively an easily measurable parameter,

b) it is an important output indicator,

c) at the feedback the knowledge of the error affects the teaching motivation.

The third reason led me to the formulation of the basic hypothesis:

Hypothesis. *It is necessary to keep the error frequency in certain limits, otherwise the teaching motivation will be substantially decreased.*

I did not deal with testing of this hypothesis, so that I do not know either the optimum limits. But I have proceeded from the suppositions that too much conscious errors cause the frustration and too much conscious good answers may cause psychical distraction that will necessitate the lack of interest in the given matter. Some knowledge determining the activation continuum [3] from the psychological distraction point of view are known in this line. Several authors [6], [7] have already dealt with the error action in the teaching process. Of course, the limitation from above, so as from below will be individual for each person. It should be thus expedient to establish the testing methods which would determine the searched limits more closely.

The construction that I will describe here enables then only to construct the programs for these pre-established limits. For illustration I have chosen here the limits 0; 1, i.e. the program permits to pass the tasks' sequence with one error maximally. If denoting the error parameter by $\boldsymbol{X}$, its value, i.e. the number of errors on the given sequence will be denoted by x. The inequalities (8) defining the set $\mathbf{M}$ have the shape

$$x \geqq 0 \, ; \quad 1 - x \geqq 0 \, . \tag{13}$$

At the construction of the effectivity degree I have proceeded from searching an admissible strategy, the realization of which has the highest probability. From the room of all the strategies $\mathbf{M}$ I have first chosen the sub-room $\mathbf{M}_n \subset \mathbf{M}$, which contains $n!$ strategies supposing solution of each sequence task.

Now the matter will be to determine the succession $\boldsymbol{\Pi}(\pi_1, \ldots, \pi_n)$, in which the student has to pass the sequence (11), i.e. to determine the succession

$$i_{\pi_1}, \ldots, i_{\pi_n} \tag{14}$$

from which the algorithm (1) is then set up. Probabilities which I needed for the construction of the measure $\boldsymbol{F}$ were obtained by the help of testing. At this testing I have achieved only relative frequencies which serve only as an approximation for the probability values. The further question will be: with which accuracy the sought probabilities may be approximated by the relative frequencies at a given extent of choice from the whole population? The Bernouilli's theorem [2] gives a statement on the convergence of the relative frequencies. In this description I am substituting the relative frequencies for the probabilities directly.

From the testing results I have determined the following frequencies:

$$\begin{aligned} & a_0 \ \ldots \ \text{number of the tested probants}\,, \\ & a_1 \ \ldots \ \text{number of correctly solved tasks } i_1\,, \\ & a_2 \ \ldots \ \text{number of correctly solved tasks } i_2\,, \\ & a_3 \ \ldots \ \text{number of correctly solved tasks } i_1 \,\&\, i_2\,, \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & a_{2^n-1} \ \text{number of correctly solved tasks } i_1 \,\&\, \ldots \,\&\, i_n\,. \end{aligned} \tag{15}$$

The frequencies (15) are arranged in the following way: to the task i_j the index 2^{j-1} is assigned and to the simultaneous occurence of the tasks $i_j \,\&\, i_k$ the sum of their indexes is assigned, i.e. index $2^{j-1} + 2^{k-1}$. It is analogously also at the occurrence of more tasks simultaneously. For instance, if the task $i_1 \,\&\, i_3 \,\&\, i_5$ is correctly solved, then that frequence indicating the number of correctly solved tasks $i_1 \,\&\, i_3 \,\&\, i_5$, has the index $2^{1-1} + 2^{3-1} + 2^{5-1} = 1 + 4 + 16 = 21$. The number of the correct solutions quoted is then indicated by the frequency a_{21}.

On the basis of the above mentioned approximations I have calculated the conditioned probabilities, which are assigned to each the stragey $\Pi \in \boldsymbol{M}$, i.e. to the n-th $\pi_1, \ldots, \pi_n$:

$$p_{\pi 1} = \frac{a_{2^{\pi 1-1}}}{a_0} = \frac{a_{01}}{a_0} = P(A_1),$$

(16)
$$p_{\pi 2} = \frac{a_{01+2^{\pi 2-1}}}{a_{01}} = \frac{a_{02}}{a_{01}} = P(A_2/A_1),$$

....

$$p_{\pi n} = \frac{a_{0n-1+2^{\pi n-1}}}{a_{0n-1}} = P(A_n/(A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_{n-1})),$$

where A_j is the phenomenon describing that the task i_j will be solved correctly.

I shall now set up the arosen conditioned probabilities (16) into the non-decreasing n-th:

(17)
$$\bar{p}_{\pi 1} \leqq \bar{p}_{\pi 2} \leqq \cdots \leqq \bar{p}_{\pi n},$$

where $\bar{p}_{\pi i} = p_{\pi j_i}$, j_i being a suitable transposition.

Should the probability approximation by relative frequency be performed with the accuracy $10^{-\omega}$, where ω is the number of correctly computed decimal places, then the probability measure, at the solution of tasks in the succession given by the formula (14) without error, will be determined by the expression

(18)
$$\bar{F}(\Pi) = \bar{p}_{\pi 1} \,.\, 10^{n\omega} + \ldots + \bar{p}_{\pi n} \,.\, 10^{\omega}.$$

The reason that led me to this construction of the measure $\bar{F}$ will be explained a little later.

Because at solving the tasks (11) only one erroneous answer is admissible, i.e. one of the tasks of the sequence (11) may be erroneously solved or unsolved at all, the most advantageous here is to consider that the error will occur in that task, the correct solution of which has the smallest probability in the succession (14). The error may be thus supposed in the j_i-th step when solving the tasks in the succession (14). The task $i_{\pi j_i}$ in the succession (14) has namely the probability of a correct solution $p_{\pi j_i} = \bar{p}_{\pi j_i}$. This probability, as it is evident from the arrangement (17), is just the smallest one. The uncorrect solution of the task $i_{\pi j_1}$ is then $1 - \bar{p}_{\pi 1}$.

So far as this probability is higher than that of the correct solution, I shall consider an erroneous solution in this step taking into account the arrangement (14).

Thus, if

(19)
$$\bar{p}_{\pi 1} < \tfrac{1}{2}$$

I shall replace the value $\bar{p}_{\pi 1}$ by the value $1 - \bar{p}_{\pi 1}$ and in the task $i_{\pi j_1}$ in the succession (14) I shall suppose an erroneous solution. This new n-th of probabilities I shall arrange again in a non-decreasing succession, i.e. the probability $1 - \bar{p}_{\pi 1}$ I shall suitably range in (17), whereby I shall obtain a new arrangement

(20) $$p_1^\pi \leqq p_2^\pi \leqq \ldots \leqq p_n^\pi .$$

The arrangement (20) is identical with the arrangement (17) there and then, if the inequality (19) is not fulfilled. Thus I shall achieve the sought degree

(21) $$F(\Pi) = p_1^\pi \,.\, 10^{n\omega} + \ldots + p_n^\pi \,.\, 10^\omega$$

that already employs those possibilites described by the inequalities (13).

The degree F, determined by the formula (21) should be still correctly normalized by dividing it by the coefficient

(22) $$\varkappa = 10^{n\omega} + \ldots + 10^\omega$$

whereby

(23) $$[F(\Pi)/\varkappa] \in \langle 0; 1 \rangle$$

would be achieved.

The normalization (23) is negligible for the purposes of this description and therefore it will be omitted in further.

Note. Importance of the construction of the measure F consists in the fact that the effectivity of the individual strategies decides first the minimum conditioned probability p_1^π. In case this minimum conditioned probability at two or more strategies is equal, then that conditioned probability decides which, due to its magnitude, is the second one in the succession, i.e. p_2^π, etc. Provided that the decision would be made by means of the conditioned probability p_1^π only, then it may happen that for some other strategy $\bar{\Pi}$, for which $p_1^\pi = p_1^{\bar{\pi}}$ would hold, the initial intention might be disturbed. It might namely happen that after having passed through the critical step, i.e. the step with the minimum conditioned probability p_1^π, or $p_1^{\bar{\pi}}$, the rest of the strategy Π with respect to its realization would be less probable than the rest of the strategy $\bar{\Pi}$. This is prevented just by the measure F given by the formula (21).

Thus, by the help of the relation (21) we shall seek for that strategy Π, for which the measure $F(\Pi)$ is the maximum one, where $\Pi \in \mathbf{M}_n$.

Let us now suppose that the strategy Π_n is already that optimum strategy in the area of strategies $\mathbf{M}_n$. Thus I have proceeded as follows:

Provided that in the first step the student will conform to the strategy Π_n, i.e. he willl solve the first task in this strategy so, as it is considered, then he will continue in this strategy also in the further task. In an opposite case it is necessary to present him a substitutional strategy. In case the first task, determined by the strategy Π_n, supposed a correct solution, but the student did not solve it correctly, then he is instructed as long as it may be considered that the aim, traced by mastering of that first task, was reached. The substitutional strategy that is presented to the student in further steps, must not more allow an erroneous solution, because the student has already caused one. The substitutional strategy is then chosen from the sub-room $\mathbf{M}_n^{\pi n, 1} \subset \mathbf{M}_n$, where $\mathbf{M}_n^{\pi n, 1}$ is the set of all the n-member strategies which have the same task in the first step, as the strategy Π_n. To each strategy from the set $\mathbf{M}_n^{\pi n, 1}$ I shall put the measure $\bar{F}(\Pi)$ determined by the formula (18), i.e. the measure that does not presume the erroneous solution in the respective strategy. However, provided that the initial strategy Π_n in its first step would suppose an erroneous solution of the respective step, but the student would not have done an error here, then the procedure is the following: in contradistinction of the previous alternative, when it was necessary to present a set of substitutional tasks to which the respective instructions were given, it is not necessary to perform any instruction here. The substitutional strategy is chosen again of the set $\mathbf{M}_n^{\pi n, 1}$, but the measure $F(\Pi)$ determined by the formula (21) is here placed to these strategies, i.e. a measure supposing one erroneous solution.

In both cases that substitutional strategy is sought, whose measure is the maximum one. In this way we may proceed in all the cases, provided that one erroneous solution would not precede. Namely, in those places, where one erroneous solution preceded, no erroneous solution may be admitted more, beacuse it might break the criterion described by the inequalities (13). In these places, instead of a substitutional strategy, is then a so called "autoregulation node". The student, who will then come into this node, cannot be regulated more by any admissible strategy without disturbing the criterion **H**. Therefore he must here decide himself, whether he will further continue, i.e. he will solve the substitutional set of tasks and will continue in that strategy, which preceded the last error. Nevertheless, he has another possibility here that enables him to study the matter from a traditionally written textbook which he will have recommended by the instruction in this so called autoregulation node.

In this way I shall set up n strategies

$$\{\Pi_n, \Pi_n^1, \ldots, \Pi_n^{n-1}\}, \tag{24}$$

and $n - 2$ autoregulation nodes

$$\{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_{n-2}\}, \tag{25}$$

which form then together an unreduced non-Markov algorithm for the sequence of tasks (11).

All the possible strategies and the so called autoregulation nodes may be plotted then into a multigraph. This multigraph may then be subjected also to a further reduction. Namely the reduction may be performed by omitting all the trajectories which, at the above mentioned testing, had a zero frequency. It means that from this multigraph we shall omit those trajectories which were not used by either one of the tested. Thus we gain a resulting reduced multigraph that already gives the resulting algorithm. This is a non-Markov algorithm and it serves as a basis for the construction of a teaching program.

Note. The reduction performed on the basis of certain frequencies of individual trajectories may be, however, performed in an other way, too. I would not like to deal with this particular problem in this article. This problem is closely connected with the mathematical statistics and with the dispersion theory.

III. CONCRETE EXAMPLE OF THE REDUCTION MULTIGRAPH CONSTRUCTION

The sequence I have tested was chosen from the matter of "Analytical geometry in plane". From this matter I have chosen that part dealing with transformation into the polar coordinates. The aim of this sequence of tasks was to verify the student's knowledge concerning the transformation of carthesian coordinates into polar ones and conversely. By means of the mentioned sequence the application abilities of the transformation formulae to transform the curves' equations were further tested. The sequence of tasks consisted here of six tasks to which one aim-task was added. This aim-task had only to verify if the sequence of six tasks forms such tasks which may fulfil the desired aim as it is required by the definition of algorithm, presented by L. N. Landa [1].

Concretely I have given the following six tasks to be tested:

i_1 a) Which polar coordinates have the points $A(\varrho, \varphi)$, if they are lying on the positive semi-axis y?
b) Where is lying the point A, whose polar coordinates are $(\varrho, 0)$?

i_2 a) Which polar coordinates have the points lying in the third quadrant?
b) In which quadrant lie the points $X(\varrho, \varphi)$ for which $\operatorname{tg} \varphi < 0$?

i_3 a) The point A has the carthesian coordinates $(1, -1)$. Write its polar coordinates.
b) Write the transformation equations used in the example a).

i_4 a) Write the carthesian coordinates of the point A, whose polar coordinates are $(\varrho = 5, \varphi = 4\pi/3)$.
b) Write the transformation equations used when solving the example a).

i_5 a) Which transformation equations can be used at the transformation of the curves' equations from the polar shape into equations in the carthesian coordinate system?
b) Write the curve equation in carthesian coordinates if its polar shape is $\varrho = 2/(1 - 3 \cos \varphi)$.

i_6 a) If the curve is given by the equation in the carthesian coordinate system, write those transformation equations which may be used at its transformation into the polar shape.
b) Transform the curve equation $2x + 3y - 7 = 0$.

For completeness I wish to mention here also the aim-task c:

a) Into the polar system plot those points having the polar coordinates $(\varrho, \varphi) = (1, 0); (1, \frac{1}{2}\pi); (2, \pi); (1, 3\pi/2)$.

b) Determine the polar coordinates of the ϱ points lying on the curve $2x + 3y - 7 = 0$, the polar coordinate of which is $\frac{1}{6}\pi; \frac{1}{3}\pi; \frac{1}{2}\pi; \pi$.

Note. After having finished the testing, I have revised the given tasks and have found out that one task concerning the definition field of the polar coordinates is still missing. On the basis of results achieved by testing I have ranged all the frequencies into the Table 1 as described in (15). From these 64 frequencies it is already possible to calculate the effectivity measure F of any from the 720 (6!) successions. As a matter of fact, each the succession is one of the strategies chosen from the set of strategies $M_6 \subset M$.

Table 1.

$a_0 = 106$	$a_{16} = 18$	$a_{32} = 27$	$a_{48} = 7$
$a_1 = 60$	$a_{17} = 13$	$a_{33} = 22$	$a_{49} = 7$
$a_2 = 40$	$a_{18} = 10$	$a_{34} = 18$	$a_{50} = 5$
$a_3 = 31$	$a_{19} = 8$	$a_{35} = 15$	$a_{51} = 5$
$a_4 = 72$	$a_{20} = 17$	$a_{36} = 23$	$a_{52} = 7$
$a_5 = 53$	$a_{21} = 15$	$a_{37} = 20$	$a_{53} = 7$
$a_6 = 35$	$a_{22} = 9$	$a_{38} = 16$	$a_{54} = 5$
$a_7 = 28$	$a_{23} = 8$	$a_{39} = 13$	$a_{55} = 5$
$a_8 = 72$	$a_{24} = 17$	$a_{40} = 26$	$a_{56} = 7$
$a_9 = 49$	$a_{25} = 14$	$a_{41} = 21$	$a_{57} = 7$
$a_{10} = 36$	$a_{26} = 10$	$a_{42} = 18$	$a_{58} = 5$
$a_{11} = 29$	$a_{27} = 8$	$a_{43} = 15$	$a_{59} = 5$
$a_{12} = 61$	$a_{28} = 16$	$a_{44} = 22$	$a_{60} = 7$
$a_{13} = 46$	$a_{29} = 14$	$a_{45} = 19$	$a_{61} = 7$
$a_{14} = 32$	$a_{30} = 9$	$a_{46} = 16$	$a_{62} = 5$
$a_{15} = 27$	$a_{31} = 8$	$a_{47} = 13$	$a_{63} = 5$

In order to clear up the construction of the measure F more closely, I shall calculate its value for the strategy $\Pi(3, 1, 2, 5, 6, 4)$. The strategy then supposes a solution of the sequence tasks in the succession $i_3, i_1, i_2, i_5, i_6, i_4$. First I shall compute the conditioned probabilities, whose computation is given by the relations (16):

$$p_{\pi 1} = \frac{a_{2^{3-1}}}{a_0} = \frac{a_4}{a_0} = \frac{72}{106} = 0{\cdot}68\,; \quad p_{\pi 4} = \frac{a_{7+2^{5-1}}}{a_7} = \frac{a_{23}}{a_7} = \frac{8}{28} = 0{\cdot}29\,,$$

$$(26) \quad p_{\pi 2} = \frac{a_{4+2^{1-1}}}{a_4} = \frac{a_5}{a_4} = \frac{53}{72} = 0{\cdot}74\,; \quad p_{\pi 5} = \frac{a_{23+2^{6-1}}}{a_{23}} = \frac{a_{55}}{a_{23}} = \frac{5}{8} = 0{\cdot}63\,;$$

$$p_{\pi 3} = \frac{a_{5+2^{2-1}}}{a_5} = \frac{a_7}{a_5} = \frac{28}{53} = 0{\cdot}53\,; \quad p_{\pi 6} = \frac{a_{55+2^{4-1}}}{a_{55}} = \frac{a_{63}}{a_{55}} = \frac{5}{5} = 1\,.$$

Now I shall arrange these conditioned probabilities into the non-decreasing 6-tuple, as described in the formula (17):

(27) $$0{\cdot}29 \leqq 0{\cdot}53 \leqq 0{\cdot}63 \leqq 0{\cdot}68 \leqq 0{\cdot}74 \leqq 1 .$$

Because the relation (14) is fulfilled for this 6-tuple, i.e. $\bar{p}_{\pi 1} = 0{\cdot}29 < \frac{1}{2}$, at the construction of the measure $\boldsymbol{F}$ it is necessary to presume an erroneous solution in this place. As $0{\cdot}29 = p_{\pi 4}$, it is necessary to presume this erroneous solution in the fourth step of the considered strategy $\boldsymbol{\Pi}$, i.e. at the solution of the task i_5. The task i_5 will be then solved erroneously with the probability $1 - 0{\cdot}29 = 0{\cdot}71$. Now it is necessary to rearrange the newly created 6-tuple again, as described by the relations (20):

(28) $$0{\cdot}53 \leqq 0{\cdot}63 \leqq 0{\cdot}68 \leqq 0{\cdot}71 \leqq 0{\cdot}74 \leqq 1 .$$

Because I am computing here the probabilities with an "accuracy" of two decimal places, is $\omega = 2$. The resulting value of the measure $\boldsymbol{F}$, as quoted in the formula (21) is then

(29) $$\boldsymbol{F}(3, 1, 2, 5, 6, 4) = 536,,368,717.500 .$$

In this way all the 720 strategies are evaluated and that one, whose measure $\boldsymbol{F}$ is the maximum, is chosen. This evaluation way performed by the automatic computer URAL-2, which determined the strategy $\boldsymbol{\Pi}_6(4, 1, 2, 5, 6, 3)$, as the optimum one. In this strategy an erroneous solution is supposed in the fourth step, i.e. at the solution of the task i_5.

In the automatic computer all the substitutional strategies were also evaluated. These strategies were chosen from the sets $\boldsymbol{M}_6^{\pi 6,1} \subset \boldsymbol{M}_6$, i.e. the substitutional strategy for the first step is chosen from the set $\boldsymbol{M}_6^4$ and without error, for the second step from the set $\boldsymbol{M}_6^1$ also without error, for the third step from the set $\boldsymbol{M}_6^2$ without error, too. Because in the fourth step of the strategy $\boldsymbol{\Pi}_6$, i.e. at the solution of the task i_5 an erroneous solution is supposed, the substitutional strategy is chosen from the set $\boldsymbol{M}_6^5$ with a possibility of erroneous solution.

The sets $\boldsymbol{M}_6^{\pi 6,1}$ are here the sub-areas of strategies, the composition of which is conditioned by the strategy $\boldsymbol{\Pi}_6(4, 1, 2, 5, 6, 3)$. It means that e.g. the sub-area of the strategies $\boldsymbol{M}_6^1$ contains all the strategies, which have common steps with the strategy $\boldsymbol{\Pi}_6$ upto and inclusive the solution of the task i_1, i. e. the strategies $\boldsymbol{\Pi}(4, 1, \ldots)$.

For the further steps of the basic strategy, i.e. to the tasks i_6 and i_3, the substitutional strategies are not more supplied, because these tasks were already preceded by an erroneous solution. That is why the so called autoregulation node appears here already instead of the substitutional strategy. It is analogous also in the substitutional strategies for the first three steps of the basic strategy.

The situation is a little changed for that substitutional strategy, in which the possibility of an erroneous solution is still admitted. In general it is namely necessary to create substitutional strategies still for this substitutional strategy, too. Because in this concrete example to the fourth task i_5 of the basic strategy Π_6 a strategy from the set M_6^5 is chosen, i.e. the strategy Π having already the first four steps firmly determined, it is not more necessary to seek for further substitutional strategies. The reason is that these strategies might contain an only step that cannot be more chosen.

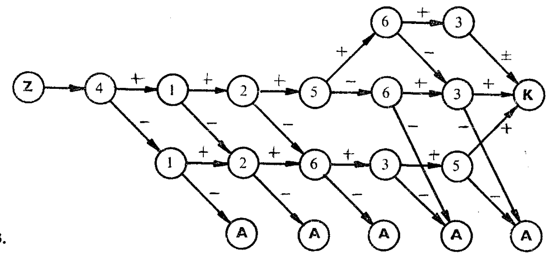


Fig. 3.

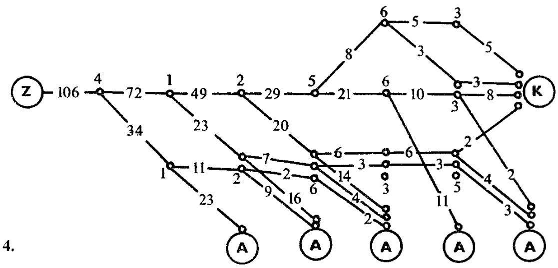


Fig. 4.

Thus it is already possible to create a multigraph plotted in Fig. 3. This multigraph gives the resulting unreduced non-Markov algorithm. The resulting algorithm is a non-Markov one, because to the same reaction in the same point does not correspond the same continuation. For instance in the basic strategy for the correct solution of the task i_2 the continuation is in the task i_5, on the other hand in the substitutional strategy with the same reaction on the same task is ranged the continuation in the task i_6. In the further stage it is possible to start with the reduction of this algorithm. To this reduction I have determined those frequencies corresponding to the trajectories used in the multigraph in Fig. 3. These frequencies are then excerpted in Fig. 4 so, as the individual sections of all the strategies were used in testing. For instance in the basic strategy Π_6 it is seen from this figure that 72 of the 106 tested probants have correctly solved the task i_4, then 49 of them have correctly solved the task i_1 etc.

From Fig. 4 it is then evident that some trajectories may be reduced, as some of their sections have the zero frequency. To be better seen which trajectories are here concerned, I have plotted them in Fig. 5. All these trajectories may be then reduced on the basis of following considerations:

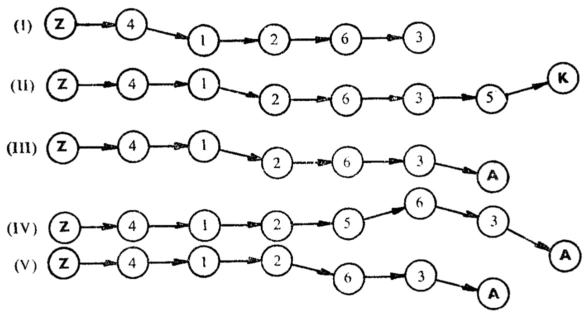


Fig. 5.

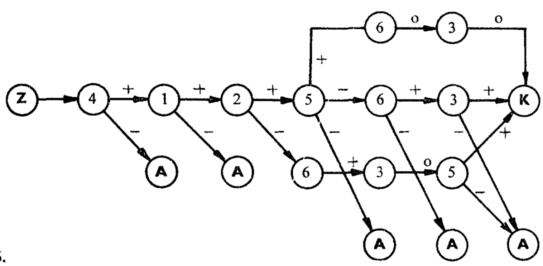


Fig. 6.

1. All students, whose reaction on the task i_4 is incorrect, will sooner or later result in so called autoregulation nodes. It means that they have not as sufficient knowledge as to pass the whole sequence successfully. That is why it is possible to rectify them immediately into this autoregulation node, as it is evident in Fig. 6.

2. In an analogous situation are those students, whose reaction on the task i_4 was correct, but the further reaction, i.e. that on the task i_1, is incorrect.

3. A little different situation occurs at those, whose reaction on the tasks i_4 and i_1, was correct, but their reaction on the task i_2 is erroneous. Should namely their further reaction, i.e. that on the task i_6, be correct, then it may be expected that their following reaction in the task i_3 will be without error, too.

4. The preceding reasons may be further employed so, that I shall omit the so called autoregulation node that is ranged to the task i_3.

5. Those students who have arrived up to the task i_5 through the trajectoty of the basic strategy and whose reaction is here correct, will pass the whole sequence with one error at most. Likewise the substitutional strategy to i_2 in the task i_3, also in further tasks it is not necessary to take their reactions into consideration. Namely the decision functions (4) do not take these reactions into consideration. Thus those tasks will arise in the multigraph in Fig. 6, from which one trajectory is emanating only.

Now it is already possible to set up a reduced non-Markov algorithm that serves as a basis for the teaching program. This resulting algorithm is graphically described as a multigraph in Fig. 6. It is still necessary to revise the programme that may be set up on the basis of this algorithm.

CONCLUSIONS

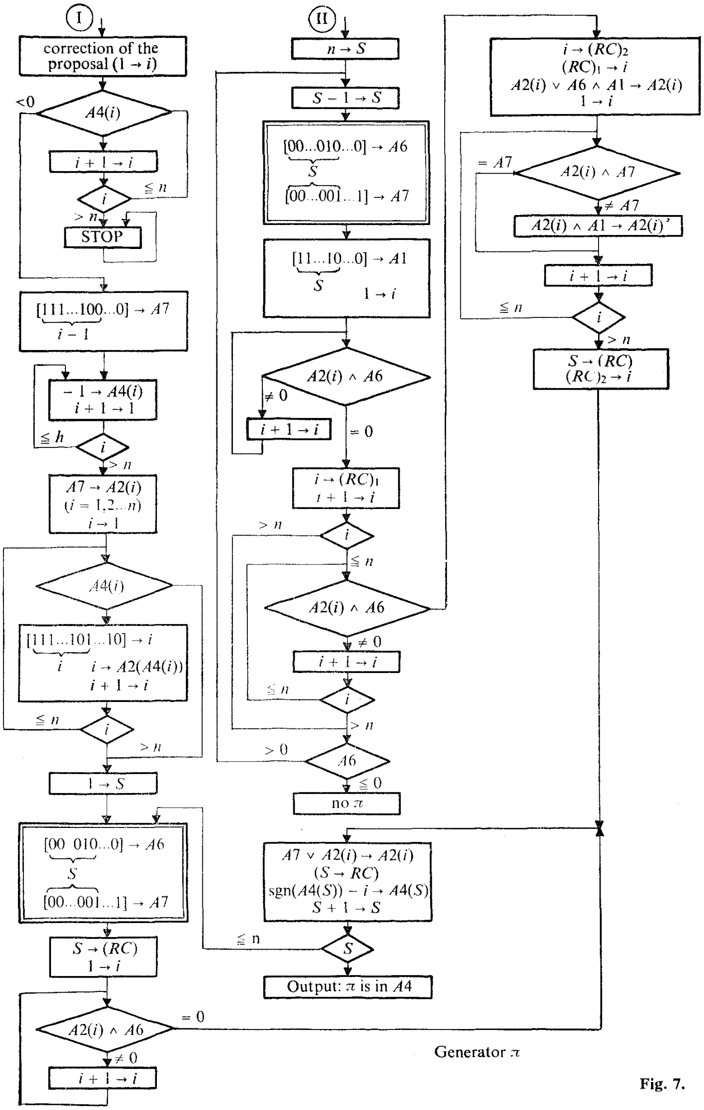
The sequence tested is a component of the prepared program on the mathematical textbook which will be determined for students of technical faculties. In the meantime we are preparing its first part containing the following chapters:

1. Introduction in the mathematical logics and theory of sets.
2. Introduction in the algebra.
3. Analytical geometry in the plane.
4. Differential calculation of the function of one variable.

On the basis of the structural analysis of this part we have prepared tests for the first year of the Faculty of Mechanical Engineering at the Technical University, Prague, which will run this year. As all the tasks in these tests are to be solved in one lot, the tested students have to margin the turn in which they have solved these tasks, accordingly. This succession will then help us in investigating the adequacies of the resulting algorithm.

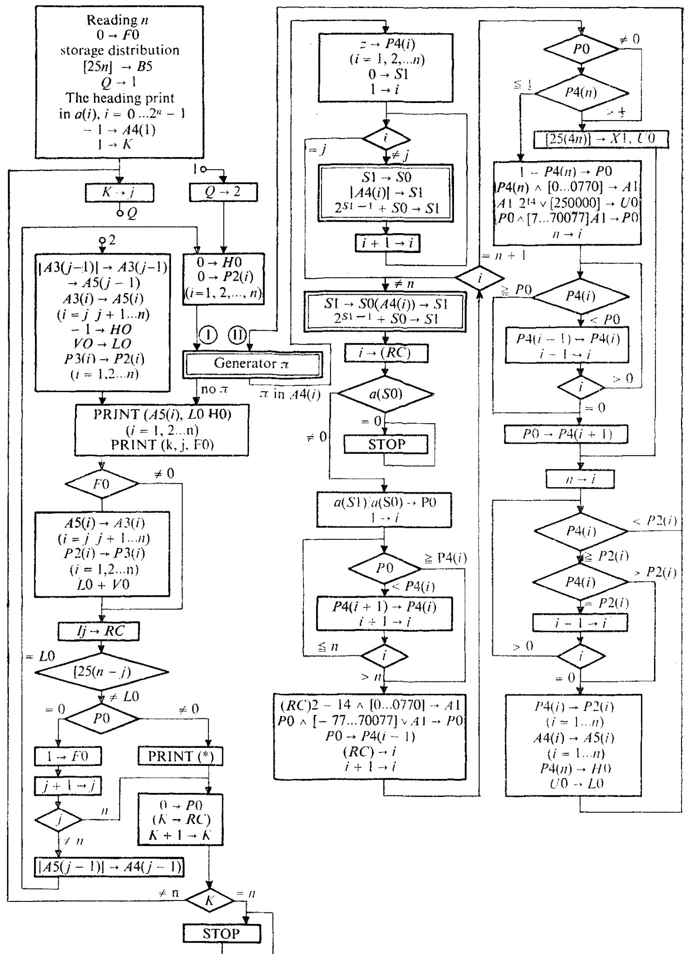
The program, eventually its block diagram, on the basis of which the computer counts the optimum algorithm, is quoted in the figures 7 and 8. This program is elaborated for sequences of 10 tasks maximally. But there is, of course, no problem to overwork it for sequences of more than ten tasks. However, here should be already used the outer computer store. When setting up the program I have, however, based on those analyses, in which the sequences of more tasks did not appear. The problem also does not consist in determining the kind of boundaries limiting the error parameter values.

Practically it has turned out that the set-up algorithm is a very natural one. After the testing of the other sequences is passed at our faculty, then it will be possible to draw more conclusions. To verify the hypothesis that I have expressed in the first



Generator π

Fig. 7.



Reading n
0 → F0
storage distribution
[25n] → B5
Q → 1
The heading print
in a(i), i = 0 ...2ⁿ − 1
− 1 → A4(1)
1 → K
K → j
Q
Q → 2
0 → H0
0 → P2(i)
(i=1, 2,..., n)
Generator π
no π
π in A4(i)
PRINT (A5(i), L0 H0)
(i = 1, 2...n)
PRINT (k, j, F0)
F0
A5(i) → A3(i)
(i = j j + 1...n)
P2(i) → P3(i)
(i = 1,2...n)
L0 + V0
Ij → RC
[25(n − j)
P0
1 → F0
j + 1 → j
PRINT (*)
0 → P0
(K → RC)
K + 1 → K
|A5(j − 1)| → A4(j − 1)
K
STOP
z → P4(i)
(i = 1, 2,...n)
0 → S1
1 → i
S1 → S0
|A4(i)| → S1
2^(S1 − 1) + S0 → S1
i + 1 → i
S1 → S0(A4(i)) → S1
2^(S1 − 1) + S0 → S1
i → (RC)
a(S0)
STOP
a(S1) a(S0) → P0
1 → i
P4(i + 1) → P4(i)
i + 1 → i
(RC)2 − 14 ∧ [0...0770] → A1
P0 ∧ [− 77...70077] ∨ A1 → P0
P0 → P4(i − 1)
(RC) → i
i + 1 → i
P4(n)
[25(4n)] → X1, U0
1 − P4(n) → P0
P4(n) ∧ [0...0770] → A1
A1 2¹⁴ ∨ [250000] → U0
P0 ∧ [7...70077]A1 → P0
n → i
P4(i)
P4(i − 1) → P4(i)
i − 1 → i
P0 → P4(i + 1)
n → i
i − 1 → i
P4(i) → P2(i)
(i = 1...n)
A4(i) → A5(i)
(i = 1...n)
P4(n) → H0
U0 → L0

Fig. 8.

chapter, it is, of course, necessary to perform a greater number of tests; nevertheless our testing may already much suggest.

I would like yet to emphasize that a whole range of other important parameters acts in the teaching process and by employing them we may arrive at a more perfect control of this process.

The way I have tried to indicate in this model is a way of a successive generalization which, due to the complexity of the whole problem, is one of the tolerable ways.

As the teaching process is mostly a finite one, the computing techniques that will considerably speed the work up, may thus be used.

(Received October 25th, 1967.)

REFERENCES

[1] L. N. Landa: Vyučování žáků metodám racionálního myšlení a problém algoritmů. (Instruction of students in the methods of rational thought and the problem of algorithms.) Na pomoc pedagogické praxi *9* (1963).

[2] Van der Waerden: Mathematische Statistik. Springer-Verlag, Berlin—Göttingen—Heidelberg 1957.

[3] J. Linhart: Psychologické problémy teorie učení. (Psychological problems of the teaching theory.) NČSAV, Praha 1965.

[4] Kemeny, Snell, Thompson: Introduction to finite mathematics. Prentice-Hall, Englewood Cliffs, 1957.

[5] R. Bellman: Dynamic programming. Princeton University Press, Princeton N. Y. 1957.

[6] V. Kulič: Psychologická problematika chyby v programovaném učení. Materiály z 1. celostátní konference o programovaném učení. P. Ú. ČSAV, Praha 1964. (The psychological problems of the error in the programmed teaching. Proceedings of the first Czechoslovak conference on the programmed teaching, issued by the Psychological Institute of the Czechoslovak Academy of Sciences, Prague, 1964.)

[7] S. L. Pressey, B. F. Skinner, N. A. Crowder: Programované učení jako světový problém. (The programmed teaching as a world problem.) SPN, Praha 1966.

[8] D. Blackwell, M. A. Girshick: Theory of games and statistical decisions. John Wiley & Sons Inc., New York 1954.

Frekvence chyby — regulátor vyučovacího algoritmu

JOZEF BRODY

Pokud chceme studovat vyučovací proces **P** z hlediska optimální regulace, potom je třeba určit vazby, v rámci kterých je třeba tento proces zefektivnit. Z tohoto důvodu je proto třeba určit tyto vazby, které lze reprezentovat r-ticí nerovností

$$
\begin{aligned}
&g_1(x_1, \ldots, x_k) \geqq 0 \,, \\
(1) \qquad &\ldots\ldots\ldots\ldots\ldots \\
&g_r(x_1, \ldots, x_k) \geqq 0 \,.
\end{aligned}
$$

Proměnné $x_1, \ldots, x_k$ jsou hodnoty parametrů $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$, které determinují vyučovací proces **P**. Aby bylo možné vyhodnotit efektivitu procesu **P**, je třeba určit kriteriální funkci $\boldsymbol{F}$, která je potom mírou efektivity procesu **P**. Optimální regulace procesu **P** spočívá potom v tom, že se určí uspořádaná k-tice hodnot $x_1, \ldots, x_k$, která splňuje nerovnosti (1) a jejíž efektivita

$$(2) \qquad \boldsymbol{F}(x_1, \ldots, x_k)$$

je maximální.

Konstrukce vazeb (1) a kriteriální funkce $\boldsymbol{F}$ je provedena na základě následující hypotézy:

Hypotéza. *Frekvenci chyby je nutné udržovat v určitých mezích, jinak se motivace učení podstatně sníží.*

Na základě této hypotézy je studován proces **P** jako proces, na který působí dominantně jenom jeden parametr frekvence chyby $\boldsymbol{X}$, jehož hodnota x je určena počtem chyb v daném úseku. Jako příklad jsem potom vzal vazby

$$(3) \qquad x \geqq 0 \,, \quad \& \quad -x \geqq -1 \,,$$

tj. na daném úseku nesmí být víc než jedna chyba. Program, který se skládal z n kroků, reprezentovaných n úkoly

$$(4) \qquad i_1, \ldots, i_n \,,$$

byl potom vyhodnocen na základě reakcí

$$(5) \qquad r_1, \ldots, r_n \,.$$

Závislosti těchto reakcí r_j na úkoly i_j popisují relativní četnosti

$$a_0, \ldots, a_{2n-1}\,. \tag{6}$$

Tyto četnosti umožňují vyhodnotit každý průchod úkoly (4). Každý z průchodů potom určuje strategii $\boldsymbol{\Pi}$. Kriteriální funkce $\boldsymbol{F}$, která je sestavena z relativních četností (6), přiřazuje každé strategii $\boldsymbol{\Pi}$ efektivitu $\boldsymbol{F}(\boldsymbol{\Pi})$. Ze všech možných strategií, kterých je celkem

$$N = \sum_{j=0}^{n} j! = 1 + 1 + 2 + 6 + \ldots + n!\,, \tag{7}$$

vybral samočinný počítač optimální. Počítač pracoval na základě programu, jehož blokové schéma je popsáno na obr. 7 a 8. Na základě tohoto projektu pracuje na FSI při ČVUT v Praze skupina pracovníků na katedrách matematiky, kteří sestavují učebnici matematiky pro 1. semestr denního studia.

Jozef Brody, prom. matematik, Katedra matematiky strojní fakulty ČVUT, Karlovo nám. 13, Praha 2.