KYBERNETIKA — VOLUME 8 (1972), NUMBER 2

A Method for Random Sampling of Well-Formed Formulas

(A Method for Random Sampling of Formulas of an Elementary Theory and Statistical Estimation of their Deducibility Equipped by a Program I)

IVAN KRAMOSIL

This paper deals with a method which enables to realize a random sampling of well-formed formulas of an appropriate formalized theory by using random numbers. The method is based on a procedure enabling to transform every sequence of random numbers into a well-formed formula.

This paper represents a rather shortened version of the first part of a report the title of which can be found in the parentheses following the title of this paper. It deals with a method which enables to realize a random sampling of well-formed formulas of an appropriate formalized theory by using random numbers. Some possibilities of application of this method are mentioned in the closing section.

First an appropriate formulation for the first-order functional calculus is described in this paper. This formulation arises when Gentzen's sequences calculus, Church's P_1-calculus and Ohama's calculus are combined together (as to Ohama's calculus see [1]). Then a procedure is constructed, transforming any sequence of indices (including random ones) into a well-formed formula of a formerly given theory based on the above formulated first-order calculus. In the theorems 1 – 3 some principal properties of this procedure are described. Questions which arise in connection with methods for a further investigation of the formulas we have obtained are not investigated in this paper and will be left for the following one. The suggested program, describing the above mentioned procedure in FORTRAN and based on the method of so-called pseudorandom numbers, which is enclosed to the report mentioned above is not enclosed to this paper because of its extent.

1. FORMULATION OF THE LOGICAL SYSTEM

Let us start with describing the logical system used in the following. It contains the following *elementary symbols*:

a) the *indeterminates* $x_1, x_2, x_3, \ldots$ (the expression *indeterminate* will be used

unless to come to misunderstanding when the expression *variable* is used below in the connection "random variable");

b) the *logical constants* $c_1, c_2, \ldots, c_{KS}$, with the indices $AA(1), \ldots, AA(KS)$ ascribed to them;

c) symbols [ ] (brackets) $\rightarrow F$.

Well-formed formulas (abbreviation w.f.f. will be used:)

a) F is w.f.f.;

b) if $(j, i_1, i_2, \ldots, i_{AA(j)})$ is a sequence of indices with $j \leqq KS$, then $c_j x_{i1} x_{i2} \ldots \ldots x_{iAA(j)}$ is a so called *elementary w.f.f.*;

c) for all indices n, j and w.f.f.'s $A_1, A_2, \ldots, A_n$ the sequence $[A_1][A_2]\ldots[A_n]$ is a w.f.f., $[x_j[A_1]\ldots[A_n]]$ is also a w.f.f.

There is the following connection between the formulation just described and the usual formulations. The symbol F represents the propositional logical constant the semantic countrepart of which is an identically false sentence (see P_1-calculus in [4]). W.f.f. $[A_1][A_2]\ldots[A_n]\ldots$ corresponds to the implication $(A_1) \supset ((A_2) \supset \ldots \ldots \supset (A_n)\ldots)$ (or $(A_1) \rightarrow ((A_2) \rightarrow \ldots (A_n)\ldots))$) in the usual notation, in correspondence with the notation introduced in [4], when the pairs of parentheses and brackets, the right of them lies at the very end of a formula, are omitted). When the brackets are to be ordered in another way, it is necessary to express them in an explicit way; so the implication $(A \supset B) \supset C$ must be written in the form $[[A][B]][C]$. The external pairs of brackets, (i.e. the pairs of brackets introduced by d)) do not change the sense of the formula. It follows that the formula $[A][F]$ has the meaning of *negation* of the formula A, i.e. $[A][F]$ corresponds to non (A), $\neg(A)$ or $\sim(A)$ in usual notations.

The formula $[x_1[A_1]\ldots[A_n]]$ corresponds to the formula $\forall x_j((A_1) \supset ((A_2) \supset \ldots \ldots \supset (A_n)\ldots))$ or $\Pi_{x_j}\{(A_1) \rightarrow ((A_2) \rightarrow \ldots \rightarrow (A_n)\ldots))\}$; so the scope of any quantifier (which is formed by a left bracket and an indeterminate) is determined by the latter left bracket and by the right bracket, corresponding to the latter left bracket.

It is well known that implication, negation and general quantifier are sufficient for the definition of existential quantifier and other propositional functors. For example, the w.f.f. $\exists_{x_i}(A)$ is expressed by $[x_1[A][F]][F]$ in the present formalization.

Let $m \geqq 0$, $n \geqq 0$ be indices, let $A_1, A_2, \ldots, A_m, B_1, \ldots, B_n$ be w.f.f.'s of our system. Then the sequence:

$$A_1, A_2, \ldots, A_m \rightarrow B_1, B_2, \ldots, B_n \tag{1}$$

will be called *logical sequence* (abbreviation l.s. will be often used). This notion was first introduced by G. Gentzen in [3].

Roughly speaking, the countrepart of the l.s. (1) is the formula $(A_1 \wedge A_2 \wedge \ldots \ldots \wedge A_m) \supset (B_1 \vee B_2 \vee \ldots \vee B_n)$ when $n > 0$; the formula non $(A_1 \wedge A_2 \wedge \ldots \ldots \wedge A_m)$ when $n = 0$, $m > 0$; the formula F when $m = n = 0$.

Let $\mathsf{AX}_1, \mathsf{AX}_2, \ldots, \mathsf{AX}_{\mathsf{IS}}$ be some w.f.f.'s without free variables, these w.f.f's will be called *axioms*. The following l.s. will be called *axiomatic* l.s.:

(I) the l.s. of the type (1) with the property that, for some indices $i \leqq m$, $j \leqq n$, $A_i = B_j$ (perhaps with external pairs of brackets);

(IIa) the l.s. of the type (1) with the property, that for some $i \leqq m$, $A_i = \mathsf{F}$ or

(IIb) for some $j \leqq n$, $k \leqq \mathsf{IS}$, $B_j = \mathsf{AX}_k$ (perhaps with an external pair of brackets).

The set T of theorems will be defined as follows:

(I) If $\mathscr{A}$ is an axiomatic l.s., then $\mathscr{A} \in T$.

(II) If $\mathscr{A}$ is an l.s. of the type (1), $\mathscr{A} \in T$, $j \leqq n$, $i \leqq m$, then $A_1, A_2, \ldots, A_{i-1}, A_{i+1}, \ldots, A_m \to B_1, \ldots, B_{j-1}, [A_i][B_j], B_{j+1}, \ldots, B_n \in T$.

(III) If $\mathscr{A}$ is an l.s. of type (1), $\mathscr{A} \in T$, and $A_1, A_2, \ldots, A_{i-1}, A_{i+1}, \ldots, A_m \to \to B_1, B_2, \ldots, B_n, B_{n+1} \in T$, then $A_1, A_2, \ldots, A_{i-1}, [B_{n+1}][A_i], A_{i+1}, \ldots, A_m \to \to B_1, B_2, \ldots, B_n \in T$.

(IV) If $\mathscr{A}$ is an l.s. of the type (1), if i, j ($j \leqq n$) are indices satisfying the property that x_i is not free in any of the w.f.f.'s $A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_{j-1}, B_{j+1}, \ldots, B_n$, if $\mathscr{A} \in T$, then

$$A_1, A_2, \ldots, A_m \to B_1, B_2, \ldots, B_{j-1}[x_i[B_j]], \quad B_{j+1}, \ldots, B_n \in T.$$

(V) If $\mathscr{A}$ is an l.s. of type (1), if $\mathscr{A} \in T$, if $\mathscr{A}'$ arises from $\mathscr{A}$ by such a substitution for an indeterminate that the new indeterminate will not occur bound at any place at which the former indeterminate occured free, then $\mathscr{A}' \in T$. The formal system just described is equivalent to the elementary formalized theory with the same axioms and logical constants (as follows from the equivalence of the classical first-order functional calculus and the Gentzen's sequential calculus). Let us describe in our formal system the elementary theory of equality: There is only one binary logical constant I, (so $\mathsf{KS} = 1$, $\mathsf{AA}(1) = 2$), there are three axioms, namely

AX_1: $[x_1[Ix_1x_1]]$ $(\forall_{x_1}(x_1 = x_1))$,

AX_2: $[x_1[x_2[Ix_1x_2][Ix_2x_1]]]$ $(\forall_{x_1}\forall_{x_2}((x_1 = x_2) \supset (x_2 = x_1)))$

AX_3: $[x_1[x_2[x_3[Ix_1x_2][Ix_2x_3][Ix_1x_3]]]]$ $(\forall_{x_1}\forall_{x_2}\forall_{x_3}((x_1 = x_2) \supset ((x_2 = x_3) \supset (x_1 = x_3))))$

(so $\mathsf{IS} = 3$; in parentheses the usual notation is given).

2. RANDOM SAMPLE OF A WELL-FORMED FORMULA

The following procedure of random sample consists of two steps; first a sequence of random numbers is chosen, then this sequence is transformed into a w.f.f.

Let M_0, K_0, N_0 be positive indices. The number of such elementary w.f.f.'s of our theory which do not contain the indeterminates $x_{N_0+1}, x_{N_0+2}, \ldots$, will be denoted by N_1, so $N_1 = \sum_{j=1}^{\mathsf{KS}} N_0^{\mathsf{AA}(j)}$. By L will be denoted the integer satisfying the condition

$10^{L-1} \leqq 2K_0 + M_0 + N_0 + N_1 + KS + 3 < 10^L$, by M_1 will be denoted the maximal index less then 10^L which is divisible by $2K_0 + M_0 + N_0 + N_1$ (this last number, i.e. $2K_0 + M_0 + N_0 + N_1$ will be denoted by N_2).

Let us suppose that there exists a mechanism which enables us to perform a random sample of mutually independent random numbers from the set of all indices less than 10^L under the condition that each of these numbers may be chosen with the same probability $1/10^L$ (zero is not excluded). The so called method of pseudo-random numbers offers a number of various algorithms which enable to simulate such a mechanism, but this matter will not be investigated here. We suppose, for our purposes, that a random variable α is defined on some probability space $(\Omega, \mathscr{A}, P)$, taking its values in the set of all indices less than 10^L and satisfying the condition:

$$P(\{\omega : \omega \in \Omega, \alpha(\omega) = j\}) = 1/10^L \text{ for every index } j, \quad 0 \leqq j < 10^L .$$

Let $\alpha_1'', \alpha_2'', \ldots$ be mutually independent random variables defined on the probability space $(\Omega, \mathscr{A}, P)$, taking their values in the same set of indices as α does, with the same distribution of probability as α. Let Z_1 be an index $(Z_1 > 0)$. By the random sequence $\{\alpha_1''(\omega), \alpha_2''(\omega), \ldots\}$ a random vector $(\alpha_1(\omega), \alpha_2(\omega), \ldots, \alpha_{Z_1}(\omega))$ will be defined in such a way that $\alpha_j(\omega)$, $j \leqq Z_1$ is the j-th component of the sequence $\{\alpha_1''(\omega), \alpha_2''(\omega), \ldots\}$ which is less than or equal to $M_1 - 1$ so that $\alpha_j(\omega) < M_1$, $j = 1, 2, \ldots, Z_1$.

The vector $(\alpha_1, \alpha_2, \ldots, \alpha_{Z_1})$ (we shall omit the argument ω when is not a danger of misunderstanding) will be transformed into a vector $(\alpha_1', \alpha_2', \ldots, \alpha_{Z_1}')$ where α_j' is equal to $\alpha_j - [\alpha_j/N_2] N_2 + 1$ ($[a]$ denotes the greatest integer $\leqq a$. As M_1 is divisible by N_2, then $\alpha_1', \alpha_2', \ldots, \alpha_{Z_1}'$ are mutually independent random variables, defined on the probability space $(\Omega, \mathscr{A}, P)$, taking their values in the set of all positive indices less than or equal to N_2 with the property:

$$P(\{\omega : \omega \in \Omega, \alpha_j'(\omega) = k\}) = 1/N_2 \quad \text{for any indices } j, k ,$$
$$1 \leqq k \leqq N_2 , \quad 1 \leqq j \leqq Z_1 .$$

Let us number the symbols of our theory and its elementary w.f.f.'s, not containing the indeterminates $x_{N_0+1}, x_{N_0+2}, \ldots$ in the following way:

To the symbol $[$ the indices $1, 2, \ldots, K_0$ are ascribed, to the symbol $]$ the indices $K_0 + 1, \ldots, 2K_0$ are ascribed, to the symbol F the indices $2K_0 + 1, \ldots, 2K_0 + M_0$ are ascribed, to the indeterminate x_j $(1 \leqq j \leqq N_0)$ the index $2K_0 + M_0 + j$ is ascribed, to the elementary w.f.f.'s the indices $2K_0 + M_0 + N_0 + 1, \ldots, N_2$ are ascribed. The tetrad (K_0, M_0, N_0, N_1) is called the *code of the theory*.

We have numbered the symbols and elementary w.f.f.'s for two reasons. First we can code any w.f.f. in the form of numerical vector, on the other hand we can understand any sequence of positive indices $(\alpha_1', \alpha_2', \ldots, \alpha_{Z_1}')$ as if it were the sequence of symbols and elementary w.f.f.'s of our theory. It follows that *we will be able to choose at random a w.f.f. if we are able to transform any vector of symbols and elementary*

w.f.f.'s into a w.f.f. We now describe such a transformation, and we suppose that $(\alpha'_1, \ldots, \alpha'_{Z_1})$ consists of symbols and elementary w.f.f.'s, not of indices.

(I) If the symbol F or an elementary w.f.f. occurs at least once among $(\alpha'_1, \alpha'_2, \ldots, \alpha'_{Z_1})$, the step (II) will be immediately applied. If it is not the case, the elementary w.f.f. with the number $2K_0 + M_0 + N_0 + 1$ is put for α'_1, and the other symbols are erased.

(II) If α'_j, $1 \leqq j \leqq Z_1$, is an elementary w.f.f., or the symbol F, α'_j will be replaced by the symbols $[\alpha'_j]$, i.e. α'_j is put into brackets.

(III) If α'_j, $1 \leqq j \leqq Z_1$, is an indeterminate, α'_j will be replaced by the symbols $[\alpha'_j$, i.e. a left bracket is inscribed immediately before any indeterminate.

(IV) Put $S = 0$, pass through whole the vector from the left to the right, and if α'_j is the left bracket, put $S + 1$ instead of S, if α'_j is the right bracket, put $S - 1$ instead of S. If the value -1 is reached by S, inscribe the left bracket in the left end of the vector, put $S = 0$ and follow as described above. If you pass through whole the vector, inscribe such a number of the right bracket on the right end of the vector, which was the last value taken by S.

(V) Pass again through the vector. If you meet with the left bracket, start with the procedure described in (IV). The right bracket by which S takes at the first time the value 0, corresponds to the left bracket by which we started and these two brackets form a pair. If there is a pair of brackets not containing an elementary w.f.f., or the symbol F, all the symbols between the letter pair of brackets including the brackets themselves are erased.

(VI) If α'_j, $2 \leqq j \leqq Z_1$, is an indeterminate, not occuring in the scope of the general quantifier formed by α'_{j-1} and α'_j, α'_j will be erased.

(VII) If $j < k \leqq Z_1$ are indices with the property that α'_{j-j} and α'_j are the left brackets, α'_{k-1} and α'_k are the right brackets, and $(\alpha'_{j-1}, \alpha'_k)$ as well as $(\alpha'_j, \alpha'_{k-1})$ are pairs of brackets in consequence of (V), then one of this pair (namely $(\alpha'_{j-1}, \alpha'_k)$) will be erased.

By $F(\alpha)$ will be denoted the transformation which ascribes to any vector of elementary w.f.f.'s and symbols $[F]$ the vector obtained from α by applying all the steps (I)–(VII).

Theorem 1. *For every vector α of elementary w.f.f.'s, or of symbols $[F]$, $F(\alpha)$ is a w.f.f.*

Proof. The induction on the number Z_1 of elements of the vector α will be used. If $Z_1 = 1$, and if the step (I) is applied, then α is F or an elementary w.f.f., by the steps (II)–(VII) α is transformed into $[\alpha]$, which is a w.f.f.

Induction. Let the statement hold for all $Z_1 \leqq n$. Let $\alpha' = (\alpha_1, \alpha_2, \ldots, \alpha_{n+1})$ be a vector satisfying the conditions. Various possibilities will be analysed.

a) α_1 *is* F *or an elementary w.f.f.*

a1) Neither any elementary w.f.f. nor F occur among $\alpha_2, \alpha_3, \ldots, \alpha_{n+1}$. For $j = 2, 3, \ldots, n+1$ the two following cases are possible:

a1a) α_j is the left bracket, then its right bracket can be found or is inscribed by (IV) right from this left bracket; there is neither any w.f.f. nor F between this pair of brackets, and so this pair of bracket is erased by (V).

a1b) α_j is an indeterminate, then the left bracket is inscribed immediately before it by (III), and this left bracket including the indeterminate α_j is then erased by the same way as in the case a1a).

It follows that only right brackets among $\alpha_2, \alpha_3, \ldots, \alpha_{n+1}$ can remain after application of the steps (I)–(V). The left brackets corresponding to them were inscribed by (IV) at the left end of the vector α'. Such pairs of brackets are superfluous external brackets and will be erased by (VII). It follows that $F(\alpha) = [\alpha_1]$, which is a w.f.f.

a2) The symbol F or an elementary w.f.f. occur among $\alpha_2, \alpha_3, \ldots, \alpha_{n+1}$.

a2a) No left bracket is inscribed by (IV) at the beginning of α'. Then all the operations performed by (II)–(VII) in the vector $(\alpha_2, \alpha_3, \ldots, \alpha_{n+1})$ are the same as if the symbol α_1 did not exist. It follows that $F(\alpha') = F(\alpha_1, \ldots, \alpha_{n+1}) = [\alpha_1] \cdot F(\alpha_2, \ldots, \alpha_{n+1})$. The later sequence is a w.f.f. because $F(\alpha_2, \ldots, \alpha_{n+1})$ is a w.f.f. by inductive assumption (when $F(\alpha_2, \ldots, \alpha_{n+1})$ is not an implication, it has external brackets by a, or by the step (III).

a2b) There is at least one left bracket inscribed at the beginning of α' by (IV) (in such a case the number of such brackets will be the same for α' as well as for the vector $(\alpha_2, \ldots, \alpha_{n+1})$). By $[_1$ the first left bracket inscribed by (IV) will be denoted, by $]_1$ the corresponding right bracket will be denoted. The existence of α_1 doesn't change the value of the index S, so the pairs of brackets in $F(\alpha')$ and $F(\alpha_2, \ldots, \alpha_{n+1})$ will be in correspondence. So $F(\alpha_2, \ldots, \alpha_{n+1})$ is expressed in the form $[_2 [_1 A]_1 B]_2$, which is a w.f.f. by the inductive assumption (the brackets $[_2]_2$ as well as one of the parts $[_1 A]_1$ or B need not occur, first symbol in A cannot be any indeterminate, because $[_1$ was inscribed by (IV), not by (III)). Clearly $F(\alpha')$ can be expressed in the form $[_2 [_1 [\alpha_1] A]_1 B]_2$. A is a sub-formula of the w.f.f. $[_2 [_1 A]_1 B]$ which has its external brackets by (II) or (III), or it is an implication. So A as well as $[\alpha_1] A$ are w.f.f.'s, $F(\alpha')$ which can be obtained from w.f.f. $F(\alpha'_2, \ldots, \alpha'_{n+1})$ by replacing one subformula for another w.f.f. is in such a way, also w.f.f.

b) α_1 *is the left bracket*

b1) No left bracket would be inscribed by (IV) at the beginning of $(\alpha_2, \ldots, \alpha_{n+1})$ during its transformating into $F(\alpha_2, \ldots, \alpha_{n+1})$. Then no left bracket will be inscribed at the beginning of α' by (IV), and the corresponding right bracket will be found at the very end of α. So these brackets will form the external pair of brackets for $F(\alpha_2, \ldots, \alpha_{n+1})$ or $F(\alpha_2, \ldots, \alpha_n)$ (in case $\alpha_{n+1} =]$). If this pair of brackets is superfluous, it will be erased by (VII), and $F(\alpha') = F(\alpha_2, \ldots, \alpha_{n+1})$ or $F(\alpha') = F(\alpha_2, \ldots, \alpha_n)$, other-

wise $F(\alpha') = [F(\alpha_2, \ldots, \alpha_{n+1})]$ or $F(\alpha') = [F(\alpha_2, \ldots, \alpha_n)]$. By induction hypothesis it follows that $F(\alpha')$ is a w.f.f.

b2) At least one left bracket would be inscribed by (IV) during the transformation of $(\alpha_2, \ldots, \alpha_{n+1})$. Let α_j, $2 \leqq j \leqq Z_1$, denote the right bracket by which the value -1 was reached by S at the first time, and so the left bracket (let us denote it by $[_1$) was inscribed. When (IV) will be applied on α', the value of S will be greater by 1 than in the case of $(\alpha_2, \ldots, \alpha_{n+1})$ till α_j is reached, then the value $\emptyset$ will be ascribed to S, and no left bracket will be inscribed, the role of $[_1$ will be played by α_1. From this moment the vectors obtained by transformation of α' and $(\alpha_2, \ldots, \alpha_{n+1})$ will be identical as well as the value of S. ($S = 0$ in both the cases). So $F(\alpha') = F(\alpha_2, \ldots, \alpha_{n+1})$, which is a w.f.f. by the induction hypothesis.

c) α_1 *is an indeterminate*

Instead of α' the vector $\alpha'' = ([_1, \alpha_2, \alpha_3, \ldots, \alpha_{n+1})$ will be considered. By *b*) $F(\alpha'')$ is a w.f.f. But $F(\alpha')$ differs from $F(\alpha'')$ only by the fact that the well-formed part of $F(\alpha'')$, lying between $[_1$ and the corresponding right bracket $]_1$, is replaced by the indeterminate α_1 followed by the same w.f.f. ($[_1$ is inscribed by (III)). So $F(\alpha')$ is also a w.f.f. (see the step a2b)).

d) α_1 *is the right bracket*

In this case the left bracket (denoted by $[_1$) will be inscribed before α_1 at the very beginning of application of (IV). The pair $([_1, \alpha_1)$ of brackets will be erased by (V), and its contemporary existence will not change the way of transformation of $(\alpha_2, \ldots, \alpha_{n+1})$, nor the value of S, so $F(\alpha') = F(\alpha_2, \ldots, \alpha_{n+1})$ is a w.f.f. by the inductive step. Q.E.D.

When $\alpha_1, \alpha_2, \ldots, \alpha_{Z_1}$ are understood as random variables, then the compound transformation $F(\alpha_1, \alpha_2, \ldots, \alpha_{Z_1})$ can be taken for a random variable defined on the probability space $(\Omega, \mathscr{A}, P)$ which takes its values in the set of w.f.f.'s of the considered theory. In order to reach better lucidity and to facilitate further manipulations, w.f.f.'s can be coded in slightly modified way, namely:

1. The indices $N_2 + 1, N_2 + 2, \ldots, N_2 + KS$ are ascribed to the functional constants $c_1, c_2, \ldots, c_{KS}$ and w.f.f.'s will be coded not by one index, but by a finite sequence of indices corresponding to the constants and indeterminates.

2. The occurences of general quantifier not binding any indeterminate in their scopes are erased.

3. The symbol $\rightarrow$ is inscribed before any w.f.f. in order to obtain a l.s. (to $\rightarrow$ the index $N_2 + KS + 2$ is ascribed, to the comma $(\ ,\ )$ $N_2 + KS + 1$ is ascribed).

Example. Let us consider the elementary theory of equality described above. Let us suppose that $K_0 = 3$, $N_0 = 5$, $M_0 = 5$, $Z_1 = 14$. There exist 25 elementary w.f.f.'s; the indices 17, 18, ..., 41 are ascribed to them in such a way that index $17 + 5(i - 1) +$

$+ j - 1$ is ascribed to $Ix_i x_j$, $i, j = 1, 2, \ldots, 5$. It follows that $N_2 = 2K_0 + M_0 + N_0 + N_1 = 41$, $M_1 = 2 \cdot 41 = 82$, $L = 2$, to $\rightarrow$ the index 44 is ascribed, to I the index 42 is ascribed.

Let us suppose that the following values have been taken by $\alpha_1, \alpha_2, \alpha_3, \ldots$:

$$17-45-91-12-23-13-06-44-87-42-05-64-19-43-09-41-18\ldots$$

After leaving the values greater than $M_1 - 1$, we obtain:

$$17-45-12-23-13-06-44-42-05-64-19-43-09-41$$

Now all the values are divided by 41, and the remainders will be enlarged by 1:

$$\begin{array}{cccccccccccccc} 18 & - & 5 & - & 13 & - & 24 & - & 14 & - & 07 & - & 04 & - & 02 & - & 06 & - & 24 & - & 20 & - & 03 & - & 10 & - & 01 \\ 1 & & 2 & & 3 & & 4 & & 5 & & 6 & & 7 & & 8 & & 9 & & 10 & & 11 & & 12 & & 13 & & 14 \end{array}$$

which corresponds to the following w.f.f.'s and symbols:

$$\begin{array}{cccccccccccccc} Ix_1x_2 &] & x_2 & Ix_2x_3 & x_3 & F &] & [&] & Ix_2x_3 & Ix_1x_4 & [& F & [\\ 1 & & 2 & 3 & 4 & & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \end{array}$$

(the small indices below the symbols and w.f.f.'s are not a part of the considered theory and serve only to better orientation in the sequence).

Clearly, the step (I) is left. By (II) we obtain:

$$\begin{array}{cccccccccccccccccccccccc} [& Ix_1x_2 &] &] & x_2 & [& Ix_2x_3 &] & x_3 & [& F &] &] &] &] & [& Ix_2x_3 &] & [& Ix_1x_4 &] & [& [& F &] & [\\ 15 & 1 & 16 & 2 & 3 & 17 & 4 & & 16 & 5 & 19 & 6 & 20 & 7 & 8 & 9 & 21 & 10 & 22 & 23 & 11 & 24 & 12 & 25 & 13 & 26 & 14 \end{array}$$

(the pairs 15–16, 17–18, ..., 25–26 of brackets were inscribed). Now the step (III) will be applied and the left brackets 27, 28 inscribed:

$$\begin{array}{cccccccccccccccccccccccc} [& Ix_1x_2 &] &] & [& x_2 & [& Ix_2x_3 &] & [& x_3 & [& F &] &] & [&] & [& Ix_2x_3 &] & [& Ix_1x_4 &] & [& [& F &] &] \\ 15 & 1 & 16 & 2 & 27 & 3 & 17 & 4 & & 16 & 28 & 5 & 19 & 6 & 20 & 7 & 8 & 9 & 21 & 10 & 22 & 23 & 11 & 24 & 12 & 25 & 13 & 26 & 14 \end{array}$$

Now we apply (IV) and the part of (V) by which the pairs of corresponding brackets are established (below any bracket the index of the corresponding bracket is inscribed). During (IV) the brackets 29, 30, 31, 32 are inscribed. We obtain:

$$\begin{array}{cccccccccccccccccccccccccccccc} [& [& Ix_1x_2 &] &] & [& x_2 & [& Ix_2x_3 &] & [& x_3 & [& F &] &] & [&] & [& Ix_2x_3 &] & [& Ix_1x_4 &] & [& [& [& F &] & [&] &] &] \\ 29 & 15 & 1 & 16 & 2 & 27 & 3 & 17 & 4 & & 16 & 23 & 5 & 19 & 6 & 20 & 7 & 8 & 9 & 21 & 10 & 22 & 23 & [11 & 24 & 12 & 25 & 13 & 25 & 14 & 30 & 31 & 32 \\ 2 & 16 & & 15 & 29 & 32 & & 16 & & & 17 & 7 & & 20 & & 19 & 28 & 9 & 8 & 22 & & 21 & 24 & & 23 & 31 & 26 & & 26 & 30 & 14 & 12 & 27 & \end{array}$$

Now the step (V) will be finished (and the pairs 8–9, 14–30 of brackets erased). Then, by (VI), the occurence of the indeterminate x_3 with index 5 will be erased and, by (VII), the pairs 29–2, 28–7, 12–31 of superfluous brackets will be erased. The result

$$\begin{array}{cccccccccccccccccccccccc} [& Ix_1x_2 &] & [& x_2 & [& Ix_2x_3 &] & [& F &] & [& Ix_2x_3 &] &] & Ix_1x_4 &] & [& F &] &] \\ 15 & 1 & 16 & 27 & 3 & 17 & 4 & 16 & 19 & 6 & 20 & 21 & 10 & 22 & 23 & 11 & 24 & 25 & 13 & 26 & 32 \end{array}$$

is a w.f.f. of the considered theory. Rather as an illustration, the code of this w.f.f. will be given, after the modifications mentioned above (after the proof of Theorem 1):

$$44-1-42-12-13-4-1-12-1-42-13-14-4-1-7-4-1-42-13-$$
$$-14-4-1-42-12-15-4-1-10-4-4-$$

3. TWO PROPERTIES OF THE RANDOM SAMPLING PROCEDURE

Theorem 2. *Let us consider a formalized theory with the code* (K_0, M_0, N_0, N_1), *let* A *be a w.f.f. of this theory not containing the indeterminates* $x_{N_0+1}, x_{N_0+2}, \ldots$ *Then there exists an index* $Z(A)$ *with the property that for every* $Z_1 \geqq Z(A)$

$$P(\{\omega : \omega \in \Omega,\ F(\alpha_1(\omega), \alpha_2(\omega), \ldots, \alpha_{Z_1}(\omega)) = A\}) > 0 .$$

Proof. Put $A'' = A$ if there is a pair of external brackets in A, put $A'' = [A]$ otherwise. By $\mathscr{A}'$ denote the sequence obtained from A'' by leaving all the pairs of brackets, containing only one occurence of w.f.f.'s, (i.e. the brackets immediately before and after an elementary w.f.f. or F are ommited). By $Z(A)$ the number of elements in $\mathscr{A}'$ will be denoted.

Let $Z_1 \geqq Z(A)$, let $\bar{A}$ denote the set of all sequences obtained from $\mathscr{A}'$ by inscribing $Z_1 - Z(A)$ brackets of the same form in one of the ends of $\mathscr{A}'$ (i.e. $\bar{A}$ contains one element, namely $\mathscr{A}'$, when $Z_1 = Z(A)$, four elements, if $Z_1 > Z(A)$). If the right brackets were inscribed at the beginning or the left brackets were inscribed at the end, then these brackets, as well as the brackets, forming the pairs with them would be erased during application of the steps (II)–(VII), because they don't contain any w.f.f.; if, the left brackets at the beginning, or the right brackets at the end were inscribed, they would be erased by (VII) as superfluous.

Then, for every $x \in \bar{A}$, $F(x) = A''$ which has the same meaning as A. By the conditions for $\alpha_1', \alpha_2', \ldots$, every vector of indices $\beta_1, \beta_2, \ldots, \beta_{Z_1}$, $1 \leqq \beta_j \leqq M_1$, will be chosen with a positive probability. It follows:

$$P(\{\omega : \omega \in \Omega,\ F(\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) = A\}) \geqq P(\{\omega : (\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in \bar{A}\}) =$$
$$= \sum_{x \in \bar{A}} P(\{\omega : (\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) = x\}) > 0 .$$

Q.E.D.

Theorem 3. *Let a formalized theory with a code* (K_0, M_0, N_0, N_1) *be considered. Let* T *be the set of all its theorems,* Z_1 *be an index,* $\alpha_1, \ldots, \alpha_{Z_1}$ *be random variables defined on a probability space* $(\Omega, \mathscr{A}, P)$ *and taking their values from the set* $\{1, 2, \ldots, N_2\}$ *of indices* $(N_2 = 2K_0 + M_0 + N_0 + N_1)$, *each with probability*

$1/N_2$. *Then the following assertion holds:*

$$P(\{\omega : F(\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in T\}) \geqq$$

$$\geqq \left(\frac{M_0+1}{N_2}\right)\left(\frac{N_1}{N_1-1}\right)\left(\frac{K_0+N_0+N_1-1}{M_0+K_0+1}\right)\left(1-\left(\frac{K_0+N_0+N_1-1}{N_2}\right)^{Z_1-1}\right) -$$

$$-\left(\frac{M_0+1}{N_2}\right)\left(\frac{N_1}{N_1-1}\right)\left(\frac{K_0+N_0}{M_0+K_0+N_1}\right)\left(1-\left(\frac{K_0+N_0}{N_2}\right)^{Z_1-1}\right) +$$

$$+\left(\frac{M_0^2}{N_1N_2}\right)\left(\frac{K_0+N_0+N_1}{M_0+K_0}\right)\left(1-\left(\frac{K_0+N_0+N_1}{N_2}\right)^{Z_1-1}\right) -$$

$$-\left(\frac{M_0^2}{N_1N_2}\right)\left(\frac{K_0+N_0}{K_0+M_0+N_1}\right)\left(1-\left(\frac{K_0+N_0}{N_2}\right)^{Z_1-1}\right)$$

which tends to

$$\left(\frac{N_1}{M_0+K_0+N_1}\right)\left(\frac{M_0+1}{M_0+K_0+1}\right)+\left(\frac{M_0}{M_0+K_0+N_1}\right)\left(\frac{M_0}{M_0+K_0}\right) \text{ if } Z_1 \to \infty .$$

Proof. By $T(n, m, k)$, $n < m \leqq Z_1$, $k \leqq N_1$, the set of all vectors $(\alpha_1, \alpha_2, \ldots, \alpha_{Z_1})$ satisfying the following conditions will be denoted:

a) α_m is the elementary w.f.f., to which the index $2K_0 + M_0 + N_0 + k$ is ascribed, it will be noted by E_k,

b) α_n is E_k or α_n is F,

c) there are no occurences of F, E_k or $\rceil$ among $\alpha_{n+1}, \ldots, \alpha_{m-1}$,

d) there are no occurences of any elementary w.f.f., F or $\rceil$ among $\alpha_{m+1}, \ldots, \alpha_{Z_1}$.

It follows, that if at least one of the conditions $n_1 \neq n_2$, $m_1 \neq m_2$, $k_1 \neq k_2$ is satisfied, the sets $T(n_1, m_1, k_1)$ and $T(m_2, n_2, k_2)$ are disjoint.

By $S(n, m)$, $n < m \leqq Z_1$, the set of all vectors $(\alpha_1, \ldots, \alpha_{Z_1})$, satisfying the following conditions will be denoted: $\alpha_m = F$, $\alpha_n = F$, there are no occurences F or $\rceil$ among $\alpha_{n+1}, \ldots, \alpha_{n-1}$, there are no occurences of F, $\rceil$ or any elementary w.f.f. among $\alpha_{m+1}, \ldots, \alpha_{Z_1}$. It follows that if at least one of the conditions $n_1 \neq n_2$, $m_1 \neq m_2$ is satisfied, the sets $S(n_1, m_1)$ and $S(n_2, m_2)$ are disjoint as well as the sets:

$$\bigcup_{n=1}^{Z_1-1}\bigcup_{m=n+1}^{Z_1}\bigcup_{k=1}^{N_1} T(n, m, k) \bigcup_{n=1}^{Z_1-1}\bigcup_{m=n+1}^{Z_1} S(n, m)$$

(their union will be denoted by R), because the last formula at the right is an elementary w.f.f. in the sequences, belonging to the first set, the F-symbol in the sequences belonging to the latter set.

Let for some indices m, n, k, $\alpha \in T(n, m, k)$. It follows that $\alpha_{m+1}, \ldots, \alpha_{Z_1}$ are indeterminates or left brackets. But all these symbols will be erased by (II)–(VII).

As it is no occurence of] among $\alpha_{n+1}, \ldots, \alpha_{m-1}$, $F(\alpha)$ could be written in the form $[B][F][A][E_k]$ (and $[B][E_k][A][E_k]$, respectively) or $F(\alpha)$ can differ from such a form only by an explicit presence of a pair of brackets the existence of which is implicitly assumed for the w.f.f. of such a form. But then $\to F(\alpha)$ is a theorem, because $\to F(\alpha)$ can be derived from a l.s. $[B], [F], [A] \to [E_k]$ (and $[B], [E_k], [A] \to$ $\to [E_k]$ respectively) by a triple use of the deductive rule (II).

Analogously, for $\alpha \in S(n, m)$, $\to F(\alpha)$ has the form $\to [B][F][A][F]$, and so it is a theorem.

Denote

$$a = P(\{\omega : \alpha_1(\omega) = F\}) = \frac{M_0}{N_2}, \quad b = P(\{\omega : \alpha_1(\omega) =]\}) =$$

$$= P(\{\omega : \alpha_1(\omega) = [\}) = \frac{K_0}{N_2},$$

$$e = P(\{\omega : \alpha_1(\omega) = E_k\}) = \frac{1}{N_2}, \quad c = N_1 \cdot e, \quad k = 1, 2, \ldots, N_1.$$

It follows:

$$P(\{\omega : \omega \in \Omega,\ F(\alpha(\omega)) \in T\}) \geqq P(\{\omega : (\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in R\}) =$$

$$= \sum_{n=1}^{Z_1-1} \sum_{m=n+1}^{Z_1} \sum_{k=1}^{N_1} P(\{\omega : (\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in T(m, n, k)\}) +$$

$$+ \sum_{n=1}^{Z_1-1} \sum_{m=n+1}^{Z_1} P(\{\omega : (\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in S(n, m)\}) =$$

$$= N_1 \sum_{n=1}^{Z_1-1} \sum_{m=n+1}^{Z_1} 1^{n-1} \left(a + \frac{c}{N_1}\right)\left(1 - a - b - \frac{c}{N_1}\right)^{m-n-1} \left(\frac{c}{N_1}\right) \cdot$$

$$\cdot (1 - a - b - c)^{Z_1-m} +$$

$$(*) \qquad + \sum_{n=1}^{Z_1-1} \sum_{m=n+1}^{Z_1} 1^{n-1} a(1 - a - b)^{m-n-1}\, a(1 - a - b - c)^{Z_1-m}$$

(the random variables $\alpha_1, \alpha_2, \ldots, \alpha_{Z_1}$ are independent). After performing summations we obtain:

$$P(\{\omega : F(\alpha_1(\omega), \ldots, \alpha_{Z_1}(\omega)) \in T\}) \geqq (*) \geqq$$

$$\geqq \left(a + \frac{c}{N_1}\right)\left(\frac{c}{c-e}\right)\left(\frac{1-a-b-e}{a+b+e}\right)(1 - (1 - a - b - e)^{Z_1-1}) -$$

$$- \left(a + \frac{c}{N_1}\right)\left(\frac{c}{c-e}\right)\left(\frac{1-a-b-c}{a+b+c}\right)(1 - (1 - a - b - c)^{Z_1-1}) +$$

$$+ \frac{a^2}{c}\left(\frac{1-a-b}{a+b}\right)(1 - (1 - a - b)^{Z_1-1}) - \frac{a^2}{c}\left(\frac{1-a-b-c}{a+b+c}\right) \cdot$$

$$\cdot (1 - (1 - a - b - c)^{Z_1-1}.$$

By substitutions for a, b, c, e and by an easy calculation we obtain the inequality contained in the statement of this theorem. As $(1 - (1 - a - b - c)^{Z_1 - 1})$, $(1 - (1 - a - b - e)^{Z_1 - 1})$ as well as $(1 - (1 - a - b)^{Z_1 - 1})$ tends to 1 when $Z_1 \to \infty$, the statement dealing with the limit follows immediately by substitutions for a, b, c, e, and by easy simplifications. Q.E.D.

As an illustration several values for the lower bound of the probability investigated in the last Theorem will be given. The same example as that given above Theorem 1 is considered.

Z_1	2	5	10	15	30	$+\infty$
value for (*)	0,09938	0,35693	0,52728	0,57252	0,59914	0,59975

4. POSSIBILITIES OF APPLICATION AND CONCLUSIVE REMARKS

Before closing this paper, some remarks should be done, in order to try to answer some questions, which possibly have arisen after having read the foregoing sections.

At first time let us consider the questions dealing with the originality of the algorithm, proposed in this paper. The basic inspiration originates from [1], namely the idea of transforming every sequence of symbols into a well-formed formula by inscribing and erasing of appropriate symbols in an appropriate way. The concrete procedure aiming to this goal and described in this paper is, however, the own author's result (maybe with the exception of the step A3 which was inspired by a similar step from [1]). Theorems 1, 2, 3, their proofs and the idea of joining such a procedure with a random number generator are also, as far as the author knows, original.

When compared with the algorithm from [1], the presented algorithm offers two advantages, namely:

1. The presented algorithm can be applied to the so called elementary theories (i.e. to the applied first-order functional calculus, if terminology from [2] is used). At the same time the algorithm from [1] was applicable only in the case of the so called "pure" first-order functional calculus which does not contain any logical constants, and which has to its disposal infinite number of infinite sequences of functional indeterminates.

2. No formula, containing only the indeterminates $x_1, \ldots, x_{N_0}$ is a priori avoided by this algorithm (c.f. Theorem 2). At the same time the procedure from [1] (supposing it was joined to a random number generator) would not be able to satisfy this requirement.

The importance of Theorem 1 is most probably beyond any doubts, because this theorem proves the proposed algorithm to be correct. However, some questions may

arise when Theorem 2 and Theorem 3 are considered, esspecially in case the role of this assertions is considered together with some practical application of the proposed algorithm. The fact deserving special attention is the following: the lower bound for the probability of choosing a theorem derived in Theorem 3 is, at the same time, a lower bound for the probability of choosing a theorem, which is, in some sense, "trivial", roughly speaking a theorem of the form "the false implies whatever you like" or "A implies A". Also the control of this lower bound, enabled by an appropriate choosing of values for the parameters M_0, N_0, K_0, represents, at the same time, the control of the lower bound of the probability that such a "trivial" theorem is chosen. The importance of this fact is then in a very close connection with the goal to which this algorithm is to serve. Several possibilities will be mentioned here.

1. The algorithm is considered to serve as a source of well-formed formulas which are submitted for further investigation. The goal is to discover a theorem which has not been known so far and which would be, in certain manner, interesting. In such a case the importance of Theorem 2 is rising, because this theorem guarantees any formula not being a priori avoided from the investigation (supposing it is not "too long" and does not contain "too many" indeterminates). Moreover, this limit number N_0 of indeterminates as well as the limit lenght (roughly Z_1) are free parameters of this algorithm. On the other hand, of course, the choice of a "trivial" theorem does not bring any new information, is useless, and hence not very desirable. In such a case it would be perhaps useful to find an upper bound for the probability of choosing a "trivial" theorem. Another solution would be the following one: to modify the algorithm in such a way that the "trivial" theorems were not considered to be the results of the algorithm; in case such a theorem were obtained it would be immediately erased and another formula would be chosen. E.g. let us consider a theory of equality and inequality, containing two binary logical constants I(equality), N(inequality), and four axioms (namely those of reflexivity, symmetry and transitivity, and the axiom by which inequality is defined in the usual way). Then a lower bound for the probability of sampling a theorem under the condition it is not a trivial one is equal to 0,6209 (in case $K_0 = M_0 = 10$, $N_0 = 5$, $Z_1 = 15$), or is equal to 0,6062 (in case $K_0 = 6$, $M_0 = 10$, $N_0 = 5$, $Z_1 = 15$).

2. The second possibility is to use the algorithm as a source (or generator) of "auxiliary" axioms sampled at random if a method of statistical estimation of deducibility of formulas in extensions sampled at random is used. In such a case an algorithm choosing only the theorems (and every theorem with a positive probability) would be very desirable. Supposing a theorem is chosen as an "auxiliary" axiom the possibility of proclaiming a non-theorem to be a theorem is avoided and, at the same time, this theorem-axiom can be useful, if it is not a trivial one, when another theorem is checked. It follows that our attitude toward the "trivial" theorems will depend on the fact which kind of error is less acceptable. If we want first to minimize the probability of proclaiming a non-theorem to be a theorem (this point of view is used in [4]) then choosing even a "trivial" theorem is desirable (or, to be more correct, more

desirable then choosing a non-theorem). Moreover, the lower bound derived in Theorem 3 enables us, in such a case, to derive a simple upper bound for the probability of proclaiming a non-theorem to be a theorem and to minimize this bound below an a priori given limit. Nevertheless, Theorem 2 will play also a very important role, because it guarantees that every theorem will be proclaimed, with a positive probability, to be a theorem. In fact, it was just the application mentioned in this paragraph and in the following one which inspired the author's attempts to derive the algorithm investigated in this paper.

3. There are, however, also other possibilities. Let us consider the situation well-known from the theory of stochastic processes and their control. We observe how a stochastic process develops (i.e. we observe its trajectory). At the same time we are involved in the process in such a way that in some time moment t_0 in the future we shall obtain a profit, smaller or higher, with respect to the stage of the process in the moment t_0 (or, more generally, with respect to the whole trajectory of the process until t_0). Let us also consider that we are given the possibility to intervene in the process at the time moments $s_1 < s_2 < \ldots < s_n < t_0$ and in such a way to influence the further development of the process. In the most simple case we are supposed to have two possible interventions, say $d_i^{(1)}$, $d_i^{(2)}$, in the moment s_i. The right intervention among them is that which guarantees the higher expected value of profit (or higher guaranteed value, if the minimax principle is applied) in the moment t_0. However, which of the two interventions is the right at the moment considered, may depend, generally, on the fact whether a number of conditions or relations hold at this moment. Therefore we can assume, in general, that the right decision making depends on the validity or non-validity of an assertion of a formalized theory. Hence, the investigated process can be understood as if it were a mechanism for random sampling of formulas which are then submitted for further decision making. At the same time, it can be easily seen that not only the "correctness" of a decision having been taken about a formula "submitted to investigation" by the process, but also the time necessary for decision making will play an important role (so that it may be possible to intervene into the process in time). A correct decision having been made too late can bring the same consequence as a wrong decision having been made in time. Sometimes the consequences may be even worse in the former case supposing we were occupied by the searching for a correct decision at the moment s_i in such a measure that we were not able to intervene at the moment s_{i+1} (or even in s_{i+2}, ..., respectively). It follows immediately that controlling such a stochastic process a statistical decision procedure will be of great value supposing this procedure guarantees that the probability for this decision to be correct is "great enough" (this depends, of course, on the actual purposes to which the procedure may serve).

Now, let us suppose we use the algorithm described in this paper as a simulator for the sequence of decision problems submitted step by step for decision making (for example, if we want to check preliminarily a decision procedure intended to be used during the control). At the same time trivial theorems can be understood as if they

were formal countreparts of such decision problems which can be easily and correctly decided. But in such a case it follows immediately that the great ratio of such "trivial" theorems is desirable. This holds particularly in the case of some possibility of a control of the probability of occuring a "trivial" decision problems during controlling the process. Therefore the importance of Theorem 3 may be now seen from quite another point of view. The author is going to investigate this matter in more details in one of the papers to follow.

4. The algorithm enables also to understand every finite sequence of symbols and elementary formulas as if it were a well-formed formula, namely the w.f.f. obtained from the considered sequence by our algorithm. This enables to build a logical calculus more desirable for the notation of and treatment with the logical problems using a computer. In such a case Theorem 2 is of principal value.

Before closing our discussion on Theorem 3 and its importance, we should like to pick out a rather useful property of the lower bound investigated there, namely the fact that this lower bound is common for all elementary theories, i.e. it is independent of their special logical constants and axioms. We should like also to mention the advantages of our algorithm if an application with a computer is considered; this algorithm is rather simple as far as the number of variables or the time necessary for its performing is considered (there are no large cycles in this algorithm).

In the following paper, which represents the part II of the report mentioned at the beginning of this paper, a procedure for statistical deducibility testing will be investigated. We shall assume there that the investigated formulas as well as the "auxiliary" axioms are chosen at random just by the algorithm investigated in this paper.

(Received March 5, 1971.)

REFERENCES

[1] Ohama Shigeo: On a formalism which makes any sequence of symbols well-formed. Nagoya Math. J. *32* (1968), 1—4.

[2] Church Alonzo: An introduction to Mathematical Logic, Part I. Princeton University Press, Princeton 1956.

[3] Gentzen Gerhardt: Untersuchungen über das Logische Schliessen. Mathem. Zeitschrift *39* (1934—5), 176—210, 405—431.

[4] Kramosil Ivan: Statistical Estimation of Deducibility in Polyadic Algebras. Kybernetika *7* (1971), 3, 181—200.

Metoda pro náhodný výběr správně vytvořených formulí

(Metoda pro náhodný výběr formulí nějaké elementární teorie a statistický odhad jejich dokazatelnosti doplněná programem I)

IVAN KRAMOSIL

V této práci je navržena metoda pro náhodný výběr z množiny správně vytvořených formulí nějaké elementární formalizované teorie s konečně mnoha funkcionálními konstantami a bez individuálních konstant.

Přirozenými čísly 1, 2, 3, ... jsou očíslovány symboly zkoumané teorie a ty z jejích elementárních formulí, které neobsahují individuální proměnné s indexy většími než předem dané N_0; nechť N_2 je nejvyšší přirozené číslo při tomto očíslování užité. Předpokládejme nyní, že máme k dispozici mechanismus pro náhodný výběr přirozených čísel takový, že alespoň N_2 přirozených čísel má kladnou pravděpodobnost být jim vybráno. Pak lze tento mechanismus snadno transformovat tak, že jsou to právě přirozená čísla $1, 2, \ldots, N_2$, která mají kladnou pravděpodobnost být takto vybrána.

Každou posloupnost takto vybraných náhodných čísel lze chápat jako posloupnost symbolů a elementárních formulí zkoumané teorie. Proto je zkonstruována procedura, která doplňováním nebo naopak vyškrtáváním určitých symbolů na určitých místech podle jistých pravidel transformuje každou posloupnost symbolů a elementárních formulí na jinou posloupnost, která už je, jak dokazuje věta 1, správně vytvořenou formulí. Přitom každá formule, neobsahující proměnné s indexy většími než N_0, má kladnou pravděpodobnost, že bude takto náhodně vybrána (věta 2). Je také uvedena určitá dolní mez pro pravděpodobnost, že uvedeným postupem bude vybrán teorém (věta 3). V závěrečném paragrafu jsou stručně uvedeny některé možnosti aplikace takového algoritmu.

Dr. Ivan Kramosil, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Vyšehradská 49, Praha 2.