ON THE RECURSIVITY OF THE FORMALIZED LANGUAGES

BY

T. RUS

In this paper a numeration notion for the words set of finite length of a free semi-group, generated by a finite symbols set V is defined. The CF-languages are studied as subsets of this free semi-group. It is proved that the set of the equivalence classes which represent the words of a CF-language in the already defined numeration is a recursive set.

At the end of the paper a parsing algorithm using numerical representation of words for CF-languages, is given.

§ 1. NUMERATIONS

It is known [1] that the numbered notion of pairs and in general of the n-tuples of natural numbers has led to a whole series of results in the recursive function theory. The study of some object sets from the point of view of their decidability seems to be easy with the help of some equivalence notions between these and some natural number sets.

DEFINITION 1.1. We call the numeration of some objects set W a mapping φ of some N_φ natural numbers set on the W set

$$\varphi : N_\varphi \to W \tag{1.1}$$

The number $n \in N_\varphi$ is called φ-number of the object $\varphi(n) \in W$ (or the number, if φ is fixed).

The N_φ set is called the set of the numeration numbers. From definition 1.1 it follows that every object of the W set has at least one number in N_φ. If every object in W has one and only one number $n \in N_\varphi$, then the numeration φ is called univocal. If N_φ is equal to the set N of all natural numbers, then the numeration is called simple.

We shall denote the system of the W set together with its numeration φ by $\langle W, \varphi \rangle$.

REV. ROUM. MATH. PURES ET APPL., TOME XIII, Nº 8, p. 1159–1172, BUCAREST, 1968

On the set N_φ we define a binary relation θ in the following way:

$$n\,\theta m \Leftrightarrow \varphi(n) = \varphi(m) \tag{1.2}$$

or in other words, two numbers $n, m \in N_\varphi$ will be found in the relation θ if they represent the numbers of the same object of W.

It is clear that θ has the following properties:

(i) $n\ \theta\ n \Leftrightarrow \varphi(n) = \varphi(n)$

(ii) $n\ \theta\ m$ and $m\ \theta\ p \Leftrightarrow \varphi(n) = \varphi(m) = \varphi(p)$.

Thus, θ is transitive.

(iii) $n\ \theta\ m \Leftrightarrow \varphi(n) = \varphi(m)$ and thus $n\ \theta\ m \Leftrightarrow m\ \theta\ n$. Hence it follows that θ is symmetrical.

So the relation θ is an equivalence relation in the N_φ set. If we denote the set of equivalence classes defined by θ in N_φ by $N_\varphi|_\theta$ then

$$\varphi : N_\varphi|_\theta \to W$$

is a 1—1 mapping.

DEFINITION 1.2. A general recursive function $f(x)$ represents a homomorphism of the system $\langle W_1, \varphi\rangle$ on the system $\langle W_2, \psi\rangle$ if $f(x)$ is a 1—1 correspondence of the natural numbers set, and if from $n, m \in N$ it follows that

$$\varphi(n) = \varphi(m) \Rightarrow \psi(f(n)) = \psi(f(m)) \tag{1.3}$$

From (1.3) it follows that the transformation

$$h : \varphi(n) \to \psi(f(n))$$

is a univocal mapping of the set W_1 on the set W_2, which we call homomorphism.

The function f is called the function of representation of this homomorphism. If we change condition (1.3) with condition (1.3′)

$$\varphi(n) = \varphi(m) \Leftrightarrow \psi(f(n)) = \psi(f(m)) \tag{1.3′}$$

then we say that f represents an isomorphism of the system $\langle W_1, \varphi\rangle$ on the system $\langle W_2, \psi\rangle$.

Let φ and ψ be two numerations of the same set W.

DEFINITION 1.3. We say that the numeration φ is reducible to the numeration $\psi (\varphi \leqslant \psi)$ if the general recursive function $f(x)$ exists so that

$$\varphi(n) = \psi(f(n))$$

Numerations φ and ψ will be called equivalents if they are reducible to each other.

DEFINITION 1.4. Numeration φ of set W is called unireducible to the numeration ψ of the same set if there is a general recursive function $f(x)$ which receives distinct values for the different values of the argument and which reduces φ to ψ.

Numerations φ and ψ are called uniequivalents if they are unireducible to each other.

DEFINITION 1.5. Numeration φ of set W will be called the numeration with effective infinite classes if there is a general recursive function $B(x, y)$ which has the properties:

(i) $\varphi(B(n, m)) = \varphi(n)$

(ii) $m \neq p \Rightarrow B(n, m) \neq B(n, p)$

DEFINITION 1.6. Numeration φ of set W is called complete if in W there is an object Λ which possesses the following properties: for each partial recursive function $F(x)$ a general recursive function $G(x)$ is so that for any $n \in N$

$$\varphi(G(n)) = \begin{cases} \varphi(F(n)) & \text{if } F(n) \text{ is definite} \\ \Lambda & \text{if } F(n) \text{ is not definite} \end{cases}$$

In all the above definitions we have considered the simple numerations, namely those numerations for which N_φ is N.

§ 2. NUMERATION OF THE FREE SEMI-GROUP OF WORDS

Let V be a finite set of symbols which we call alphabet. We denote by W the set of all the symbol strings of finite length from V. The elements of the W set will be called words. If we denote the operation of concatenation of words by $\circ$ then it is obvious that $(W, \circ)$ is a free semi-group generated by V. If we suppose that the empty word Λ whose length $l(\Lambda)$ is zero, $l(\Lambda) = 0$, belongs to W then the semi-group $(W, \circ)$ is with identity.

DEFINITION 2.1. We shall call a subset $L \subseteq W$ primitive recursive, recursive or recursively enumerable if the set of its corresponding numbers in a definite numeration φ for W is primitive recursive, recursive or recursively enumerable.

From definition 2.1. follow the properties:

(i) Every finite set of words is primitive-recursive.

(ii) The reunion and the intersection of the finite number of the primitive recursive set of words, recursive or recursively enumerable are of the same type sets.

(iii) If a set of words L and its complement relative to W, $W-L$ are recursively enumerable then they are recursive.

DEFINITION 2.2. We say that in the alphabet V is given an n-places function of words F if for any n-tuple of words $(x_1, x_2, \ldots, x_n)$ in W corresponds in a unique way a word $F(x_1, x_2, \ldots, x_n) \in W$.

An n-places numeric function f is called the representation function of F in the numeration φ if

$$F(\varphi(m_1), \varphi(m_2), \ldots, \varphi(m_n)) = \varphi(f(m_1, m_2, \ldots, m_n)) \tag{2.1}$$

for any $m_1, m_2, \ldots, m_n$ in N.

If function f is the representation function of the function F in the numeration φ then the following equalities take place:

$$F(x_1, x_2, \ldots, x_n) = \varphi(f(\mathfrak{N}(x_1), \mathfrak{N}(x_2), \ldots, \mathfrak{N}(x_n))) \tag{2.2}$$

$$f(m_1, m_2, \ldots, m_n) = \mathfrak{N}(F(\varphi(m_1), \varphi(m_2), \ldots, \varphi(m_n))) \tag{2.3}$$

where by $\mathfrak{N}(x)$ we have denoted the number of the word x in the numeration φ.

The function F is called primitive recursive, general recursive or partial recursive if its function of representation is the same.

Let the alphabet V be decomposed so that

$$V = V_N \cup V_T,\ V_N \cap V_T = \varnothing,\ V_N \neq \varnothing,\ V_T \neq \varnothing$$

where $V_N = \{A_1, A_2, \ldots, A_p\}$ and $V_T = \{a_1, a_2, \ldots, a_t\}$. We define an ordered relation R_i, $i = 1, 2$ in the sets V_N and V_T in the following way:

$$a_i R_1 a_j \Leftrightarrow i \leqslant j$$

$$A_i R_2 A_j \Leftrightarrow i \leqslant j \tag{2.4}$$

DEFINITION 2.3. A word $x \in W$ is called a segment if $x = x_1 x_2 \ldots x_n$, $x_j \in V$ $j = 1, 2, \ldots n$, and if $x_1 R_i x_2 R_i \ldots x_{n-1} R_i x_n$.

It follows immediately that if x is a segment in W then either $x \in W(V_T)$ or $x \in W(V_N)$.

We now consider the function $\delta : N \to V$ defined as:

$$\delta(n) = \begin{cases} a_i & \text{if } n = p(i) \text{ and div } (i, 2) = 0 \\ A_i & \text{if } n = p(i) \text{ and div } (i, 2) = 1 \\ \delta(p(i)) = \delta(p(i + kl)) \text{ where } l = \text{Card } (V), & \text{if } n = p(i + kl) \\ \Lambda & \text{in the other cases.} \end{cases}$$

By $p(0), p(1), p(2), \ldots$ we have denoted the prime numbers sequence. It is clear that $\delta : N \to V$ transforms the prime numbers sequence in the V alphabet, so that if $p(i)$ is the ith prime number and i is odd then by δ the ith letter in V_T corresponds to it in the order defined by R_1 in V_T, and if i is even the ith letter in V_N corresponds to it in the order defined by R_2 in V_N.

If $i >$ Car (V) then a symbol is chosen from V_N or V_T according to whether rest $(i,$ Card $(V))$ is even or odd.

Remark. If $p(i)$ is denoted by p_i then we can write the equality $\delta(p(i)) = \delta(p(i + kl))$ as well as the equality $\delta(p_i) = \delta_{(pi+kl)}$. If x is a segment in W we shall denote in the following by $\omega(x)$ its right limit and by $\nu(x)$ its left limit.

LEMMA 1. For any word $x \in W$ there are the words $\rho_1, \rho_2, \ldots \rho_k$ which are segments in W so that

$$x = \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k$$

where $\circ$ is the semi-group operation, the decomposition

$$x = \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k$$

being unique.

Proof. Let $x \in W$ be a word and we suppose that we have

$$x = \rho \circ x_1$$

where ρ is a segment. This decomposition is immediate because we consider ρ as being composed only of the first letter which has different occurrences of Λ in the word x. We now consider the letter a_{i_1} which is the beginning of the word x_1. Then it occurs one of the following possibilities:

(a) $\omega(\rho) R_1 a_{i_1}$, (b) $\omega(\rho) \overline{R}_1 a_{i_1}$

(c) $\omega(\rho) R_2 a_{i_1}$, (d) $\omega(\rho) \overline{R}_2 a_{i_1}$

where by $\overline{R}$ we have denoted the negation of the relation R.

In case (a) we have $\rho_1 = \rho a_{i_1}$ a segment and $x = \rho_1 \circ x_1'$. With the word x_1' we act in the same way as with the word x_1.

In case (b) a_{i_1} represents the beginning of the other segment and therefore $x = \rho_1 \circ \rho_2 \circ x_2$ where $\rho_2 = a_{i_1}$. With x_2 we act in the same manner as x_1 in case (a). Cases (c) and (d) are treated in the same manner.

Applying the above procedure in a finite number of steps we arrive at the decomposition of x in the form

$$x = \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k$$

where $\rho_1, \rho_2, \ldots, \rho_k$ are segments.

Supposing now that this decomposition is not unique there would be another decomposition of x in the form

$$x = q_1 \circ q_2 \circ \ldots \circ q_s$$

different from the first, $q_1, q_2, \ldots, q_s$ being segments. Then it is clear that the equality:

$\rho_1 \circ \rho_2 \circ \ldots \circ \rho_k = q_1 \circ q_2 \circ \ldots \circ q_s$ takes place.

Because $\rho_1, \rho_2, \ldots, \rho_k$ and $q_1, q_2, \ldots, q_s$ are segments it follows that $\rho_1 = a_{j_1} \rho_1'$, $q_1 = a_{j_1} q_1'$ and therefore ρ_1 and q_1 are at the same time segments in W in relation to R_1 or R_2.

If we now delete the letter a_{j_1} from the word x it is obtained the word x_1 so that

$$x_1 = \rho_1' \circ \rho_2 \circ \ldots \circ \rho_k$$

$$x_1 = \rho_1' \circ q_2 \circ \ldots \circ q_s$$

Repeating now the above procedure for an r number of steps, where $r = l(\rho_1) < l(q_1)$ we find the following equalities:

$$x_r = \rho_2 \circ \ldots \circ \rho_k$$

$$x_r = q_1^{(r)} \circ q_2 \circ \ldots \circ q_s$$

or

$$a_{j_2} \rho_2' \circ \ldots \circ \rho_k = a_{j_2} q_1^{(r+1)} \circ q_2 \circ \ldots \circ q_s$$

Let $b = \omega(\rho_1)$ be the right limit of ρ_1. Then it follows that $b R_1 a_{j_2}$ or $b R_2 a_{j_2}$. Hence ρ_1 and ρ_2 are not the different segments. In an analogous way we can prove that q_1 and q_2 are not the different segments if we suppose that $r = l(q_1) < l(\rho_1)$. This contradiction proves that the decomposition

$$x = \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k$$

where $\rho_1, \rho_2, \ldots, \rho_k$ are segments in W, is unique.

Let $n \in N$ be a natural number and its decomposition the following:

$$n = p_0^{\alpha_0} p_1^{\alpha_1} \dots p_k^{\alpha_k}$$

where $p_0, p_1, \dots, p_k$ are the first k prime numbers and $\alpha_0, \alpha_1, \dots, \alpha_k$ are their exponents in the decomposition in the prime factors of the n number with the condition $\alpha_k \neq 0$.

We denote by $\mu(p_i, n)$ the smallest prime number which appears in the decomposition of n on condition that its exponent is different from zero. Then its exponent $i(n)$ is given by the expression

$$i(n) = \mu_i(\overline{S}g\ (\mathrm{ex}\ (i, n) = 0) \tag{2.5}$$

where by $\mathrm{ex}(i, n)$ we have denoted the exponent of p_j in the decomposition of n.

It results that the equality

$$\mu(p_i, n) = \mu_m\ (|\pi(m) - (p(i\ (n)) + 1)| = 0) \text{ has taken place.} \tag{2.6}$$

We shall build the function $\varphi : N \to W$ which will represent a numeration of the set W in the following way:

$$\varphi(1) = \Lambda$$

$$\varphi(n+1) = \delta(\mu(p_i, n))^{\mathrm{ex}\,(i(n), n)} \circ \varphi\left(\frac{n+1}{\mu(p_i, n)^{\mathrm{ex}\,(i(n), n)}}\right). \tag{2.7}$$

It is clear that

$$\delta(\mu(p_i, n))^{\mathrm{ex}\,(i(n), n)} : N \to V^{\mathrm{ex}\,(i(n), n)}$$

THEOREM 1. For any word $x \in W$ there exists a number $n \in N$ so that $\varphi(n) = x$.

Proof. If $x = \Lambda$ it is obvious from the definition of φ that 1 is this number, because $\varphi(1) = \Lambda$.

Let us suppose that $x \neq \Lambda$ and x is a segment. Then let $x = a_{i_1} a_{i_2} \dots a_{ik}$ and $a_{i_1} R_1 a_{i_2} R_1 \dots a_{ik-1} R_1 a_{ik}$ and let $n = p_{i_1} p_{i_2} \dots p_{ik}$ be a natural number so that

$$\mathrm{div}\ (i_s, 2) = 0,\ s = 1, 2, \dots, k;\ i_1 \leqslant i_2 \leqslant \dots \leqslant i_k$$

Then from the definition of φ we have

$$\varphi(n) = \delta(p_{i_1}) \circ \delta(p_{i_2}) \circ \dots \circ \delta(p_{ik})$$

Because $\delta(p_{i_1}) = \delta(p(i_1))$ and because $\delta(p(i_1)) = a_{i_1}$ we have

$$\varphi(n) = a_{i_1} a_{i_2} \dots a_{ik}$$

If x is a segment and if we have $x = a_{i_1} a_{i_2} \dots a_{ik}$ and $a_{i_1} R_2 a_{i_2} R_2 \dots a_{ik-1} R_2 a_{ik}$ then we choose the number n so that $n = p_{i_1} p_{i_2} \dots p_{i_k}$, $\mathrm{div}\ (i_s, 2) = 1$, $s = 1, 2 \dots k$; $i_1 \leqslant i_2 \leqslant \dots \leqslant i_k$

In this case we have $\varphi(n) = A_{i_1} A_{i_2} \dots A_{ik}$.

Let $x \in W$ be a word and its decomposition of the form $x = \rho_1 \circ \rho_2 \circ \dots \circ \rho_k$ where $\rho_1, \rho_2, \dots, \rho_k$ are segments. We denote by $n_1, n_2, \dots, n_k$ the correspondent numbers of these segments, where

$$n_1 = p_{11} p_{12} \dots p_{1s_1}$$
$$n_2 = p_{21} p_{22} \dots p_{2s_2}$$
$$\dots\dots\dots\dots$$
$$n_k = p_{k_1} p_{k_2} \dots p_{ks_k}$$

We consider the number

$$m = n_1 \cdot p_{21+l} p_{22+1} \dots p_{2s_2+l} \dots p_{k_1+(k-1)l} p_{k_2+(k-1)l} \dots p_{ks_k+(k-1)l}$$

From the definition of φ it immediately follows that

$$\varphi(m) = \rho_1 \circ \rho_2 \circ \dots \circ \rho_k$$

which proves the theorem 1.

We now define an equivalence relation θ in N in the following way:

Let $n_1, n_2 \in N$ be two natural numbers, and

$$n_1 = p_{i_1}^{\alpha_1} \cdot p_{i_0}^{\alpha_2} \dots p_{ik}^{\alpha_k}$$
$$n_2 = p_{j_1}^{\alpha_1} \cdot p_{j_2}^{\alpha_2} \dots p_{jk}^{\alpha_k}$$

We shall say that n_1 is equivalent to n_2 and we denote this by $n_1 \sim n_2$ **iff** $i_s \equiv j_s \pmod{l}$, $s = 1, 2, \dots, k$, where l is the cardinal of the set V.

THEOREM 2. If $n_1\ m \in N$ and $n \sim m$ then $\varphi(n) = \varphi(m)$.

Proof. From the definition of φ we have

$$\varphi(n) = \delta(\mu(p_i, n))^{\mathrm{ex}\,(i(n), n)} \circ \varphi\left(\frac{n}{\mu(p_i, n)^{\mathrm{ex}\,(i(n), n)}}\right)$$

$$\varphi(m) = \delta(\mu(p_i, m))^{\mathrm{ex}\,(i(m), m)} \circ \varphi\left(\frac{m}{\mu(p_i, m)^{\mathrm{ex}(i(m), m)}}\right)$$

Because $n \sim m$ it follows that

$$\mathrm{ex}\,(i\,(n), n) = \mathrm{ex}\,(i(m), m)$$
$$i\,(n) \equiv i(m)\ (\mathrm{mod.}\ l)$$

But according to the δ definition

$$\delta(p\,(i(n))) = \delta(p\,(i(m)))$$

In this way the equality

$$\delta(\mu(p_i, n))^{\mathrm{ex}\,(i(n), n)} = \delta(\mu(p_i, m))^{\mathrm{ex}\,(i(m), m)} \text{ will take place.}$$

Thus, in a finite number of steps we shall arrive at the equality

$$\varphi(n) = \tau(m)$$

which proves theorem 2.

We denote by $N/\sim$ the set of the equivalence classes of set N according to the relation $\sim$ or in other words the quotient set according to the relation $\sim$.

Remark. If $\varphi(n) = \varphi(m)$ then $n \sim m$.

THEOREM 3. The function φ establishes a 1—1 correspondence of the set $N/\sim$ on W set.

For the demonstration of this theorem it is sufficient to show that φ is an injection and at the same time a surjection of $N/\sim$ on W. The first part of this statement results immediately from theorem 2 and the second part results from theorem 1 and from the unique representation of a number n in the following way

$$n = p_{i_1}^{\alpha_1} p_{i_2}^{\alpha_2} \dots p_{i_k}^{\alpha_k}$$

We now construct the inverse function of the function

$$\varphi : N/\sim \to W$$

which we denote by $\mathfrak{N}$ and which attaches to one word $x \in W$ the smallest number from the equivalence class of $N/\sim$ with the property $\varphi(n) = x$.

We shall make the following notations:

$$\mathfrak{N}(a_{i_j}) = p_{i_j} = p(i_j)$$

$$T_1(\mathfrak{N}(a_{i_j})) = p_{i_j} + l = p(i_j + l)$$

$$T_k(\mathfrak{N}(a_{i_j})) = p_{i_j} + kl = p(i_j + kl)$$

In consequence, for every k, T_k is a primitive recursive function.

Let $x \in W$ be a word and $x = \rho_1 \circ \rho_2 \circ \dots \circ \rho_k$ its decomposition according to lemma 1. Then we shall have

$$\mathfrak{N}(\rho_i) = \mathfrak{N}(a_{i_1}) \cdot \mathfrak{N}(a_{i_2}) \dots \mathfrak{N}(a_{ik_i})$$

$$\mathfrak{N}(x) = \mathfrak{N}(\rho_1)\, T_1(\mathfrak{N}(\rho_2)) \dots T_{k-1}(\mathfrak{N}(\rho_k))$$

where

$$\rho_i = a_{i_1} a_{i_2} \dots a_{ik_i} \quad T_1(\mathfrak{N}(\rho_i)) = T_1(\mathfrak{N}(a_{i_1})) \dots T_1(\mathfrak{N})\, a_{ik_i}))$$

It is easy to observe that the equalities:

$$\varphi(\mathfrak{N}(x)) = x$$

$$\mathfrak{N}(\varphi(n)) = n \text{ take place}$$

In this way the functions φ and $\mathfrak{N}$ are called the inverse functions one to the other, according to the equivalence relation $\sim$ defined in N.

§ 3. NUMERATION OF A FORMALIZED LANGUAGE DEFINITE IN W

Let W be the free semi-group generated by the set $V = V_N \cup V_T$, $V_N \cap V_T = \emptyset$, $V_N \neq \emptyset$, $V_T \neq \emptyset$, and let G be a grammar in W. It is known [2] that

$$G = (W, V_N, V_T, \mathfrak{Q})$$

where

$$\mathfrak{Q} = \{(\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_n, \beta_n)\}$$

$$\alpha_i, \beta_i \in W,\ i = 1, 2, \dots n$$

We shall consider the case of *CF*-grammars, that is to say for which the set $\mathfrak{Q}$ has the following property:

$$\alpha_i = A_i \in V_N,\ \beta_i \in W,\ \beta_i \neq \Lambda,\ i = 1, 2, \dots, n$$

The set $\mathfrak{Q}$ induces in W a system of binary relations, noted with Q_i, $i = 1, 2, \dots, n$, $n = \text{Card}\,(\mathfrak{Q})$ and defined in the following way: if $x, y \in W$ then $xQ_iy \Leftrightarrow x = x_1 A_i x_2$ & $y = x_1 \beta_i x_2$ & $(A_i, \beta_i) \in \mathfrak{Q}$.

We shall say that Q_i is a binary relation defined in W by (A_i, β_i).

If xQ_iy then we shall denote this by $x \vdash Q_i y$.

If $x, y \in W$ and if there is a sequence of words $x_0, x_1, \dots, x_p$ so that

$$x_0 = x$$

$$x_i \vdash Q_{i_i} x_{i+1},\ i = 0, 1, \dots, p-1$$

$$x_p = y \tag{3.1}$$

then we say that y is a derivation of x and we denote this by

$$x \vdash\!\!=_Q y\ ,\ Q = \prod_{i=0}^{p-1} Q_{j_i}$$

where $Q = \prod_{i=0}^{p-1} Q_{j_i}$ is the usual product of the relations which intervene in (3.1).

A *CF*-language is defined as the greatest subset so that we have

$$L = \{x \in W(V_T) \mid (\exists A_i)\ (A_i \in V_N)\ \&\ (A_i \vdash\!\!=_Q x)\}$$

Let L be a *CF*-language. We shall study this language from the point of view of its recursivity in the sense defined at the beginning.

For this we shall first see what the induced relation is in N by $x, y \in W$ and $x \vdash Q_i\ y$. From the definition of Q_i we have in this case

$$x = x_1 A_i x_2$$

$$y = x_1 \beta_i x_2$$

We denote by $\mathfrak{N}(A_i)$ and by $\mathfrak{N}(\beta_i)$ the numbers of A_i respectively β_i in the numeration φ defined in § 2. Then the equalities:

$$\mathfrak{N}(x) = \mathfrak{N}(x_1)\, T_1(\mathfrak{N}(A_i))\, T_2(\mathfrak{N}(x_2))$$

$$\mathfrak{N}(y) = \mathfrak{N}(x_1)\, T_1(\mathfrak{N}(\beta_i))\, T_{s+1}(\mathfrak{N}(x_2)) \tag{3.2}$$

are obvious, where s is the number of segments by which β_i is formed.

In this way if $x \vdash_{Q_i} y$ then

$$\mathfrak{N}(y) = \mathfrak{N}(x)\, \frac{T_1(\mathfrak{N}(\beta_i))\, T_{s+1}(\mathfrak{N}(x_2))}{T_1(\mathfrak{N}(A_i))\, T_2(\mathfrak{N}(x_2))} \tag{3.3}$$

We shall show now that the reciprocal statement occurs. If between two numbers $\mathfrak{N}(x)$ and $\mathfrak{N}(y)$ the relation (3.3) takes place then these are the corresponding numbers in the numeration φ to two words, $x, y \in W$ so that $x \vdash_{Q_i} y$.

For this we shall proceed in the following way:

$$\mathfrak{N}(y) = \mathfrak{N}(x)\, \frac{T_1(\mathfrak{N}(\beta_i)) \cdot T_{s+1}(\mathfrak{N}(x_2))}{T_1(\mathfrak{N}(A_i)) \cdot T_2(\mathfrak{N}(x_2))}$$

$$\frac{\mathfrak{N}(y)}{\mathfrak{N}(x)} = \frac{T_1(\mathfrak{N}(\beta_i)) \cdot T_{s+1}(\mathfrak{N}(x_2))}{T_1(\mathfrak{N}(A_i))\, T_2(\mathfrak{N}(x_2))}$$

Thus

$$\mathfrak{N}(y) = k_1 \cdot T_1(\mathfrak{N}(\beta_i))\, T_{s+1}(\mathfrak{N}(x_2))$$

$$\mathfrak{N}(x) = k_1 T_1(\mathfrak{N}(A_i)) \cdot T_2(\mathfrak{N}(x_2))$$

Applying now the function φ to the numbers $\mathfrak{N}(y)$ and $\mathfrak{N}(x)$ we shal have

$$\varphi(\mathfrak{N}(y)) = x_1 \beta_i x_2$$

$$\varphi(\mathfrak{N}(x)) = x_1 A_i x_2$$

Hence

$$x \vdash_{Q_i} y.$$

Remark. Both in the above proof and in relation (3.3) we have used the hypothesis that x_1 is a segment and also that k_1 is the smallets number of the equivalence class to which it belongs.

We now consider the general case $x = x_1 A_i x_2$, $y = x_1 \beta_i x_2$ the words x_1, x_2 being any word. Then, according to lemma 1

$$x_1 = \rho_1 \circ \rho_2 \circ \ldots \circ \rho_k$$

and thus

$$\mathfrak{N}(x_1) = \mathfrak{N}(\rho_1) \cdot T_1(\mathfrak{N}(\rho_2)) \ldots T_{k-1}(\mathfrak{N}(\rho_k))$$

In this case we have

$$\mathfrak{N}(y) = \mathfrak{N}(x)\, \frac{T_k(\mathfrak{N}(\beta_i)) \cdot T_{k+s+1}(\mathfrak{N}(x_2))}{T_k(\mathfrak{N}(A_i))\, T_{k+1}(\mathfrak{N}(x_2))} \tag{3.4}$$

THEOREM 4. If L is a *CF*-language in W then the set $\mathfrak{N}(L)|_\sim$ of the equivalence classes which represent the words of language L in the numeration φ is a recursive set.

Proof. Let L be a *CF*-language in W. Then by the definition

$$L = \{x \in W(V_T) \mid (\exists A_i)\ (A_i \in V_N \ \&\ A_i \models_Q x)\}$$

Thus, if $x \in L$ then $\mathfrak{N}(x) = p_{i_1}^{\alpha_1} p_{i_2}^{\alpha_2} \ldots p_{i_k}^{\alpha_k}$

$$\operatorname{div}(i_s, 2) = 0, \quad s = 1, 2, \ldots, k$$

We now consider the function

$$\chi_1(n) = \sum_{i=0}^{n} \operatorname{ex}(i, n) \operatorname{div}(i, 2) \tag{3.5}$$

We can immediately see that $\chi_1(n)$ is a primitive recursive function and if $\varphi(n) \in L$ then $\chi_1(n) = 0$. Otherwise $\chi_1(n)$ may be zero if $\varphi(n) \in W(V_T)$ or $\chi_1(n) \neq 0$ if $\varphi_1(n) \notin W(V_T)$.

We shall now construct a recursive function $\sigma(n, m)$ with the following property:

$$\sigma(n, m) = \begin{cases} \mathfrak{N}(A_i) \text{ if } A_i \in V_N \text{ exists so that } \varphi(n) \in L \text{ and} \\ \qquad A_i \models_Q \varphi(n),\ Q = \prod_{i-1}^{n} Q_i \\ 0 \qquad \text{otherwise} \end{cases} \tag{3.6}$$

For this we shall proceed in the following way: let the recursive function be

$$\chi_2(n) = \begin{cases} \mu_i(\operatorname{div}(n, \mathfrak{N}(\beta_i)) = 1) \text{ if such an } i \text{ exists}, 1 \leqslant i \leqslant n_1 \\ \qquad n_1 = \operatorname{Card}(\mathfrak{A}), \\ 0 \qquad \text{otherwise} \end{cases} \tag{3.7}$$

Because the set $\mathfrak{A}$ is a finite one, and $\chi_2(n)$ is definite all over the N it follows that $\chi_2(n)$ is a recursive function. We shall make the following notation

$$k(p_i^\alpha, p_j^\beta) = \begin{cases} p_s^{\alpha-\beta} \text{ if } p_i \sim p_j \text{ and } \alpha > \beta,\ s = \min(i, j) \\ 0 \qquad \text{otherwise} \end{cases} \tag{3.8}$$

Let n, m, be the numbers so that

$$n = p_{i_1}^{\alpha_1} p_{i_2}^{\alpha_2} \ldots p_{i,k}^{\alpha k}$$

$$m = p_{j_1}^{\beta_1} p_{j_2}^{\beta_2} \ldots p_{j_s}^{\beta_s}$$

where $s \leqslant k$. If the factors $p_{r_1}, p_{r_2}, \ldots, p_{r_s}$

in decomposition of n exist so that

$$p_{r_1} \sim p_{j_1} \qquad i_1 \leqslant r_1 \leqslant i_k$$
$$p_{r_2} \sim p_{j_2} \qquad i_1 \leqslant r_2 \leqslant i_k$$
$$\cdots\cdots\cdots\cdots$$
$$p_{r_s} \sim p_{j_s} \qquad i_1 \leqslant r_s \leqslant i_k$$

with the exponents $\alpha_{r_1}, \alpha_{r_2}, \ldots, \alpha_{r_s}$ different from zero if the exponents $\beta_{j_1}, \beta_{j_2}, \ldots, \beta_{j_s}$ are different from zero, and $r_1 \leqslant r_2 \leqslant \ldots \leqslant r_s$ then we denote by $K(n, m)$ the function defined by the equality

$$K(n,m) = \prod_{q=1}^{s} k(p_{r_q}^{\alpha_{r_q}}, p_{j_q}^{\beta_{r_q}}) \left[\frac{n}{p_{r_1}^{\alpha_{r_1}} p_{r_2}^{\alpha_{r_2}} \ldots p_{r_s}^{\alpha_{r_s}}} \right] \tag{3.9}$$

It is clear that

$$K(n, m) = \begin{cases} p \text{ where } n = pm_1,\ m \sim m_1 \text{ if such a } p \text{ exists} \\ 0 \text{ otherwise} \end{cases}$$

We put by definition $K(0, O) = 0$, $0(n, 0) = n$.

We now define the function $O(n, 1)$ in the following way:

$$O(n, 1) = K(\mathfrak{N}(\varphi(n)) \cdot \operatorname{sg}(\chi_2(n)) \cdot \mathfrak{N}(A_{\chi_2(n)}), \operatorname{Sg}(\chi_2(n))\, \mathfrak{N}(\beta_{\chi_2(n)})) \tag{3.10}$$

Thus, if $\chi_2(n) = 0$ then $O(n, 1) = 0$. The function $O(n, m)$ will be defined by the equality:

$$O(n, m) = O(O(n_1\ m - 1), 1) \tag{3.11}$$

We now consider the function

$$\psi(n) = \mu_m(C(n, m) \in \{0, p_2, p_4, \ldots, p_{2_p}\}) \tag{3.12}$$

where by $\{p_2, p_4, \ldots, p_{2_p}\}$ we denoted φ-numbers of the symbols from V_N. Because $O(n, m)$ is a recursive function and V_N is a finite set it is clear that $\psi(n)$ is a recursive function. Then the following equality is clear:

$$O(n, \psi(n)) = \begin{cases} \mathfrak{N}(A_i) \text{ if } A_i \in V_N \text{ exists so that } A_i \models_Q \varphi(n) \\ \qquad \text{and } Q = \prod_{i=1}^{\psi(n)} Q_{\chi_2 o (n_1 m - i))} \\ 0 \qquad \text{otherwise} \end{cases}$$

This function represents exactly the function (3.6) for the case $m = \psi(n)$.

The characteristic function of the set $\mathfrak{N}(L)/\sim$ in this way is given by the equality

$$\chi(n) = \bar{\operatorname{Sg}}(\bar{\operatorname{Sg}}(\chi_1(n)) \cdot \operatorname{Sg}(O(n, \psi(n)))) \tag{3.13}$$

which is a recursive function.

Thus the set $\mathfrak{N}(L)/\sim$ is a recursive set.

Result. The word set of a *CF*-language is recursive in the definite sense. Thus, every *CF*-language is a decidable subset in the free semigroup in which this language is defined.

The construction of the characteristic function $\chi(n)$ of the set $\mathfrak{N}(L)/\sim$ has important applications in the construction of parsing algorithms for programming languages.

We now consider the following algorithm for the parsing of a formalized language.

(1) Let L be a *CF*-language and let x be a word, $x \in W(V_T)$ We shall define the number $\mathfrak{N}(x)$

(2) In the second step of the algorithm the values of the function $o(\mathfrak{N}(x), i)$ for $i = 1, 2, \ldots$ are computed.

In a finite number of steps we shall find.

(a) $$o(\mathfrak{N}(x), m) = 0 \Rightarrow x \notin L$$

(b) $$o(\mathfrak{N}(x), m) = p(2k) \Rightarrow x \in L$$

and

$$A_k \models_Q x,\ Q = \prod_{i=1}^{m} Q_{\chi_2(o(\mathfrak{N}(x), m-i))}$$

where

$$A_k \in V_N,\ V_N = \{A_1, A_2, \ldots, A_k, \ldots, A_p\}$$

For instance, we have:

$$V_N = \{A, B\},\ V_T = \{s, 0, 1\}$$
$$\mathfrak{X} = \{(A, AB), (A,s), (B, BB), (B, 0), (B, 1)\}$$

The prime numbers sequence which we consider is:

$$p(1) = 2,\ p(2) = 3,\ p(3) = 5,\ p(4) = 7,\ p(5) = 11,$$

and so on.

Then we have:

$$\mathfrak{N}(A) = 3 \quad \mathfrak{N}(B) = 7$$
$$\mathfrak{N}(s) = 2, \quad \mathfrak{N}(0) = 5, \quad \mathfrak{N}(1) = 11$$

because AR_2B and $sR_10R\,1$

For our grammar we have:

1. (A, AB) $\mathfrak{N}(A) = 3$ $\mathfrak{N}(AB) = 3 - 7$
2. (A, s) $\mathfrak{N}(A) = 3$ $\mathfrak{N}(s) = 2$
3. (B, BB) $\mathfrak{N}(B) = 7$ $\mathfrak{N}(BB) = 7^2$
4. $(B, 0)$ $\mathfrak{N}(B) = 7$ $\mathfrak{N}(0) = 5$
5. $(B, 1)$ $\mathfrak{N}(B) = 7$ $\mathfrak{N}(1) = 11$

We now consider the word $x = s\,01$ and we want to parse this word.

According to the above algorithm we have:

$$\mathfrak{N}(s01) = 2.5.11$$

$$\begin{cases} \chi_2(2.5.11) = 2, \ \mathrm{sg}(\chi_2(2.5.11)) = 1 \\ \sigma(2.5.11,1) = K(2.5.11.3,2) = 3.5.11 \end{cases}$$

$$\begin{cases} \chi_2(3.5.11) = 4, \ \mathrm{sg}(\chi_2(3.5.11)) = 1 \\ \sigma(2.5.11,2) = \sigma(3.5.11,1) = K(3.5.11.7,5) = 3.7.11 \end{cases}$$

$$\begin{cases} \chi_2(3.7.11) = 1, \ \mathrm{sg}(\chi_2(3.7.11)) = 1 \\ \sigma(2.5.11,3) = \sigma(3.7.11,1) = K(3.7.11.3,3.7) = 3.11 \end{cases}$$

$$\begin{cases} \chi_2(3.11) = 5, \ \mathrm{sg}(\chi_2(3.11)) = 1 \\ \sigma(2.5.11,4) = \sigma(3.11,1) = K(3.11.7,11) = 3.7 \end{cases}$$

$$\begin{cases} \chi_2(3.7) = 1, \ \mathrm{sg}(\chi_2(3,7)) = 1 \\ \sigma(2.5.11,5) = \sigma(3.7,1) = K(3.7.3, \ 3.7) = 3 \end{cases}$$

$$3 = \mathfrak{N}(A).$$

Hence we have

$$A \vdash_{Q_1} AB1 \vdash_{Q_5} A1 \vdash_{Q_1} AB1 \vdash_{Q_4} A01 \vdash_{Q_2} s01, \ s01 \in L(G)$$

Received June 14, 1967

Calculus Institute
Academy of the Socialist Republic
of Romania

REFERENCES

1. А. И. Мальцев, *Алгоритмы и рекурсивные функции*, Изд. "Наука". Moscow, 1965.
2. T. Rus, *Tratarea algebrică a limbajelor formalizate*. St. Cercet. Mat., Bucharest, 1967, **19**, 2, 259–272.