MATHEMATICS

PAAVO TURAKAINEN

On characterization of recursively enumerable languages in terms of linear languages and VW-grammars

Communicated by Prof. A. van Wijngaarden at the meeting of June 18, 1977

University of Oulu, Department of Mathematics, Oulu, Finland

ABSTRACT

It is proved that for any alphabet Σ there exist a homomorphism h, a deterministic minimal linear language L_1 and a linear language L_2 such that every recursively enumerable language L over Σ is of the form $L = h(L_1 \cap L_2 \cup R_L)$ for some regular language R_L depending on L. Some other homomorphic characterizations are also presented. As an application of them it is proved that for every alphabet Σ there exists a deterministic linear language M_0 such that every λ-free recursively enumerable language over Σ is generated by a strict normal van Wijngaarden grammar in which one metavariable denotes M_0 and all other metavariables denote regular deterministic linear languages or regular languages of the form X^+ (X an alphabet).

1. HOMOMORPHIC CHARACTERIZATIONS

We first recall some definitions. A language is called *deterministic linear* iff it is generated by a linear grammar $G = (V, \Sigma, X_0, F)$ such that all productions in F are of the two forms (i) $X \to aYP$, (ii) $X \to a$, $a \in \Sigma$, $Y \in V - \Sigma$, $P \in \Sigma^*$, and for any $X \in V - \Sigma$ and $a \in \Sigma$ there is at most one production of type $X \to aP$, $P \in V^*$, in F. If, in addition, $V - \Sigma = \{X_0\}$ and $X_0 \to d$ for some terminal letter d is the only production of type (ii) and d does not occur in any other production, then L is called *deterministic minimal linear*. If a language is regular and deterministic linear, it is called regular deterministic linear.

The mirror image and the length of any word P will be denoted by

miP and $|P|$, respectively. Symbols Σ, Δ, A, B, Z, V and W stand for alphabets. The size of any alphabet Σ is denoted by $|\Sigma|$.

THEOREM 1. For any recursively enumerable language $L \subseteq \Sigma^*$ there exists a deterministic minimal linear language

$$L_1 = \{Pdmig(P) \mid P \in \Sigma_1^*\} \ \ (g: \Sigma_1 \to \Delta_1 \text{ is a bijection})$$

such that

$$L = h[h_1(L_1) \cap h_2(L_1c) \cap \Sigma^* c \Delta^*]$$

where $c \in \Delta$, $\Sigma \cap \Delta = \phi$ and h, h_1, h_2 are homomorphisms such that $h(a) = a$, $h(x) = \lambda$ for every $a \in \Sigma$ and $x \in \Delta$.

PROOF. L is generated by a grammar $G = (V, \Sigma, Y_0, F)$ such that the left sides of all productions $p \in F$ are in $(V - \Sigma)^+$. For each $z \in \Sigma$ and $p \in F$, let $\bar{z}$, $\bar{\beta}_z$, a_p, $\bar{a}_p$, b_p and $\bar{b}_p$ be new symbols. Let c and d be new symbols, too. Define a deterministic minimal linear grammar G_1 as follows. The terminal alphabet is the union of V and of the new symbols introduced above. X_0 is the only nonterminal letter, and the productions are

$$\begin{aligned}
&X_0 \to aX_0\bar{\beta}_a,\ X_0 \to \bar{a}X_0\bar{a} \text{ for every } a \in \Sigma,\\
&X_0 \to xX_0x \text{ for every } x \in V - \Sigma,\\
&X_0 \to cX_0c,\ X_0 \to d,\\
&X_0 \to a_pX_0b_p,\ X_0 \to \bar{a}_pX_0\bar{b}_p \text{ for every } p \in F.
\end{aligned}$$

Clearly $L_1 = L(G_1)$ is of the required form, Σ_1 is the union of V and of all symbols c, $\bar{a}$, a_p, $\bar{a}_p$, and Δ_1 is the union of $V - \Sigma$ and of all symbols $\bar{a}$, $\bar{\beta}_a$, b_p, $\bar{b}_p$, c.

The homomorphism h_1 is defined as follows:

(i) $h_1(\bar{\beta}_a) = \lambda$, $h_1(d) = dY_0c$ (Y_0 was the initial letter of G);

(ii) If $p: P \to Q$ is in F, then $h_1(a_p) = h_1(\bar{a}_p) = P$ and $h_1(b_p) = h_1(\bar{b}_p) = mi\bar{Q}$ where $\bar{Q}$ is obtained from Q by replacing each $a \in \Sigma$ by $\bar{a}$;

(iii) $h_1(x) = x$ otherwise.

Now we find that the language

$$h_1(L_1) \cap \Sigma^* c((V - \Sigma) \cup \bar{\Sigma} \cup \{c, d\})^* \tag{1}$$

contains only words of the form

$$P_1c\bar{P}_3c \ldots c\bar{P}_{2k-1}dY_0cmi\bar{P}_{2k-2}c \ldots cmi\bar{P}_2c \tag{2}$$

where $P_1 \in \Sigma^*$ and for each i, $P_{2i+1} \overset{*}{\Rightarrow} P_{2i}$. Conversely, if $P_1 \in \Sigma^*$ and for each i, $P_{2i+1} = P_{2i}$ or $P_{2i+1} \Rightarrow P_{2i}$, then the corresponding word (2) is in the language (1).

The homomorphism h_2 is defined as follows:

(i) $h_2(\bar{\beta}_a) = \bar{a}$;

(ii) If $p: P \to Q$ is in F, then $h_2(a_p) = Q$, $h_2(b_p) = miP$, $h_2(\bar{a}_p) = \bar{Q}$ and $h_2(\bar{b}_p) = mi\bar{P}$;

(iii) $h_2(x) = x$ otherwise.

For this homomorphism we find that the language

(3) $$h_2(L_1c) \cap \Sigma^*c((V-\Sigma) \cup \bar{\Sigma} \cup \{c, d\})^*$$

contains only words of the form

(4) $$P_1c\bar{P}_3c \dots c\bar{P}_{2k-1}dmi\bar{P}_{2k}c \dots cmi\bar{P}_2c$$

where $P_1 \in \Sigma^*$ and for each i, $P_{2i} \overset{*}{\Rightarrow} P_{2i-1}$. Conversely, if $P_1 \in \Sigma^*$ and for each i, $P_{2i} = P_{2i-1}$ or $P_{2i} \Rightarrow P_{2i-1}$, then the language (3) contains the corresponding word (4).

The above considerations imply that in the intersection of the languages (1) and (3) the words P_1 are exactly those terminal words for which there is a derivation $Y_0 \overset{*}{\Rightarrow} P_1$ in G. Thus we find that L is obtained from this intersection by a homomorphism which erases the symbols belonging to the alphabet $\Delta = (V-\Sigma) \cup \bar{\Sigma} \cup \{c, d\}$. The proof is thus complete.

Theorem 1 contains the result of Baker and Book (1974) that every recursively enumerable language is a homomorphic image of an intersection $L_1' \cap L_2'$ of two linear languages L_1' and L_2'. We shall improve this result by showing that for any fixed alphabet Σ, the homomorphism and the language L_1' may be assumed to be independent of the considered language L over Σ, and that L_1' is a deterministic minimal linear language. We first prove the following sharpened form of Theorem 1, which will be the main tool in our considerations concerning VW-grammars.

THEOREM 2. For any alphabet Σ there exist an alphabet A ($\Sigma \cap A = \phi$), a deterministic minimal linear language

$$L_1 = \{Pdmig(P) \mid P \in \Sigma_1^*\} \ (g: \Sigma_1 \to \Delta_1 \text{ is a bijection})$$

and three homomorphisms h, h_1, h_2 such that the following holds:

Every recursively enumerable language L over Σ is of the form

$$L = h[h_1(L_1) \cap h_2(L_1c) \cap R_L] \quad (c \in A)$$

for some regular language $R_L \subseteq \Sigma^*A^+$ depending on L. Moreover $h(a) = a$ and $h(x) = \lambda$ for every $a \in \Sigma$ and $x \in A$.

PROOF. Let $\Sigma = \{x_1, \dots, x_k\}$ be fixed. Let $\alpha = x_{k+1}$, $\beta = x_{k+2}$, $\gamma = x_{k+3}$, $c = x_{k+4}$, $d = x_{k+5}$, 0 and 1 be new letters. For each x_i ($i = 1, \dots, k+3$) let $\bar{x}_i = x_{k+5+i}$ and $\bar{0} = x_{2k+9}$, $\bar{1} = x_{2k+10}$ be new letters. For any grammar $G = (V, \Sigma, X_0, F)$ having Σ as the terminal alphabet, we may assume that $X_0 = x_{2k+11}$ and $V - \Sigma \subset \{x_{2k+i} \mid i > 10\}$. All letters x_i are now encoded in the alphabet $\{0, 1\}$ by defining $f(x_i) = 10^i1$ ($i = 1, 2, \dots$). Every word $P = x_{i_1}x_{i_2} \dots x_{i_r} \in V^+$ is encoded by defining $f(P) = f(x_{i_1})f(x_{i_2}) \dots f(x_{i_r})$. If

$P_1 \to Q_1, \ldots, P_n \to Q_n$ are the productions of G in some order, we define a Gödel-word $f(G)$ for G by

$$f(G) = \gamma f(P_1)\beta f(Q_1)\alpha f(P_2)\beta f(Q_2)\alpha \ldots \alpha f(P_n)\beta f(Q_n)\alpha.$$

Then the language

$$U_0 = \bigcup_G \{Pf(G) \mid P \in L(G)\} \subseteq \Sigma^* \{0, 1, \alpha, \beta, \gamma\}^+$$

is recursively enumerable. (The union is taken over all grammars having Σ as the terminal alphabet.)

We now apply Theorem 1 to the language U_0. Thus

$$(5) \qquad U_0 = h[h_1(L_1) \cap h_2(L_1 c) \cap \Sigma_0^* c \Delta^*]$$

where $\Sigma_0 = \Sigma \cup \{0, 1, \alpha, \beta, \gamma\}$, Δ consists of $\overline{\Sigma}_0$, c, d and of some nonterminal letters x_{2k+i} with $i > 10$, $\Sigma_0 \cap \Delta = \phi$ and $h(a) = a$, $h(x) = \lambda$ for every $a \in \Sigma_0$, $x \in \Delta$.

Let $L = L(G)$ be any recursively enumerable language over Σ. Then clearly

$$(6) \qquad L = h_3[h_1(L_1) \cap h_2(L_1 c) \cap \Sigma^* f(G) c \Delta^*]$$

where $h_3(a) = a$ and $h_3(x) = \lambda$ for every $a \in \Sigma$ and $x \in A = \Delta \cup \{0, 1, \alpha, \beta, \gamma\}$. This proves the theorem, because in the above formula only the language $R_L = \Sigma^* f(G) c \Delta^* \subseteq \Sigma^* A^+$ depends on L.

COROLLARY. Let Σ be any alphabet and $h(0) = h(1) = \lambda$, and $h(a) = a$ for every $a \in \Sigma$. There exist two linear languages L_1' and L_2' such that the following holds:

Every recursively enumerable language L over Σ can be expressed in the forms

$$L = h(L_1' \cap L_2' \cap R_L) = (L_1' \cap L_2')/R_L'$$

for some regular languages $R_L \subseteq \Sigma^* \{0, 1\}^+$ and $R_L' \subseteq \{0, 1\}^+$ depending on L.

PROOF. Consider the formulas (5) and (6). As we mentioned in the previous proof, Δ consists of $\overline{\Sigma}_0$, c, d and of some nonterminal letters x_i with $i > 2k + 10$. By using the encoding introduced above, we conclude by (6) that

$$L = h_4(L_1' \cap L_2' \cap \Sigma^* f'(G) f(c) \{f(b) \mid b \in \Delta\}^*)$$

where L_1' and L_2' are obtained from $h_1(L_1)$ and $h_2(L_1 c) \cap \Sigma^* \gamma \{0, 1, \alpha, \beta\}^* c \Delta^*$ by the homomorphism which replaces each letter of $\Delta \cup \{\alpha, \beta, \gamma\}$ by its encoding; $f'(G)$ is obtained from $f(G)$ by replacing α, β and γ by their encodings; $h_4(a) = a$ for every $a \in \Sigma$ and $h_4(0) = h_4(1) = \lambda$. We also find that

$$L = (L_1' \cap L_2')/f'(G) f(c) \{f(b) \mid b \in \Delta\}^*.$$

This completes the proof. (Note that $R_L = \Sigma^* R_L'$.)

THEOREM 3. For any alphabet Σ there exist an alphabet B, a deterministic minimal linear language

(7) $$L_1=\{Pdmig(P) \mid P \in \Sigma_1^*\} \ (g: \Sigma_1 \to \Delta_1 \text{ is a bijection}),$$

a linear language L_2 and a homomorphism h such that the following holds:

Every recursively enumerable language L over Σ can be expressed in the forms

(8) $$L=h(L_1 \cap L_2 \cap R_L)=(L_1 \cap L_2)/R_L'$$

where $R_L \subseteq \Sigma^* B^+$ and $R_L' \subseteq B^+$ are regular languages depending on L. Moreover, $h(a)=a$ and $h(b)=\lambda$ for every $a \in \Sigma$ and $b \in B$.

PROOF. Let $U_0=L(G)$, $G=(V, \Sigma_0, Y_0, F)$, $\Sigma_0=\Sigma \cup \{0, 1, \alpha, \beta, \gamma\}$, where U_0 is the language defined in the proof of Theorem 2. By considering the proof of Theorem 1 we find that in the formula for L the part $h_1(L_1) \cap \Sigma^* c \Delta^*$ can be replaced by $h_1(L_1 \cap \Sigma^* c \Delta_2^*)$ where Δ_2 consists of $V-\Sigma$ and of the letters c, d, $\bar{\beta}_a$, $\bar{a}$, $\bar{a}_p$ and $\bar{b}_p$. Consequently, if we choose $L=U_0$, we have (Δ_2 is now determined by Σ_0)

$$U_0=h[h_1(L_1 \cap \Sigma_0^* c \Delta_2^*) \cap h_2(L_1 c)].$$

This implies that if $h'=hh_1$, then

$$U_0=h'(L_1 \cap L_2' \cap \Sigma_0^* c \Delta_2^*)$$

where $L_2'=h_1^{-1}(h_2(L_1 c))$ being linear. Clearly $h'(a)=a$ and $h'(b)=\lambda$ for every $a \in \Sigma_0$ and $b \in \Delta_2$.

Let $L=L(G_1)$ be any recursively enumerable language over Σ. Then

(9) $$L=h''(L_1 \cap L_2' \cap \Sigma^* f(G_1) c \Delta_2^*)$$

where $h''(a)=a$ and $h''(x)=\lambda$ for every $a \in \Sigma$ and $x \in B=\Delta_2 \cup \{0, 1, \alpha, \beta, \gamma\}$. If we now replace L_2' by the language

$$L_2=L_2' \cap \Sigma^* \gamma \{0, 1, \alpha, \beta\}^* c \Delta_2^*,$$

then (8) holds when we choose

$$R_L=\Sigma^* f(G_1) c \Delta_2^*, \ R_L'=f(G_1) c \Delta_2^*.$$

In the above considerations Σ was fixed, but if we want to use encodings and inverse homomorphisms, we conclude by formula (9) that the following holds:

COROLLARY. The family of all recursively enumerable languages is the full principal semi-*AFL* generated by the language $L_1 \cap L_2$ where L_1 is a deterministic minimal linear language of the form (7) and L_2 is a linear language of the form $L_2=h_1^{-1}(h_2(L_1 c))$.

In the rest part of the paper we apply Theorems 3 and 2 to van Wijngaarden grammars (shortly VW-grammars). For the definitions and terminology we refer to Greibach (1974) and van Leeuwen (1977). Our results are closely related to those of van Leeuwen. His starting point was much more complicated than Theorem 2 and, consequently, the proofs are hard to present in a concise form.

In the following theorems "metavariable" means "metavariable appearing in at least one hyper-rule".

By using Theorem 3 and modifying the construction of Greibach (1974) we easily obtain the following

THEOREM 4. For any alphabet Σ there exist an alphabet B and a linear language L_0 such that the following holds:

Every λ-free recursively enumerable language L over Σ is generated by a strict normal VW-grammar G in which one metavariable denotes L_0 and all other metavariables denote Σ^+ or B^+ or regular deterministic linear languages only one of which depends on L. Moreover, the hyper-rules of G are independent of L.

PROOF. By Theorem 3, any λ-free recursively enumerable language $L \subseteq \Sigma^+$ is of the form $L = h(L_1 \cap L_2 \cap R_L)$ where $R_L \subseteq \Sigma^+ B^+$ is a regular language, L_1 and L_2 are linear languages independent of L, and $h(a) = a$, $h(b) = \lambda$ for every $a \in \Sigma$, $b \in B$.

Let β_1, β_2 and c_a $(a \in \Sigma)$ be new letters. Let S, T, C, D and E_a $(a \in \Sigma)$ be metavariables denoting the languages

$$L_S = \Sigma^+,\ L_T = B^+,\ L_C = \beta_1 L_1 \cup L_2,\ L_D = \beta_1 R_L \beta_2,$$
$$L_{E_a} = \beta_1 \Sigma^* a B^+ \beta_2 c_a \ (a \in \Sigma).$$

The only protovariable is σ, and the hyper-rules are

$$\sigma \to S\langle SaT\rangle,\ \sigma \to \langle aT\rangle \text{ for every } a \in \Sigma,$$
$$\langle C\rangle \to \langle \beta_1 C\rangle,\ \langle C\rangle \to \langle C\beta_2\rangle,$$
$$\langle D\rangle \to \langle Dc_a\rangle \text{ for every } a \in \Sigma,$$
$$\langle E_a\rangle \to a \text{ for every } a \in \Sigma.$$

A derivation of G leads to a terminal word iff it is of the form

$$\sigma \Rightarrow P\langle PaQ\rangle \Rightarrow P\langle \beta_1 PaQ\rangle \Rightarrow P\langle \beta_1 PaQ\beta_2\rangle \Rightarrow P\langle \beta_1 PaQ\beta_2 c_a\rangle \Rightarrow Pa$$

where $P \in \Sigma^*$, $a \in \Sigma$, $Q \in B^+$, and $PaQ \in L_2$, $\beta_1 PaQ \in \beta_1 L_1$, $\beta_1 PaQ\beta_2 \in \beta_1 R_L \beta_2$, i.e. $PaQ \in L_1 \cap L_2 \cap R_L$ and $Pa \in L$. Thus G generates L and is of the required type with $L_0 = \beta_1 L_1 \cup L_2$.

In the previous theorem $L_0 = \beta_1 L_1 \cup L_2$ where L_1 is a deterministic minimal linear language and L_2 is linear. It will be seen later that L_0 may be assumed to be deterministic linear.

By slightly modifying the definitions of G_1, h_1 and h_2 in the proof of Theorem 1 we may assume in Theorem 2 that Σ, A, Σ_1 and $\Delta_1 = g(\Sigma_1)$ are disjoint. Then by using renaming of Σ_1 and Δ_1, say Σ_2 and Δ_2, we can replace $h_1(L_1)$ by $h_1(L_2)$. In addition, we replace d in L_1 and L_2 by new letters $\S_1$ and $\S_2$. Since h, h_1 and h_2 are now acting on languages over disjoint alphabets, we may replace h_1 and h_2 by h. Thus we have the following

LEMMA 1. For any alphabet Σ there exist disjoint alphabets Σ_1, Σ_2, Δ_1, Δ_2 and A such that if

$$L_i = \{P\S_i mig(P) \mid P \in \Sigma_i^*,\ g\colon \Sigma_i \to \Delta_i \text{ is a bijection}\}\ (i = 1, 2)$$

then the following holds:

Every recursively enumerable language L over Σ is of the form

$$L = h[h(L_1 c_1) \cap h(L_2) \cap R_L]$$

for some regular language $R_L \subseteq \Sigma^* A^+$ depending on L. Here $c_1 \in \Delta_1$, $h(c_1) = c \in A$, and $h(a) = a$, $h(x) = \lambda$ for every $a \in \Sigma$, $x \in A$.

In what follows we apply this lemma to the left derivatives L_v, $|v| = 5$, of L and prove the following

THEOREM 5. For any alphabet Σ there exist two alphabets V and W and a linear language

$$M = \bigcup_{i=1}^{2} \{P\S_i mig(P) \mid P \in V^*\} \cup \{P\,\$\,mih(P) \mid P \in W^*\}$$

(the union of three deterministic minimal linear languages)

such that the following holds:

Every λ-free recursively enumerable language L over Σ is generated by a strict normal VW-grammar in which one metavariable denotes M and all other metavariables denote regular deterministic linear languages or regular languages of the form Z^+ (Z an alphabet).

PROOF. Let Σ be fixed and let $L \subseteq \Sigma^+$ be any λ-free recursively enumerable language. We write L in the form

$$L = L_0 \bigcup \cup\, vL_v$$

where $L_0 = \{w \in L,\ |w| \leq 5\}$, $L_v = \{w \mid vw \in L,\ w \neq \lambda\}$, and the union is taken over all v's with $|v| = 5$.

Lemma 1 is now applied to every L_v. We first introduce for each $v \in \Sigma^5$ three new letters a_v, b_v and $g(a_v)$ for which we define $h(a_v) = b_v$ and $h(b_v) = \lambda$. Then by Lemma 1,

$$L_v = h[h(a_v L_1 c_1) \cap h(a_v L_2) \cap b_v R_v]$$

for any L_v with $v \in \Sigma^5$. Here $R_v \subseteq \Sigma^+ A^+$ is a regular language depending on L_v, and $h(a)=a$, $h(x)=\lambda$ for every $a \in \Sigma$, $x \in A$. Denote

$$V = \Sigma_1 \cup \Sigma_2 \cup \{a_v | v \in \Sigma^5\},$$
$$W = V \cup \Sigma \cup \Delta_1 \cup \Delta_2 \cup \{b_v | v \in \Sigma^5\} \cup \{\S_1, \S_2\} \cup A.$$

Now the language M in the theorem is completely defined.

The required VW-grammar for L is constructed as follows. There is only one protovariable σ, and the hyper-rules are defined in the following manner:

(i) $\sigma \to w$ for each $w \in L_0$,

(ii) For any $v = a_1^{(v)} a_2^{(v)} \dots a_5^{(v)} \in \Sigma^5$,
$\sigma \to \langle T_v \$ S \rangle \langle a_v C c_1 \$ T_v \rangle \langle a_v D \$ T_v \rangle \langle a_v C g(a_v) \rangle \langle a_v D g(a_v) \rangle S$,
$\langle X_i^{(v)} \rangle \to a_i^{(v)} \qquad (i = 1, 2, 3, 4, 5)$;

(iii) $\langle X_M \rangle \to \langle X_M \delta \rangle$ (δ is a new letter).

The languages corresponding to the metavariables appearing in these rules are

$$L_{X_M} = M,\ L_S = \Sigma^+,\ L_{T_v} = (miR_v)b_v,$$
$$L_C = (\Sigma_1 \cup \Delta_1 \cup \{\S_1\})^+,$$
$$L_D = (\Sigma_2 \cup \Delta_2 \cup \{\S_2\})^+,$$
$$L(X_1^{(v)}) = A^+ \Sigma^+ b_v \$ \Sigma^+ \delta,$$
$$L(X_2^{(v)}) = a_v L_C c_1 \$ A^+ \Sigma^+ b_v \delta,$$
$$L(X_3^{(v)}) = a_v L_D \$ A^+ \Sigma^+ b_v \delta,$$
$$L(X_4^{(v)}) = a_v L_C g(a_v) \delta,$$
$$L(X_5^{(v)}) = a_v L_D g(a_v) \delta.$$

Consider an arbitrary derivation ending to a terminal word

$$a_1 a_2 \dots a_5 w,\ w \neq \lambda.$$

We have to show that, for some $v = a_1^{(v)} \dots a_5^{(v)}$, $a_1 a_2 a_3 a_4 a_5$ is equal to v and $w \in L_v$. There is a word $v \in \Sigma^5$ such that the derivation is of the form

$$\sigma \overset{*}{\Rightarrow} \langle P \$ w \rangle \langle a_v Q_1 c_1 \$ P \rangle \langle a_v Q_2 \$ P \rangle \langle a_v Q_1 g(a_v) \rangle \langle a_v Q_2 g(a_v) \rangle w$$
$$\overset{*}{\Rightarrow} a_1 a_2 a_3 a_4 a_5 w$$

where

$$w \in \Sigma^+,\ P \in (miR_v)b_v,\ Q_i \in (\Sigma_i \cup \Delta_i \cup \{\S_i\})^+ \quad (i = 1, 2)$$

and

$$\langle P \$ w \rangle \Rightarrow \langle P \$ w \delta \rangle \Rightarrow a_1,\ \dots,\ \langle a_v Q_2 g(a_v) \rangle \Rightarrow \langle a_v Q_2 g(a_v) \delta \rangle \Rightarrow a_5.$$

This is possible only if the words $P \$ w$, $a_v Q_1 c_1 \$ P$, $a_v Q_2 \$ P$, $a_v Q_1 g(a_v)$ and $a_v Q_2 g(a_v)$ are in M. By the definition of M this implies that

$$w = mih(P),\ P = mih(a_v Q_1 c_1) = mih(a_v Q_2),\ Q_1 \in L_1,\ Q_2 \in L_2.$$

Consequently,

$$w = h'(miP),\ miP = h(a_v Q_1 c_1) = h(a_v Q_2) \in b_v R_v$$

where $h'(x) = mih(x) = h(x)$ for each letter x belonging to

$$\Sigma \cup A \cup \{b_u |\, u \in \Sigma^5\}.$$

Thus $w \in L_v$. Since $w \neq \lambda$, it follows that $Q_1 \neq \S_1$ and $Q_2 \neq \S_2$. Therefore, $Q_i \in \Sigma_i^+ \S_i \Delta_i^+$ $(i = 1, 2)$. This implies that $a_i = a_i^{(v)}$ for $i = 2, 3, 4, 5$. Since P contains the letter b_v, we finally conclude that $a_1 = a_1^{(v)}$. Hence

$$a_1 a_2 a_3 a_4 a_5 w \in L.$$

Since $h' = h$ in the alphabet $\Sigma \cup A \cup \{b_v |\, v \in \Sigma^5\}$, it is obvious that every word of L is generated by our VW-grammar. This completes the proof, because the grammar is of the required form.

In Theorem 4 the only nonregular language L_0 was linear. In the following theorem we show that even a deterministic linear language can be used.

THEOREM 6. For any alphabet Σ there exists a deterministic linear language M_0 such that the following holds:

Every λ-free recursively enumerable language L over Σ is generated by a strict normal VW-grammar in which one metavariable denotes M_0 and all other metavariables denote regular deterministic linear languages or regular languages of the form Z^+ (Z an alphabet).

PROOF. In the previous proof we replace M by the language

$$M_0 = \bigcup_{i=1}^{2} \varrho_i \{P \S_i mig(P) |\, P \in V^*\} \cup \{P \,\$\, mih(P) |\, P \in W^*\}$$

where ϱ's are new letters. This language is deterministic linear. In the hyper-rules (ii) we replace $\langle a_v C g(a_v) \rangle$ and $\langle a_v D g(a_v) \rangle$ by $\langle \varrho_1 a_v C g(a_v) \rangle$ and $\langle \varrho_2 a_v D g(a_v) \rangle$, respectively. Finally, ϱ_1 and ϱ_2 are written in front of the expressions of $L(X_4^{(v)})$ and $L(X_5^{(v)})$.

REFERENCES

1. Baker, B. S. and R. V. Book – Reversal-bounded multipushdown machines, J. Comp. Syst. Sci. **8**, 315–332 (1974).
2. Greibach, S. A. – Some restrictions on W-grammars, Int. J. Comp. Inf. Sci. **3**, 289–327 (1974).
3. Leeuwen, J. van – Recursively enumerable languages and van Wijngaarden grammars, Indag. Math. **39**, 29–39 (1977).
4. Wijngaarden, A. van – The generative power of two-level grammars, Proc. 2nd Colloq. on Automata, Languages and Programming, Saarbrücken, Springer Lecture Notes in Computer Science **14**, 9–16 (1974).