Kybernetika

Karel Čulík II

n-ary grammars and the description of mapping of languages

Kybernetika, Vol. 6 (1970), No. 2, (99)–117

Persistent URL: http://dml.cz/dmlcz/124904

KYBERNETIKA ČÍSLO 2, ROČNÍK 6/1970

n-ary Grammars and the Description of Mapping of Languages*

KAREL ČULÍK II

The n-ary grammars generating relations (set of n-tuples of words) are introduced. Chomsky's classification is generalized for them and closure, projective and other properties of different classes of relations generated by n-ary grammars are studied. The n-ary grammars are used for describing mappings of languages (e.g. translation) and for the classification of their complexity.

INTRODUCTION

Up to now, the phrase-structure grammars have been used for describing languages, i.e. sets of words over some alphabet. There already exists a number of different generalisations of phrase structure grammars, we introduce another, namely the n-ary grammars which generate relations, i.e. sets of n-tuples of words over some alphabets. An n-ary grammar is a system of n terminal alphabets, n nonterminal alphabets, a set of productions and an initial n-tuple of nonterminals. Each production is an n-tuple of common productions or empty places. Three different types of generation of relations by n-ary grammars are introduced and their properties and relations are investigated. Chomsky's classification is generalized for n-ary grammars and closure, projective and other properties of different classes of relations are studied.

The binary grammars are used to describe alphabetical mappings, particularly translations, from one language to another. The types β or γ of generation are particularly convenient for this application. The translations definable by binary grammars include well-translation [1] or syntax-directed translation [7] as simple cases. The classification of complexity of mappings by the type of grammar defining them are introduced and several examples are given.

* This research was partially supported by the National Research Council of Canada, grant A-4096.

For basic terminology and notation see S. Ginsburg [4] and C. C. Elgot, J. E. Mezei [3]. Here we consider them only briefly and with certain modifications:

Let Σ be an alphabet, Σ^* a set of all words over the alphabet including the empty word ε; $(\Sigma^*)^n$ is the set of all n-tuples of words over Σ.

Concatenation of n-tuples: if u, v are in $(\Sigma^*)^n$, $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ then $uv = (u_1v_1, \ldots, u_nv_n)$.

Kleene's operations for relations (sets of n-tuples of words):

if $R \subset (\Sigma^*)^n$, $S \subset (\Sigma^*)^n$ then $R \cup S$ is the common set union;

if $R \subset (\Sigma^*)^n$, $S \subset (\Sigma^*)^n$ then concatenation $RS = \{uv \mid u \text{ in } R, v \text{ in } S\}$;

if $R \subset (\Sigma^*)^n$ then iteration $R^* = R^0 \cup R \cup RR \cup RRR \cup \ldots$, where

$$R^0 = \underbrace{(\varepsilon, \varepsilon, \ldots, \varepsilon)}_{n\text{-times}}$$

The domain of a binary relation R is denoted by dom R, dom $R = \{x \mid (x, y) \text{ is in } R\}$.

Let $G = (V_T, V_N, P, S)$ be a phrase structure grammar in the sense of [4], $L(G)$ the language generated by the grammar G. Let $V = V_T \cup V_N$.

A general phrase structure grammar is said to be of the type 0.

A grammar $G = (V_T, V_N, P, S)$ is called context sensitive (type 1) if each production in P is of the form $uAv \to uyv$ where A is in V_N, u, v are in V_N^* and y is in $V^* - \{\varepsilon\}$.

A grammar $G = (V_T, V_N, P, S)$ is called context-free (type 2) if each production in P is of the form $A \to v$ where A is in V_N and v is in V^*.

A grammar $G = (V_T, V_N, P, S)$ is called finite state (type 3) if each production in P is of the form $A \to vB$ or $A \to v$, where A, B are in V_N and v is in V_T^*.

A finite state grammar $G = (V_T, \{S\}, P, S)$ (one nonterminal symbol) is said to be of type 4.

n-ARY GRAMMARS

Definition. An n-ary grammar G is a system $(V_T^1 \times V_T^2 \times \ldots \times V_T^n, V_N^1 \times V_N^2 \times \ldots \times V_N^n, P, (S_1, S_2, \ldots, S_n))$, where $V_T^i \cap V_N^i = \emptyset$, $i = 1, 2, \ldots, n$,

(i) V_T^i, $i = 1, 2, \ldots, n$ are alphabets (of terminal symbols);
(ii) V_N^i, $i = 1, 2, \ldots, n$ are alphabets (of nonterminal symbols);

(iii) P is a finite nonempty set of n-ary productions. An n-ary production is an n-tuple $(q_1, q_2, \ldots, q_n)$, where q_i is either empty or a pair (x, y), where x is in $V_N^{i*} - \{\varepsilon\}$, y is in V^{i*}. V^i denotes here and throughout the whole paper the union $V_N^i \cup V_T^i$.

The pair (x, y) is a production in the usual way and is normally written in the form $x \to y$.

(iv) $(S_1, S_2, \ldots, S_n)$ is the initial n-tuple, S_i is in V_N^i for $i = 1, 2, \ldots, n$.

Note. A n-ary production $(q_1, q_2, \ldots, q_n)$ will also be written in the form $u \to v$, where $u = (u_1, u_2, \ldots, u_n)$, $v = (v_1, v_2, \ldots, v_n)$, $u_i \to v_i = q_i$ for $q_i \neq \emptyset$, $u_i = v_i = \varepsilon$ for $q_i = \emptyset$.

Definition. Let G be an n-ary grammar. Let the n-tuples u, v be in $V^{1*} \times V^{2*} \times \times \ldots \times V^{n*}$. We write $u \stackrel{\alpha}{\Rightarrow} v$ if there exist n-tuples x, y, z in $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$, w in $V_N^{1*} \times V_N^{2*} \times \ldots \times V_N^{n*}$, such that $u = ywz$, $v = yxz$ and $w \to x$ is in P.

Let relation $\stackrel{\alpha}{\Rightarrow}{}^*$ be the reflexive and transitive closure of $\stackrel{\alpha}{\Rightarrow}$. If $u \stackrel{\alpha}{\Rightarrow}{}^* v$ then there exists the sequence $u_0, u_1, \ldots, u_r$ $r \geqq 0$ such that $u_0 = u$, $u_r = v$ and $u_0 \stackrel{\alpha}{\Rightarrow} u_1 \stackrel{\alpha}{\Rightarrow} \stackrel{\alpha}{\Rightarrow} \ldots \stackrel{\alpha}{\Rightarrow} u_r$. The sequence is called derivation of v (from u).

Definition. Let $\stackrel{\beta}{\Rightarrow}{}^*$ be the minimal reflexive and transitive binary relation on $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$ closed under the following procedure for obtaining new members from ones already admitted:

if for u, v in $V^{1*} \times \ldots \times V^{n*}$ there exist y, z, t, s in $V^{1*} \times \ldots \times V^{n*}$ and x in $V_T^{1*} \times \ldots \times V_T^{n*}$ such that $u = ysz$, $v = yxz$, $s \to t$ in P and $t \stackrel{\beta}{\Rightarrow}{}^* x$ then $u \stackrel{\beta}{\Rightarrow}{}^* v$.

The length of $u \stackrel{\beta}{\Rightarrow}{}^* v$ is said to be the minimal necessary number of applications of the given procedure when proving that $u \stackrel{\beta}{\Rightarrow}{}^* v$ by definition.

Example 1. $G = (\{a, c\} \times \{b, d\}, \{S, A\} \times \{S, B\}, P, (S, S))$ where P is the set of productions:

(1) $\quad (S, S) \to (SA, SB)$;

(2) $\quad (S, S) \to (A, B)$;

(3) $\quad (A, B) \to (a, b)$;

(4) $\quad (A, B) \to (c, d)$;

by the reflexivity of $\stackrel{\beta}{\Rightarrow}{}^*$

(5) $\quad (a, b) \stackrel{\beta}{\Rightarrow}{}^* (a, b)$;

(6) $\quad (c, d) \stackrel{\beta}{\Rightarrow}{}^* (c, d)$;

from (3) and (5)

(7) $\quad (A, B) \stackrel{\beta}{\Rightarrow}{}^* (a, b)$;

from (4) and (6)

(8) $$(SA, SB) \overset{\beta}{\Rightarrow}{}^* (Sc, Sd)\,;$$

from (2) and (7)

(9) $$(Sc, Sd) \overset{\beta}{\Rightarrow}{}^* (ac, bd)\,;$$

from (8) and (9) by transitivity

(10) $$(SA, SB) \overset{\beta}{\Rightarrow}{}^* (ac, bd)\,;$$

from (1) and (10)

(11) $$(S, S) \overset{\beta}{\Rightarrow}{}^* (ac, bd)\,.$$

Obviously 4 is the minimal number of applications of the procedure for getting new members when proving (11), thus the length of (11) is 4.

Definition. Let $G = (V_T^1 \times V_T^2 \times \ldots \times V_T^n, V_N^1 \times V_N^2 \times \ldots \times V_N^n, P, (S_1, S_2, \ldots \ldots, S_n))$, then the relation (set of n-tuples of words)

$$\{w \in V_1^{T*} \times V_T^{2*} \times \ldots \times V_T^{n*} |\, (S_1, S_2, \ldots, S_n) \overset{\alpha}{\Rightarrow}{}^* w\}$$
$$(\{w \in V_T^{1*} \times V_T^{2*} \times \ldots \times V_T^{n*} |\, (S_1, S_2, \ldots, S_n) \overset{\beta}{\Rightarrow}{}^* w\})$$

is said to be α-degenerated (β-generated) by the n-ary grammar G and is denoted by $R^\alpha(G)$ $(R^\beta(G))$.

Example 2. For binary grammar G from example 1 it is $R^\alpha(G) = \{(u, v)|\, u \in \in \{a, c\} \,.\, \{a, c\}^*, v \in \{b, d\} \,.\, \{b, d\}^*$, the number of occurrences of a in u (c in u) is equal to the number of occurrences of b in v (d in v)$\}$;
$R^\beta(G) = \{(x_1 x_2 \ldots x_n, y_1 y_2 \ldots y_n)|\, n \geqq 1, x_i \in \{a, c\}$, if $x_i = a$ then $y_i = b$ else $y_i = d$, $i = 1, 2, \ldots, n\}$.

The β-generation is of principal importance in the application of binary grammars to the description of the translation of languages, as will be shown later. We now show the relation between α-generation and β-generation which will clarify the meaning of the latter. Let $u_0 \overset{\alpha}{\Rightarrow} u_1 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} u_m$ be the derivation of u_m (from u_0). Then there exists (for each $i = 0, 1, 2, \ldots, m - 1$) y_i, z_i such that $u_i = y_i x_i z_i$, $u_{i+1} = y_i w_i z_i$ and $x_i \to w_i$ is in P. The occurrences of symbols in w_i are said to be created in $(i + 1)$-th step of the derivation of u_m. The occurrences of symbols in u_0 are said to be created in the zero step of the derivation. The derivation of u_m is said to satisfy condition β if any production used in the derivation does not replace nonterminal symbols created in different steps of the derivation, i.e. more precisely if in the above notation for each $i = 1, \ldots, m - 1$ there exist j_i, $0 \leq j_i < i$, a_i and b_i, such that $w_{j_i} = a_i x_i b_i$ and $y_{j_i} a_i b_i z_{j_i} \overset{\alpha}{\Rightarrow}{}^* y_i z_i$.

Theorem 1. *Let G be an n-ary grammar,* $G = (V_T^1 \times V_T^2 \times \ldots \times V_T^n, V_N^1 \times V_N^2 \times \times \ldots \times V_N^n, P, (S_1, S_2, \ldots, S_n))$. *Let u be in* $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$, *v be in* $V_N^{1*} \times \times V_T^{2*} \times \ldots \times V_T^{n*}$, *then* $u \overset{\beta}{\Rightarrow}{}^* v$ *iff there exists a derivation of v fron m satisfying condition* β.

Proof. 1. Let us assume that the derivation $u_0 \overset{\alpha}{\Rightarrow} u_1 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} u_m$ satisfies condition β. We prove by induction on the length of the derivation that $u_0 \overset{\beta}{\Rightarrow}{}^* u_m$. If the derivation is of length 1 then trivially $u_0 \overset{\beta}{\Rightarrow}{}^* u_1$. Otherwise let the assertion hold for all derivations of length m. As $u_0 \overset{\alpha}{\Rightarrow} u_1$ then there exist x, z, w in $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$ and y in $V_N^{1*} \times V_N^{2*} \times \ldots \times V_N^{n*}$ such that $u_0 = xyz$, $u_1 = xwz$ and $y \to w$ is in P. The derivation $u_1 \overset{\alpha}{\Rightarrow} u_2 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} u_m$ satisfies condition β and thus by the induction assumption it follows that $u_1 \overset{\beta}{\Rightarrow}{}^* u_m$. From condition β it follows further that there exist r, s, t in $V_T^{1*} \times V_T^{2*} \times \ldots \times V_T^{n*}$ such that $u_m = rst$ and $w \overset{\beta}{\Rightarrow}{}^* s$. By the definition of $\overset{\beta}{\Rightarrow}{}^*$ it follows that $u_0 \overset{\beta}{\Rightarrow}{}^* u_1$ and thus $u_0 \overset{\beta}{\Rightarrow}{}^* u_m$.

2. Let u be in $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$, v be in $V_T^{1*} \times V_T^{2*} \times \ldots \times V_T^{n*}$, $u \overset{\beta}{\Rightarrow}{}^* v$. We prove by induction on the length of $u \overset{\beta}{\Rightarrow}{}^* v$ that there exists the derivation of v from u satisfying the condition β. If the length of $u \overset{\beta}{\Rightarrow}{}^* v$ is one then $u \overset{\alpha}{\Rightarrow} v$ and condition β is obviously fulfilled. Otherwise, let the assertion hold when the length is $m - 1$. Let $u \overset{\beta}{\Rightarrow}{}^* v$ be of length m. Then there exist s, t, y, z in $V^{1*} \times V^{2*} \times \ldots \times \times V^{n*}$, x in $V_T^{1*} \times V_T^{2*} \times \ldots \times V_T^{n*}$ such that $u = ysz$, $w = yxz$, $s \to t$ in P and $t \overset{\beta}{\Rightarrow}{}^* x$, $t \overset{\beta}{\Rightarrow}{}^* x$ and $w \overset{\beta}{\Rightarrow}{}^* v$ are both of length less than m. By the inductive assumption there exist derivations satisfying condition β:

$$t = x_0 \overset{\alpha}{\Rightarrow} x_1 \overset{\alpha}{\Rightarrow} x_2 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} x_p = x$$

and

$$w = v_0 \overset{\alpha}{\Rightarrow} v_1 \overset{\alpha}{\Rightarrow} v_2 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} v_q = v\,.$$

Thus $U = ysz \overset{\alpha}{\Rightarrow} ytz \overset{\alpha}{\Rightarrow} yx_1z \overset{\alpha}{\Rightarrow} yx_2z \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} yxz \overset{\alpha}{\Rightarrow} v_1 \overset{\alpha}{\Rightarrow} v_2 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} v_q = v$ is the derivation of v from u satisfying condition β.

For the phrase structure (unary) grammars obviously $\overset{\alpha}{\Rightarrow}$ is the usual $\Rightarrow$. In the case of a context-free grammar only one nonterminal symbol is substituted in each step of a derivation and therefore condition β is always satisfied. We have the following:

Corollary 1. *Let G be a context-free grammar, then* $R^\beta(G) = R^\alpha(G) = L(G)$.

Example 3. We give an example of a context-sensitive grammar G for which $R^\beta(G) \neq R^\alpha(G)$.

$$G = (\{a, b, c, d\}, \quad \{S, A, B, C\}, \quad \{S \to ABC, AB \to aB, BC \to bC, A \to d, C \to c\}, S).$$

Obviously, $R^\alpha(G) = \{abc, dbc\}$, $R^\beta(G) = \{dbc\}$.

Definition. Let G be an n-ary grammar $(V_T^1 \times V_T^2 \times \ldots \times V_T^n, V_N^1 \times V_N^2 \times \ldots \times V_N^n, P, (S_1, S_2, \ldots, S_n))$. Then the (unary) grammars $G_i = (V_T^i, V_N^i, P_i, S_i)$ where $P_i = \{q_i \mid (q_1, q_2, \ldots, q_i, \ldots, q_n) \in P, \ q_i \neq \emptyset\}$ for $i = 1, 2, \ldots, n$ are said to be the partial grammars of the n-ary grammar G.

Theorem 2. *Let G be an arbitrary n-ary grammar then*

$$R^\beta(G) \subset R^\alpha(G) \subset L(G_1) \times L(G_2) \times \ldots \times L(G_n),$$

where $G_1, G_2, \ldots, G_n$ are the partial grammars of the n-ary grammar G. There exists a binary grammar for which both inclusions are proper.

Proof. 1. Validity of the first inclusion follows immediately from the theorem 1.

2. Let $u^1, u^2, \ldots, u^t$ be a derivation of the n-tuple u^t in $R^\alpha(G)$. Let $u^k = (u_1^k, u_2^k, \ldots, u_n^k)$ for $k = 1, 2, \ldots, t$ (particularly $u^1 = (S_1, S_2, \ldots, S_n)$). Then leaving out the repeating occurrences of the same members in the sequence $u_i^1, u_i^2, \ldots, u_i^t$ we get the derivation of u_i^t in the grammar G_i. Thus u_i^t is in $L(G_i)$ and the second inclusion is valid.

3. We give an example of a binary grammar for which both inclusions are proper.

Let $G = (\{a, b\} \times \{a, b\}, \{S_1, A, B\} \times \{S_2, A\}, \{(S_1 \to A, S_2 \to A), (A \to aA, A \to aA), (A \to a, A \to a), (S_1 \to B, \emptyset), (B \to b, S_2 \to b)\}, (S_1, S_2))$.
It is obvious that $R^\beta(G) = \{(a^n, a^n) \mid n = 1, 2, \ldots\}$,
$R^\alpha(G) = \{(a^n, a^n) \mid n = 1, 2, \ldots\} \cup \{(b, b)\}$,
$L(G_1) \times L(G_2) = \{(a^m, a^n), (a^m b), (b, a^n), (b, b) \mid m, n = 1, 2, \ldots\}$.

Definition. The n-ary grammar G is said to be of type $(i_1, i_2, \ldots, i_n)$ if its partial grammars $G_1, G_2, \ldots, G_n$ are of types $i_1, i_2, \ldots, i_n$ respectively. The class of n-ary grammars of type $(i_1, i_2, \ldots, i_n)$ is denoted by $G_{i_1, i_2, \ldots, i_n}$. The class of relations α-generated (β-generated) by the n-ary grammars of type $(i_1, i_2, \ldots, i_n)$ is denoted by $\mathbf{R}^\alpha_{i_1, i_2, \ldots, i_n}$ ($\mathbf{R}^\beta_{i_1, i_2, \ldots, i_n}$).

Theorem 3. *The classes $\mathbf{R}^\alpha_{2,2,\ldots,2}$ and $\mathbf{R}^\alpha_{1,1,\ldots,1}$ are closed under union and concatenation.*

Proof. Let R_1, R_2 be in $\mathbf{R}^{\alpha}_{2,2,\ldots,2}$ (in $\mathbf{R}^{\alpha}_{1,1,\ldots,1}$), $R_1 = R^{\alpha}(G_1)$, $R_2 = R^{\alpha}(G_2)$;
G_1, G_2 in $\mathbf{G}_{2,2,\ldots,2}(\mathbf{G}_{1,1,\ldots,1})$;

$$G_1 = (V_T^1 \times V_T^2 \times \ldots \times V_T^n,\ V_N^1 \times V_N^2 \times \ldots \times V_N^n,\ P_1, (S_1^1, S_1^2, \ldots, S_1^n))\,,$$

$$G_2 = (W_T^1 \times W_T^2 \times \ldots \times W_T^n, W_N^1 \times W_N^2 \times \ldots \times W_N^n, P_2, (S_2^1, S_2^2, \ldots, S_2^n))\,.$$

Let $V_N^i \cap W_N^i = \emptyset$ and S^i not be in $V_N^i \cup W_N^i$ for $i = 1, 2, \ldots, n$.

1. We construct the n-ary grammar G_3, $G_3 = (V_T^1 \cup W_T^1 \times V_T^2 \cup W_T^2 \times \ldots \times V_T^n \cup W_T^n, V_N^1 \cup W_N^1 \cup \{S^1\} \times V_N^2 \cup W_N^2 \cup \{S^2\} \times \ldots \times V_N^n \cup W_N^n \cup \{S^n\}, P_1 \cup P_2 \cup \{(S^1 \to S_1^1, S^2 \to S_1^2, \ldots, S^n \to S_1^n), (S^1 \to S_2^1, S_2 \to S_2^2, \ldots, S^n \to S_2^n), (S^1, S^2, \ldots, S^n)\})$.

If G_1, G_2 are in $\mathbf{G}_{2,2,\ldots,2}$ (in $\mathbf{G}_{1,1,\ldots,1}$) then G_3 is in $\mathbf{G}_{2,2,\ldots,2}$ (in $\mathbf{G}_{1,1,\ldots,1}$). It is obvious that $R^{\alpha}(G_3) = R_1 \cup R_2$, thus both $\mathbf{R}^{\alpha}_{2,2,\ldots,2}$ and $\mathbf{R}^{\alpha}_{1,1,\ldots,1}$ are closed under union.

2. We construct the n-ary grammar G_4, $G_4 = (V_T^1 \cup W_T^1 \times V_T^2 \cup W_T^2 \times \ldots \times V_T^n \cup W_T^n, V_N^1 \cup W_N^1 \{S^1\} \times V_N^2 \cup W_N^2 \cup \{S^2\} \times \ldots \times V_N^n \cup W_N^n \cup \{S^n\}, P_1 \cup P_2 \cup \{(S^1 \to S_1^1 S_2^1,\ S^2 \to S_1^2 S_2^2, \ldots S^n \to S_1^n S_2^n)\},\ (S^1, S^2, \ldots, S^n))$. If G_1, G_2 are in $\mathbf{G}_{2,2,\ldots,2}$ (in $\mathbf{G}_{1,1,\ldots,1}$) then G_4 is in $\mathbf{G}_{2,2,\ldots,2}$ (in $\mathbf{G}_{1,1,\ldots,1}$).

It is obvious that $R^{\alpha}(G_4) = R_1 R_2$, and thus both $\mathbf{R}^{\alpha}_{2,2,\ldots,2}$ and $\mathbf{R}^{\alpha}_{1,\ldots,1}$ are closed under concatenation.

Theorem 4. *The classes $\mathbf{R}^{\beta}_{2,2,\ldots,2}$ and $\mathbf{R}^{\beta}_{1,1,\ldots,1}$ are closed under union, concatenation and the class $\mathbf{R}^{\beta}_{2,2,\ldots,2}$ under iteration.*

Proof. Let R_1, R_2 be in $\mathbf{R}^{\beta}_{2,2,\ldots,2}$ ($\mathbf{R}^{\beta}_{1,1,\ldots,1}$), $R_1 = R^{\beta}(G_1)$, $R_2 = R^{\beta}(G_2)$; G_1, G_2 in $\mathbf{G}_{2,2,\ldots,2}$ (in $\mathbf{G}_{1,1,\ldots,1}$);

$$G_1 = (V_T^1 \times V_F^2 \times \ldots \times V_T^n,\ V_N^1 \times V_N^2 \times \ldots \times V_N^n,\ P_1, (S_1^1, S_1^2, \ldots, S_1^n))\,,$$

$$G_2 = (W_T^1 \times W_T^2 \times \ldots \times W_T^n, W_N^1 \times W_N^2 \times \ldots \times W_N^n, P_2, (S_2^1, S_2^2, \ldots, S_2^n))\,.$$

Let $V_N^i \cap W_N^i = \emptyset$ and S^i, Q^i not be in $V_N^i \cup W_N^i$ for $i = 1, 2, \ldots, n$.

1. We construct the grammars G_3, G_4 in the same way as in the proof of theorem 3. If the derivations of u in grammar G_1 and v in G_2 fulfil condition (1) then the derivations of u, v in G_3 and of uv in G_4 also fulfil condition (1) and therefore $R^{\beta}(G_3) = R_1 \cup R_2$, $R^{\beta}(G_4) = R_1 R_2$. Thus the classes $\mathbf{R}^{\beta}_{2,2,\ldots,2}$ and $\mathbf{R}^{\beta}_{1,1,\ldots,1}$ are closed under union and concatenation.

2. For any R_1 we construct the n-ary grammar G_5,
$G_5 = (V_T^1 \times V_T^2 \times \ldots \times V_T^n, (V_N^1 \cup \{S^1\}) \times (V_N^2 \cup \{S^2\}) \times \ldots \times (V_N^n \cup \{S^n\}), P_1 \cup \{(S^1 \to \varepsilon, S^2 \to \varepsilon, \ldots, S^n \to \varepsilon), (S^1 \to S_1^1 S^1, S^2 \to S_1^2 S^2, \ldots, S^n \to S_1^n S^n)\}, (S^1, S^2, \ldots, S^n))$. We shall show that $R^{\beta}(G_5) = R_1^*$. Let $(S_1^1, S_1^2, \ldots, S_1^n)$ be denoted by S_1 and $(S^1, S^2, \ldots, S^n)$ by S.

a) Let u be in R_1^*, then there exists a sequence of n-tuples in $V_N^{1*} \times V_N^{2*} \times \ldots \times V_T^{n*}$ $u_1, u_2, \ldots, u_m$ ($m \geqq 0$) such that u_k is in $R^\beta(G_1)$ for $k = 1, 2, \ldots, m$ and $u = u_1, u_2, \ldots, u_m$ (in the case $m = 0$, $u = \varepsilon$). Then $S_1 \overset{\beta}{\Rightarrow}{}^* u^k$ and using it and the production $S \to S_1S$ we have $S \overset{\beta}{\Rightarrow}{}^* u_1\ S \overset{\beta}{\Rightarrow}{}^* u_1u_2\ S \overset{\beta}{\Rightarrow}{}^* \ldots \overset{\beta}{\Rightarrow}{}^* u_1u_2 \ldots u_mS$ and finally using the production $S \to (\varepsilon, \varepsilon, \ldots, \varepsilon)$ we have $S \overset{\beta}{\Rightarrow}{}^* u_1u_2 \ldots u_m = u$. Thus $R^* \subset R^\beta(G_5)$.

b) Let u be in $R^\beta(G_5)$. Then there exists a derivation $S \overset{\alpha}{\Rightarrow}{}^* u$, which fulfils condition (1). The fulfilling of condition (1) enables us to rearrange the derivation of u as $S \overset{\alpha}{\Rightarrow} S_1S \overset{\alpha}{\Rightarrow}{}^* u_1S \overset{\alpha}{\Rightarrow} u_1S_1S \overset{\alpha}{\Rightarrow}{}^* u_1u_2\ S \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} u_1u_2 \ldots u_m\ S \overset{\alpha}{\Rightarrow} u_1u_2 \ldots u_m$. The steps of the derivation show that $S_1 \overset{\alpha}{\Rightarrow}{}^* u_k$ for $k = 1, 2, \ldots, m$. Since they also fulfil condition (1) $S_1 \overset{\beta}{\Rightarrow}{}^* u_k$. Thus u_k is in R_1 and as $u = u_1u_2 \ldots u_k$, u is in R_1^* and $R^\beta(G_5) \subset R_1^*$.

Note. The question of whether the class $\mathbf{R}^\alpha_{2,2,\ldots,2}$ is closed under iteration is open.

Definition. A binary grammar G is called linear if G is in $\mathbf{G}_{2,2}$ and both its partial grammars G_1, G_2 are linear (see [4]).

Example 4. Let G be an arbitrary linear binary grammar. We construct the binary grammar G' such that $(R^\alpha(G))^*$ is in $\mathbf{R}^\alpha_{2,2}$. The construction is not as simple as that for the β-generation but nevertheless G' exists for any linear grammar.

Let $G = (V_T^1 \times V_T^2, V_N^1 \times V_N^2, P, (S_1, S_2))$. Assuming X, Y, Z is not in $V_N^1 \cup V_N^2$ we construct $G' = (V_T^1 \times V_T^2, (V_N^1 \cup \{X, Y, Z\}) \times (V_N^2 \cup \{X, Y, Z\}), P' \cup \{(Z \to S_1Y, Z \to S_2Y), (Z \to \varepsilon, Z \to \varepsilon), (X \to \varepsilon, Y \to Z), (Y \to Z, X \to \varepsilon), (Z, Z))$ where P' is created from P in such a way that each unary production, involved in an n-ary production in P, of the form $A \to w$, $w \in V_T^{i*}$ (with terminal right side) is rewritten as $A \to wX$ (nonterminal X is added from the right); e.g. if $(A \to aB, C \to ab)$ is in P then $(A \to aB, C \to abX)$ is in P'. It is not difficult to show that $R^\alpha(G') = (R^\alpha(G))^*$.

Further, we shall formulate and prove some results only for binary grammars but their generalization for n-ary grammars is mostly only a formal matter.

Theorem 5. *If G is a linear binary grammar, then there exists a linear binary grammar G' such that $R^\beta(G') = R^\alpha(G)$ and a linear binary grammar G'' such that $R^\alpha(G'') = R^\beta(G)$. Moreover if G is in $\mathbf{G}_{3,3}$ then also both G', G'' are in $\mathbf{G}_{3,3}$.*

Proof. Let $G = (V_T^1 \cup V_T^2, V_N^1 \cup V_N^2, P, (S_1, S_2))$.

1. We construct the linear binary grammar G', $G' = (V_T^1 \times V_T^2, V_N^1 \times V_N^2, P \cup P', (S_1, S_2))$, where the set P' is created as follows:

(i) if $(q_1, \emptyset)$ is in P and A is in V_N^2 then $(q_1, A \to A)$ is in P';

(ii) if $(\emptyset, q_2)$ is in P and A is in V_N^1 then $(A \to A, q_2)$ is in P'.

The productions from P' do not increase the α-generative power of G, thus $R^\alpha(G') = R^\alpha(G)$. However, substituting the productions from P' for some productions of the type $(q_1, \emptyset)$ or $(\emptyset, q_2)$, each derivation $(S_1, S_2) \overset{\alpha}{\Rightarrow}{}^* (w_1, w_2)$ in the linear grammar G, (w_1 in V_T^1, w_2 in V_T^2) can obviously be rewritten in such a way that it fulfils condition β and by

Theorem 1 $(S_1, S_2) \overset{\beta}{\Rightarrow}{}^* (w_1, w_2)$. Thus $R^\alpha(G') \subset R^\beta(G')$. By
Theorem 2 $R^\beta(G') \subset R^\alpha(G')$ and thus $R^\alpha(G) = R^\beta(G')$.

2. We construct the linear binary grammar G'', $G'' = (V_T^1 \times V_T^2, (V_N^1 \cup W_N^1) \times (V_N^2 \cup W_N^2), P'', (S_1, S_2))$, where $W_N^1 = \{A' \mid A \text{ in } V_N^1\}$, $W_N^2 = A' \mid A \text{ in } V_N^2\}$ (we assume $W_N^1 \cap V_N^1 = \emptyset$ and $W_N^2 \cap V_N^2 = \emptyset$) and P'' is created from P as follows:

(i) each production in P of the form $(A \to t_1Bt_2, \emptyset)$ where A, B are in V_N^1 and t_1, t_2 are in V_T^{1*}, is replaced by two productions $(A \to t_1B't_2, \emptyset)$ and $(A' \to t_1B't_2, \emptyset)$;

(ii) each production of the form $(\emptyset, A \to t_1Bt_2)$, where A, B are in V_N^2, t_1, t_2 in V_T^{2*} is replaced by two productions $(\emptyset, A \to t_1B't_2)$, $(\emptyset, A' \to t_1B't_2)$;

(iii) for each production of the form $(A \to t, \emptyset)$, where A is in V_N^1, t in V_N^{1*} another production $(A' \to t, \emptyset)$ is added;

(iv) for each production of the form $(\emptyset, A \to t)$, where A is in V_N^2, t in V_T^{2*} another production $(\emptyset, A' \to t)$ is added.

Obviously $R^\beta(G'') \subset R^\beta(G)$. If $(S_1, S_2) \overset{\beta}{\Rightarrow}{}^* (w_1, w_2)$ in G then there exists a derivation $(S_1, S_2) = u_1 \overset{\alpha}{\Rightarrow} u_2 \overset{\alpha}{\Rightarrow} \ldots \overset{\alpha}{\Rightarrow} u_n = (w_1, w_2)$ satisfying condition β. It is possible only if for some r, $1 \leqq r \leqq n$ the derivation $u_1 \overset{\alpha}{\Rightarrow}{}^* u_r$ uses only productions of the type (q_1, q_2), $q_1 \neq \emptyset$, $q_2 \neq \emptyset$, and the derivation $u_r \overset{\alpha}{\Rightarrow}{}^* u_n$ uses only productions of the type $(q_1, \emptyset)$ or $(\emptyset, q_2)$. Let u be in $V^1 \times V^2$ then the pair obtained from u by replacing all occurrences of each A in $V_N^1 \cup V_N^2$ by the corresponding A' in $W_N^1 \cup W_N^2$ is denoted by u'. Then obviously $u_r \overset{\alpha}{\Rightarrow} u'_{r+1} \overset{\alpha}{\Rightarrow} u'_{r+2} \overset{\alpha}{\Rightarrow} \ldots u'_{n-1} \overset{\alpha}{\Rightarrow} u_n$ and therefore $u_1 \Rightarrow^* u_n$ is a derivation in G'' satisfying condition β. Thus $R^\beta(G) \subset R^\beta(G'')$ and therefore $R^\beta(G) = R^\beta(G'')$. Since no productions of the form (q_1, q_2), $q_1 \neq \emptyset$, $q_2 \neq \emptyset$ contain a symbol from $W_N^1 \cup W_N^2$, no production of the type $(q_1, \emptyset)$ or $(\emptyset, q_2)$ has a symbol from $V_N^1 \cup V_N^2$ on the right side and the grammar is linear (the number of nonterminals cannot increase), each derivation in G'' fulfils condition β and $R^\alpha(G'') = R^\alpha(G)$. Thus $R^\alpha(G'') = R^\beta(G)$.

Corollary 2. $R_{3,3}^\alpha = R_{3,3}^\beta$.

Example 5. For an individual grammar in $\mathbf{G}_{3,3}$ the generated relations need not be equal. For instance, let $G = (\{a, c\} \times \{b\}, \{S_1, C\} \times \{S_2\}, \{(S_1 \to a, S_2 \to b), (S_1 \to cC, \emptyset), (C \to cS_1, \emptyset)\}, (S_1, S_2))$. Obviously, $R^\alpha(G) = \{(c^{2n}a, b) \mid n = 1, 2, \ldots$ but $R^\beta(G) = \{(a, b)\}$.

In [3] the notation of transduction is introduced. Transductions are n-ary relations, which are definable by means of the n-tape finite automata (NDA in [3]). It is shown there that the class of transductions is equal to the class of regular relations, i.e. relations obtainable from finite relations by means of a finite number of the Kleene's operations. Now we show, that the class of transductions is equal to $\mathbf{R}_{3,3}$.

Definition. Let $\mathbf{M}_n$ denote the class of n-tape automata over an alphabet (NDA in [3]). A in $\mathbf{M}_n$ is the system (S, v, s_0, F), where S is a finite set (of states); v is a finite set $v \subset S \times (\Sigma^*)^n \times S$ (multivalued next-state function); s_0 in S is the initial state and $F \subset S$ is a set of terminal states. The n-tuple of words u in $(\Sigma^*)^n$ is said to be accepted by the automaton A if there exists a sequence of states $s_0, s_1, \ldots, s_t$ such that

(i) $(s_{i-1}, u_i, s_i) \in v$ for $i = 1, 2, \ldots, t$;
(ii) $u = u_1 u_2 \ldots u_t$;
(iii) $s_t \in F$.

Let the set of n-tuples of words (relations) accepted by the automaton A be denoted by $R(A)$. The class of relations defined by automata in $\mathbf{M}_n$ is denoted by $R(\mathbf{M}_n)$; $R(\mathbf{M}_n) = \{R(A) \mid A \in \mathbf{M}_n\}$. In [3] such relations are called transductions.

Theorem 6. $\mathbf{R}^\alpha_{3,3,\ldots,3} = R(\mathbf{M}_n)$. *

Proof. 1. Let G be in $\mathbf{G}_{3,3,\ldots,3}$, $G = (V_T^1, V_N^1 \times \ldots \times V_T^n \times \ldots \times V_N^n, P, (S_1, \ldots, S_n))$. We shall construct the automaton A in $\mathbf{M}_n$, such that $R(A) = R^\alpha(G)$.

$A = ((V_N^1 \cup \{e\}) \times \ldots \times (V_N^n \cup \{e\}), v, (S_1, \ldots, S_n), (e, e, \ldots, e))$, where e is the empty symbol and v is defined as follows:

Let $A_i, B_i \in V_N^i \cup \{e\}$, $v_i \in V_T^{i*}$, then $((A_1, A_2, \ldots, A_n), (v_1, v_2, \ldots, v_n), (B_1, B_2, \ldots \ldots, B_n))$ is in v if and only if there exists q in P such that $q = (q_1, q_2, \ldots, q_n)$ where one of the following conditions holds for all $i = 1, 2, \ldots$:

(i) $A_i, B_1 \in V_N^i$, $A_i \to v_i B_i$;
(ii) $A_i \in V_N$, $B_i = e$, $A_i \to v_i$;
(iii) $A_i, B_i \in V_N \cup \{e\}$, $A_i = B_i$, $q_i = \emptyset$.

We need to prove that $R(A) = R^\alpha(G)$. According to the definition of the n-tuples accepted by the automaton, u belongs to $R(A)$ iff there exists a sequence of states $s_0, \ldots, s_t \in (V_N^1 \cup \{e\}) \times \ldots \times (V_n^1 \cup \{e\})$ such that the following conditions hold:

(i) $(s_{i-1}, u_i, s_i) \in v$ for $i = 1, 2, \ldots, t$;
(ii) $u = u_1 u_2 \ldots u_t$:
(iii) $s_t = (e, e, \ldots, e)$.

* An equivalent theorem was also proved independently by J. Král (On Multiple Grammars. Kybernetika 5 (1969), 1, 60—85).

The sequence of states $s_0, s_1, \ldots, s_t$ fulfils conditions (i), (ii) iff $s_{i-1} \overset{\alpha}{\Rightarrow} u_i s_i$ for $i = 1, 2, \ldots, t$, therefore $s_0 = (S_1, \ldots, S_n) \overset{\alpha}{\Rightarrow}{}^* u_1 \ldots u_t = u$. Thus u is in $R(A)$ iff u is in $R^\alpha(G)$.

2. Let A in $\boldsymbol{M}_n$ be an automaton over the alphabet Σ, $A = (S, v, s_0, F)$. We construct the n-ary grammar $G = (\underbrace{\Sigma \times \Sigma \ldots \times \Sigma}_{n\text{-times}}, \underbrace{S \times \ldots \times S}_{n\text{-times}}, P, (\underbrace{s_0, \ldots, s_0}_{n\text{-times}}))$ where P is chosen as follows:

(i) if $(s, (v_1, v_2, \ldots, v_n), s')$ is in v, $s \neq s'$, then $(s \to v_1 s', s \to v_2 s', \ldots, s \to v_n s') \in P$ and if in addition $s' \in F$ then also $(s \to v_1, s \to v_2, \ldots, s_n \to v_n) \in P$.

(ii) if $(s, (v_1, v_2, \ldots, v_n), s) \in v$ then $q \in P$ where $q = (q_1, q_2, \ldots, q_n)$, $q_i = \emptyset$ for $v_i = e$, $q_i = s \to v_i s$ for $v_i \neq e$. If in addition $s \in F$ then also $q' \in P$, $q' = (q'_1, q'_2, \ldots \ldots, q'_n)$ where $q'_i = \emptyset$ for $v_i = e$, $q_i = s \to v_i$ for $v_i \neq e$.

It is obvious that $R^\alpha(G) = R(A)$.

From Theorem 6 and the results in [3] it immediately follows:

Corollary 3. *The class $\boldsymbol{R}^\alpha_{3,3,\ldots,3}$ $(= \boldsymbol{R}^\beta_{3,3,\ldots,3})$ is closed under Kleene's operations, i.e. under union, concatenation and iteration. It is closed under intersection and complementation.*

Corollary 4. *The class $\{\text{dom } R \mid R \in \boldsymbol{R}_{3,3}\}$ is the class of regular sets.*

Note. Let us consider the class of relations α-generated by n-ary grammars in $\boldsymbol{G}_{3,3,\ldots,3}$ satisfying the condition: if q is in P, $q = (q_1, q_2, \ldots, q_n)$ then either q_i is of the form $A \to uB$ for all $i = 1, 2, \ldots, n$ (A, B in V^i_N, u in V^{i*}_N) or q_i is of the form $A \to u$ for all $i = 1, 2, \ldots, n$(A, B in V^i_N, u in V^{i*}_T). It is not difficult to show that this class is equal to the class $R(\boldsymbol{S}_n)$ from [2].

Unlike the regular sets, we have the following result for context-free languages.

Theorem 7. *The class of context-free languages ($\boldsymbol{L}_{CF}$) is a proper subset of the class $\{\text{dom } R \mid R \text{ in } \boldsymbol{R}^\alpha_{2,2}\}$.*

Proof. Obviously $\boldsymbol{L}_{CF} \subset \{\text{dom } R \mid R \text{ in } \boldsymbol{R}^\alpha_{2,2}\}$. The following example shows that this inclusion is proper.

Let $G = (\{a, b\} \times \{b\}, \{S_1, B\} \times \{S_2\}, \{(S_1 \to aS_1a, S_2 \to S_2S_2), (S_1 \to B, S_2 \to b), (B \to bB, S_2 \to b), (B \to b, S_2 \to b)\}, (S_1, S_2))$.

In the derivation of a pair in $R^\alpha(G)$ the productions must obviously occur in the following order and number:

1. n applications of the production $(S_1 \to aS_1a, S_2 \to S_2S_2)$, $n = 1, 2, \ldots$ gives $(a^n S_1 a^n, S_2^{n+1})$;
2. One application of the production $(S_1 \to B, S_2 \to b)$ gives $(a^n B a^n, b S_2^n)$;

3. $n-1$ applications of the production $(B \to bB, S_2 \to b)$ gives $(a^n b^{n-1} Ba^n, b^n S_2)$;

4. One application of the production $(B \to b, S_2 \to b)$, gives $(a^n b^n a^n, b^{n+1})$. The symbols of the second word in the pair may be arbitrarily permutated after step 2 or step 3, but it does not change the terminal result.

Thus $\{a^n b^n a^n \mid n = 1, 2, \ldots\}$ is in $\{\text{dom } R \mid R \text{ in } \mathbf{R}^{\alpha}_{2,2}\}$ but it is well known [4] that $\{a^n b^n a^n \mid n = 1, 2, \ldots\}$ is not in $\mathbf{L}_{\mathrm{CF}}$.

Theorem 8. $\{\text{dom } R \mid R \text{ in } \mathbf{R}^{\beta}_{2,2}\} = \mathbf{L}_{\mathrm{CF}}$.

Proof. Obviously $\mathbf{L}_{\mathrm{CF}} \subset \{\text{dom } R \mid R \text{ in } \mathbf{R}^{\beta}_{2,2}$. To prove the reverse inclusion we must show that for every binary grammar G in $\mathbf{G}_{2,2}$ $G = (V^1_T \times V^2_T, V^1_N \times V^2_N, P, (S_1, S_2))$ the dom $R^{\beta}(G)$ is in $\mathbf{L}_{\mathrm{CF}}$. We construct the context-free grammar G' such that $L(G') = \text{dom } R^{\beta}(G)$.

Let $G' = (V^1_T, W_N, P', (S_1, S_2))$ where $W_N = (V^1_N \cup \{\emptyset\}) \times (V^2_N \cup \{\emptyset\})$, P' is defined in the following way:

(i) If $(A_1 \to u_1, A_2 \to u_2)$ is in P, $u_1 = v_0 B_1 v_1 B_2 \ldots B_m v_m$, $u_2 = w_0 C_1 w_1 C_2 \ldots \ldots C_n w_n$ where $A_1, B_1, \ldots, B_m$ is in V^1_N, $A_2, C_1, C_2, \ldots, C_n$ is in V^2_N; $v_0, v_1, \ldots, v_m$ is in V^{1*}_T; $w_0, w_1, \ldots, w_n$ is in V^{2*}_T then P' contains all productions of the form $(A_1, A_2) \to \to v_0(B_1, D_1)\, v_1(B_2, D_2) \ldots (B_m, D_m)\, v_m\, (\emptyset, D_{m+1})\, (\emptyset, D_{m+2}) \ldots (\emptyset, D_p)$ where $m \leqq \leqq p \leqq m + n$, D_i is in $V^2_N \cup \{\emptyset\}$ for $1 \leqq i \leqq m$ and in V^2_N for $m + 1 \leqq i \leqq p$; and $D_1, D_2, \ldots, D_p$ is a permutation of $C_1, C_2, \ldots, C_n, \underbrace{\emptyset, \emptyset, \ldots, \emptyset}_{(p-n)\text{-times}}$.

(ii) If $(A_1 \to u_1, \emptyset)$ is in P, $u_1 = v_0 B_1 v_1 B_2 \ldots B_m v_m$ then $(A_1, \emptyset) \to v_0(B_1, \emptyset) \times \times v_1 (B_2, \emptyset) \ldots (B_m, \emptyset)\, v_m$ is in P'.

(iii) If $(\emptyset, A_2 \to u_2)$ is in P, $u_2 = w_0 C_1 w_1 C_2 \ldots C_n w_n$ then $(\emptyset, A_2) \to (\emptyset, C_1) \,.\, .\, (\emptyset, C_2) \ldots (\emptyset, C_n)$ is in P' for $n \geqq 1$, $(\emptyset, A_2) \to \varepsilon$ is in P' for $n = 0$ (u_2 in V^{2*}_T).

We will prove now that $L(G') = \text{dom } R^{\beta}(G)$. By $\overset{\beta}{\Rightarrow}{}^*$ the derivation in grammar G is understood in the following. A derivation in grammar G' is denoted by $\overset{G'}{\Rightarrow}{}^*$. Let μ be a homomorphism from W^*_N into V^1_N defined as follows: $\mu(\varepsilon) = \varepsilon$, $\mu((A, B)) = = A$ if $A = \emptyset$, $\mu((\emptyset, B)) = \varepsilon$ for each A in V^1_N and B in V^2_N, $\mu(uv) = \mu(u)\, \mu(v)$ for each u, v in W^*.

1. We prove by induction on the length of the derivation $u \overset{\beta}{\Rightarrow}{}^* v$ that if $u \overset{\beta}{\Rightarrow}{}^* v$, where $u = (u_1, u_2)$, $v = (v_1, v_2)$, u_1 in V^{1*}, u_2 in V^{2*}, v_1 in V^{1*}_T, v_2 in V^{2*}_T, then there exists $\bar{u}_1$ in W_N such that $\mu(\bar{u}_1) = u_1$ and $\bar{u}_1 \overset{G'}{\Rightarrow} v_1$.

Suppose that the inductive hypothesis is true for all derivations of length $n - 1$ or less. Since $u \overset{\beta}{\Rightarrow}{}^* v$ then by definition of $\overset{\beta}{\Rightarrow}{}^*$ there exist $s \to t$ in P, x in $V^1_T \times V^2_T$ such that $u = ysz$, $v = yxz$ and $t \overset{\beta}{\Rightarrow}{}^* x$. Because of the inductive assumption there exists $\bar{t}_1$ such that $\mu(\bar{t}_1) = t_1$, $s = (s_1, s_2)$, $t = (t_1, t_2)$, $x = (x_1, x_2)$ and $\bar{t}_1 \overset{G'}{\Rightarrow}{}^* x_1$. According to the construction of G' $\bar{s}_1 \to \bar{t}_1$ is in P' and letting $y = (y_1, y_2)$, $z =$

$= (z_1, z_2)$ we have $y_1\bar{s}_1z_1 \overset{G'}{\Rightarrow} y_1\bar{t}_1z_1$. Thus $\bar{u}_1 \overset{G'}{\Rightarrow} v_1$ where $\bar{u}_1 = y_1\bar{s}_1z_1, (\bar{u}_1) = y_1s_1z_1$ and $u = (y_1s_1z_1, y_2s_2z_2)$. Thus $R^\beta(G) \subset L(G')$.

2. We prove by induction on the length of derivation that for $\bar{u}_1$ in W_N, v_1 in V_T^1, $\bar{u}_1 \overset{G'}{\Rightarrow}{}^* v_1$ there exist u_2 in W_N and v_2 in V_T^2 such that $(u_1, u_2) \overset{\beta}{\Rightarrow}{}^* (v_1, v_2)$ where $u_1 = \mu(\bar{u}_1)$. For a derivation of length one there exists a production $\bar{s}_1 \to x_1$ in P' such that $\bar{u}_1 = y_1\bar{s}_1z_1$, $v_1 = y_1x_1z_1$ and y_1, z_1 are in V_T^1. By the construction of P' there exists a production $(s_1, s_2) \to (x_1, x_2)$ in P such that $\mu(\bar{s}_1) = s_1$. Thus the inductive hypothesis is true for $n = 1$. Suppose that it holds for all derivations of length $n - 1$ or less. Since $\bar{u}_1 \overset{G'}{\Rightarrow}{}^* v_1$ there exist x_1, y_1, z_1 in V_T^1, $\bar{A}_1$ in W_N such that $\bar{u}_1 = y_1\bar{A}_1z_1$, $v_1 = y_1x_1z_1$, $\bar{A}_1 \to \bar{t}_1$ in P' and $\bar{t}_1 \overset{G'}{\Rightarrow}{}^* x_1$. By the inductive hypothesis there exist $t = (t_1, t_2)$, $x = (x_1, x_2)$ such that $\mu(\bar{t}_1) = t_1$ and $(t_1, t_2) \Rightarrow^* (x_1, x_2)$. By the construction of grammar G' there is a production $(A_1, A_2) \to (t_1, t_2)$ in P such that $\mu(\bar{A}_1) = A_1$ and $\mu(\bar{t}_1) = t_1$. Thus $u \Rightarrow^* v$ where $u = (y_1A_1z_1, A_2)$, $v = (y_1x_1z_1, x_2)$ and $\mu(u_1) = y_1A_1z_1$. Thus $L(G') \subset R^\beta(G)$.

Theorems 7 and 8 give:

Corollary 5. *There exists a relation R, R is in $\mathbf{R}^\alpha_{2,2}$ but not in $\mathbf{R}^\beta_{2,2}$.*

Example 6. Let R be in $\mathbf{R}^\alpha_{2,\ldots,2}$ then for $n \geqq 2$ all n-tuples are not obtainable by means of left-most derivations as in the case $n = 1$. For instance let $G = (\{a, b\} \times \{c\}, \{S_1, A\} \times \{S_2, B, C\}, \{(S_1 \to AA, S_2 \to B), (A \to a, B \to C), (A \to b, C \to c)\}, (S_1, S_2))$. Obviously $R^\alpha(G) = \{(a, b, c), (b, a, c)\}$. Both pairs in $R^\alpha(G)$ are derived using successively all productions of the grammar G (each once); the pair (ab, c) as the left-most derivation, the pair (ba, c) as the right-most derivation.

For context-free n-ary grammars (G in $\mathbf{G}_{2,\ldots,2}$) we introduce a third type of generation.

Definition. Let $\overset{\gamma}{\Rightarrow}{}^*$ be the minimal reflexive and transitive binary relation on $V^{1*} \times V^{2*} \times \ldots \times V^{n*}$ closed under following procedure for getting new members: if for v, w in $V^{1*} \times \ldots \times V^{n*}$ there exist A_i in V_N^i, u_i in V^i, y_i in V_T^{i*} and $k_i \geqq 1$ for $i = 1, 2, \ldots, n$ and s_i^j in V^{1*} for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k_i$ such that

$$\begin{aligned}
v &= (v_1, v_2, \ldots, v_n)\ ; \quad w = (w_1, w_2, \ldots, w_n)\ ;\\
v_1 &= s_1^1A_1s^2A_1 \ldots s_1^{k_1-1}A_1s_1^{k_1}\ ;\\
&\ \vdots\\
v_n &= s_n^1A_ns_n^2A_n \ldots s_n^{k_n-1}A_ns_n^{k_n}\ ;\\
w_1 &= s_1^1y_1s_1^2y_1 \ldots s_1^{k_1-1}y_1s_1^{k_1}\ ;\\
&\ \vdots\\
w_n &= s_n^1y_ns_n^2y_n \ldots s_n^{k_n-1}y_ns_n^{k_n}\ ;\\
&\ \vdots\\
(A_1, A_2, &\ldots, A_n) \to (u_1, u_2, \ldots, u_n) \text{ in } P\ ;
\end{aligned}$$

 and

$$(u_1, u_2, \ldots, u_n) \stackrel{\gamma}{\Rightarrow}{}^* (y_1, y_2, \ldots, y_n)$$

then $v \stackrel{\gamma}{\Rightarrow}{}^* w$.

The length of $u \stackrel{\gamma}{\Rightarrow}{}^* v$ is said to be the minimal necessary number of applications of the given procedure when proving that $u \stackrel{\gamma}{\Rightarrow}{}^* v$ by definition.

Lemma 1. *Let G be a context-free grammar, $G = (V_T, V_N, P, S)$ (G in $\mathbf{G}_2$) then $x \stackrel{\gamma}{\Rightarrow}{}^* v$ implies $x \stackrel{\gamma}{\Rightarrow}{}^* y$.*

Proof. We prove the lemma by the induction on the length n of $x \stackrel{\gamma}{\Rightarrow}{}^* y$. For $n = 1$ the assertion of the lemma obviously holds, let us assume that it holds for $r < n$.

Let the length of $x \stackrel{\gamma}{\Rightarrow}{}^* y$ be n. By the definition of γ-derivation we can write

$$x \stackrel{\gamma}{\Rightarrow}{}^* w \stackrel{\gamma}{\Rightarrow}{}^* y$$

and

$$x = s^1As^2A \ldots As^{k-1}As^k,$$

$$w = s^1ts^2A \ \ldots ts^{k-1}ts^k$$

where $z \stackrel{\gamma}{\Rightarrow}{}^* t$ and $A \to z$ is in P. Both $w \stackrel{\gamma}{\Rightarrow}{}^* y$ and $z \stackrel{\gamma}{\Rightarrow}{}^* t$ are of lengths less than n and therefore $w \stackrel{\beta}{\Rightarrow}{}^* y$ and $z \stackrel{\beta}{\Rightarrow}{}^* t$. Thus $x = s^1As^2A \ldots As^{k-1}As^k \stackrel{\beta}{\Rightarrow}{}^* s^1ts^2A \ldots \ldots As^{k-1}As^k \stackrel{\beta}{\Rightarrow}{}^* s^1ts^2t \ldots As^{k-1}As^k \stackrel{\beta}{\Rightarrow}{}^* \ldots \stackrel{\beta}{\Rightarrow}{}^* s^1ts^2t \ldots ts^{k-1}ts^k \stackrel{\beta}{\Rightarrow}{}^* y$.

Corollary 6. *Let G be a context-free grammar (G in $\mathbf{G}_2$), then $R^\gamma(G) \subset R^\beta(G)$ $(=R^\alpha(G) = L(G))$*

Lemma 2. *Let $G = (V_T^1 \times V_T^2 \times \ldots \times V_T^n, V_N^1 \times V_N^2 \times \ldots \times V_N^n, P, (S_1, S_2))$ be in $\mathbf{G}_{2,2,\ldots,2}$ and for all $q = (q_1, q_2, \ldots, q_n)$ in P there does not exist q_i of the form $A \to w$ where $w = w^1Bw^2Bw^3$, w^1, w^2, w^3 in V^{i*}, A, B in V_N^i.(On the right sides of unary productions there are all different nonterminal symbols.) Then $R^\beta(G) = R^\gamma(G)$.*

Proof is obvious.

Corollary 7. *For G in $\mathbf{G}_{3,3,\ldots,3}$ $R^\gamma(G) = R^\beta(G)$ and therefore $\mathbf{R}^\gamma_{3,3,\ldots,3} = \mathbf{R}^\beta_{3,3,\ldots,3}$ $(= \mathbf{R}^\alpha_{2,2,\ldots,2} = \mathbf{M}_n)$.*

Theorem 9. $\mathbf{R}^\beta_{2,2,\ldots,2} \subset \mathbf{R}^\gamma_{2,2,\ldots,2}$.

Proof. By Lemma 2 it is sufficient to show that for every n-ary grammar G there exists a grammar G', with different non-terminals on the right sides of each unary production such that $R^\beta(G') = R^\beta(G)$. The binary grammar G' is constructed from the grammar G as follows:

If any nonterminal symbol A occurs more than once in the right side of any unary production q_i ($(q_1, q_2, \ldots, q_n) \in P$) then we attach different subscripts to all occurrences of A. These subscripted nonterminals are added to the set V_N' and for each production $(q_1, q_2, \ldots, q_n)$ in which the left side of any q_i can be subscripted, the new productions are added for all the combinations of subscripts in both the left and the right sides of $q_1, q_2, \ldots, q_n$. Obviously, $R^\beta(G) = R^\beta(G')$.

Example 7. This is an example showing that even for the class $\mathbf{G}_2$ (context-free grammars) the inclusion $R_2^\beta \subset R_2^\gamma$ is proper. Let $G = (\{a, b\}, \{S, A\}, \{S \to AA, A \to aAb, A \to ab\}, S)$. Obviously, $R^\gamma(G) = \{a^n b^n a^n b^n \mid n = 1, 2, \ldots\}$ and it is a well-known example of a language which is not context-free.

MAPPINGS DESCRIBED BY GRAMMARS

A mapping from Σ_1^* into subsets of Σ_2^* is given by every binary relation $R, R \subset \Sigma_1^* \times \times \Sigma_2^*$. Thus by means of binary grammars a large class of mappings from Σ_1^* into subsets of Σ_2^* (e.g. transducions, translations [7]) can be described.

In the case that a binary grammar G has no productions of form $(\emptyset, q_2)$ and is of such a type, that for its partial grammar G_1 there exists an effective syntactic analysis procedure, then the mapping f given by the grammar G is given effectively. By the following procedure for each x in Σ_1^* the set $f(x)$, empty if x is not in the domain of f, can be found: All derivations of x in the partial grammar G_1 are found. Then for each unary derivation all such binary derivations are created, each of them using binary productions whose nonempty left components are the productions used in the common derivation. From the assumption that no productions in G has the form $(\emptyset, q_2)$ follows that to each unary derivation, there are only a finite number of corresponding binary derivations and the described procedure is actually effective.

Note. By the existence of an effective analysis procedure for partial grammar G_1, we mean that for each word w in Σ_1^* there exist only finite number of derivations in G_1 all of which can be found effectively.

Definition. Let f be a mapping from Σ_1^* into subsets of Σ_2^*, then the binary grammar G is said to ξ-induce (ξ in $\{\alpha, \beta, \gamma\}$) the mapping f if $f(x) = \{y \mid (x, y) \text{ in } R^\xi(G)\}$ for any x in Σ_1^*.

The binary grammar G is said to ξ-realize (ξ in $\{\alpha, \beta, \gamma\}$) the mapping f if $f(x) = = \{y \mid (x, y) \text{ in } R^\xi(G)\}$ for x in D_f, where $D_f = \{x \mid f(x) \neq \emptyset\}$ is the domain of mapping f.

Note the same terminology for multitape automata as in [2] and the notions "transduction" and "translation" as in [7]. We will call translation the mapping from a language L_1 into a language L_2 which preserves the meaning of all sentences.

Example 8. We write the binary grammar G β-inducing the translation of a simple arithmetic expression from common notation into reverse Polish notation.

$G = (\{a, b, c, +, *\} \times \{a, b, c,], [, +, *\}, \{P, T, E\} \times \{P, T, E\}, \{(P \to a, P \to a), (P \to b, P \to b), (P \to c, P \to c), (P \to [E], P \to E), (T \to P, T \to P), (T \to T*P, T \to TP*), (E \to T, E \to T), (E \to E + T, E \to ET+)\}, (E, E))$. The use of β-induction in this example is essential, for e.g. $(E, E) \overset{\alpha}{\Rightarrow} (E + T, ET+) \overset{\alpha}{\Rightarrow} (T + T, TT+) \overset{\alpha}{\Rightarrow} (T*P + T, TTP*+) \overset{\alpha}{\Rightarrow} (P*P + T, TPP*+) \overset{\alpha}{\Rightarrow} (P*P + P, PPP*+) \overset{\alpha}{\Rightarrow} (a*P + P, PaP*+) \overset{\alpha}{\Rightarrow} \overset{\alpha}{\Rightarrow} (a*b + P, baP*+) \overset{\alpha}{\Rightarrow} (a*b + c, bac*+)$ and $bac*+$ is not the translation of $a*b+c$ into reverse Polish notation.

Example 9. We write the binary grammar G γ-inducing the translation of simplified ALGOL-60 for statements into FORTRAN. The rules for writing numbers (N) and variables (V) in both languages are strongly simplified only for brevity of the example.

As the comma is a terminal symbol, the symbol "|" will be used instead of the comma to separate different elements of sets and components of pairs. The symbol $\neq$ indicates the end of a line in FORTRAN-programs.

$G = (V_T^1 \times V_T^2, V_N^1 \times V_N^2, P, (Q, Q))$ where

$V_T^1 = \{A \mid B \mid C \mid + \mid - \mid \times \mid \uparrow \mid := \mid \textbf{for} \mid \textbf{step} \mid \textbf{until} \mid \textbf{do} \mid \textbf{begin} \mid \textbf{end} \mid ;\}$,

$V_T^2 = \{A \mid B \mid C \mid 0 \mid 1 \mid + \mid - \mid * \mid = \mid \mathrm{DO} \mid \mathrm{CONTINUE} \mid (\mid) \mid , \mid \neq \mid \}$,

$V_N^1 = \{Q \mid V \mid E_1 \mid E_2 \mid E_3 \mid E \mid W \mid R \mid S \mid T \mid P \mid L \mid N\}$,

$V_N^2 = \{Q \mid V \mid E_1 \mid E_2 \mid E_3 \mid E \mid W \mid R \mid S \mid T \mid P \mid L \mid V_1 \mid V_2 \mid V_3 \mid N\}$.

and P consists of the productions

$$(Q \to \textbf{for } V := E_1 \textbf{ step } E_2 \textbf{ until } E_3 \textbf{ do } W \mid Q \to V_1 = E_1 \neq V_2 = E_3 \neq V_3 = \\ = E_2 \neq \mathrm{DO}(N)\, V = V_1, V_2, V_3 \neq W \neq N\, \mathrm{CONTINUE})$$

$(W \to S \mid W \to S)$	$(\emptyset \mid V_2 \to N)$
$(W \to \textbf{begin } R \mid W \to R)$	$(\emptyset \mid V_3 \to N)$
$(R \to S \textbf{ end} \mid R \to S)$	$(E_1 \to E \mid E_1 \to E)$
$(R \to S ; R \mid R \to S \neq R)$	$(E_2 \to E \mid E_2 \to E)$
$(S \to V := E \mid S \to V = E)$	$(E_3 \to E \mid E_3 \to E)$
$(E \to PTE \mid E \to PTE)$	$(N \to 1 \mid N \to 1)$
$(T \to + \mid T \to +)$	$(N \to 0 \mid N \to 0)$
$(T \to - \mid T \to -)$	$(N \to N0 \mid N \to N0)$
$(T \to \times \mid T \to *)$	$(N \to N1 \mid N \to N1)$
$(T \to \uparrow \mid T \to **)$	$(V \to L \mid V \to L)$
$(P \to V \mid P \to V)$	$(L \to A \mid L \to A)$
$(P \to N \mid P \to N)$	$(L \to B \mid L \to B)$
$(\emptyset \mid V_1 \to N)$	$(L \to C \mid L \to C)$

Note. It is possible to show that well-translations of languages [1] and syntax-directed translation [7] are a special case of the mappings β-induced by binary grammars of type (2,2).

The complexity of a mapping can be classified by the type of a binary grammar inducing (realizing) it.

Definition. Let a mapping f be ξ-induced (ξ in $\{\alpha, \beta, \gamma\}$) by a binary grammar of type (i, j). Then f is said to be of type $\mathbf{I}^{\xi}_{i,j}$. Similarly the type of a grammar realizing the mapping could be considered. For $i, j \geqq 3$ there is no difference between types with α or β and we will omit the Greek letter in this case. In the following we show the types of some commonly known mappings and we give some examples.

Obviously, a homomorphism is of type $\mathbf{I}_{4,4}$, sequential mappings realized by generalized sequential machine ([5]) will now be considered.

Theorem 10. *Any sequential mapping f is both of type $\mathbf{I}_{4,3}$ and type $\mathbf{I}_{3,4}$.*

Proof. Let the mapping f, having input alphabet X and output alphabet Y, be realized by the generalized sequential machine $(K, X, Y, \delta, \lambda, p_1)$, where K is a set of states, δ is a next-state function (from $K \times X$ into K), λ is an output function (from $K \times X$ into Y^*) and p_1 is the initial state.

1. We construct the binary grammar $G = (X \times Y, \{W\} \times K, P(W, p_1))$, where W is an arbitrary symbol and the set of productions P is given as follows: for each a in X and p_i in K the productions $(W \to aW,\ p_i \to \delta(p_i, a)\, \lambda(p_i, a))$ and $(W \to a,\ p_i \to \lambda(p_i, a))$ are in P. Obviously, (x, y) is in $R^{\alpha}(G)$ iff $y = f(x)$.

2. We construct the binary grammar $G' = (X \times Y, K \times \{W\}, P', (p_1, W))$, where W is a new symbol and the set of productions P' is given as follows: for each a in X and p_i in K the productions $(p_i \to a\, \delta(p_i, a),\ W \to \lambda(p_i, a)\, W)$ and $(p_i \to \lambda(p_i, a),\ W \to a)$ are in P'. Obviously, (x, y) is in $R^{\alpha}(G)$ iff $y = f(x)$.

Note. The sequential mappings are only a subclass of the mappings of type $\mathbf{I}_{4,3}$ ($\mathbf{I}_{3,4}$). For instance the inverse mappings of the sequential mappings are also of type $\mathbf{I}_{4,3}$ ($\mathbf{I}_{3,4}$).

Note. By Theorem 6 a mapping is of type $\mathbf{I}_{3,3}$ iff there exists a transduction R [3] such that $f(x) = \{y \mid (x, y) \text{ in } R\}$ for all x.

Example 10. The binary grammar G, from the proof of Theorem 7, α-induces a mapping f, the domain of which "is more complex" than the mapping itself (by interchanging the two partial grammars to give G', it follows trivially that G' α-induces a mapping the range of which "is more complex" than the mapping itself). By Theorem 8 it follows that the use of α-generation is essential in Example 10.

Example 11. The mapping f from Example 10 can be realized (but not induced) by the binary grammar G in $\mathbf{G}_{3,3}$.

Let $G = (\{a, b\} \times \{a, b\}, \{S, B, C\} \times \{S, C\}, \{(S \to aS, S \to aS), (S \to bB, S \to bC), (B \to bB, \emptyset) (B \to aC, C \to aC), (C \to aC, C \to aC), (B \to a, C \to a), (C \to a, C \to a), (S, S))$.

It is not difficult to see that

$$R^{\alpha}(G) = \{(a^m b^n a^p, a^m b a^p) \mid m = 0, 1, \dots ; n = 1, 2, \dots ; p = 1, 2, \dots\}$$

and thus f is α-realized by G.

Note. The complexity of the domain or/and the range of a mapping influences the necessary complexity of the binary grammar only if we consider the mapping induced (but not realized) by the grammar. It is trivial that the identical mapping of any domain over Σ onto itself is realized (but not induced) by the binary grammar G in $\mathbf{G}_{4,4}$, $G = (\Sigma \times \Sigma, \{S\} \times \{S\}, (S \to aS, S \to aS), (S \to a, S \to a)$, for all a in Σ, $(S, S))$.

We can also describe more general mappings by means of n-ary grammars (for $n > 2$). We give the following example.

Example 12. In [2] the mapping f is realized by means of a multi-tape automaton. The mapping f maps the pair of words over $\{a, b, c, \dots, z, \oplus, *\} \times \{a, b, c, \dots, a, \oplus\}$ into words over $\{a, b, c, \dots, z\}$. We get the output for a pair of words when taking the first word in the pair, omit the subwords between the $(2k - 1)$-th and $2k$-th occurrences of the symbol * for $k = 1, 2, \dots$, and replacing the subwords between the $(2k - 1)$-th and $2k$-th occurrences of the symbol $\oplus$ by the subwords of the second word of the pair between the $(k - 1)$-th and k-th occurrences of the symbol $\oplus$. (We assume the zero occurrence of $\oplus$ at the beginning of the second word.) In [2] the mapping f is described formally by means of a multi-tape automaton. We now write the ternary grammar G which β-induces f.

Let $\Sigma = \{a, b, c, \dots, z\}$ then $G = ((\{\oplus, *\} \cup \Sigma) \times (\Sigma \cup \{\oplus\}), \{(S_1 \to S_1 D, S_2 \to S_2, S_3 \to S_3 D), (S_1 \to S_1 {*C*}, S_2 \to S_2, S_3 \to S_3), (S_1 \to S_1 \oplus P \oplus, S_2 \to S_2 B \oplus, S_3 \to S_3 B), (C \to CA, \emptyset, \emptyset), (A \to \xi, \emptyset, \emptyset)$ for ξ in Σ, $(C \to \xi, \emptyset, \emptyset), (\emptyset, B \to BA, B \to BA), (\emptyset, A \to \xi, A \to \xi)$ for ξ in Σ, $(\emptyset, B \to \varepsilon, B \to \varepsilon), (D \to DA, \emptyset, D \to DA), (A \to \xi, \emptyset, A \to \xi), (D \to \varepsilon, \emptyset, D \to \varepsilon), (S_1 \to \varepsilon, S_2 \to \varepsilon, S_3 \to \varepsilon)\})$.

Acknowledgement. The author is grateful to Derick Wood for his comments on an earlier draft of this paper.

(Received March 10th, 1969.)

[1] Čulík, K.: Well-translatable grammars and ALGOL-like languages. In Steel, T. B. Jr. (Ed.): Proc. IFIP Working Conference on Formal Language Description Languages. North-Holland, Amsterdam 1966, pp. 76—85.
[2] Čulík II, K.: Sequential Machines with several input and output tapes. Kybernetika *3* (1967), 6, 540—559.
[3] Elgot, C. C. and Mezei, J. E.: On finite relation defined by generalized automata. IBM J. of Research and Development *9* (Jan. 1965), 1, 47—68.
[4] Ginsburg, S.: The Mathematical Theory of Context-Free Languages. McGraw-Hill, New York 1966.
[5] Ginsburg, S. and Rose, G. F.: A Characterization of Machine Mappings. The Canadian Journal of Mathematics *18* (1966), 381—388.
[6] Glushkov, V. M.: Introduction to Cybernetics. Academic Press, New York 1966.
[7] Lewis II, P. M. and Stearns, R. E.: Syntax-Directed Transduction. J. ACM *15* (July 1968), 3, 465—488.

VÝTAH

n-ární gramatiky a popis zobrazení jazyků

KAREL ČULÍK II

Frázové gramatiky byly doposud používány pro popis jazyků, tj. množin slov nad nějakou abecedou. Existuje řada jejich zobecnění, ale všechny jsou určeny rovněž k popisu jazyků. Zavádíme zde jiné zobecnění frazových gramatik — n-ární gramatiky, které generují relace, tj. množiny n-tic slov nad danými abecedami. n-ární gramatika je systém n terminálních abeced, n nonterminálních abeced, množiny pravidel a výchozí n-tice nonterminálních symbolů. Každé pravidlo je n-tice obyčejných pravidel nebo prázdných míst. Jsou zavedeny tři různé způsoby generování relací n-árními gramatikami a jsou zkoumány jejich vlastnosti a vzájemné vztahy. Chomského klasifikace je zobecněna pro n-ární gramatiky a jsou vyšetřovány uzávěrové a jiné vlastnosti různých tříd. Binární gramatiky jsou použity k popisu zobrazení jazyků, zvláště překladu z jazyku do jazyku. Překlady, které lze popsat binárními gramatikami, zahrnují dobrou přeložitelnost [1] a syntakticky řízený překlad [7] jako jednoduché případy. Je zavedena klasifikace zobrazení jazyků podle typu gramatiky, která ho popisuje a je uvedeno několik příkladů.

RNDr. Karel Čulík II, CSc., Centrum numerické matematiky Karlovy university, Malostranské nám. 25, Praha 1.

Department of Applied Analysis and Computer Science, University of Waterloo, Waterloo. Ontario, Canada.