INFORMATION AND CONTROL 14, 252–277 (1969)

Transformational Classes of Grammars

C. P. Schnorr

Institut für Angewandte Mathematik, Universität Saarbrücken, Germany

Given two Chomsky grammars G and $\bar{G}$, a homomorphism φ from G to $\bar{G}$ is, roughly speaking, a map which assigns to every derivation of G a derivation of $\bar{G}$ in such a manner that φ is uniquely determined by its restriction to the set of productions of G. Two grammars are contained in the same transformational class, if the one can be transformed into the other by a sequence of homomorphisms. If two grammars are related in such a manner, then there are two relations, one concerning the words of the languages generated and the other regarding the derivations of these words. We establish several classifications of context-free grammars in transformational classes which are recursively solvable.

1. INTRODUCTION

Chomsky grammars are special devices for generating languages. They are usually considered to be equivalent, if they produce the same language. However, the information included in a grammar concerns the words generated as well as the distinct derivations for a word itself. The concept pursued in this paper is to examine relations of grammars which are stronger than the common notion of equivalence. For example, we consider a classification of grammars, where two grammars fall under the same class, if they generate the same language and if, in addition, there is a certain correspondence regarding their derivations. Furthermore we will investigate relationships of grammars whose languages on the one hand do not coincide, but which are related on the other hand by a certain correspondence as regards the derivations of the grammars. This correspondence of the derivations of two related grammars will be specified by the term of homomorphism of grammars which is a central notion in our paper.

The concept of homomorphism is based on a characterization of a Chomsky grammar introduced by G. Hotz (1966). He describes the set of derivations of a grammar as a free category with an additional monoid multiplication and establishes the notion of a homomorphism of grammars. It is the idea of G. Hotz (1968a and 1968b) to develop this notion into a relationship concerning grammars. Without this inducement this paper would never have been written.

In Chapter 2 we will summarize the presentation of derivations given by G. Hotz. We will develop this concept as far as it will be used in our further considerations. Our concept of transformational classes will appear in Chapter 3. We will formulate two transformational problems relative to a certain class of homomorphisms. Given two grammars, find an algorithm to determine whether the one can be transformed into the other by a sequence of surjective (bijective) homomorphisms of the considered class. In Chapter 4 we will present the central tool in our considerations, i.e. an axiomatic property for a class of homomorphisms, the reduction property. This property provides a considerable simplification of our problems. Examples of classes of homomorphisms which satisfy the reduction property will be given in the Chapters 5 and 6. Finally in the Chapters 7 and 8 we will apply the generalized finite automata theory to examine the two transformational problems as regards context-free grammars. We will establish several classes of homomorphisms for which the first transformational problem is solvable, when restricting ourselves to context-free grammars.

It is assumed that the reader is familiar with the common concept of a Chomsky grammar and knows the definition of a category (Mitchell, 1965).

2. A FORMAL SYMBOLISM FOR THE DERIVATIONS OF CHOMSKY GRAMMARS

Let $G = (O, T, S, P)$ be a Chomsky grammar with the alphabet O, the terminal alphabet $T \subset O$, the set of initial symbols $S \subset O$ and the set $P \subset O^* \times O^*$[1] of productions.

It has been observed by numerous authors that the common description of a derivation as a sequence of words is ambiguous, since, for example, the productions are not specified. On the other hand this description distinguishes derivations which are not essentially different. As far as we know, it was M. Paul (1962) who used for the first time canonical derivations to specify essentially different derivations. Two derivations are called similar by Griffiths (1968), if the one can be obtained from the other by trivial rearrangement within the sequence of

[1] For any set X the free monoid on X is denoted by X^*.

productions applied. In the following we shall simply write derivation, when we mean an equivalence class of similar derivations. Later on we will develop a precise definition of this notion.

The notation $w \xrightarrow{\alpha} v$ or $\alpha: w \to v$ means that α is a derivation from w to v, i.e. an equivalence class of similar derivations. We denote $D_0(\alpha) = w$, $D_1(\alpha) = v$, D_0, D_1 being domain and codomain functions. For $w \in O^*$ let id_w be the derivation $\mathrm{id}_w : w \to w$, where no production is applied.

The total set of derivations generated by some Chomsky grammar is called a derivation system ($\mathfrak{DS}$).

Let us consider the structure of a $\mathfrak{DS}$ in a more detailed manner. There are two binary operations on a $\mathfrak{DS}$ M.

The one operation is the juxtaposition of derivations.

If $v_1 \xrightarrow{\alpha_1} w_1$, $v_2 \xrightarrow{\alpha_2} w_2$ are derivations, then the derivation $\alpha_1 \times \alpha_2 : v_1v_2 \to w_1w_2$ is equivalently described by the sequences $v_1v_2 \xrightarrow{\text{applying } \alpha_1} w_1v_2 \xrightarrow{\text{applying } \alpha_2} w_1w_2$ and $v_1v_2 \xrightarrow{\text{applying } \alpha_2} v_1w_2 \xrightarrow{\text{applying } \alpha_1} w_1w_2$ respectively. (2.1)

The operation $\times$ is associative and defined for any elements of M, i.e. $\times$ is a monoid multiplication.

In addition to this we have the postposition of derivations.

Let $\alpha: v \to w$, $\beta: w \to u$ be derivations, then $\beta \circ \alpha : v \to u$ is defined to be the composition $v \xrightarrow{\alpha} w \xrightarrow{\beta} u$. (2.2)

The operation $\circ$ is associative and $\beta \circ \alpha$ is defined, if and only if $D_0(\beta) = D_1(\alpha)$.

The following relations are easy to verify.

$$\begin{aligned} \alpha \circ \mathrm{id}_{D_0(\alpha)} &= \alpha, \\ \mathrm{id}_{D_1(\alpha)} \circ \alpha &= \alpha; \end{aligned} \tag{2.3}$$

$$(\alpha_1 \circ \beta_1) \times (\alpha_2 \circ \beta_2) = (\alpha_1 \times \alpha_2) \circ (\beta_1 \times \beta_2) \tag{2.4}$$

if $$D_0(\alpha_i) = D_1(\beta_1) \qquad i = 1, 2.$$

The meaning of (2.4) is outlined by the following sketch.

$$\begin{pmatrix}\beta_1\\ \circ\\ \beta_2\end{pmatrix} \times \begin{pmatrix}\beta_2\\ \circ\\ \alpha_2\end{pmatrix} = \begin{matrix}(\beta_1 \times \beta_2)\\ \circ \\ (\alpha_1 \times \alpha_2)\end{matrix}$$

FIG. 1

A precise definition of a derivation can now be derived from the following two propositions stated without proof. Let M be a $\mathfrak{DS}$ with the set of productions P and the alphabet O, E being the set of identities on elements of O, $E = \{\mathrm{id}_a \mid a \in O\}$.

(2.5) PROPOSITION. *Every derivation* $\alpha \in M$ *can be written as a product of elements in* $P \cup E$ *relative to the operations* $\circ$ *and* $\times$.

(2.6) PROPOSITION. *Two products of elements in* $P \cup E$ *relative to the operations* $\circ$ *and* $\times$ *represent the same derivation,* iff *the one can be transformed into the other by the relations* (2.3) and (2.4).

Let us consider an example which illustrates (2.6). If $\alpha_1 : v_1 \to w_1$, $\alpha_2 : v_2 \to w_2$ are derivations, then we can transform $\alpha_1 \times \alpha_2 : v_1 v_2 \to w_1 w_2$ as follows:

$$\begin{aligned} (\mathrm{id}_{w_1} \times \alpha_2) \circ (\alpha_1 \times \mathrm{id}_{v_2}) &= (\mathrm{id}_{w_1} \circ \alpha_1) \times (\alpha_2 \circ \mathrm{id}_{v_2}) \\ &= \alpha_1 \times \alpha_2 = (\alpha_1 \circ \mathrm{id}_{v_1}) \times (\mathrm{id}_{w_2} \circ \alpha_2) \\ &= (\alpha_1 \times \mathrm{id}_{w_2}) \circ (\mathrm{id}_{v_1} \times \alpha_2) \end{aligned}$$

Furthermore we state two representation theorems for derivations which will be useful in some applications.

(2.7) PROPOSITION (Hotz, 1966). *Every derivation* $\alpha \in M$, *not being an identity, can be written* $\alpha = \beta_1 \circ \beta_2 \circ \cdots \circ \beta_n$ *with* $\beta_i \in E^* \times P \times E^{*2}$ $i = 1, 2, \cdots, n$.

$\beta_1 \circ \beta_2 \circ \cdots \circ \beta_n$ is called a sequential product. A pair (β_i, β_{i+1}) of two consecutive factors in a sequential product is defined to be canonical, if the factors can be written $\beta_\mu = \mathrm{id}_{u_\mu} \times \alpha_\mu \times \mathrm{id}_{v_\mu}$ for $\mu = i, i+1$ such that

$$|u_{i+1}| < |u_i| + |D_0(\alpha_i)|^3.$$

[2] For $A \subset M$ the closure of A relative to $\times$ is denoted by A^*.

[3] For a word w the length is denoted by $|w|$.

A sequential product is defined to be canonical, if every pair of consecutive factors is canonical.

(2.8) PROPOSITION (Hotz, 1966). *If M has either of the following properties, and if $\alpha \in M$ is not an identity, then there is a unique canonical product which represents α.*

(i) $D_0(\alpha) \neq \Lambda$[4] $\quad (\alpha \in P)$,

(ii) $D_1(\alpha) \neq \Lambda \quad (\alpha \in P)$.

The number of factors in a sequential product is an invariant for each derivation α and called its length, $L(\alpha)$. Naturally we define $L(\mathrm{id}_w) = 0$ for all words w.

(2.10) DEFINITION. Given two $\mathfrak{DS}$'s M_1 and M_2 with the alphabets O_1, O_2 resp. and the sets of productions P_1, P_2 resp., a homomorphism $\varphi: M_1 \to M_2$ consists of:

H1 a homomorphism $\quad \varphi: O_1^* \to O_2^*$

H2 a map $\quad \varphi: M_1 \to M_2$

The following axioms are postulated.

H3 $\quad \varphi(\alpha_1 \times \alpha_2) = \varphi(\alpha_1) \times \varphi(\alpha_2) \qquad \alpha_1, \alpha_2 \in M_1$

H4 $\quad \varphi(\alpha_1 \circ \alpha_2) = \varphi(\alpha_1) \circ \varphi(\alpha_2) \qquad \alpha_r, \alpha_2 \in M_1$

H5 the commutativity relation holds in the diagram

$$\begin{array}{ccc} M_1 & \xrightarrow{D_0,\, D_1} & O_1^* \\ \varphi \downarrow & & \downarrow \varphi \\ M_2 & \xrightarrow{D_0,\, D_1} & O_2^* \end{array}$$

(2.11) PROPOSITION (Hotz, 1966). *A homomorphism $\varphi: M_1 \to M_2$ is uniquely determined by a homomorphism $\varphi: O_1^* \to O_2^*$ and a map $\varphi: P_1 \to M_2$ such that commutativity holds in the diagram*

$$\begin{array}{ccc} P_1 & \xrightarrow{D_0,\, D_1} & O_1^* \\ \varphi \downarrow & & \downarrow \varphi \\ M_2 & \xrightarrow{D_0,\, D_1} & O_2^* \end{array}$$

[4] Λ denotes the empty word.

Proof $\varphi:P_1 \to M_2$ and $\varphi:O_1^* \to O_2^*$ are extended to a homomorphism $\varphi:M_1 \to M_2$ with the recursive definition

$$\varphi(\mathrm{id}_w) = \mathrm{id}_{\varphi(w)}$$

$$\varphi(\alpha_1 \times \alpha_2) = \varphi(\alpha_1) \times \varphi(\alpha_2)$$

$$\varphi(\alpha_1 \circ \alpha_2) = \varphi(\alpha_1) \circ \varphi(\alpha_2)$$

The relations (2.3) and (2.4) ensure that $\varphi:M_1 \to M_2$ is uniquely determined by these formulas.

A more detailed description of the structure of a $\mathfrak{DS}$ can be found in (Hotz, 1966). A $\mathfrak{DS}$ can be characterized as a category with an additional monoid multiplication, called X-category, which is free in the sense suggested in (2.11).

Next we will use this concept to define homomorphisms of grammars.

(2.11) Definition. Given two grammars $G = (O, T, S, P)$ and $\bar{G} = (\bar{O}, \bar{T}, \bar{S}, \bar{P})$, a homomorphism $\varphi:G \to \bar{G}$ is defined to be a homomorphism $\varphi:M \to \bar{M}$ relative to the associated $\mathfrak{DS}$'s which satisfies

(i) $\varphi(S) \subset \bar{S}$

(ii) $\varphi(T) \subset \bar{T}^*$

To motivate the conditions (i) and (ii) we consider the language L_G generated by a grammar G which can be written in our terminology $L_G = \{D_1(\alpha) \in T^* \mid \alpha \in M, D_0(\alpha) \in S\}$. The set of derivations which produce words in L_G will be denoted by M_G. $M_G = \{\alpha \in M \mid D_0(\alpha) \in S. D_1(\alpha) \in T^*\}$. Then the conditions (i) and (ii) are so chosen as to ensure $\varphi(M_G) \subset M_{\bar{G}}$ and hence $\varphi(L_G) \subset L_{\bar{G}}$.

Our formalization of a derivation yields a precise definition of the *ambiguity* $\langle w, G\rangle$ of a word w relative to a grammar G which is defined by $\langle w, G\rangle = |\{\alpha \in M_G \mid D_1(\alpha) = w\}|$.[5]

A homomorphism $\varphi:G \to \bar{G}$ is called *surjective*, if and only if $\varphi(M^G) = M_{\bar{G}}$. From the surjectivity of φ we infer $\varphi(L_G) = L_{\bar{G}}$. As regards the ambiguity we have

$$\langle w, \bar{G}\rangle \leqq \sum_{\varphi(v)=w} \langle v, G\rangle \qquad (w \in \bar{T}^*).$$

We shall call a homomorphism $\varphi:G \to \bar{G}$ *injective*, if the restriction

[5] If X is a set, $| X |$ denotes the cardinality of X.

$\varphi/M_G : M_G \to M_{\bar{G}}$ is injective. It follows from the injectivity of φ that

$$\langle w, \bar{G}\rangle \geqq \sum_{\varphi(v)=w} \langle v, G\rangle \qquad (w \in \bar{T}^*).$$

A homomorphism is called *bijective*, if it is surjective as well as injective.

In the special case that $\varphi: G \to \bar{G}$ is a homomorphism of two grammars with the same terminal alphabet $T = \bar{T}$ and φ is identical in T we infer from

the surjectivity of φ: $\langle w, \bar{G}\rangle \leqq \langle w, G\rangle \qquad (w \in T^*)$,

from the injectivity of φ: $\langle w, \bar{G}\rangle \geqq \langle w, G\rangle \qquad (w \in T^*)$,

and from the bijectivity of φ: $\langle w, \bar{G}\rangle = \langle w, G\rangle \qquad (w \in T^*)$.

Evidently homomorphisms can be composed in a natural way. A homomorphism $\varphi: G \to \bar{G}$ is called an *isomorphism*, if there is a homomorphism $\tau: \bar{G} \to G$ such that $\varphi\tau = \mathrm{id}_{\bar{G}}$ and $\tau\varphi = \mathrm{id}_{\bar{G}}$.

To avoid misunderstanding we emphasize that bijective homomorphisms are, in general, not isomorphisms, but on the other hand each isomorphism is bijective.

A homomorphism $\varphi: G \to \bar{G}$ is called an *inclusion* and is denoted by $G \overset{\varphi}{\subset} \bar{G}$, if it is induced by inclusions $P \subset \bar{P}$ and $O \subset \bar{O}$. G is called a *subgrammar* of $\bar{G}$.

3. TRANSFORMATIONAL CLASSES OF GRAMMARS

Two grammars G and $\bar{G}$ are usually called equivalent, if they produce the same language. It is a well-known theorem by Bar-Hillel, Perles and Shamir (1961) that this equivalence is not decidable even for context-free grammars. On the other hand there is a strong interest in a comparison of grammars relative to their generative power. This is the reason, why we introduce another type of an equivalence relation which is derived from the notion of homomorphism.

(3.1) Definition. A category $\mathcal{C}$ is called a category of grammars, if the objects of $\mathcal{C}$ are grammars, and if the morphisms from G to $\bar{G}$ are homomorphisms from G to $\bar{G}$. The composition law of $\mathcal{C}$ is to be the composition of homomorphisms.

Our concept of classifying grammars now appears in the following definition.

(3.2) Definition. The grammars G and $\bar{G}$ are to be equivalent relative to the category $\mathcal{C}$ of grammars, iff there exists for some n a sequence $G \leftarrow G_1 \rightarrow G_2 \leftarrow \cdots \leftarrow G_{2n+1} \rightarrow \bar{G}$ of surjective homomorphisms in $\mathcal{C}$.

We take into account that an identity in a category of grammars is always surjective. From this it is obvious that (3.2) induces an equivalence relation. An equivalence class of grammars relative to $\mathcal{C}$ is to be called a transformational class relative to $\mathcal{C}$.

To motivate this concept let $\mathcal{C}^T$ be the category whose set of objects is the set of all grammars with the distinguished terminal alphabet T and whose morphisms are all those homomorphisms of the above grammars which are constant in T. The equivalence relative to $\mathcal{C}^T$, which falls under the equivalence by reduction in (Hotz, 1968b), implies the common equivalence, i.e. grammars that are equivalent relative to $\mathcal{C}^T$ generate the same language.

We now state what we shall call the first transformational problem.

Let $\mathcal{C}$ be a category of grammars. Given two grammars G and $\bar{G}$, find an algorithm to decide whether G and $\bar{G}$ are equivalent relative to $\mathcal{C}$.

If G and $\bar{G}$ are grammars, equivalent relative to $\mathcal{C}$, there arises another problem as regards the ambiguity of words, i.e. whether a derivation in M_G corresponds to a unique derivation in $M_{\bar{G}}$ Let us be more precise.

(3.3) Definition. The grammars G and $\bar{G}$ are to be strictly equivalent relative to the category $\mathcal{C}$ of grammars, if and only if there exists for some n a sequence

$$G \leftarrow G_1 \rightarrow G_2 \leftarrow \cdots \leftarrow G_{2n+1} \rightarrow \bar{G}$$

of bijective homomorphisms in $\mathcal{C}$.

The equivalence relation, given hereby, yields the second transformational problem:

Let $\mathcal{C}$ be a category of grammars. Given two grammars G and $\bar{G}$, find an algorithm to decide whether G and $\bar{G}$ are strictly equivalent relative to $\mathcal{C}$.

4. REDUCTION PROPERTY

At first we consider some structures in categories which will be useful for our concept.

(4.1) Definition. A diagram $A_1 \xleftarrow{g^1} A \xrightarrow{g^2} A_2$ in a category $\mathcal{C}$ is called a product diagram for A_1 and A_2, if for every commutative diagram

$$\begin{tikzcd}
& \bar{A} \arrow[dl, "f^1"'] \arrow[dr, "f^2"] & \\
A_1 & & A_2 \\
& A \arrow[ul, "g^1"] \arrow[ur, "g^2"'] &
\end{tikzcd}$$

there exists a unique morphism $f: \bar{A} \to A$ such that $g^1 f = f^1$ and $g^2 f = f^2$.

If for every pair of objects A_1 and A_2 in $\mathcal{C}$ there exists a product diagram in $\mathcal{C}$, then we shall simply say that $\mathcal{C}$ has products.

(4.2) Definition. Given two morphisms $f^1: A_1 \to A, f^2: A_2 \to A$ with a common codomain A, a commutative diagram

$$\begin{tikzcd}
& A' \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\
A_1 \arrow[dr, "f^1"'] & & A_2 \arrow[dl, "f^2"] \\
& A &
\end{tikzcd}$$

is called a pullback for f^1 and f^2, if for every pair of morphisms $\bar{g}^i: \bar{A} \to A_i$ $i = 1, 2$ such that $f^1\bar{g}^1 = f^2\bar{g}^2$, there exists a unique morphism $h: \bar{A} \to A'$ such that $g^i h = \bar{g}^i$ for $i = 1, 2$.

If for every pair of morphisms f^1 and f^2 in a category $\mathcal{C}$ with a common codomain there exists a pullback for f^1 and f^2 in $\mathcal{C}$, then we shall say that $\mathcal{C}$ has pullbacks.

It follows from the definition that products and pullbacks are unique up to isomorphisms, if they exist.

We will now formulate the crucial property for categories of grammars which will simplify the transformational problems.

(4.3) Definition. A category $\mathcal{C}$ of grammars is said to have the reduction property, if

R1 $\quad\mathcal{C}$ has products and pullbacks.

$$\begin{tikzcd}
& A' \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\
A_1 \arrow[dr, "f^1"'] & & A_2 \arrow[dl, "f^2"] \\
& A &
\end{tikzcd}$$

R2 relative to the pullback

in $\mathcal{C}$, if f^1 is surjective, then so is g^2.

(4.4) THEOREM. *Two grammars G and $\bar{G}$ are equivalent relative to a category which has the reduction property, if and only if the homomorphisms in the product diagram of G and $\bar{G}$ are surjective.*

Proof. Let $G \leftarrow G_1 \rightarrow G_2 \leftarrow \cdots G_{2n} \leftarrow G_{2n+1} \rightarrow \bar{G}$ be a sequence of surjective homomorphisms. Suppose that $n > 0$. Then we can reduce it to a sequence with a smaller n. Consider the diagram

$$\begin{tikzcd} & & G \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & & \\ & G_1 \arrow[dl, "f^1"'] \arrow[dr, "f^2"'] & & G_3 \arrow[dl, "f^3"] \arrow[dr, "f^4"] & \\ G & & G_2 & & G_4 \end{tikzcd}$$

where

$$\begin{tikzcd} & G \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\ G_1 \arrow[dr, "f^2"'] & & G_3 \arrow[dl, "f^3"] \\ & G_2 & \end{tikzcd}$$

is the pullback diagram for f^2 and f^3. It follows from (R2) that g^1 and g^2 are surjective. Hence the above sequence can be reduced to

$$G \xleftarrow{f^1g^1} G \xrightarrow{f^4g^2} G_4 \leftarrow \cdots .$$

This proves that there is a sequence of two surjective homomorphisms.

$$G \xleftarrow{f^1} G \xrightarrow{f^2} \bar{G}.$$

Because of (R1) we can complete the product diagram $G \xleftarrow{g^1} G_0 \xrightarrow{g^2} \bar{G}$ to a commutative diagram

$$\begin{tikzcd} & G \arrow[dl, "f^1"'] \arrow[dr, "f^2"] \arrow[d, dashed, "g"] & \\ G & G_0 \arrow[l, "g^1"] \arrow[r, "g^2"'] & \bar{G} \end{tikzcd}$$

From the surjectivity of f^1 and f^2 it follows that the homomorphisms g_1 and g_2 are subjective, too. This proves the one implication of the theorem, whereas the other is trivial. The reduction property is by no means a necessary condition for theorem (4.4). In fact, the properties of the pullback are not fully utilized in the above proof. The reduction property is so interesting, since it can be established in a systematic manner for numerous categories of grammars.

5. THE CATEGORY $\mathcal{C}_r$ OF REDUCING HOMOMORPHISMS

We will consider a category of grammars which is of special interest for our purposes.

Let $\mathcal{C}_r$ be the category whose set of objects is the set of all Chomsky grammars and whose set of morphisms is the set of all reducing homomorphisms in the following sense

A homomorphism $\varphi: G \to \bar{G}$ is to be called reducing, if the conditions (5.1) and (5.2) are satisfied.

$$\varphi(a) \in \bar{O} \cup \{\Lambda\} \qquad (a \in O), \tag{5.1}$$

$$\varphi(\alpha) \in \bar{P} \cup \bar{E}^* \qquad (\alpha \in P). \tag{5.2}$$

Condition (5.1) asserts that φ reduces the length of words. Because of (5.2) φ also reduces the length of derivations

To avoid trivialities we assume that there is no production (w, w) in any Chomsky grammar. To obtain homogeneous notations we will write (w, w) for id_w.

(5.3) THEOREM. *The category $\mathcal{C}_r$ has products.*

Proof. Let $G_i = (O_i, T_i, S_i, P_i)$ $i = 1, 2$ be any Chomsky grammars. We construct the product $G_1 \xleftarrow{g^1} G \xrightarrow{g^2} G_2$ $(G = (O, T, S, P))$ as follows. Let $\bar{O}_i = O_i \cup \{\Lambda\}$ and $\bar{T}_i = T_i \cup \{\Lambda\}$ and $\bar{P}_i = P_i \cup E_i^*$ for $i = 1, 2$.

$$O \underset{\text{def}}{=} \bar{O}_r \times \bar{O}_2 - \{(\Lambda, \Lambda)\}, \qquad T \underset{\text{def}}{=} \bar{T}_1 \times \bar{T}_2 - \{(\Lambda, \Lambda)\},$$

$$S \underset{\text{def}}{=} S_1 \times S_2$$

We use the projections $p_i: O^* \to O_i^*$, which are given by $p_i(a_1, a_2) = a_i$ *f.a.* $(a_1, a_2) \in O$, $i = 1, 2$, to define

$$P = \{(w, v) \in O^* \times O^* \mid ((p_1(w), p_1(v)), (p_2(w), p_2(v))) \in \bar{P}_1 \times \bar{P}_2 - E_1^* \times E_2^*\}$$

It is easy to verify that the finiteness of P follows from the finiteness of P_1 and P_2.

The projections p_i induce homomorphisms

$$g^1:G \to G_i \qquad (i = 1, 2)$$

by the formulas

$$g^1(a) = p_i(a) \qquad (a \in O, \quad i = 1, 2),$$
$$g^1(w, v) = (p_i(w), p_i(v)) \qquad ((w, v) \in P, \quad i = 1, 2).$$

To show that $G_1 \xleftarrow{g^1} G \xrightarrow{g^2} G_2$ is a product we assume a commutative diagram

$$\begin{tikzcd} & G_3 \arrow[dl, "f^1"'] \arrow[dr, "f^2"] & \\ G_1 & G \arrow[l, dashed, "g^1"] \arrow[r, dashed, "g^2"'] & G_2 \end{tikzcd}$$

Suppose in addition that commutativity holds in

(D)
$$\begin{tikzcd} & G_3 \arrow[dl, "f^1"'] \arrow[dr, "f^2"] \arrow[d, "g"] & \\ G_1 & G \arrow[l, dashed, "g^1"] \arrow[r, dashed, "g^2"'] & G_2 \end{tikzcd}$$

This implies

$$g^i g(a) = f^i(a) \qquad a \in O_3, \quad i = 1, 2. \tag{5.4}$$

Because of (5.4) $g:O_3^* \to O^*$ is uniquely determined. And the relation

$$g(w, v) = (g(w), g(v)) \qquad (w, v) \in P_3, \tag{5.5}$$

which holds for any reducing homomorphism $g:G_3 \to G$, proves that $g:G_3 \to G$ is uniquely determined by the commutativity of (D).

(5.6) THEOREM. *The category $\mathcal{C}_r$ has pullbacks.*

Proof. Let $G_1 \xrightarrow{f^1} \tilde{G} \xleftarrow{f^2} G_2$ be any diagram in $\mathcal{C}_r$ and let $G_1 \xleftarrow{h^1} G \xrightarrow{h^2} G_2$ be the product diagram constructed above. We construct the pullback diagram

(D1)
$$\begin{tikzcd} & \bar{G} \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\ G_1 \arrow[dr, "f^1"'] & & G_2 \arrow[dl, "f^2"] \\ & \tilde{G} & \end{tikzcd}$$

$\bar{G} = (\bar{O}, \bar{T}, \bar{S}, \bar{P})$ is to be a subgrammar of G. Using the above representation of G we define.

$$\bar{O} = \{(a_1, a_2) \in O \subset \bar{O}_1 \times \bar{O}_2 \mid f^1(a_1) = f^2(a_2)\}$$

$$\bar{T} = T \cap \bar{O}, \qquad \bar{S} = S \cap \bar{O}, \qquad \bar{P} = Pn\ \bar{O}^* \times \bar{O}^*.$$

Evidently there is a natural inclusion $\bar{G} \overset{j}{\subset} G$.

The homomorphisms g^i are to be the compositions $h^i j = g^i$, $i = 1, 2$. It is easy to verify that (D1) is commutative. To prove that (D1) is a pullback diagram we consider any commutative diagram

(D2)
$$\begin{tikzcd}
 & G' \arrow[dl, dashed, "g'^1"'] \arrow[dr, dashed, "g'^2"] & \\
G_1 \arrow[dr, dashed, "f^1"'] & \bar{G} \arrow[l, dashed, "g^1"'] \arrow[r, dashed, "g^2"] & G_2 \arrow[dl, dashed, "f^2"] \\
 & G &
\end{tikzcd}$$

The argument is quite the same as in the proof of theorem (5.3) concerning the product diagram.

If there is a homomorphism $g' : G' \to \bar{G}$ such that commutativity holds in the completed diagram (D2), then g' is uniquely determined by the relations:

$$g^i g'(a) = g'^i(a) \qquad (a \in O'),$$

$$g'(w, v) = (g'(w), g'(v)) \qquad ((w, v) \in P').$$

and these relations can be regarded as a definition of an appropriate g'.

(5.7) Corollary.

If

$$\begin{tikzcd}
 & \bar{G} \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\
G_1 \arrow[dr, "f^1"'] & & G_2 \arrow[dl, "f^2"] \\
 & \tilde{G} &
\end{tikzcd}$$

is a pullback in $\mathcal{C}_r$, *then for every pair of derivations* α_i *of* G_i *such that* $f^1(\alpha_1) = f^2(\alpha_2)$, *there exists some derivation* α *of* $\bar{G}$ *such that* $g^i(\alpha) = \alpha_i$ $i = 1, 2$.

Proof. More precisely, we assert that for any pair of words $w, v \in \bar{O}^*$ such that $g^i(w) = D_0(\alpha_i)$ and $g^i(v) = D_1(\alpha_i)$ for $i = 1, 2$, there exists a derivation α of $\bar{G}$ such that

$$D_0(\alpha) = w, D_1(\alpha) = v \quad \text{and} \quad g^i(\alpha) = \alpha_i \quad \text{for} \quad i = 1, 2.$$

If $(\alpha_1, \alpha_2) \in (P_1 \cup E_1^*) \times (P_2 \cup E_2^*) - E_1^* \times E_2^*$ and $f^1(\alpha_1) = f^2(\alpha_2)$, then we infer from the construction of the pullback that there exists $\alpha \in \bar{P}$ satisfying $g^i(\alpha) = \alpha_i$ for $i = 1, 2$, $D_0(\alpha) = w$ and $D_1(\alpha) = v$.

Let us denote $[P_i] = E_i^* \times P_i \times E_i^* \cup E_i^*$ for $i = 1, 2$. The assertion then follows in the case

$$(\alpha_1, \alpha_2) \in [P_1] \times [P_2] - E_1^* \times E_2^*.$$

To prove the general case we take into consideration that, if $\varphi : G \to G'$ is a reducing homomorphism and if $\varphi(\delta) = \beta_1 \circ \beta_2$, there exists a (possibly not unique) decomposition $\delta = \delta_1 \circ \delta_2$ with $\varphi(\delta_i) = \beta_i$ for $i = 1, 2$. By this argument there exist decompositions for arbitrary α_i

$$\alpha_i = \alpha_1^i \circ \alpha_2^i \circ \cdots \circ \alpha_n^i \qquad (i = 1, 2)$$

such that

$$(\alpha_\mu^1, \alpha_\mu^2) \in [P_1] \times [P_2] - E_1^* \times E_2^* \qquad \mu = 1, 2, \cdot\cdot, n,$$

and

$$f^1(\alpha_\mu^1) = f^2(\alpha_\mu^2) \qquad \mu = 1, 2, \cdot\cdot, n.$$

The assertion now follows immediately by induction on n.

(5.8) COROLLARY.
If

$$\begin{tikzcd} & \bar{G} \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\ G_1 \arrow[dr, "f^1"'] & & G_2 \arrow[dl, "f^2"] \\ & G & \end{tikzcd}$$

is a pullback in $\mathfrak{C}_r$ *and if* f^1 *is surjective, then* g^2 *is surjective, too.*

Proof. Let $\alpha_2 \in M_{G_2}$. Because of the surjectivity of f^1 there exists $\alpha_1 \in M_{G_1}$ such that $f^1(\alpha_1) = f^2(\alpha_2)$. Because of Corollary (5.7) there exists a derivation α of $\bar{G}$ such that $g^i(\alpha) = \alpha_i$ for $i = 1, 2$. From $g^i(\alpha) \in M_{\bar{G}_i}$ for $i = 1, 2$ it follows that $\alpha \in M_{\bar{G}}$. Hence g^2 is surjective.

Summarizing the theorems and corollaries of this chapter we have proved:

(5.9) THEOREM. *The category $\mathcal{C}_r$ has the reducing property.*

6. SUBCATEGORIES OF $\mathcal{C}_r$

The equivalence relative to $\mathcal{C}_r$ represents an extremely weak relationship between grammars There are only two equivalence classes relative to $\mathcal{C}_r$. The one consists of all grammars which generate the empty language. Moreover it is not difficult to show that each other grammar is equivalent to the grammar which consists of a single symbol s and the single production (s, Λ). However, there are many subcategories of $\mathcal{C}_r$ which likewise have the reduction property and provide non-trivial equivalences of grammars. We give two examples.

A homomorphism $\varphi: G \to \bar{G}$ is to be called strongly length-preserving, if φ maps symbols on symbols and productions on productions, i.e. $\varphi(a) \in \bar{O}$ for all a $\in O$ and $\varphi(\alpha) \in \bar{P}$ for all $\alpha \in P$.

Let $\mathcal{C}_1$ be the category whose set of objects is the set of all Chomsky grammars and whose set of morphisms is the set of all strongly length-preserving homomorphisms.

Furthermore we consider the category $\mathcal{C}_1^T$ whose set of objects is the set of all Chomsky grammars with the distinguished terminal alphabet T and whose morphisms are all those morphisms of $\mathcal{C}_1$ which are identical in T.

The equivalence relative to $\mathcal{C}_1^T$ is a very strong relationship and implies that the languages generated by equivalent grammars coincide.

It is easy to verify that the categories $\mathcal{C}_1$ and $\mathcal{C}_1^T$ have the reduction property. The proofs differ only slightly from that given concerning $\mathcal{C}_r$ and are to be left to the reader. We will only state how the products in $\mathcal{C}_1$ and $\mathcal{C}_1^T$ can be represented. Let $G_1 \xleftarrow{g^1} G \xrightarrow{g^2} G_2$ be the product in $\mathcal{C}_r$ constructed above.

The product $G_1 \xleftarrow{f^1} \bar{G} \xrightarrow{f^2} G_2$ in $\mathcal{C}_1$ can be constructed as follows: $\bar{O} = O_1 \times O_2 \subset O$, $\bar{T} = T_1 \times T_2 \subset T$, $\bar{S} = S_1 \times S_2 = S$,

$$\bar{P} = \{\alpha \in P \mid g^i(\alpha) \in P_i \quad \text{for} \quad i = 1, 2, \quad \text{and} \quad \alpha \in \bar{O}^* \times \bar{O}^*\}.$$

There is a natural inclusion $\bar{G} \overset{j}{\subset} G$ and the homomorphisms f^i are to be the compositions $f^i = g^i j$ for $i = 1, 2$.

As regards the product $G_1 \overset{h^1}{\leftarrow} \tilde{G} \overset{h^2}{\rightarrow} G_2$ in $\mathcal{C}_1^{\tilde{T}}$ we know that $T_1 = T_2 = \tilde{T}$. We define $\tilde{O} = O_1 \times O_2 - \tilde{T} \times \tilde{T} \cup \tilde{T}$. We identify an element $(a, a) \in O_1 \times O_2$ with a so that $\tilde{O}$ is a subset of the alphabet O of G. $\tilde{S} = S$ and $\tilde{P} = \{\alpha \in \bar{P} \mid \alpha \in \tilde{O}^* \times \tilde{O}^*\}$ complete the definition of G. Again there is an inclusion $\tilde{G} \overset{\tau}{\supset} G$ and the homomorphisms h^i are to be the compositions $h^i = g^i \tau$ for $i = 1, 2$.

So products in $\mathcal{C}_1$ and $\mathcal{C}_1^T$ respectively are, roughly speaking, restrictions of the product in $\mathcal{C}_r$. On the other hand every such restriction of a product in $\mathcal{C}_r$ can be considered as a product in some subcategory of $\mathcal{C}_r$ which has the reduction property.

7. CONTEXT-FREE GRAMMARS

In the preceding chapters no sufficient conditions are derived as to decide the transformational problems. We will now, in a few words, discuss the situation as regards context-sensitive grammars and then consider the essential case of context-free grammars.

A Chomsky grammars $G = (O, T, S, P)$ is called context-sensitive[6], if $1 \leqq |D_0(\alpha)| \leqq |D_1(\alpha)|$ for all $\alpha \in P$. G is called context-free,[7] if $P \subset O \times O^*$.

It is shown in (Schnorr, 1967) that there is no algorithm to decide, given a homomorphism $\varphi : G \rightarrow \bar{G}$ of context-sensitive grammars, whether φ is surjective. And there is no algorithm, even if φ is strictly length-preserving and $\varphi(P) = \bar{P}$. From this it is not difficult to show that the first transformational problem is not recursively solvable regarding the categories $\mathcal{C}_1$ and $\mathcal{C}_1^T$ respectively, when restricting ourselves to context-sensitive grammars. On the other hand the situation is quite favorable concerning context-free grammars.

Algorithms which decide whether a strictly length-preserving homomorphism of context-free grammars is surjective and injective respectively can be derived from the generalized finite automata theory which

[6] Our definition differs slightly from that of Chomsky in that we allow all length increasing rules.

[7] This definition is an insignificant extension of that given by Chomsky, since we allow erasing rules and rules which have terminals on the left side.

is developed in papers of J. B. Wright, J. W. Thatcher and S. Eilenberg. Let M be the $\mathfrak{DS}$ of some context-free grammar.

(7.1) Definition. **A finite automaton $\mathfrak{A}$ relative to the $\mathfrak{DS}$ M is a tripel $\mathfrak{A} = (A, h, a_0)$, where**

(i) A is a finite set of states.
(ii) $a_0 \in A$ is the initial state.
(iii) h is a function which assigns to each production $\alpha \in P$ a map $h_\alpha : A^{|D_1(\alpha)|} \to A^{|D_0(\alpha)|}$. h is called the direct transition function.

h is extended to the transition function which is defined for all $\alpha \in M$. Postulating that the notation $h_\alpha(X)$ always includes $X \in A^{|D_1(\alpha)|}$, h is recursively defined by:

$$h_{\alpha_1 \times \alpha_2}(X_1 X_2) = h_{\alpha_1}(X_1) h_{\alpha_2}(X_2)$$
$$h_{\alpha \circ \beta}(X) = h_\beta(h_\alpha(X))$$

and naturally

$$h_{\mathrm{id}_w}(X) = X$$

So the inputs of the automaton are derivations in M. Each input derivation $\alpha \in M$ produces an output word $g(\alpha) \in A^{|D_0(\alpha)|}$ defined as follows:

$$g(\alpha) = h_\alpha(a_0^{|D_1(\alpha)|}).$$

For any set of final states $A_F \subset A$, the set of derivations recognized by, or the behaviour of, the automaton is defined to be

$$bh_{\mathfrak{A}}(A_F) = \{\alpha \in M \mid g(\alpha) \in A_F\}.$$

There is no essential difference to the concept proposed by J. W. Thatcher (1967), characterizing derivation trees of context-free grammars. We merely profit by the characterization of derivations given in Chapter 2.

(7.2) Proposition. *There is an algorithm for determining, given a strictly length-preserving homomorphism of context-free grammars $\varphi : G \to \bar{G}$, whether φ is surjective.*

This proposition is a consequence of (Thatcher, 1967) Theorems 2, 3 and Proposition 1. We will only sketch the proof and later on we will use the following construction to decide whether a homomorphism is injective.

Proof. We construct an automaton relative to the $\mathfrak{DS}$ $\bar{M}$ associated with $\bar{G}$ which accepts $M_{\bar{G}}$ as well as $\varphi(M_G)$, if appropriate final states have been chosen. We assume that $G = (O, T, S, P)$ and $\bar{G} = (\bar{O}, \bar{T}, \bar{S}, \bar{P})$ have disjoint alphabets, $O \cap \bar{O} = \varnothing$. $\mathfrak{A} = (A, h, a_0)$ is defined as follows.

$$\text{(I)} \qquad A = 2^{O \cup \bar{O}}, a_0 = T \cup \bar{T}$$

To specify the direct transition function we introduce the following notations (a) and (b). Let Y be a set then $(2^Y)^*$ and $2^{(Y^*)}$ are both monoids and there is an injective homomorphism

$$\text{(a)} \qquad \pi_Y : (2^Y)^* \to 2^{(Y^*)} \qquad \text{given by}$$
$$\pi_Y(X) \mapsto X \qquad \text{for all } X \in 2^Y.$$

In the following proof we will write $\pi = \pi_{(O \cup \bar{O})}$.

(b) For $\alpha \in \bar{M}$, $X \in A^{|D_1(\alpha)|}$, and $w \in O^{|D_0(\alpha)|}$

$$R(w, \alpha, X) = \{\beta \in M \mid \varphi(\beta) = \alpha, D_1(\beta) \in \pi(X), D_0(\beta) = w\}.$$

The notation $R(w, \alpha, X)$ shall always imply $X \in A^{|D_1(\alpha)|}$ and $w \in O^{|D_0(\alpha)|}$. For $\alpha \in \bar{P}$ the direct transition function is to be defined by:

$$\text{(II)} \qquad h_\alpha(X) \cap \bar{O} = \begin{cases} \varnothing & \text{if } D_1(\alpha) \notin \pi(X) \\ \{D_0(\alpha)\} & \text{if } D_1(\alpha) \in \pi(X) \end{cases}$$

$$\text{(III)} \qquad h_\alpha(X) \cap O = \{w \in O \mid R(w, \alpha, X) \neq \varnothing\}.$$

We will profit by the following properties of the transition function:

$$\text{(II}'\text{)} \quad \text{for all} \quad \alpha \in \bar{M}, \pi h_\alpha(X) \cap \bar{O}^* = \begin{cases} \varnothing & \text{if } D_1(\alpha) \in \pi(X) \\ \{D_0(\alpha)\} & \text{if } D_1(\alpha) \in \pi(X); \end{cases}$$

$$\text{(III}'\text{)} \quad \text{for all} \quad \alpha \in \bar{M}, \pi h_\alpha(X) \cap O^* = \{w \in O^* \mid R(w, \alpha, X) \neq \varnothing\}.$$

The proof of (II′) is very easy and will be left to the reader. We wil prove (III′) by induction on the length of α.

For $\alpha \in \bar{M}$ there exists either a decomposition $\alpha = \alpha_1 \times \alpha_2$ or α can be written $\alpha = \alpha_2 \circ \alpha_1$ with $\alpha_1 \in \bar{P}$.

If $\alpha = \alpha_1 \times \alpha_2$, we get with an appropriate decomposition $X = X_1X_2$:

$$\pi h_{\alpha_1 \times \alpha_2}(X) = \pi(h_{\alpha_1}(X_1) \cdot h_{\alpha_2}(X)) = \pi h_{\alpha_1}(X_1) \cdot \pi h_{\alpha_2}(X_2).$$

By the induction hypothesis applied to α_1 and α_2 we obtain:

$$\pi h_{\alpha_1 \times \alpha_2}(X) \cap O^* = \{u_1 \mid R(u_1, \alpha_1, X_1) \neq \emptyset\} \cdot \{u_2 \mid R(u_2, \alpha_2, X_2) \neq \emptyset\}$$
$$= \{u_1 u_2 \mid R(u_1 u_2, \alpha_1 \times \alpha_2, X_1 X_2) \neq \emptyset\}.$$

If $\alpha = \alpha_2 \circ \alpha_1$ with $\alpha_1 \in \bar{P}$, it follows by (III):

$$\pi h_{\alpha_2 \circ \alpha_1}(X) \cap O = \pi h_{\alpha_1}(h_{\alpha_2}(X)) \cap O = \{u_1 \mid R(u_1, \alpha_1, h_{\alpha_2}(X)) \neq \emptyset\}.$$

This means, if $u_1 \in \pi h_{\alpha_2 \circ \alpha_1}(X) \cap O$, then there exists $\beta_1 \in M$ with $\varphi(\beta_1) = \alpha_1$, $D_0(\beta_1) = u_1$, and $D_1(\beta_1) \in \pi h_{\alpha_2}(X)$. By the induction hypothesis applied to α_2 it follows $R(D_1(\beta_1), \alpha_2, X) \neq \emptyset$. This means that there exists $\beta_2 \in M$ with $D_0(\beta_2) = D_1(\beta_1)$, $\varphi(\beta_2) = \alpha_2$, and $D_1(\beta_2) \in \pi(X)$. Hence $\beta_2 \circ \beta_1 \in R(u_1, \alpha_2 \circ \alpha_1, X) \neq \emptyset$. To show the converse inclusion let $R(u_1, \alpha_2 \circ \alpha_1, X) \neq \emptyset$ and take $\beta \in R(u_1, \alpha_2 \circ \alpha_1, X)$. Then there is a decomposition $\beta = \beta_2 \circ \beta_1$ such that $\varphi(\beta_i) = \alpha_i$. It follows $\beta_2 \in R(D_1(\beta_1), \alpha_2, X)$. From the induction hypothesis applied to α_2 we infer $D_1(\beta_1) \in \pi h_{\alpha_2}(X)$. From $\beta_1 \in R(u_1, \alpha_1, h_{\alpha_2}(X))$ and from the definition of h (III) we conclude that $u_1 \in \pi h_{\alpha_1}(h_{\alpha_2}(X)) = \pi h_{\alpha_2 \circ \alpha_1}(X)$. Thus, by induction we have established property (III′) which combined with (II′) implies:

($\overline{\text{II}}$) for all $\alpha \in \bar{M}$,

$$\pi g(\alpha) \cap \bar{O}^* = \begin{cases} \emptyset & \text{if } D_1(\alpha) \notin \bar{T}^* \\ \{D_0(\alpha)\} & \text{if } D_1(\alpha) \in \bar{T}^* \end{cases}$$

($\overline{\text{III}}$) for all $\alpha \in \bar{M}$, $\pi g(\alpha) \cap O^*$

$$= \left\{ w \in O \,\middle|\, \begin{matrix} \text{ex. } \beta \in M, D_0(\beta) = w \\ D_1(\beta) \in T^*, \text{ and } \varphi(\beta) = \alpha \end{matrix} \right\}$$

We now choose the sets of final states A_1 and A_2 with the definitions

$$A_1 = \{H \subset O \cup \bar{O} \mid H \cap \bar{S} \neq \bar{O}\}$$
$$A_2 = \{H \subset O \cup \bar{O} \mid H \cap S \neq \emptyset\}$$

It is now easy to verify by ($\overline{\text{II}}$) and ($\overline{\text{III}}$) respectively that

$$bh_a(A_1) = M_{\bar{G}},$$
$$bh_a(A_2) = \varphi(M_G).$$

To determine whether φ is surjective we consider the set $\bar{A} = A_1 \cup A_2 -$

$A_1 \cap A_2$. Because of $M_{\bar{G}} \cup \varphi(M_G) - M_{\bar{G}} \cap \varphi(M_G) = bh_{\mathfrak{A}}(\bar{A})$ φ is surjective, if and only if $bh_{\mathfrak{A}}(\bar{A}) = \emptyset$. (7.2) follows now from the fact that there is an effective procedure for determining whether or not $bh_{\mathfrak{A}}(\bar{A}) = \emptyset$, (Thatcher and Wright, 1968) Theorem 7.

The presupposition in (7.2) that φ is to be strictly length-preserving can be dropped, but we will restrict ourselves within the purpose of this paper to reducing homomorphisms.

(7.3) COROLLARY. *Proposition (7.2) holds also as regards reducing homomorphisms.*

Proof. Let $\varphi: G \to \bar{G}$ be a reducing homomorphism of context-free grammars. We construct a strictly length-preserving homomorphism $\varphi': G' \to \bar{G}$ which satisfies $\varphi(M_G) = \varphi'(M_{G'})$. We define $G' = (O', T', S', P')$ by

$$O' = \{a \in O \mid \varphi(a) \neq \Lambda\}, \qquad T' = T \cap O', \qquad S' = S,$$

and using the homomorphism $p: O^* \to O'^*$ which is given by

$$p(a) = a \quad (a \in O'), \qquad p(a) = \Lambda \quad (a \in O - O'),$$

$$P' = \{(pD_0(\delta), pD_1(\delta)) \mid \delta \in M, \varphi(\delta) \in \bar{P}\}$$

It is easy to verify that the strictly length-preserving homomorphism $\varphi': G' \to G$ which is induced by $\varphi'(a) = \varphi(a)\, a \in O'$, has the required property $\varphi(M_G) = \varphi'(M_{G'})$.

This corollary and Theorem (4.4) now lead to the main theorem of this paper.

(7.4) THEOREM. *There is an effective procedure for determining, given two context-free grammars, whether they are in the same transformational class relative to $\mathfrak{C}_1$ and $\mathfrak{C}_1^T$ respectively.*

Example. Let $\mathfrak{C}_1^T$ be the category in consideration. For simplicity reasons the example relates to linear grammars, i.e. $P \subset (O - T) \times T^*(O - T)\, T^* \cup (O - T) \times T^*$. The grammars are illustrated by diagrams whose edges are either nonterminal symbols or Λ and the arrows stand for productions. $u_1 \underset{i}{\to} v_1$ means a production (u_1, w_1) with $w_1 \in T^* v_1 T^*$,i.e. $u_1 \underset{i}{\to} v_1$ stands for a so-called terminal production, iff $v_1 = \Lambda$. The indices i and j below the arrows in $u_1 \underset{i}{\to} v_1$, $u_2 \underset{j}{\to} v_2$ have to coincide, iff the productions α_1, α_2 resp. which the arrows represent differ only in nonterminal symbols, i.e. $\alpha_1 = (u_1, yv_1w)$, $\alpha_2 = (u_2, yv_2w)$.

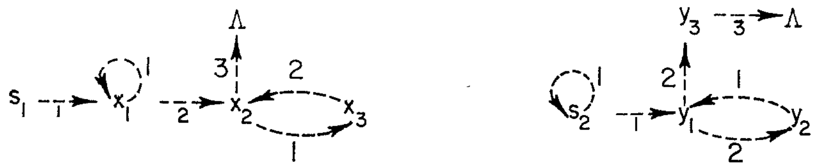

FIG. 2

G_1 with the initial symbol s_1. G_2 with the initial symbol s_2.

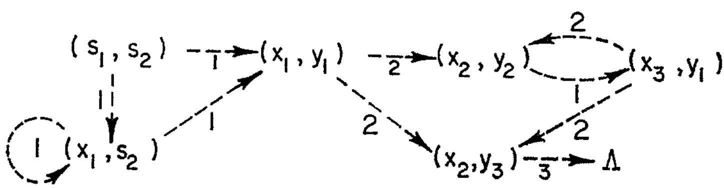

FIG. 3

The product grammar of G_1 and G_2 with the initial symbol (s_1, s_2).

The reader will immediately verify, without applying the algorithm (7.2), that the homomorphisms of the associated product diagram are surjective as well as injective, i.e. G_1 and G_2 are strictly equivalent relative to $\mathfrak{C}_1^T$.

8. THE SECOND TRANSFORMATIONAL PROBLEM

The second transformational problem seems to be more complicated than the first one. We will restrict our investigations to the category $\mathfrak{C}_1$ and presuppose throughout this chapter that every production of any grammar satisfies $D_0(\alpha) \neq \Lambda$. Thus we can profit by the Proposition (2.8). Nevertheless we shall not completely solve the problem, but merely transform it into another which seems to be simpler.

At first we will strengthen corollary (5.7) as regards the category $\mathfrak{C}_1$.

(8.1) PROPOSITION. *If*

$$\begin{tikzcd} & \bar{G} \arrow[dl, dashed, "g^1"'] \arrow[dr, dashed, "g^2"] & \\ G_1 \arrow[dr, "f^1"'] & & G_2 \arrow[dl, dashed, "f^2"] \\ & G & \end{tikzcd}$$

is a pullback in $\mathfrak{C}_1$, *then for every pair of derivations* α_1 *of* G_i *such that* $f^1(\alpha_1) = f^2(\alpha_2)$, *there exists a unique derivation* α *of* G *such that* $g^i(\alpha) = \alpha_i$ *for* $i = 1, 2$.

Proof. Consider the following specification of the above pullback to be stated without proof.

$G_1 \xleftarrow{g^1} \bar{G} \xrightarrow{g^2} G_2$ is to be defined by:

$$\bar{O} = \{(a_1, a_2) \in O_1 \times O_2 \mid f^1(a_1) = f^2(a_2)\},$$

$$\bar{T} = \bar{O} \cap T_1 \times T_2, \qquad \bar{S} = S_1 \times S_2,$$

and using the projections $p_i: \bar{O}^* \to O_i^*$

$$\bar{P} = \{(w, v) \in \bar{O} \times \bar{O} \mid (p_i(w), p_i(v)) \in P_i \quad i = 1, 2\}$$

The homomorphisms $g^i: \bar{G} \to G_i$ are induced by the projections p_i by the formulas:

$$g^i(a) = p_i(a) \qquad (a \in \bar{O})$$

$$g^i(w, v) = (p_i(w), p_i(v)) \qquad ((w, v) \in \bar{P}).$$

If $a_i \in O_i$ for $i = 1, 2$ and $f^1(a_1) = f^2(a_2)$, then there exists a unique $a \in \bar{O}$ such that $g^i(a) = a_i$ for $i = 1, 2$, namely, $a = (a_1, a_2)$. Moreover, if $\alpha_i \in P_i$ for $i = 1, 2$ and $f^1(\alpha_1) = f^2(\alpha_2)$, then there exists a unique $\alpha \in \bar{P}$ such that $g^i(\alpha) = \alpha_i$ for $i = 1, 2$. The assertion of (8.1) is now obvious, if $\alpha_i \in E_i^* \times P \times E_i^*$. For arbitrary α_i let $\alpha_i = \alpha_1^i \circ \alpha_2^i \circ \cdots \circ \alpha_{n_i}^i$ for $i = 1, 2$ be the canonical representations of α_i. Since the homomorphisms f^i are strictly length-preserving it follows that

$$f^i(\alpha_i) = f^i(\alpha_1^i) \circ f^i(\alpha_2^i) \circ \cdots \circ f^i(\alpha_{n_i}^i) \qquad i = 1, 2$$

are canonical representations, too. We infer from (2.8) that $n_1 = n_2 = n$ and $f^1(\alpha_\mu^1) = f^2(\alpha_\mu^2)$ $\mu = 1, \cdots, n$. The assertion now follows immediately by induction on n.

The proposition (8.1) implies the following corollary analogous to (5.8).

(8.2) Corollary. *Relative to the pullback*

$$\begin{tikzcd} & \bar{G} \arrow[dl, "g^1"'] \arrow[dr, "g^2"] & \\ G_1 \arrow[dr, "f^1"'] & & G_2 \arrow[dl, "f^2"] \\ & G & \end{tikzcd}$$

in $\mathcal{C}_1$, *if* f^1 *is injective, then so is* g^2.

Proof. In contradiction to the assertion we assume that there exist two distinct derivations α, $\bar{\alpha} \in M_{\bar{G}}$ such that $g^2(\alpha) = g^2(\bar{\alpha})$. Since $g^1(\alpha) = g^1(\bar{\alpha})$ would be in contradiction to the term 'unique' in (8.1), it follows $g^1(\alpha) \neq g^1(\bar{\alpha})$. The commutativity of the pullback ensures $f^1g^1(\alpha) = f^1g^1(\bar{\alpha})$, contrary to the hypothesis that f^1 is injective.

(8.3) THEOREM. *Let $\mathcal{C}$ be a subcategory of $\mathcal{C}_1$ which has the reduction property. Two grammars G_1 and G_2 are strictly equivalent relative to $\mathcal{C}$, if and only if there exists a diagram $G_1 \leftarrow G \rightarrow G_2$ of bijective homomorphisms in $\mathcal{C}$.*

Obviously (8.2) holds for any subcategory of $\mathcal{C}_1$ having pullbacks. From this we infer that an arbitrary chain of bijective homomorphisms in $\mathcal{C}$ can be reduced to a chain consisting of two homomorphisms only.

Theorem (8.3) gives a hint to solve the second transformational problem in $\mathcal{C}$. Let $G_1 \xleftarrow{f^1} G \xrightarrow{f^2} G_2$ consist of bijective homomorphisms and let $G_1 \xleftarrow{h^1} \bar{G} \xrightarrow{h^2} G_2$ be the product in $\mathcal{C}$, then it follows that f is injective relative to the commutative diagram

$$\begin{tikzcd} & G \arrow[dl, dashed, "f^1"'] \arrow[d, dashed, "f"] \arrow[dr, dashed, "f^2"] & \\ G_1 & \bar{G} \arrow[l, dashed, "h^1"] \arrow[r, dashed, "h^2"'] & G_2 \end{tikzcd}$$

The homomorphisms h^i provide a one to one correspondence of $f(M_G)$ and M_{G_i} for $i = 1, 2$.

When restricting on context-free grammars, $f(M_G)$ is recognizable. On the other hand, given a recognizable set $R \subset M_{\bar{G}}$ it is easy to construct an injective homomorphism $\varphi: G' \rightarrow \bar{G}$ such that $\varphi(M_{G'}) = R$. Hence the second transformational problem relative to context-free grammars amounts to search for a recognizable set $R \subset M_{\bar{G}}$ so that the homomorphisms h^i provide a one to one correspondence of R and M_{G_i} for $i = 1, 2$. In doing so we can profit by the following theorem.

(8.4) THEOREM. *There is an effwxricw peoxwsuew doe swrwemininf, given a strictly length-preserving homomorphism $\varphi: G \rightarrow \bar{G}$ of context-free grammars, whether φ is injective.*

Proof. We construct an automaton relative to the $\mathfrak{DS}$ $\bar{M}$ associated with $\bar{G}$ which accepts exactly those derivations in $M_{\bar{G}}$ which have more than one inverse image in M_G. $\mathfrak{A} = (A, h, a_0)$ is defined as follows.

Let the set O' be in one to one correspondence with the alphabet O of G. a' is to be the corresponding element of $a \in O$.

(I) $$A = 2^{O \cup O'}, \qquad a_0 = T.$$

To specify the direct transition function we will use the homomorphism $\pi: (2^{(O \cup O')})^* \to 2^{((O \cup O')^*)}$ (7.2) (a) as well as the notation (7.2) (b): for $\alpha \in M$, $X \in A^{|D_1(\alpha)|}$ and $w \in O^{|D_1(\alpha)|}$,

$$R(w, \alpha, X) = \{\beta \in M \mid D_0(\beta) = w, D_1(\beta) \in \pi(X), \varphi(\beta) = \alpha\}.$$

The transition function is now given by the relations (II)–(III).

(II) for $\alpha \in \bar{P}$, $h_\alpha(X) \cap O = \{a \in O \mid R(a, \alpha, X) \neq \emptyset\}$;

(III) a' is to be contained in $h_\alpha(X) \cap O'$, iff either of the following relations holds:

(a) $|R(a, \alpha, X)| > 1$,

(b) there exists $\beta \in R(a, \alpha, X)$ such that $D_0(\beta)$ can be written $D_0(\beta) = wuv$ with $wu'v \in \pi(X)$.

The proof of (8.4) now depends on the following properties of the transition function.

(II′) for all $\alpha \in \bar{M}$, $\pi h_\alpha(X) \cap O^* = \{w \in O^* \mid R(w, \alpha, X) \neq \emptyset\}$;

(III′) for all $\alpha \in \bar{M}$ with $D_0(\alpha) \in \bar{O}$,

$$g(\alpha) \cap O' = \{a' \in O' \mid |R(a, \alpha, T^{|D_1(\alpha)|})| > 1\}.$$

Because of the similarity of (II) and (II′) above with (III) and (III′) in the proof of (7.2) it should be obvious that (II′) above holds by the same reasons which proved (III′) (7.2) to be true.

Both inclusions of (III′) are proved by induction on the length of α. „$\subset$" Let $\alpha = \alpha_2 \circ \alpha_1$ with $\alpha_1 \in \bar{P}$. If $a' \in g(\alpha)$, then $a' \in h_{\alpha_1} g(\alpha_2)$ and by the definition of the direct transition function either (a) or (b) holds in (III) above. At first we assume that (a) holds. This means $|R(a, \alpha_1, g(\alpha_2))| > 1$. Take $\beta_1, \beta_2 \in R(a, \alpha_1, g(\alpha_2))$ with $\beta_1 \neq \beta_2$. Because of $D_1(\beta_i) \in \pi g(\alpha_2)$ we infer from (II′) that $R(D_1(\beta_i), \alpha_2, T^{|D_1(\alpha)|}) \neq \emptyset$. Take $\gamma_i \in R(D_1(\beta_i), \alpha_2, T^{|D_1(\alpha)|})$. It is easy to verify that $\gamma_i \circ \beta_i \in R(a, \alpha, T^{|D_1(\alpha)|})$ for $i = 1, 2$. Hence $|R(a, \alpha, T^{|D_1(\alpha)|})| > 1$.

Supposing now that $a' \in h_{\alpha_1} g(\alpha_2)$ and (b) holds in (III) we know that there exists $\beta \in R(a, \alpha_1, g(\alpha_2))$ such that $D_1(\beta)$ can be written

$D_1(\beta) = w_1w_2 \cdots w_n$ with $w_i \in O$ and $w_1 \cdots w_{r-1}w_r'w_{r+1} \cdots w_n \in \pi g(\alpha_2)$. Because of $D_1(\beta) \in \pi g(\alpha_2)$ it follows from (II′) that $R(D_1(\beta), \alpha_2, T^{|D_1(\alpha)|}) \neq \emptyset$. Look at the decompositions $\pi g(\alpha_2) = \prod_{i=1}^n g(\delta_i)$, $R(D_1(\beta), \alpha_2, T^{|D_1(\alpha_2)|}) = \prod_{i=1}^n \times R(w_i, \delta_i, T^{|D_1(\delta_i)|})$ with $\alpha_2 = \delta_1 \times \cdots \times \delta_n$ and $D_0(\delta_i) \in \bar{O}$ From $w_r' \in g(\delta_r)$ and from the induction hypothesis applied to δ_r we infer $|R(w_r, \delta_r, T^{|D_1(\delta_r)|})| > 1$. Hence $|R(D_1(\beta), \alpha_2, T^{|D_1(\alpha)|})| > 1$. From this and $\beta \in R(a, \alpha_1, g(\alpha_2))$ we conclude $|R(a, \alpha_2 \circ \alpha_1, T^{|D_1(\alpha)|})| > 1$. It remains the converse inclusion of (III′) to be proved.

„$\supset$" Suppose that $\alpha = \alpha_2 \circ \alpha_1$ with $\alpha_1 \in \bar{P}$ and $|R(a, \alpha, T^{|D_1(\alpha)|})| > 1$. Take $\gamma, \beta \in R(a, \alpha, T^{|D_1(\alpha)|})$ with $\beta \neq \gamma$. There are decompositions $\beta = \beta_2 \circ \beta_1$ and $\gamma = \gamma_2 \circ \gamma_1$ such that $\varphi(\beta_i) = \varphi(\gamma_i) = \alpha_i$. If $\beta_1 \neq \gamma_1$, then $|R(a, \alpha_1\ g(\alpha_2))| > 1$; hence by (III) (a) $a' \in h_{\alpha_1}(g(\alpha_2)) = g(\alpha)$.

Let now $\beta_1 = \gamma_1$ and $\beta_2 \neq \gamma_2$. We denote $w = D_1(\beta_1)$. Look at the decomposition

$$R(w, \alpha_2, T^{|D_1(\alpha)|}) = \prod_{i=1}^{n} \times R(w_i, \delta_i, T^{|D_1(\delta_i)|}).$$

with $w = w_1w_2 \cdots w_n$, $\alpha_2 = \delta_1 \times \cdots \times \delta_n$, and $D_0(\delta_i) \in \bar{O}$. Because of $\beta_2, \gamma_2 \in R(w, \alpha_2, T^{|D_1(\alpha)|})$ there exists r such that $|R(w_r, \delta_r, T^{|D_1(\delta_r)|})| > 1$. From the induction hypothesis applied to δ_r we infer $w_r' \in g(\delta_r)$. It is now easy to verify that the relation (III)(b) holds as regards $\beta_1 \in R(a, \alpha_1, g(\alpha_2))$; hence $a' \in g(a)$. Thus we have proved (III′).

Taking into account the relations (II′) and (III′) it is not difficult to choose a set of final states A_F such that $bh_G(A_F) = \{\alpha \in M_{\bar{G}} \mid \cdot \mid \varphi^{-1}(\alpha) \cap M_G \mid > 1\}$. Because φ is injective, if and only if $bh_G(A_F) = \emptyset$, Theorem (8.4) follows from the fact that there is an effective solution of the emptiness problem (Thatcher and Wright, 1968, Theorem 7).

It is easy to verify that A_F can be chosen as follows: A subset $B \subset O \cup O'$ is to be an element of A_F, iff either of the following relations hold:

(i) $B \cap O' \neq \emptyset$,

(ii) $|B \cap S| > 1$.

RECEIVED: September 30, 1968

REFERENCES

Bar-Hillel, Y., Perles, M., and Shamir, E. (1961), On formal properties of simple phrase structure grammars. *Zeit. für Phonetik, Sprachw. und Kommunik.* **14**, 143–172.

Chomsky, N., and Schützenberger, M. P. (1963), The algebraic theory of context-free languages. In "Computer Programming and Formal Systems," Braffort P., and Hirschberg, P. Eds. North-Holland Publishing Co., Amsterdam, 118–161.

Eilenberg, S., and Wright, J. B. (1967), Automata in general algebras. *Inform. and Control.* **11**, 452–471.

Ginsburg, S. (1966), "The mathematical theory of context-free languages." McGraw-Hill Book Company, New York.

Griffiths, T. V., Some remarks on derivations in general rewriting systems. *Inform. and Control* **12**, 27–54.

Hotz, G. (1966), Eindeutigkeit und Mehrdeutigkeit formaler Sprachen. *EIK.* **2**, 235–247.

Hotz, G. (1968a), Übertragung automatentheoretischer Sätze auf Chomskysprachen. Main lecture, Scientific meeting, GAMM Praha 1.–5.4. To appear in *Computing.*

Hotz, G. (1968b), Reduktionssätze über eine Klasse formaler Sprachen mit endlich vielen Zuständen. *Math. Zeitschrift.* **104**, 205–221.

Mitchell, B. (1965), "Theory of Categories." Academic Press, New York.

Schnorr, C. P., Reguläre Untergrammatiken. 4. Colloquium über Automatentheory, München 5.–6.10.1967.

Schnorr, C. P., Vier Entscheidbarkeitsprobleme für kontext-sensitive Sprachen. To appear in *Computing.*

Schnorr, C. P., and Walter, H., Pullbackkonstruktionen in Semi-Thue-Systemen. To appear in *EIK.*

Thatcher, J. W. (1967), Characterizing derivation trees of context-free grammars through a generalization of finite automata theory. *J. Comp. System Sciences* **1**, 317–323.

Thatcher, J. W., and Wright, J. B. (1968), Generalized finite automata theory with an application to a decision problem of second order logic. *Mathematical System Theory.* **2**, 57–83.

Paul, M. (1962), Zur Struktur formaler Sprachen. Dissertation Universität Mainz.

Eickel, J., Paul, M., Bauer, F. L., and Samuelson, K. (1963), A Syntax controlled generator of formal language processors. *Comm. ACM* **6**, 451–455.