AD 668 087

DERIVATION-BOUNDED LANGUAGES

Seymour Ginsburg, et al

System Development Corporation
Santa Monica, California

8 January 1968

TM-738/041/00

Derivation-Bounded Languages

Seymour Ginsburg   Edwin H. Spanier

SCIENTIFIC REPORT NO. 14

8 January 1968

Contract Monitor Thomas V. Griffiths
Data Sciences Labortory
Monitored by: R. Swanson, SRIR (AFOSR)

Supported by SDC's Independent Research Program;
Contract F1962867C0008, Programming (Algorithmic) Languages, Project No. 5632,
Task No. 563205; and Grant No. AF-AFOSR- 1203-67

Prepared for
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH, USAF
BEDFORD, MASSACHUSETTS
and
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
OFFICE OF AEROSPACE RESEARCH, USAF
ARLINGTON, VIRGINIA

42

Work Unit No. 56320501

AFCRL-68-0080

TM-738/041/00

SDC
TM SERIES

The work reported herein was supported by SDC and Contract F1962867C0008, Programming (Algorithmic) Languages; Project No. 5632, Task No. 563205; Work Unit No. 56320501; and Grant No. AF-AFOSR-1203-67.

SCIENTIFIC REPORT NO. 14

Derivation-Bounded Languages

by
Seymour Ginsburg*
Edwin H. Spanier**

8 January 1968

Monitored by: Contract Monitor
Thomas V. Griffiths
Data Sciences Laboratory
R. Swanson, SRIR (AFOSR)

Prepared for
AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH, USAF
BEDFORD, MASSACHUSETTS
and
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
OFFICE OF AEROSPACE RESEARCH, USAF
ARLINGTON, VIRGINIA

SYSTEM
DEVELOPMENT
CORPORATION
2500 COLORADO AVE.
SANTA MONICA
CALIFORNIA
90406

*System Development Corporation and University of Southern California

**University of California at Berkeley and Consultant, System Development Corporation

DERIVATION-BOUNDED LANGUAGES*

INTRODUCTION

In [6] the authors studied sets generated by the imposition of certain restrictions on the use of the rewriting rules in a phrase-structure grammar. The present paper is related to [6] in that again we examine sets generated by derivations in a phrase-structure grammar which are restricted. In particular, call a derivation "k-bounded" if each word in the derivation contains at most k occurrences of nonterminals. For a given grammar G and a positive integer k, let $L_k(G)$ denote those words in the language generated by G which have at least one k-bounded derivation. Such sets $L_k(G)$ are called "derivation bounded" and are the objects of study in the paper.

A nonterminal bounded grammar [1] is a context-free grammar G for which there exists a positive integer k such that <u>every</u> derivation in G is k-bounded. Since such grammars define the family of ultralinear languages [5], every ultralinear language is a derivation-bounded set (but not conversely). Thus the definition of derivation-bounded set extends that of ultralinear language in two ways. Firstly, arbitrary phrase-structure grammars (not just context-free grammars) are considered. And secondly, the set of all words generated by <u>some</u> k-bounded derivation is considered.

The main result is that every derivation-bounded set is a context-free language. In case G is a context-free grammar, it is not surprising that $L_k(G)$ is a context-free language for every k. (In fact, a simpler argument can be given in this case than the one given in the paper for an arbitrary phrase-structure grammar.) It is somewhat unexpected, however, that for every <u>phrase-structure</u> grammar G and every integer k, the set $L_k(G)$ is context-free.

---

*Research sponsored in part by the Air Force Cambridge Research Laboratories, Office of Aerospace Research, USAF, under contract F1962867C0008, and by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under AFOSR Grant No. AF-AFOSR-1203-67.

ABSTRACT

A derivation in a phrase-structure grammar is said to be k-bounded if each word in the derivation contains at most k occurrences of nonterminals. A set L is said to be derivation bounded if there exists a phrase-structure grammar G and a positive integer k such that L is the set of words in the language generated by G which have some k-bounded derivation. The main result is that every derivation-bounded set is a context-free language. Various characterizations of the derivation-bounded languages are then given. For example, the derivation-bounded languages coincide with the standard matching-choice sets discussed by Yntema. They also coincide with the smallest family of sets containing the linear context-free languages and closed under arbitrary substitution.

Other interesting results give equivalent characterizations of the family of derivation-bounded languages and suggest that this is a natural family of languages.

The paper is divided into four sections. In section one the family of derivation-bounded sets is introduced and its study reduced to the study of the sets generated by k-bounded derivations in a "weighted context-free grammar." (A "weighted context-free grammar" is a context-free grammar in which every nonterminal is assigned a positive integer as its weight. A "k-bounded derivation" in such a grammar is a derivation in each word of which the sum of the weights of all the occurrences of nonterminals is bounded by k.)

In section two a certain family of derivations, the family of "standard derivations," in a weighted context-free grammar is examined. It is shown that every word in the set generated by a weighted context-free grammar is generated by some standard derivation. Furthermore, a unique factorization of standard derivations as composites of "minimal" standard derivations is established.

In section three notation and terminology are introduced which provide a description of the "weights" of various subwords of words in standard derivations. The main result of the section is a technical one relating various weights in a standard derivation.

In section four the technical results of the earlier sections are used to construct a context-free grammar generating $L_k(G)$. Thus the main result, namely that $L_k(G)$ is context free for every G and every k, is obtained. Various characterizations of the family of derivation-bounded languages are then presented. One of these characterizes the family as generated by "nonexpansive" context-free grammars, a concept introduced elsewhere [9] in another connection. A consequence of this characterization is that the family is a proper subfamily of the family of all context-free languages. Another characterization shows the family as the smallest family containing all linear languages and closed under arbitrary substitution. As a consequence of this characterization, it follows that the family is a (full) AFL. This, in turn, implies that it is undecidable whether an arbitrary context-free language is derivation bounded.

Section 1. Derivation-bounded sets

In an earlier paper [5] we discussed the family of languages generated by nonterminal bounded grammars. Such grammars $G = (V,\Sigma,P,\sigma)$ are context-free[1] and have the property that there exists a positive integer k such that if $\sigma = w_0 \Rightarrow \ldots \Rightarrow w_t$, with w_t in Σ^*, then each w_i contains at most k occurrences of elements of $V-\Sigma$. In the present paper we extend this family of sets in two ways: (a) We allow G to be an arbitrary phrase-structure grammar, and (b) we consider those words w in Σ^* for which there is at least one derivation as described above. In this section we reduce consideration of such phrase-structure grammars to consideration of "weighted" context-free grammars.

[1] The reader is referred to [3] for all undefined terms and symbolism.

Definition. Let $G = (V,\Sigma,P,\sigma)$ be a phrase-structure grammar and k a positive integer. A derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$ is said to be k-bounded if each w_i contains at most k occurrences of elements of $V-\Sigma$. Let $L_k(G)$ be the set of those words w in L(G) for which there exists a k-bounded derivation

$$\sigma = w_0 \Rightarrow \ldots \Rightarrow w_t = w.$$

A set $L \subseteq \Sigma^*$ is said to be derivation bounded if $L = L_k(G)$ for some phrase-structure grammar G and some positive integer k.

It is clear that $L_1(G) \subseteq L_2(G) \subseteq \ldots \subseteq L(G)$ and that $L(G) = \bigcup_{k=1}^{\infty} L_k(G)$.

Example. Let $G = (V,\Sigma,P,\sigma)$, where $\Sigma = \{a,b\}$, $V = \Sigma \cup \{\sigma\}$, and $P = \{\sigma \to a\sigma\sigma, \sigma \to b\}$. Let $\{w_n\}_{n \geq 1}$ be the sequence of words defined inductively by $w_1 = b$ and $w_{i+1} = aw_iw_i$. Then it is easy to see that for all $n > 1$, w_n is in $L_n(G) - L_{n-1}(G)$. Therefore, $L_n(G) \neq L(G)$ for each n (and also $L_n(G) \neq L_m(G)$ for $n \neq m$). Additionally, $L_n(G)$ is finite for every $n \geq 1$.

Observe that if G is nonterminal bounded,(2) then there exists k such that $L(G) = L_k(G)$. This raises the question as to whether every derivation-bounded set is ultralinear. An example is now provided to show that this is not so. It was noted in [5] that the context-free language $L(G) = (L'c)^*$ is not ultralinear, where $G = (V,\Sigma,P,\sigma)$, $\Sigma = \{a,b,c,\}$, $V = \Sigma \cup \{\sigma,\xi\}$, $P = \{\sigma \to \xi c\sigma, \sigma \to \epsilon, \xi \to a\xi b, \xi \to \epsilon\}$, and $L' = \{a^n b^n / n \geq 0\}$. Since every leftmost

(2) A context-free grammar $G = (V,\Sigma,P,\sigma)$ is said to be nonterminal bounded [1] if there exists a positive integer m such that for any word w in V^* such that $\sigma \overset{*}{\Rightarrow} w$ there exists at most m occurrences of symbols of $V-\Sigma$ in w. A set generated by some nonterminal bounded grammar is called an ultralinear language [5].

derivation in G is 2-bounded, $L_2(G) = L(G)$. Thus, $L_2(G)$ is a derivation-bounded set which is not ultralinear.

We now introduce an auxiliary concept with which we shall be concerned.

<u>Definition</u>. A <u>weighted context-free grammar</u> is a 5-tuple $G = (V,\Sigma,P,\sigma,\rho)$, where (V,Σ,P,σ) is a context-free grammar and ρ, called the <u>weight</u> function, is a mapping of $V-\Sigma$ into the positive integers.

The relations $\Rightarrow$ and $\overset{*}{\Rightarrow}$ are defined for a weighted context-free grammar as in a context-free grammar, and L(G) is defined as the set $\{w \text{ in } \Sigma^*/\sigma \overset{*}{\Rightarrow} w\}$.

<u>Definition</u>. Let $G = (V,\Sigma,P,\sigma,\rho)$ be a weighted context-free grammar. Let $\rho(a) = 0$ for each a in Σ, $\rho(\epsilon) = 0$, and for each w in V^*, let $\rho(w)$, called the <u>weight</u> of w, be $\sum_{i=1}^{r} \rho(x_i)$ where $w = x_1 \ldots x_r$, each x_i in V.

Each context-free grammar may be regarded as a weighted context-free grammar in which each element of $V-\Sigma$ has weight one. In this case, the weight of each word is the number of occurrences of variables in it.

<u>Definition</u>. Let $G = (V,\Sigma,P,\sigma,\rho)$ be a weighted context-free grammar and k a positive integer. A derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$ in G is said to be <u>k-bounded</u> if $\rho(w_i) \leq k$ for each w_i. Let $L_k(G)$ be the set of those words w in L(G) for which there exists a k-bounded derivation $\sigma = w_0 \Rightarrow \ldots \Rightarrow w_t = w$.

It is clear that

$$L_1(G) \subseteq L_2(G) \subseteq \ldots \subseteq L(G)$$

and that $L(G) = \bigcup_{k=1}^{\infty} L_k(G)$. Furthermore, $L_k(G)$ as defined for weighted context-free grammars generalizes the previously defined $L_k(G)$ for context-free grammars.

Our interest in weighted context-free grammars is due to the following result.

<u>Lemma 1.1</u>. For every phrase-structure grammar G and every positive integer k there is a weighted context-free grammar G_k' such that $L_k(G) = L_k(G_k')$.

<u>Proof</u>. Let $G = (V,\Sigma,P,\sigma)$. For each non-ϵ word u in $(V-\Sigma)^*$ of length $\le k$, let ξ_u be a distinct symbol and let $V' = \Sigma \cup \{\xi_u / \text{all } u\}$. Define a weight function ρ by $\rho(\xi_u) = |u|$ for each ξ_u. Let P' consist of the following productions (where j, (u,u_1), v_1, and (w_1,w_2) denote arbitrary positive integers, pairs of non-ϵ words in $(V-\Sigma)^*$ of length $\le k$, non-ϵ words in Σ^*, and arbitrary pairs of words in $(V-\Sigma)^*$ such that $|w_1uw_2| \le k$ respectively):

(1) If $u \to \epsilon$ is in P, let $\xi_{w_1uw_2} \to \xi_{w_1w_2}$ be in P'.

(2) If $u \to v_1$ is in P, let $\xi_{w_1uw_2} \to \xi_{w_1}v_1\xi_{w_2}$ be in P'.

(3) If $u \to u_1$ is in P and $|w_1u_1w_2| \le k$, let $\xi_{w_1uw_2} \to \xi_{w_1u_1w_2}$ be in P'.

(4) If $u \to v_1u_1\ldots v_ju_jv_{j+1}$ is in P, let $\xi_{w_1uw_2} \to \xi_{w_1}v_1\xi_{u_1}v_2\ldots\xi_{u_j}v_{j+1}\xi_{w_2}$ be in P'.

(5) If $u \to u_1v_2\ldots u_j$ $(j \ge 2)$ is in P and $|w_1u_1| \le k$ and $|u_jw_2| \le k$, let $\xi_{w_1uw_2} \to \xi_{w_1u_1}v_2\xi_{u_2}v_3\ldots v_j\xi_{u_jw_2}$ be in P'.

(6) If $u \to u_1v_2\ldots u_jv_{j+1}$ is in P and $|w_1u_1| \le k$, let $\xi_{w_1uw_2} \to \xi_{w_1u_1}v_2\xi_{u_2}\ldots\xi_{u_j}v_{j+1}\xi_{w_2}$ be in P'.

(7) If $u \to v_1u_1v_2\ldots u_j$ is in P and $|u_jw_2| \le k$, let $\xi_{w_1uw_2} \to \xi_{w_1}v_1\xi_{u_1}v_2\ldots v_j\xi_{u_jw_2}$ be in P'.

Then $G_k' = (V',\Sigma,P',\sigma,\rho)$ is a weighted context-free grammar and, as is easily seen, $L_k(G) = L_k(G_k')$.

Note that the weighted context-free grammar G_k' has the property that there is a one to one correspondence between all derivations in G_k' and those derivation in G in which each subword of variables in each word of the derivation is of length at most k. In particular, $L(G_k')$ is the set of words in Σ^* which can be derived in G (from σ) by such derivations. This yields

Corollary. Let $G = (V,\Sigma,P,\sigma)$ be a phrase-structure grammar and k a positive integer. The set of all words w in Σ^* with the following property is a context-free language: There exists a derivation $\sigma = w_0 \Rightarrow \ldots \Rightarrow w_t = w$ such that each word of $(V-\Sigma)^*$ which is a subword of some w_i is of length at most k.

It seems reasonable to consider nonterminal bounded phrase-structure grammars. (3) If G is such a phrase-structure grammar and k is such that $\sigma \overset{*}{\Rightarrow} w$ in G implies w contains at most k occurrences of variables, then the grammar G_k' is also nonterminal bounded and $L(G) = L(G_k')$. Therefore languages generated by nonterminal bounded phrase-structure grammars are ultralinear.

Remark. One could extend the notion of a phrase-structure grammar to that of a weighted phrase-structure grammar in the obvious way. With this extension, Lemma 1.1 would be valid if the phrase-structure grammar G in its statement were replaced by a weighted phrase-structure grammar G. The lemma has been stated for the case when the weights are one since our primary interest in this paper is in phrase-structure and context-free grammars, not weighted phrase-structure grammars. We consider weighted context-free grammars only as a tool in studying phrase-structure grammars.

---

(3) The definition of a phrase-structure grammar being nonterminal bounded is the same as for a context-free grammar.

Section 2. Standard derivations

In this section we study derivations in weighted context-free grammars. We first define the concept of a "standard derivation." Then we show that it is possible to rearrange the productions used in any derivation of a terminal word to obtain a standard derivation of that word. Furthermore, if the original derivation is k-bounded for some $k \geq 1$, then so is the rearranged one. (Thus the set $L_k(G)$ consists of words derivable by some k-bounded standard derivation in G.) Finally, we establish a unique factorization of standard derivations as composites of "minimal" standard derivations.

Let $G = (V,\Sigma,P,\sigma,\rho)$ be a weighted context-free grammar and let w_o be a word in V^* with a factorization $w_o = v_1 u_1^{(o)} v_2 \ldots v_s u_s^{(o)} v_{s+1}$, where $s \geq 1$, v_j is in Σ^* for $1 \leq j \leq s+1$, and $u_j^{(o)}$ is in V^* for $1 \leq j \leq s$. As in known [3], given a derivation

$$w_o \Rightarrow \ldots \Rightarrow w_t$$

there are induced factorizations $w_i = v_1 u_1^{(i)} v_2 \ldots v_s u_s^{(i)} v_{s+1}$ such that for each $i, 0 \leq i \leq t-1$, and j, $0 \leq j \leq s$, either $u_j^{(i)} = u_j^{(i+1)}$ or $u_j^{(i)} \Rightarrow u_j^{(i+1)}$. Thus for each j, the distinct words $u_j^{(i)}$ form a corresponding derivation $u_j^{(o)} \Rightarrow \ldots \Rightarrow u_j^{(t)}$.

<u>Definition</u>. Let $w_o \Rightarrow \ldots \Rightarrow w_t$ be a derivation such that $\rho(w_t) < \rho(w_o)$. Let $w_o = v_1 \xi_1 v_2 \ldots v_s \xi_s v_{s+1}$ be a factorization such that $s \geq 1$, v_j is in Σ^* for $1 \leq j \leq s+1$, and ξ_j is in $V-\Sigma$ for $1 \leq j \leq s$. For $0 \leq i \leq t$ let $w_i = v_1 u_1^{(i)} v_2 \ldots v_s u_s^{(i)} v_{s+1}$ be the induced factorization. Since $\rho(w_t) < \rho(w_o)$, there exist integers j, $1 \leq j \leq s$, such that $\rho(u_j^{(t)}) < \rho(\xi_j)$. Hence there is a smallest integer i_o, $0 < i_o \leq t$, such that for some j_o, $1 \leq j_o \leq s$, $\rho(u_{j_o}^{(i_o)}) < \rho(\xi_{j_o})$. In this

case ξ_{j_o} is called the <u>distinguished variable</u> of w_o in the derivation[4] and w_{i_o} is called the <u>distinguisher</u> of w_o in the derivation.

In a derivation $w_o \Rightarrow \ldots \Rightarrow w_t$, with $\rho(w_t) < \rho(w_o)$, if r is an integer such that $\rho(w_r) < \rho(w_o)$, then the distinguisher of w_o (under a given factorization) in the derivation $w_o \Rightarrow \ldots \Rightarrow w_t$ is in the subderivation $w_o \Rightarrow \ldots \Rightarrow w_r$ and is the distinguisher of w_o in this subderivation. Furthermore, the distinguished variable of w_o in the original derivation is also the distinguished variable of w_o in the subderivation.

<u>Definition</u>. Let $w_o \Rightarrow \ldots \Rightarrow w_t, t \geq 1$, be a derivation, with $\rho(w_t) < \rho(w_i)$ for each i, $0 \leq i < t$. The derivation is said to be <u>standard at w_i</u> for $0 \leq i < t$ if w_{i+1} is obtained from w_i by applying a production to the distinguished variable of w_i in the subderivation $w_i \Rightarrow \ldots \Rightarrow w_t$.[5] The derivation is said to be <u>standard</u> if it is standard at w_i for each i, $0 \leq i < t$.

It follows from the definition that if $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t$, $t \geq 2$, is a standard derivation, then $w_1 \Rightarrow \ldots \Rightarrow w_t$ is a standard derivation.

We now show how to rearrange the productions in an arbitrary derivation of a terminal word to obtain a standard derivation of the same word.

<u>Lemma 2.1</u>. Let $w_o \Rightarrow \ldots \Rightarrow w_t, t \geq 1$, be a derivation with $\rho(w_t) < \rho(w_i)$ for all $i, 0 \leq i < t$. Suppose that for some r, $0 \leq r < t$, and all i, $0 \leq i < r$, the derivation is standard at w_i. Suppose that $\rho(w_i) \geq \rho(w_o)$ for each $i, 0 \leq i \leq r$.

---

[4] To be precise, we should actually say that this occurrence of ξ_{j_o} is the occurrence of the distinguished variable of w_o in the derivation.

[5] For each $i < t$, w_i has a unique factorization in the definition of distinguished variable of w_i.

Let $w_o = u\xi u'$, where u and u' are in V^* and ξ is the distinguished variable of w_o. Then for each i, $0 \le i \le r$, $w_i = uu^{(i)}u'$ for some $u^{(i)}$ and the distinguished variable of w_r is in $u^{(r)}$.

Proof. The proof is by induction on r. If $r = 0$, then there is nothing to prove. Assume that $r > 0$ and the result is valid for r-1. Then $w_i = uu^{(i)}u'$ for all i, $0 \le i \le r-1$, and the distinguished variable of w_{r-1} is in $u^{(r-1)}$. Since the derivation is standard at w_{r-1}, $w_r = uu^{(r)}u'$, where $u^{(r-1)} \Rightarrow u^{(r)}$. It only remains to verify that the distinguished variable of w_r is in $u^{(r)}$.

Assume the distinguished variable of w_r is not in $u^{(r)}$ but is in u or u'. Let $w_o = v_1\xi_1 \ldots v_s\xi_s v_{s+1}$, $s \ge 1$, each v_i in Σ^*, and each ξ_j in $V-\Sigma$. Let ξ_k be the distinguished variable of w_o. For each i, let $w_i = v_1u_1^{(i)} \ldots v_su_s^{(i)}v_{s+1}$ be the induced factorization. Then $w_r = v_1u_1^{(r)} \ldots v_su_s^{(r)}v_{s+1}$, with $u_k^{(r)} = u^{(r)}$ and $u_i^{(r)} = \xi_i$ for $i \ne k$. By assumption, the distinguished variable of w_r is ξ_m for some $m, m \ne k$. Let w_q be the distinguisher of w_r. Then $\rho(\xi_m) > \rho(u_m^{(q)})$, $\rho(\xi_j) \le \rho(u_j^{(q)})$ for $j \ne m, j \ne k$, and $\rho(u_k^{(r)}) \le \rho(u_k^{(q)})$. Furthermore, for all $n, r < n < q$, $\rho(\xi_m) \le \rho(u_m^{(n)})$, $\rho(\xi_j) \le \rho(u_j^{(n)})$ for $j \ne m, j \ne k$, and $\rho(u_k^{(r)}) \le \rho(u_k^{(n)})$. Thus w_q is the distinguisher of w_o, and ξ_m, $m \ne k$, is the distinguished variable of w_o. This is a contradiction.

Lemma 2.2. Let $w_o \Rightarrow \ldots \Rightarrow w_t$, w_t in Σ^*, be a derivation such that for some r, $0 \le r < t$, the derivation is standard at w_i for all i, $0 \le i < r$. Then a derivation $w_r = w_r' \Rightarrow \ldots \Rightarrow w_t' = w_t$ can be found, using the same productions with the same frequency as in $w_r \Rightarrow \ldots \Rightarrow w_t$, such that $w_i' = w_i$ for $0 \le i \le r$ and the derivation $w_o' \Rightarrow \ldots \Rightarrow w_t'$ is standard at w_i for all i, $0 \le i \le r$. Furthermore, if, for some k, the derivation $w_o \Rightarrow \ldots \Rightarrow w_t$ is k-bounded, then the derivation $w_o' \Rightarrow \ldots \Rightarrow w_t'$ is also k-bounded.

Proof. Let $w_r = v_1\xi_1 \ldots v_s\xi_s v_{s+1}$, with ξ_j in $V-\Sigma$ for $1 \le j \le s$ and v_j in Σ^* for $1 \le j \le s+1$. For $r \le i \le t$, let $w_i = v_1 u_1^{(i)} \ldots v_s u_s^{(i)} v_{s+1}$ be the induced factorization. Let ξ_{j_o} be the distinguished variable of w_r and let w_q, where $r < q \le t$, be the distinguisher of w_r in $w_r \Rightarrow \ldots \Rightarrow w_t$. Using the productions occurring in $w_r \Rightarrow \ldots \Rightarrow w_q$ and with the same frequency in the same relative order where possible, we obtain an integer p, $r < p \le q$, and a derivation $w_r = w_r' \Rightarrow \ldots \Rightarrow w_q' = w_q$ with the following property: For $r \le i \le q$ let $w_i' = v_1 u_1'^{(i)} \ldots v_s u_s'^{(i)} v_{s+1}$ be the induced factorization. Then the productions are first applied to some variable in $u_{j_o}'^{(i)}$ for all i, $r \le i < p$, and then applied to variables in $u_j^{(i)}$, $j \ne j_o$, for all i, $p \le i < q$.

Let $w_i' = w_i$ if $i \le r$ or $i > q$. We shall show that $w_o' \Rightarrow \ldots \Rightarrow w_t'$ has the desired properties. To prove that the derivation $w_o' \Rightarrow \ldots \Rightarrow w_t'$ is standard at w_i' for all i, $0 \le i \le r$, it suffices to verify that for each i, $0 \le i \le r$, the distinguished variable of w_i' in the derivation $w_i' \Rightarrow \ldots \Rightarrow w_t'$ is also the distinguished variable of w_i in the derivation $w_i \Rightarrow \ldots \Rightarrow w_t$. Thus let i be given, $0 \le i \le r$. Let w_{i_o}' be the distinguisher of w_i' in the derivation $w_i' \Rightarrow \ldots \Rightarrow w_t'$. Several cases arise.

(α) Suppose $i < i_o \le r$. Since $w_j' = w_j$ for $0 \le j \le r$, w_{i_o} obviously is the distinguisher of w_i in $w_i \Rightarrow \ldots \Rightarrow w_t$, and w_i and w_i' have the same distinguished variable.

(β) Suppose $r < i_o < p$. This case cannot occur because of the way the w_j, $r < j < p$, have been defined.

(γ) Suppose $i_o \ge p$. Since $i_o \ge p$, $\rho(w_j) \ge \rho(w_i')$ for each j, $i \le j \le r$. Let $w_i' = w_i = u\xi u'$ where ξ is the distinguished variable of w_i. By Lemma 2.1

applied to the derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$, $w_j = uu^{(j)}u'$ for each j, $i \leq j \leq r$, and the distinguished variable of w_r is in $u^{(r)}$. From the way the derivation $w_r' \Rightarrow \ldots \Rightarrow w_p'$ was defined, it follows that $w_p' = uu^{(p)}u'$ for some $u^{(p)}$. Suppose $i_o = p$. Then $\rho(u^{(p)}) < \rho(\xi)$, so that ξ is the distinguished variable of w_i'. Suppose $i_o \geq q$. Since $w_j = w_j'$ for $j \geq q$, $w_{i_o}' = w_{i_o}$ and ξ is the distinguished variable in $w_i' \Rightarrow \ldots \Rightarrow w_t'$. Suppose $p < i_o < q$ and ξ is not the distinguished variable of w_i'. Then $w_i' = y\upsilon y'$, with υ the distinguished variable of w_i'. Either $y\upsilon$ is an initial subword of u or $\upsilon y'$ is a terminal subword of u'. Assume the former, an analogous argument holding if the latter. Then $w_r = y\upsilon z$ for some z, and the distinguished variable ξ_{j_o} of w_r is in z. Let $w_{i_o}' = y_1 y_2 y_1'$ be the induced factorization by $w_i' = y\upsilon y'$ in the derivation $w_i' \Rightarrow \ldots \Rightarrow w_t'$. Then $\rho(y_2) < \rho(\upsilon)$. From the definition of p and the fact that w_q is the distinguisher of w_r, it follows that the production applied at w_{q-1} is also applied at w_{p-1}'. Since the productions occurring in $w_r \Rightarrow \ldots \Rightarrow w_q$, with the same frequency and in the same relative order, are used to derive $w_r' \Rightarrow \ldots \Rightarrow w_q'$, there exists some i_1, $r < i_1 < q$, such that $w_{i_1} = y_1 y_2 z_2$ for some z_2, is the induced factorization by $w_r = y\upsilon z$. Since $\rho(\upsilon) < \rho(y_2)$, w_q cannot be the distinguisher of w_r, a contradiction.

To show that the new derivation is k-bounded if the original one is, we prove that for each $i, 0 \leq i \leq t$, there is $k(i), 0 \leq k(i) \leq t$, such that $\rho(w_i') \leq \rho(w_{k(i)})$. In case $i \leq r$ or $i \geq q$, let $k(i) = i$. In case $r < i < q$, let $k(i)$, $r < k(i) < q$, be the integer such that the occurrence of the production used to obtain w_i' from w_{i-1}' is the occurrence of the production used to obtain $w_{k(i)}$ from $w_{k(i)-1}$. If $r < i \leq p$, then $u_{j_o}^{(k(i))} = u_{j_o}'^{(i)}$ and, for $j \neq j_o$, $u_j'^{(i)} =$

ξ_j. Since ξ_{j_o} is the distinguished variable of w_r and w_q is the distinguisher of w_r, $\rho(u_j^{(k(i))}) \geq \rho(\xi_j) = \rho(u_j'^{(i)})$ for $j \neq j_o$. In this case, therefore, $\rho(w_i')$ $\leq \rho(w_{k(i)})$. If $p < i < q$, then $u_{j_o}'^{(i)} = u_{j_o}^{(q)}$, whence $\rho(u_{j_o}'^{(i)}) = \rho(u_{j_o}^{(q)})$ $< \rho(u_{j_o}^{(k(i))})$, and $u_j'^{(i)} = u_j^{(k(i))}$ for $j \neq j_o$. Therefore $\rho(w_i') < \rho(w_{k(i)})$ in this case, and the proof is complete.

Using the previous lemma we obtain

Theorem 2.1. Let $w_o \Rightarrow \ldots \Rightarrow w_t$ be a derivation with w_t in Σ^*. Then there is a standard derivation $w_o = w_o' \Rightarrow \ldots \Rightarrow w_t' = w_t$. Furthermore, if, for some k, the original derivation is k-bounded, then the new derivation is k-bounded.

Proof. By Lemma 2.2 and induction on r, $0 \leq r < t$, it follows that there exists a derivation $w_o = w_o^{(r)} \Rightarrow \ldots \Rightarrow w_t^{(r)} = w_t$, k-bounded if $w_o \Rightarrow \ldots \Rightarrow w_t$ is k-bounded, which is standard at w_i for each i, $0 \leq i \leq r$. Then $w_o^{(t-1)} \Rightarrow \ldots \Rightarrow w_t^{(t-1)}$ is a standard derivation satisfying the theorem.

Given a weighted context-free grammar, the last result shows that in considering derivations of terminal words, we may restrict ourselves to standard derivations. The next result shows that standard derivations can be composed to yield a standard derivation.

Theorem 2.2. Let $w_o \Rightarrow \ldots \Rightarrow w_t$ be a derivation such that $\rho(w_t) < \rho(w_j)$ for all j, $0 \leq j < t$. Let r be an integer, $0 < r < t$, such that $\rho(w_r) < \rho(w_i)$ for all i, $0 \leq i < r$. Then the derivation is a standard derivation if and only if both of the subderivations $w_o \Rightarrow \ldots \Rightarrow w_r$ and $w_r \Rightarrow \ldots \Rightarrow w_t$ are standard derivations.

Proof. Since $\rho(w_r) < \rho(w_i)$ for each i, $0 \leq i < r$, it follows that for each k, $0 \leq k < r$, the distinguished variable of w_k in the derivation $w_k \Rightarrow \ldots \Rightarrow w_t$ is also the distinguished variable of w_k in the subderivation $w_k \Rightarrow \ldots \Rightarrow w_r$.

Therefore for each k, $0 \le k < t$, the derivation $w_0 \Rightarrow \ldots \Rightarrow w_t$ is standard at w_k if and only if either (a) $k < r$ and the subderivation $w_0 \Rightarrow \ldots \Rightarrow w_r$ is standard at w_k, or (b) $k \ge r$ and the subderivation $w_r \Rightarrow \ldots \Rightarrow w_t$ is standard at w_k. This gives the result.

Definition. A derivation $w_0 \Rightarrow \ldots \Rightarrow w_t$ is called a standard block if it is a standard derivation such that $\rho(w_i) \ge \rho(w_0)$ for each $i, 0 \le i < t$.

Suppose $w_0 \Rightarrow \ldots \Rightarrow w_t$ is a standard block. Then w_t is the distinguisher of w_0. Futhermore, for each j, $0 < j < t$, since $\rho(w_j) \ge \rho(w_0)$, the subderivation $w_0 \Rightarrow \ldots \Rightarrow w_j$ is not a standard derivation. Hence a standard block is a "minimal" standard derivation.

It follows from Theorem 2.2 that a composite of standard derivations is a standard derivation. We shall prove that each standard derivation is a unique composite of standard blocks.

Lemma 2.3. Let $w_0 \Rightarrow \ldots \Rightarrow w_t$ be a standard derivation and let w_q be the distinguisher of w_0. Then q is the unique integer such that $w_0 \Rightarrow \ldots \Rightarrow w_q$ is a standard block.

Proof. Since $w_0 \Rightarrow \ldots \Rightarrow w_t$ is standard, q is the smallest integer such that $\rho(w_q) < \rho(w_0)$. By Theorem 2.2, $w_0 \Rightarrow \ldots \Rightarrow w_q$ is a standard derivation. Clearly it is a standard block. To complete the proof it suffices to show that there is no other standard block $w_0 \Rightarrow \ldots \Rightarrow w_{q'}$. Thus suppose that $w_0 \Rightarrow \ldots \Rightarrow w_{q'}$ is a standard block. Then $\rho(w_{q'}) < \rho(w_0)$, so that $q \le q'$ by the minimality of q. Since $\rho(w_i) \ge \rho(w_0)$ for $0 \le i < q'$, $q' \le q$. Therefore $q' = q$, so there is no other standard block.

Theorem 2.3. Each standard derivation is a unique composite of standard blocks.

<u>Proof</u>. Let $w_o \Rightarrow \ldots \Rightarrow w_t$ be a standard derivation and let $w_{m_o}, w_{m_1}, \ldots, w_{m_r} = w_t$ be defined as follows. Let $w_{m_o} = w_o$ and for each $j \geq 1$ let w_{m_j} be the distinguisher of $w_{m_{j-1}}$. By induction, Lemma 2.3, and Theorem 2.2, it follows that $w_{m_r} = w_t$ for some r, and that

$$w_{m_o} \Rightarrow \ldots \Rightarrow w_{m_1}, \ \ldots, \ w_{m_{j-1}} \Rightarrow \ldots \Rightarrow w_{m_j}, \ \ldots, \ w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r}$$

is a unique factorization of $w_o \Rightarrow \ldots \Rightarrow w_t$ into standard blocks.

If we regard a standard block as a "prime" (or "irreducible") standard derivation, then Theorem 2.3 gives a unique factorization of any standard derivation into "primes". The next result concerns how "primes" multiply.

<u>Lemma 2.4</u>. Let $w_o = w_o' w_o''$. A derivation $w_o \Rightarrow \ldots \Rightarrow w_t$ is a standard block if and only if either (a) there is a standard block $w_o' \Rightarrow \ldots \Rightarrow w_t'$ such that $w_i = w_i' w_o''$ for each i, $0 \leq i \leq t$, or (b) there is a standard block $w_o'' \Rightarrow \ldots \Rightarrow w_t''$ such that $w_i = w_o' w_i''$ for each i, $0 \leq i \leq t$.

<u>Proof</u>. Let $w_o \Rightarrow \ldots \Rightarrow w_t$ be a standard block and let $w_i = w_i' w_i''$ be the induced factorization for $0 \leq i \leq t$. By Lemma 2.1, if the distinguished variable of w_o is in w_o' (or in w_o''), then the distinguished variable of w_i for $0 \leq i < t$ is in w_i' (or in w_i''). Therefore $w_i'' = w_o''$ (or $w_i' = w_o'$) for $0 \leq i \leq t$, so that $w_o' \Rightarrow \ldots \Rightarrow w_t'$ (or $w_o'' \Rightarrow \ldots \Rightarrow w_t''$) is a standard block such that $w_i = w_i' w_o''$ (or $w_i = w_o' w_i''$) for $0 \leq i \leq t$. Hence either (a) or (b) holds.

Conversely, if either (a) or (b) is satisfied, it is trivial that $w_o \Rightarrow \ldots \Rightarrow w_t$ is a standard block.

<u>Theorem 2.4</u>. Let $w_o = x_{o,1} \ldots x_{o,k}$, where $x_{o,j}$ is in V^* for each j, $1 \leq j \leq k$. A derivation $w_o \Rightarrow \ldots \Rightarrow w_t$ is a standard block if and only if there exist j,

$1 \le j \le k$, and a standard block $x_{o,j} \Rightarrow \ldots \Rightarrow x_{t,j}$ such that $w_i = x_{o,1} \ldots x_{o,j-1}$ $x_{i,j}\, x_{o,j+1} \ldots x_{o,k}$ for each i, $0 \le i \le t$.

Proof. Since the "if" is obvious, it suffices to show the "only if". The argument is by induction on k. For $k = 1$, the "only if" is clearly true. Assume that $k > 1$ and that the "only if" is valid for k-1. Let $w_o' = x_{o,1}$ and $w_o'' = x_{o,2} \ldots x_{o,k}$. By Lemma 2.4, either there exists a standard block $w_o' \Rightarrow \ldots \Rightarrow w_t'$ such that $w_i = w_i'\, w_o'' = w_i'\, x_{o,2} \ldots x_{o,k}$ or there exists a standard block $w_o'' \Rightarrow \ldots \Rightarrow w_t''$ such that $w_i = w_o'\, w_i'' = x_{o,1} w_i''$. In the former case, let $j = 1$ and $x_{i,1} = w_i'$, thereby obtaining the result. Consider the latter case. By the induction hypothesis on the standard block $w_o'' \Rightarrow \ldots \Rightarrow w_t''$, where $w_o'' = x_{o,2} \ldots x_{o,k}$, there exist j, $2 \le j \le k$, and a standard block $x_{o,j} \Rightarrow \ldots \Rightarrow x_{t,j}$ with the desired properties. This completes the induction and the proof.

Corollary. Let $w_o = v_1 \xi_1 \ldots v_s \xi_s v_{s+1}$, $s \ge 1$, each v_i in Σ^* and each ξ_j in $V-\Sigma$. The standard blocks of a standard derivation $w_o \Rightarrow \ldots \Rightarrow w_t$ give rise, in the obvious way, to the standard blocks of the induced derivation $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow$ w_{j,t_j} for each ξ_j.

Proof. Let $w_i \Rightarrow \ldots \Rightarrow w_{i'}$ be a standard block and $w_i = v_1 u_1^{(i)} \ldots v_s u_s^{(i)} v_{s+1}$ the induced factorization. By Theorem 2.4, with $k = 2s+1$, there exist j, $1 \le j \le s$, and a standard block $u_j^{(i)} \Rightarrow \ldots \Rightarrow u_j^{(i')}$ such that $w_q = v_1 u_1^{(i)} \ldots u_{j-1}^{(i)} v_j u_j^{(q)} v_{j+1}$ $\ldots v_s u_s^{(i)} v_{s+1}$ for each q, $i \le q \le i'$. This gives the result.

Section 3. Control functions

In this section we introduce notation and concepts that allow us to consider standard derivations starting with $w = v_1 \xi_1 \ldots v_s \xi_s v_{s+1}$ in which the weights of the words derived from each ξ_j, $1 \le j \le s$, satisfy suitable restrictions.

<u>Definition</u>. Given $s \ge 1$, denote by $\beta = (\beta(1), \ldots, \beta(s))$ an s-tuple of non-negative integers. Let $\le$ be the partial order in the set of such s-tuples defined by $\beta \le \beta'$ if $\beta(j) \le \beta'(j)$ for each j, $1 \le j \le s$. The <u>weight</u> $|\beta|$ of β is defined as the integer $\sum_{j=1}^{s} \beta(j)$.

Thus $|\beta| \ge 0$, and $|\beta| = 0$ if and only if $\beta = (0, \ldots, 0)$. Furthermore, $\beta \le \beta'$ implies $|\beta| \le |\beta'|$, and if $\beta \le \beta'$ then $\beta < \beta'$ if and only if $|\beta| < |\beta'|$. In a weighted context-free grammar s-tuples arise in the following natural manner.

<u>Definition</u>. Let $G = (V, \Sigma, P, \sigma, \rho)$ be a weighted context-free grammar. Given a word w in V^* and a factorization $w = v_1u_1 \ldots v_su_sv_{s+1}$, where $s \ge 1$, v_j is in Σ^* for $1 \le j \le s+1$, and u_j is in V^* for $1 \le j \le s$, let $\beta(w)$ be the s-tuple $(\rho(u_1), \ldots, \rho(u_s))$.

Note that if $w = v_1u_1 \ldots v_su_sv_{s+1}$ is as above and if $w \Rightarrow \ldots \Rightarrow w'$ is a derivation in G, then there is an induced factorization $w' = v_1u_1' \ldots v_su_s'v_{s+1}$ which defines an s-tuple $\beta(w') = (\rho(u_1'), \ldots, \rho(u_s'))$. In particular, given a derivation $w_0 \Rightarrow \ldots \Rightarrow w_t$ in G and a factorization of $w_0 = v_1u_1 \ldots v_su_sv_{s+1}$ there is defined a sequence $\beta(w_0), \ldots, \beta(w_t)$ of s-tuples. Restrictions on the weights of subwords of the w_i will be expressed in terms of $\beta(w_i)$ using the partial ordering of s-tuples.

<u>Definition</u>. Given an s-tuple β, $s \ge 1$, a <u>β-chain</u> B is a simply ordered sequence of s-tuples $\beta_0 > \ldots > \beta_r$ for some $r \ge 0$ such that $\beta_0 = \beta$, $\beta_r = (0, \ldots, 0)$, and for each i, $1 \le i \le r$, β_{i-1} and β_i differ in exactly one coordinate.

Given β, it is obvious that there are only finitely many β-chains. If B is a β-chain $\beta_0, \ldots, \beta_r$ of r+1 terms, then $r \le |\beta|$ and r is at least as large as the number of nonzero coordinates of β.

In case $s = 1$, we denote the 1-tuple (a) by a and observe that $|a| = a$. Then an a-chain A is a sequence of integers $a = a_0 > a_1 > \ldots \quad a_r = 0$. We shall use A to denote an a-chain and B to denote a β-chain, where β is an s-tuple, $s \geq 1$.

Definition. Let $w_0 \Rightarrow \ldots \Rightarrow w_t$ be a standard derivation, with w_t in Σ^*. Let $w_0 = v_1 \xi_1 \ldots v_s \xi_s v_{s+1}$, $s \geq 1$, be a factorization with v_j in Σ^* for $1 \leq j \leq s+1$ and ξ_j in $V-\Sigma$ for $1 \leq j \leq s$. Let $m_0 = 0 < m_1 < \ldots < m_r = t$ be such that the subderivation $w_{m_{j-1}} \Rightarrow \ldots \Rightarrow w_{m_j}$ is a standard block for each j, $1 \leq j \leq r$ (as given in Theorem 2.3). Then the $\beta(w_0)$-chain $\beta_0, \ldots, \beta_r$, where $\beta_j = \beta(w_{m_j})$ for $0 \leq j \leq r$, is called the $\beta(w_0)$-chain defined by the standard derivation.

Note that if $\beta_0, \ldots, \beta_r$ is the $\beta(w_0)$-chain defined by a standard derivation $w_0 \Rightarrow \ldots \Rightarrow w_t$, then the sequence $m_0 = 0 < m_1 < \ldots < m_r = t$ has the property that for each j, m_j is the smallest integer i with $0 \leq i \leq t$ such that $\beta_j = \beta(w_i)$.

Definition. Given an s-tuple β, a β-control function is an ordered pair (B,f), where B is a β-chain $\beta = \beta_0, \ldots, \beta_r$, and f is a function from B to the nonnegative integers such that $f(\beta_0) = |\beta_0|$ and $f(\beta_i) \leq f(\beta_j)$ whenever $i \leq j$.

Obviously f can be regarded as the sequence of integers $|\beta| = f(\beta_0) \leq \ldots \leq f(\beta_r)$.

Given an s-tuple β and a positive integer M, there obviously are only a finite number of β-control functions (B,f) such that $f((0, \ldots, 0)) \leq M$.

We now use β-control functions in conjunction with standard derivations in a particular way.

<u>Definition</u>. Let $w_0 \Rightarrow \ldots \Rightarrow w_t$ be a standard derivation with w_t in Σ^*. Let $w_0 = v_1\xi_1 \ldots v_s\xi_s v_{s+1}$, $s \geq 1$, be a factorization, with v_j in Σ^* for $1 \leq j \leq s+1$ and ξ_j in $V-\Sigma$ for $1 \leq j \leq s$. The derivation is said to be <u>controlled</u> by a $\beta(w_0)$-control function (B,f) if B is the $\beta(w_0)$-chain $\beta_0, \ldots, \beta_r$ defined by the derivation, and if for each j, $0 \leq j \leq r$, the subderivation $w_0 \Rightarrow \ldots \Rightarrow w_{m_j}$ is $f(\beta_j)$-bounded.

Since $f(\beta_{j-1}) \leq f(\beta_j)$ for $1 \leq j \leq r$ and $\rho(w_0) = f(\beta_0)$, the derivation is controlled by (B,f) if and only if $\rho(w_i) \leq f(\beta_j)$ for all $1 \leq j \leq r$ and all i, $m_{j-1} \leq i \leq m_j$.

Given a standard derivation $w_0 \Rightarrow \ldots \Rightarrow w_t$ with w_0 and w_t as above let B be the $\beta(w_0)$-chain defined by the derivation. Let (B,f) be the $\beta(w_0)$-control function defined by $f(\beta_j) = \max\,\{\rho(w_i)/\ 0 \leq i \leq m_j\}$. Then the derivation is controlled by (B,f), and, if it is controlled by any $\beta(w_0)$-control function (B',f'), then $B = B'$ and $f(\beta_j) \leq f'(\beta_j)$ for all j, $0 \leq j \leq r$.

Given a weighted context-free grammar G and a positive integer k, in the next section we shall construct a context-free grammar G' such that $L(G') = L_k(G)$. The variables of G' will be ordered pairs $(\xi, (A,f))$, where ξ is a variable of G and (A,f) is a $\rho(\xi)$-control function such that $f(a_i) \leq k$ for all i. Corresponding to a production $\xi \to v_1\xi_1 \ldots v_s\xi_s v_{s+1}$ of G there will be productions

$$(\xi,(A,f)) \to v_1(\xi_1,(A^{(1)},f^{(1)})) \ldots v_s(\xi_s,(A^{(s)},f^{(s)}))v_{s+1}$$

of G'. We shall introduce another auxiliary concept, the notion of "domination", in order to make the relation between the sequence $\{(A^{(j)}, f^{(j)})\}_{1 \leq j < s}$ and (A,f) precise.

Assume that $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t$ is a standard derivation, with $w_o = \xi$, w_t in Σ^*, and $\xi \to v_1\xi_1 \ldots v_s\xi_s v_{s+1} = w_1$, $s \geq 1$, each v_i in Σ^*, and each ξ_j in $V-\Sigma$. Suppose that $w_1 \Rightarrow \ldots \Rightarrow w_{m_1}$, ..., $w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r}$ are the standard blocks of the standard derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$. If $\rho(w_1) \geq \rho(\xi)$, then $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_{m_1}, \ldots, w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r}$ are the standard blocks of $w_o \Rightarrow \ldots \Rightarrow w_t$. If $\rho(w_1) < \rho(\xi)$, then $w_o \Rightarrow w_1$, $w_1 \Rightarrow \ldots \Rightarrow w_{m_1}$, ..., $w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r}$ are the standard blocks of $w_o \Rightarrow \ldots \Rightarrow w_t$. This fact will be useful and motivates the following definition.

Definition. Let (A,f) be an a-control function, with A having r+1 elements, $r \geq 1$. A β-control function (B,g) is dominated by (A,f) if either

(1) B has r+1 elements, $|\beta| \geq a$, and $|\beta_i| = a_i$ and $g(\beta_i) \leq f(a_i)$ for all $i > 0$;

or (2) B has r elements, and $|\beta_i| = a_{i+1}$ and $g(\beta_i) \leq f(a_{i+1})$ for all $i \geq 0$.

For a given s-tuple β and an a-control function (A,f), there are only finitely many (possibly none) β-control functions (B,g) dominated by (A,f). The next result shows how the concept of domination is used.

Lemma 3.1. Let $w_o = \xi \to v_1\xi_1 \ldots v_s\xi_s v_{s+1} = w_1$ be a production in a weighted context-free grammar $G = (V,\Sigma,P,\sigma,\rho)$, with $s \geq 1$, each v_i in Σ^*, and each ξ_j in $V-\Sigma$. Given w in Σ^*, a derivation $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t = w$ is a standard derivation controlled by a $\rho(\xi)$-control function (A,f) if and only if $w_1 \Rightarrow \ldots \Rightarrow w_t$ is a standard derivation controlled by some $\beta(w_1)$-control function (B,g) dominated by (A,f).

Proof. Assume that $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t = w$ is a standard derivation controlled by (A,f). Let B be the $\beta(w_1)$-chain defined by the standard derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$. Thus B is the sequence of s-tuples $\beta_o, \ldots, \beta_r$, where $w_1 = w_{m_o} \Rightarrow \ldots \Rightarrow w_{m_1}$, $w_{m_{r-1}}$

$\Rightarrow \ldots \Rightarrow w_{m_r} = w_t$ are the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$, and $\beta_j = \beta(w_{m_j})$ for each j. Let (B,g) be the $\beta(w_1)$-control function defined by $g(\beta_j) = \max\{\rho(w_i)/1 \le i \le m_j\}$. As noted earlier, $w_1 \Rightarrow \ldots \Rightarrow w_t$ is controlled by (B,g). To show that (B,g) is dominated by (A,f), we distinguish two cases.

(α) Suppose $\rho(w_1) \ge \rho(w_o)$. Then the standard blocks of the two standard derivations are equal in number, and are identical except for the beginning of the first one. Thus $|\beta_j| = a_j$ and $g(\beta_j) \le f(a_j)$ for $0 < j \le r$, so that (B,g) is dominated by (A,f).

(β) Suppose $\rho(w_1) < \rho(w_o)$. Then A has r+1 terms. Furthermore, the standard blocks of $w_o \Rightarrow \ldots \Rightarrow w_t$ consist of $w_o \Rightarrow w_1$ and the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$. Thus $a_1 = \rho(w_1)$, so that $|\beta_j| = a_{j+1}$ and $g(\beta_j) \le f(a_{j+1})$ for all j, $0 \le j \le r$, so that (B,g) is dominated by (A,f).

To see the converse, assume $w_1 \Rightarrow \ldots \Rightarrow w_t$ is a standard derivation controlled by some $\beta(w_1)$-control function (B,g) dominated by some $\rho(\xi)$-control function (A,f). Then $w_o \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t$ is a standard derivation whose standard blocks are related to the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$ as described in (α) and (β) above. If $\rho(w_1) \ge \rho(w_o)$, then $|\beta_j| = a_j$ and $g(\beta_j) \le f(a_j)$ for all j, $0 < j \le r$. Then $\rho(w_o) < \rho(w_1) < g(\beta_j) \le f(a_j)$ and $\rho(w_i) \le g(\beta_j) \le f(a_j)$ for $1 \le i \le m_j$ and $1 \le j \le r$. Thus $\rho(w_i) \le f(a_j)$ for all w_i in the j-th standard block of $w_o \Rightarrow \ldots \Rightarrow w_t$. If $\rho(w_1) < \rho(w_o)$, then $|\beta_j| = a_{j+1}$ and $g(\beta_j) \le f(a_{j+1})$ for all j, $0 \le j \le r$. Thus $\rho(w_1) < \rho(w_o) = f(a_o) \le f(a_1)$, so that $\rho(w_i) \le f(a_1)$ for all w_i in the first standard block of $w_o \Rightarrow \ldots \Rightarrow w_t$. For all w_i in the j-th standard block of $w_1 \Rightarrow \ldots \Rightarrow w_t$, $1 \le j \le r$, thus all w_i in the j+1-st standard block of $w_o \Rightarrow \ldots \Rightarrow w_t$, we have $\rho(w_i) \le g(\beta_j) \le f(a_{j+1})$. Hence $w_o \Rightarrow \ldots \Rightarrow w_t$ is controlled by (A,f).

The preceding lemma gives one necessary and sufficient condition for the existence of a standard derivation

$$v_1\xi_1 \dots v_s\xi_s v_{s+1} = w_1 \Rightarrow \dots \Rightarrow w_t = w$$

controlled by some $\beta(w_1)$-control function. The next lemma gives another such condition. The two conditions together will then be used to describe the productions in the grammar G' we are seeking.

Definition. Let (B,g) be a β-control function, where B is a β-chain $\beta_o, \dots, \beta_r$ of s-tuples. For each j, $1 \le j \le s$, let $A^{(j)}$ be the chain $a_o^{(j)}, \dots, a_{r_j}^{(j)}$ which consists of the distinct j-th coordinates of $\beta_o, \dots, \beta_r$. (Thus $\sum_j r_j = r$.) For each j and each i, $0 \le i \le r_j$, let $i(j)$ be the smallest integer such that the j-th coordinate of $\beta_{i(j)}$ is $a_i^{(j)}$. For each j, $1 \le j \le s$, let $f^{(j)}$ be the function on $A^{(j)}$ defined by

$$f^{(j)}(a_i^{(j)}) = g(\beta_{i(j)}) - \sum_{j' \ne j} \beta_{i(j)}(j').\text{[6]}$$

Then $(A^{(j)}, f^{(j)})$ is said to be determined by (B,g).

Since $|\beta_{i(j)}| \le g(\beta_{i(j)})$, $f^{(j)}(a_i^{(j)}) \ge 0$. For $i' \le i$, since $g(\beta_{i'(j)}) \le g(\beta_{i(j)})$ and $\beta_{i'(j)} \ge \beta_{i(j)}$, it follows that $f^{(j)}(a_{i'}^{(j)}) \le f^{(j)}(a_i^{(j)})$. Furthermore, $f^{(j)}(a_o^{(j)}) = |\beta_o| - \sum_{j' \ne j} \beta_o(j') = \beta_o(j) = a_o^{(j)}$. Thus $(A^{(j)}, f^{(j)})$ is an $a_o^{(j)}$- control function. In case $s=1$, it is clear that $(A^{(1)}, f^{(1)}) = (B,g)$.

Lemma 3.2. Let B be the β_o-chain $\beta_o, \dots, \beta_r$ and let $w_1 = v_1\xi_1 \dots v_s\xi_s v_{s+1}$ be given, with $s \ge 1$, each v_i in Σ^*, and each ξ_j in $V-\Sigma$. There exists a standard

---

[6] $\beta_{i(j)}(j')$ is the j'-th coordinate of $\beta_{i(j)}$.

derivation $w_1 \Rightarrow \ldots \Rightarrow w_t = w$ of a word w in Σ^* controlled by the $\beta(w_1)$-control function (B,g) if and only if $w = v_1 u_1' \ldots v_s u_s' v_{s+1}$ for some $u_1', \ldots, u_s'$ in Σ^* and for each j, $1 \le j \le s$, there exists a standard derivation $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j}$, with $t_j \le t-1$, controlled by the $\rho(\xi_j)$-control function $(A^{(j)}, f^{(j)})$ determined by (B,g).

Proof. Let $v_1\xi_1 \ldots v_s\xi_s v_{s+1} = w_1 \Rightarrow \ldots \Rightarrow w_t$ be a standard derivation of w in Σ^* controlled by the $\beta(w_1)$-control function (B,g). For each i, $1 \le i \le t$, let $w_i = v_1 u_1^{(i)} \ldots v_s u_s^{(i)} v_{s+1}$ be the induced factorization, with $u_k^{(1)} = \xi_k$ for each k. Let $w_1 = w_{m_o} \Rightarrow \ldots \Rightarrow w_{m_1}, \ldots, w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r} = w_t$ be the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$. Thus $\beta_i = \beta(w_{m_i})$ for each i, $0 \le i \le r$. For each j, $1 \le j \le s$, let $(A^{(j)}, f^{(j)})$ be the control function determined by (B,g), $A^{(j)}$ the sequence $a_o^{(j)}, \ldots, a_r^{(j)}$. We use the notation of the preceding definition. Also, let $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j} = u_j^{(t)}$ be the induced derivation of $\xi_j \overset{*}{\Rightarrow} u_j^{(t)}$. By the corollary to Theorem 2.4, the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$ give rise, in an obvious manner, to the standard blocks of $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j}$. Thus $A^{(j)}$ is defined by $w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j}$, and $t_j \le t-1$.

Let j be an integer, $1 \le j \le s$. Given i, $0 \le i \le r_j$, let i' be the smallest integer such that $a_i^{(j)} = \rho(w_{j,i'})$. Since $A^{(j)}$ is defined by $w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j}$, the integer i' exists. From the correspondance of the standard blocks, it follows that $m_{i(j)}$ is the smallest integer n such that $a_i^{(j)} = \rho(w_{j,i'}) = \rho(u_n^{(j)})$ and that $w_{j,i'} = u_j^{m_{i(j)}}$. Thus, for each k, $0 \le k \le i'$, there exists k', $1 \le k' \le m_{i(j)}$, such that $w_{j,k} = u_j^{(k')}$. Since $k' \le m_{i(j)}$ and $w_1 \Rightarrow \ldots \Rightarrow w_t$ is controlled by (B,g), it follows that $\rho(w_{k'}) \le g(\beta_{i(j)})$. Then

$$\rho(w_{j,k}) + \sum_{j'\neq j} \rho(u_{j'}^{(k')}) = \rho(w_{k'}) \le g(\beta_{i(j)}) = f^{(j)}(a_1^{(j)}) + \sum_{j'\neq j} \beta_{i(j)}(j').$$

Since the derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$ is standard, $\rho(u_j^{(k')}) \ge \beta_{i(j)}(j')$ for $j' \neq j$. Then $\rho(w_{j,k}) \le f^{(j)}(a_1^{(j)})$. Hence the induced derivation $\xi_j \Rightarrow \ldots \Rightarrow u_j^{(t)} = u_j'$ is controlled by $(A^{(j)}, f^{(j)})$.

Consider the converse. For each j, $1 \le j \le s$, let $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j} = u_j'$ be a standard derivation controlled by $(A^{(j)}, f^{(j)})$, where $A^{(j)}$ is the sequence $a_o^{(j)}, \ldots, a_{r_j}^{(j)}$. We obtain a standard derivation $w_1 \Rightarrow \ldots \Rightarrow w_t$ as follows. For each k, $0 \le k \le r$, there exists a unique j, $1 \le j \le s$, and unique i, $0 \le i \le r_j$, such that $k = i(j)$. Let $w_1 \Rightarrow \ldots \Rightarrow w_t$ be the standard derivation such that the $i(j)$-th standard block, $i(j) \ge 1$, corresponds to the i-th standard block of the derivation $\xi_j = w_{j,o} \Rightarrow \ldots \Rightarrow w_{j,t_j} = u_j'$. To complete the proof, it suffices to show that $w_1 \Rightarrow \ldots \Rightarrow w_t$ is controlled by (B,g). It is clear (using induction on k for B_k and the definition of the derivation $w_1 \Rightarrow \ldots \Rightarrow w_k$) that the $\beta(w_1)$-chain defined by the derivation is B itself. Let $w_1 = w_{m_o} \Rightarrow \ldots \Rightarrow w_{m_1}, \ldots, w_{m_{r-1}} \Rightarrow \ldots \Rightarrow w_{m_r} = w_t$ be the standard blocks of $w_1 \Rightarrow \ldots \Rightarrow w_t$. Thus $\beta_i = \beta(w_{m_i})$ for each i, $0 \le i \le r$. Let k and i' be given, $0 \le k \le r$ and $1 \le i' \le m_k$. By induction on k we now prove that $\rho(w_{i'}) \le g(\beta_k)$, thereby completing the proof. If $k = 0$, then $m_k = 1$. The induction is thus started since $\rho(w_1) = g(\beta_o)$. Assume $k > 0$. If $i' \le m_{k-1}$, then, by induction,

$$\rho(w_{i'}) \le g(\beta_{k-1}) \le g(\beta_k).$$

Suppose $i' > m_{k-1}$. There exists a unique j, $1 \le j \le s$, and unique i, $0 \le i \le r_j$, such that $k = i(j)$. Then $w_{i'} = v_1 u_1^{(i')} \ldots v_s u_s^{(i')} v_{s+1}$, where $u_j^{(i')}$ is in the i-th standard block of the derivation $\xi_j \Rightarrow \ldots \Rightarrow u'_j$, and $\rho(u_j^{(i')}) = \beta_{i(j)}(j')$ for $j' \neq j$. Since the derivation $\xi_j \Rightarrow \ldots \Rightarrow u'_j$ is controlled by $(A^{(j)}, f^{(j)})$,

$$\rho(u_j^{(i')}) \leq f^{(j)}(a_i^{(j)}) = g(\beta_{i(j)}) - \sum_{j' \neq j} \beta_{i(j)}(j').$$

Therefore

$$\rho(w_{i'}) = \rho(u_j^{(i')}) + \sum_{j' \neq j} \rho(u_{j'}^{(i')}) \leq g(\beta_{i(j)}) = g(\beta_k)$$

and the induction is extended.

Combining Lemmas 3.1 and 3.2, we obtain the following main result of this section.

__Theorem 3.1.__ Let $w_0 = \xi \to v_1\xi_1 \ldots v_s\xi_s v_{s+1} = w_1$ be a production in a weighted context-free grammar $G = (V, \Sigma, P, \sigma, \rho)$, with $s \geq 1$, each v_i in Σ^*, and each ξ_j in $V-\Sigma$. Given w in Σ^* there exists a standard derivation $w_0 \Rightarrow w_1 \Rightarrow \ldots \Rightarrow w_t = w$ controlled by a $\rho(\xi)$-control function (A,f) if and only if $w = v_1u'_1 \ldots v_su'_sv_{s+1}$ for some $u'_1, \ldots, u'_s$ in Σ^* and for each j, $1 \leq j \leq s$, there exists a standard derivation $\xi_j \Rightarrow \ldots \Rightarrow u'_j$, of length $< t$, controlled by $(A^{(j)}, f^{(j)})$, where $\{(A^{(j)}, f^{(j)})/1 \leq j \leq s\}$ is determined by some $\beta(w_1)$-control function (B,g) dominated by (A,f).

Section 4. Nonexpansive grammars

In this section we finally prove our main result, namely that every derivation-bounded set is context free. The proof will actually show that each derivation-bounded set can be generated by a special type of context-free grammar called "nonexpansive". From this it will follow that the family of all derivation-bounded sets is a proper subfamily of the context-free languages. We also give several characterizations of the family of all derivation-bounded sets.

In order to prove our main result we need the following technical lemma.

Lemma 4.1. Let $G = (V,\Sigma,P,\sigma,\rho)$ be a weighted context-free grammar and let k be a positive integer. Let S be the set of all pairs $(\xi, (A,f))$, where ξ is in $V-\Sigma$ and (A,f) is a $\rho(\xi)$-control function with $f(0) \leq k$. Let $G' = (V',\Sigma,P',\sigma')$, where $V' = \Sigma \cup S \cup \{\sigma'\}$, σ' being a new symbol, and P' consists of the following productions.

(i) $\sigma' \to (\sigma, (A,f))$ for each $(\sigma, (A,f))$ in S.

(ii) $(\xi,(A,f)) \to v_1(\xi_1,(A^{(1)},f^{(1)}))\ldots v_s(\xi_s,(A^{(s)}, f^{(s)}))\, v_{s+1}$ for each production $\xi \to v_1\xi_1\ldots v_s\xi_s v_{s+1}$ in P, where $s\geq 1$, each v_i is in Σ^*, each ξ_j is in $V-\Sigma$, and the set $\{(A^{(j)}, f^{(j)})/1 \leq j \leq s\}$ is determined by some $[\rho(\xi_1)+\ldots+\rho(\xi_s)]$-control function dominated by (A,f).

(iii) $(\xi,(A,f)) \to v$ for each production $\xi \to v$ in P, v in Σ^*, and each (A,f) where A is the sequence $a_0 = \rho(\xi)$, $a_1 = 0$.

Then G' is a context-free grammar and for each $(\xi,(A,f))$ in S, the set $\{w \text{ in } \Sigma^*/ (\xi,(A,f)) \overset{*}{\Rightarrow} w \text{ in } G'\}$ coincides with the set of all w in Σ^* for which there exists a standard derivation $\xi \Rightarrow\ldots\Rightarrow w$ in G controlled by (A,f).

Proof. Obviously S is finite, so that G' is context-free.

We first prove that for each derivation

$$(\xi,(A,f)) = w_0' \Rightarrow\ldots\Rightarrow w_t' = w$$

in G' of a word w in Σ^* there exists a standard derivation $\xi = w_0 \Rightarrow\ldots\Rightarrow w_{t'} = w$ in G controlled by (A,f). Suppose $t = 1$. Then $w_1' = w$ and $(\xi, (A,f)) \to w$ is in P'. This is possible if and only if $\xi \to w$ is in P and A is the sequence $a_0 =$

$\rho(\xi)$, $a_1 = 0$. Since $\xi \Rightarrow w$ is a standard derivation controlled by (A,f), the result is valid for $t = 1$. Using induction, assume $t > 1$ and the result valid for derivations of length $< t$. Suppose that $(\xi,(A,f)) = w'_0 \Rightarrow w'_1 \Rightarrow \ldots \Rightarrow w'_t = w$ is a derivation of length t. Then $w'_0 \rightarrow w'_1$ is a production in P', say

$$(\xi,(A,f)) \rightarrow v_1(\xi_1,(A^1, f^{(1)}))\ldots v_s(\xi_s, (A^{(s)}, f^{(s)}))\, v_{s+1} = w'_1,$$

with $s \geq 1$ since $t > 1$. By (ii), $\xi \rightarrow v_1\xi_1\ldots v_s\xi_s v_{s+1}$ is in P and the set $\{(A^{(j)}, f^{(j)})/1 \leq j \leq s\}$ is determined by some $[\rho(\xi_1)+\ldots+\rho(\xi_s)]$-control function dominated by (A,f). There exist $u'_1, \ldots, u'_s$ in Σ^* such that $w = v_1u'_1\ldots v_s u'_s v_{s+1}$ and, for each j, $1 \leq j \leq s$, $(\xi_j, (A^{(j)}, f^{(j)})) \Rightarrow \ldots \Rightarrow u'_j$. Since each of these derivations has length $< t$, by induction, for each j, $1 \leq j \leq s$, there exists in $\mathcal{G}$ a standard derivation $\xi_j \Rightarrow \ldots \Rightarrow u'_j$ controlled by $(A^{(j)}, f^{(j)})$. By Theorem 3.1, there exists a standard derivation $\xi = w_0 \Rightarrow \ldots \Rightarrow w_t$, controlled by (A,f). Thus the induction is extended.

We now prove that for each standard derivation $\xi = w_0 \Rightarrow \ldots \Rightarrow w_t = w$ in G, w in Σ^*, controlled by (A,f), $(\xi,(A,f)) \overset{*}{\Rightarrow} w$ in G'. Suppose $t = 1$. Then $\xi \rightarrow w$ is in P and A is the sequence $a_0 = \rho(\xi)$, $a_1 = 0$. Therefore $(\xi,(A,f)) \rightarrow w$ is in P'. Using induction, assume $t > 1$ and the result is valid for each standard derivation of length $< t$. Let $\xi = w_0 \Rightarrow \ldots \Rightarrow w_t = w$ be a standard derivation in G, of length t, controlled by (A,f). Then $\xi \rightarrow w_1$ is in P and w_1 is of the form $v_1\xi_1\ldots v_s\xi_s v_{s+1}$, where $s \geq 1$, each v_i is in Σ^*, and each ξ_j is in $V-\Sigma$. By Theorem 3.1, $w = v_1u'_1\ldots v_s u'_s v_{s+1}$ for some $u'_1,\ldots,u'_s$ in Σ^* and for each j, $1 \leq j \leq s$, there exists a standard derivation $\xi_j \Rightarrow \ldots \Rightarrow u'_j$, of length $< t$, controlled by $(A^{(j)}, f^{(j)})$, where $\{(A^{(j)}, f^{(j)})/ 1 \leq j \leq s\}$ is determined by some $[\rho(\xi_1)+\ldots+\rho(\xi_s)]$-control function dominated by (A,f). By induction,

$(\xi_j,(A^{(j)},f^{(j)})) \overset{*}{\Rightarrow} u'_j$ in G' for each j, $1 \le j \le s$. By (ii),

$$(\xi,(A,f)) \rightarrow v_1(\xi_1,(A^{(1)},f^{(1)}))\ldots v_s\ (\xi_s,\ (A^{(s)},\ f^{(s)}))\ v_{s+1}$$

is in P'. Thus

$$\begin{aligned}(\xi,(A,f)) \rightarrow v_1(\xi_1,(A^{(1)},f^{(1)}))\ldots v_s(\xi_s,(A^{(s)},f^{(s)}))\ v_{s+1} \\ \overset{*}{\Rightarrow} v_1u'_1\ \ldots v_su'_sv_{s+1} \\ = w,\end{aligned}$$

completing the induction and the lemma.

We are now ready for the proof of the main result.

<u>Theorem 4.1</u>. Every derivation-bounded set is context free, that is, $L_k(G)$ is a context-free language for every phrase-structure grammar G and every positive k.

<u>Proof</u>. By Lemma 1.1, there exists a weighted context-free grammar $G'=(V',\Sigma,P',\ \sigma',\rho')$ such that $L_k(G) = L_k(G')$. Let G'' be the context-free grammar defined in Lemma 4.1 by means of G' and k. From the definition of G'', $L(G'')$ is the set of all words w in Σ^* for which there exists a standard derivation $\sigma'\Rightarrow\ldots\Rightarrow w$ in G' controlled by some $\rho'(\sigma')$-control function (A,f) with $f(0) \le k$. This set coincides with $L_k(G')$ by Theorem 2.1, since every standard derivation $\sigma'\Rightarrow\ldots\Rightarrow w$ which is k-bounded is controlled by some $\rho'(\sigma')$-control function (A,f) with $f(0) \le k$. Therefore $L_k(G) = L(G'')$, so that $L_k(G)$ is context free.

The proof of Theorem 4.1 for the special case of context-free grammars is equivalent to the argument for weighted context-free grammars in which the weight of every variable is one. The demonstration for this special case is much simpler than that for the general case since it does not require the machinery of control functions developed in section three.

Observe that the family of ultralinear languages is a subfamily of the family of derivation-bounded languages,(7) and as noted in section one, the inclusion is proper. In fact, it is undecidable whether or not a derivation-bounded language is ultralinear. (For let U be the family of ultralinear languages over an alphabet Σ containing at least two elements and let c be a new symbol. It follows from Theorem 4.2 below that $\{(Lc)^*/L \text{ in } U\}$ is a family of derivation-bounded languages. By the proof of Theorem 4.2.2 in [3], it is undecidable whether or not an arbitrary L in U is regular. It is shown in [5] that for an arbitrary context-free language $L \subseteq \Sigma^*$, thus for L in U, $(Lc)^*$ is ultralinear if and only if L is regular. Thus for L in U, it is undecidable whether or not $(Lc)^*$, which is derivation bounded, is ultralinear.) In the balance of this paper, we shall study characterizations and properties of derivation-bounded languages.

We now introduce a special class of grammars which characterize the derivation-bounded languages. These grammars have also been considered by Yntema [9] in her investigation of "standard matching-choice sets".

Definition. A context-free grammar $G = (V,\Sigma,P,\sigma)$ is called nonexpansive, if, for every ξ in $V-\Sigma$ and w in V^*, $\xi \overset{*}{\Rightarrow} w$ implies w does not contain two occurrences of ξ.

Lemma 4.2. For each derivation-bounded language L there exists a nonexpansive context-free grammar G' such that $L = L(G')$.

---

(7) Since the derivation-bounded sets are now known to be context-free languages, we call them derivation-bounded languages.

Proof. Since L is derivation bounded, $L = L_k(G)$ for some weighted context-free grammar $G = (V,\Sigma,P,\sigma,\rho)$. Let $G' = (V',\Sigma,P',\sigma')$ be the context-free grammar defined in Lemma 4.1 It suffices to prove that G' is nonexpansive.

Let ρ' be the function on $V'-\Sigma$ defined by $\rho'(\sigma') = k$ and $\rho'((\xi,(A,f))) = f(0)$ for each $(\xi,(A,f))$ in V'. Since $f(0) \geq \rho(\xi) > 0$, ρ' is a function from $V'-\Sigma$ to the positive integers. Thus $(V',\Sigma,P',\sigma',\rho')$ is a weighted context-free grammar. To prove G' nonexpansive, it suffices to show that if $\xi' = w'_o \Rightarrow \ldots \Rightarrow w'_t = w$ is an arbitrary derivation in G', where ξ' is in $V'-\Sigma$ and w is in V'^*, then w can contain no variables of weight $> \rho'(\xi')$ and at most one variable of weight $\rho'(\xi')$. Since the only productions involving σ' are of the form $\sigma' \rightarrow (\sigma,(A,f))$, there is no loss in assuming $\xi' \neq \sigma'$, i.e., there is no loss in assuming ξ' is of the form $(\xi,(A,f))$. Suppose $t = 1$. Then the derivation is $\xi' \Rightarrow w'_1 = w$. Either w'_1 is in Σ^*, in which case the result is true, or else $w' = v_1(\xi_1,(A^{(1)},f^{(1)}))\ldots v_s(\xi_s,(A^{(s)},f^{(s)}))$, where $s \geq 1$, each v_i is in Σ^*, and $\{(A^{(j)},f^{(j)})/1 \leq j \leq s\}$ is determined by a control function (B,g) dominated by (A,f). In the latter case, $g(0) \leq f(0) = \rho'((\xi,(A,f)))$. Now for each j and i, $1 \leq j \leq s$ and $0 \leq i \leq r_j$, $f^{(j)}(a_i^{(j)}) = g(\beta_{i(j)}) - \sum_{j' \neq j} \beta_{i(j)}(j')$. Thus there exists some j_o, $1 \leq j_o \leq s$, such that $f^{(j_o)}(0) = g(0) - \sum_{j' \neq j} 0 = g(0)$. For $j \neq j_o$,

$$
\begin{aligned}
f^{(j)}(0) &= f^{(j)}(a_{r_j}^{(j)}) \\
&= g(\beta_{r_j(j)}) - \sum_{j' \neq j} \beta_{r_j(j)}(j') \\
&\leq g(0) - \sum_{j' \neq j} \beta_{r_j(j)}(j') \\
&\leq g(0) - \beta_{r_j(j)}(j_o) \\
&< g(0).
\end{aligned}
$$

Thus $\rho'(\xi_{j_o},(A^{(j_o)},f^{(j_o)})) \leq \rho'(\xi')$ and, for $j \neq j_o$, $\rho'(\xi_j,(A^{(j)},f^{(j)})) < \rho'(\xi')$. Therefore the result is true in this case. Continuing by induction, suppose the result is true for all derivations of length $< t, t > 1$. Consider $\xi' = w_o' \Rightarrow \ldots \Rightarrow w_t' = w$ in G'. By induction, w_{t-1}' can contain no variable of weight $> \rho'(\xi')$, and at most one variable of weight $\rho'(\xi')$. Now w_t' is obtained by applying a production $\upsilon' \rightarrow z'$ to a variable υ' in w_{t-1}'. By induction, z' can contain no variable of weight $> \rho(\upsilon')$, thus none of weight $> \rho'(\xi') \geq \rho'(\upsilon')$, and can contain at most one variable of weight $\rho'(\upsilon') \leq \rho'(\xi')$. Thus $w_t' = w$ can contain no variable of weight $> \rho'(\xi')$ and at most one variable of weight $\rho'(\xi')$. Hence the induction is extended and the proof is complete.

<u>Remark.</u> It was shown in [9] that there exist context-free languages generated by no nonexpansive grammar. (In fact, the set $L \subseteq \{a,b\}^*$ of all words w with the following two properties is such a language: (i) The number of occurrences of a in w equals the number of occurrences of b in w; and (ii) For each initial subword w' of w, the number of occurrences of a in w' is greater than or equal to the number of occurrences of b in w'.) From Lemma 4.2 it follows that the family of derivation-bounded languages is a proper subfamily of the family of context-free languages.

We now present several characterizations of derivation-bounded languages.

<u>Theorem 4.2</u>. Given a set $L \subseteq \Sigma^*$, the following statements are equivalent:

(1) L is a derivation-bounded language.

(2) $L = L(G)$ for some nonexpansive context-free grammar G.

(3) L belongs to the smallest family of sets containing all

the linear languages and closed under arbitrary substitution of sets in the family for letters.

(4) $L = L(G) = L_k(G)$ for some context-free grammar G and some positive integer k.

<u>Proof</u>. By Lemma 4.2, (1) $\Rightarrow$ (2). The implication (4) $\Rightarrow$ (1) is trivial. Thus there only remain proofs of (2) $\Rightarrow$ (3) and (3) $\Rightarrow$ (4). Let $\mathcal{L}_3$ and $\mathcal{L}_4$ be the families satisfying conditions (3) and (4) respectively.

Consider (2) $\Rightarrow$ (3). We shall prove that L(G) is in $\mathcal{L}_3$ for each nonexpansive context-free grammar $G = (V,\Sigma,P,\sigma)$. To this end, suppose that $V-\Sigma$ contains just one element, i.e., just σ. Since G is nonexpansive, it is linear. Thus L(G) is in $\mathcal{L}_3$. Continuing by induction assume that $V-\Sigma$ contains $n > 1$ elements and that the result is valid for all nonexpansive grammars with $< n$ variables. Without loss of generality, we may assume that G is reduced.[8] (For otherwise, as noted in [2;3], there exists a reduced grammar $\overline{G} = (\overline{V},\Sigma,\overline{P},\sigma)$, with $\overline{V} \subseteq V$ and $\overline{P} \subseteq P$, such that $L(G) = L(\overline{G})$. Clearly $\overline{G}$ is also nonexpansive.) Let $H \subseteq V-\Sigma$ be the set of all variables ξ such that $\xi \overset{*}{\Rightarrow} u_1\sigma u_2$ for some u_1 and u_2 in V^*. Obviously σ is in H. Let $G' = (V,\Sigma',P',\sigma)$, where $\Sigma' = V-H$ and P' consists of all productions $\xi \to w$ in P such that ξ is in H. Since G is nonexpansive, G' is a linear grammar. (For suppose otherwise. Then there exists a production $\xi \to$

---

[8] A context-free grammar $G = (V,\Sigma,P,\sigma)$ is said to be <u>reduced</u> if for each variable ξ, (i) there exist u and v in V^* such that $\sigma \overset{*}{\Rightarrow} u\xi v$, and (ii) there exists w in Σ^* such that $\xi \overset{*}{\Rightarrow} w$.

$u_1\gamma_1u_2\gamma_2u_3$, where ξ, γ_1, and γ_2 are in H and u_1, u_2, and u_3 are in V^*. Since γ_1 and γ_2 are in H and G is reduced, there exist u_4, $\overline{u}_4$, u_5, $\overline{u}_5$, u_6, $\overline{u}_6$, u_7, and $\overline{u}_7$ in V^* such that $\gamma_1 \overset{*}{\Rightarrow} u_4\sigma u_5 \overset{*}{\Rightarrow} \overline{u}_4\xi\overline{u}_5$ and $\gamma_2 \overset{*}{\Rightarrow} \overline{u}_6\sigma u_7 \overset{*}{\Rightarrow} \overline{u}_6\xi\overline{u}_7$. Then $\xi \overset{*}{\Rightarrow} u_1\overline{u}_4\xi\overline{u}_5u_2\overline{u}_6\xi\overline{u}_7u_3$, contradicting G being nonexpansive.) If $H=V-\Sigma$, then $G=G'$ is linear, so that L(G) is in $\mathcal{L}_3$. Suppose $H \neq V-\Sigma$, so that $\Sigma \subseteq \Sigma'$, $\Sigma'-\Sigma \neq \emptyset$, and $P-P' \neq \emptyset$. For each ξ in $\Sigma'-\Sigma$ let $G_\xi = (\Sigma', \Sigma, P-P', \xi)$. Then G_ξ is a nonexpansive grammar with fewer than n variables. (Note that if $\nu \to u_1\gamma u_2$ is in $P-P'$ for some γ in H, then $\gamma \overset{*}{\Rightarrow} u_3\sigma u_4$ for some u_3 and u_4. Thus $\nu \overset{*}{\Rightarrow} u_1u_3\sigma u_4u_2$, so that ν is in H, a contradiction. Hence, each producation in $P-P'$ involves only symbols in Σ', so that G_ξ is a context-free grammar. Since G is nonexpansive and $P-P' \subseteq P$, G_ξ is nonexpansive. Since σ is not in Σ', G_ξ has fewer than n variables.) Therefore $L(G_\xi)$ is in $\mathcal{L}_3$ by induction. Since each finite set is a linear language, $\{a\}$ is in $\mathcal{L}_3$ for each a in Σ. Let τ be the substitution mapping defined by $\tau(a) = \{a\}$ for each a in Σ, and $\tau(\xi) = L(G_\xi)$ for each ξ in H. Obviously $L(G) = \tau(L(G'))$. Thus L(G) is in $\mathcal{L}_3$, so that (2) $\Rightarrow$ (3).

Consider (3) $\Rightarrow$ (4). Obviously it suffices to prove $\mathcal{L}_3 \subseteq \mathcal{L}_4$. Since $\mathcal{L}_4$ contains every linear language, it therefore suffices to show that $\mathcal{L}_4$ is closed under arbitrary substitution of sets in $\mathcal{L}_4$ for letters. Assume that L is in $\mathcal{L}_4$ and that τ is a substitution of sets in $\mathcal{L}_4$ for letters. Thus, for each a in Σ, $\tau(a) \subseteq \Sigma_a{}^*$ is in $\mathcal{L}_4$. Then there exists a context-free grammar $G = (V,\Sigma,P,\sigma)$ such that $L = L(G) = L_k(G)$ for some $k\geq 1$. Also, for each a in Σ, there exists a context-free grammar $G_a = (V_a, \Sigma_a, P_a, \sigma_a)$ such that $\tau(a) = L(G_a) = L_{k_a}(G_a)$ for some $k_a > 1$. Without loss of generality, we may assume that

$(V_a-\Sigma_a) \cap V = \phi$ and $(V_a-\Sigma_a) \cap (V_b-\Sigma_b) = \phi$ for all a and b in Σ, $a \neq b$. Let h be the homomorphism on V^* defined by $h(\xi) = \xi$ for ξ in $V-\Sigma$ and $h(a) = \sigma_a$ for a in Σ. Let $G' = (V', \Sigma', P', \sigma)$, where $V' = V \cup \bigcup_{a \text{ in } \Sigma} V_a$, $\Sigma' = \bigcup_{a \text{ in } \Sigma} \Sigma_a$, and

$$P' = \bigcup_{a \text{ in } \Sigma} P_a \cup \{\xi \to h(z)/\xi \to z \text{ in } P\}.$$

Clearly G' is a context-free grammar such that $\tau(L) = L(G')$. Let $m = \max \{|z|/\xi \to z \text{ in } P\}$, $m' = \max \{k_a/a \text{ in } \Sigma\}$, and $n = mm' + k$. To complete the proof, it suffices to show that $L(G') = L_n(G')$.

Clearly $L_n(G') \subseteq L(G')$. To see the reverse containment, let w' be in $L(G')$. Then there exists w in $L(G)$ such that $\tau(w) = w'$. Since $L(G) = L_k(G)$, there exists a derivation $\sigma \Rightarrow \ldots \Rightarrow w$ in G which is k-bounded. Then $\sigma = h(\sigma) \Rightarrow \ldots \Rightarrow h(w)$ in G' is k-bounded. For each a in Σ_a, $L(G_a) = L_{k_a}(G_a)$. Thus, for each word u in $L(G_a)$, a in Σ, there exists a k_a-bounded derivation $\sigma_a \overset{*}{\Rightarrow} u$ in G_a, thus in G'. These derivations give rise, in the obvious manner, to a derivation

$$(*) \quad \sigma \Rightarrow \ldots \Rightarrow w \Rightarrow \ldots \Rightarrow w'$$

in G'. Rearrange the application of the productions used in (*) so that whenever a symbol in $\{\sigma_a/a \text{ in } \Sigma\}$ in introduced by some production in $\overline{P} = \{\xi \to h(z)/\xi \to z \text{ in } P\}$, no other production of $\overline{P}$ is applied until all occurrences of elements in $\{\sigma_a/a \text{ in } \Sigma\}$ have been replaced by words in $\bigcup_a L(G_a)$. This yields a new derivation $\sigma \Rightarrow \ldots \Rightarrow w'$ in G'. Each word in the new derivation can contain at most m occurrences of symbols in $\{\sigma_a/a \text{ in } \Sigma\}$, and therefore at most mm' occurrences of symbols in $\bigcup_a (V_a-\Sigma_a)$. Since each word in the new derivation has at most k occurrences of symbols in $V'-\Sigma'$, the new derivation is $(mm' + k)$-bounded. Thus $L(G') = L_n(G')$ and the proof is complete.

Remarks. (1) The proof that the family of context-free languages is closed under intersection with regular sets [2;3] can be readily adopted to show that the family of languages generated by nonexpansive grammars, thus the family of derivation-bounded languages, is closed under intersection with regular sets. Since this family is closed under substitution, it follows from a result in [6] that it is an abstract family of languages (AFL) as defined in [4] which is closed under arbitrary homomorphism, i.e., is a full AFL.

(2) It is shown in [7] that for any AFL $\mathcal{L}$ properly contained in the context-free languages, it is undecidable whether a context-free language belongs to $\mathcal{L}$. Thus it is undecidable whether a context-free language is derivation bounded.

(3) It can be shown that any AFL closed under arbitrary substitution and containing the context-free language $L = \{ww^R / w \text{ in } \{a,b\}^*\}$ contains all linear, hence all derivation-bounded, languages. Therefore the family of derivation-bounded languages is the smallest AFL closed under substitution and containing L.

BIBLIOGRAPHY

[1] R.B. Banerji: "Phrase structure languages, finite machines and channel capacity," Information and Control, Vol. 6, 1963, p. 153-162.

[2] Y. Bar-Hillel, M Perles, and E. Shamir: "On formal properties of simple phrase structure grammars," Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung, Vol. 14, 1961, p. 143-172.

[3] S. Ginsburg: "The Mathematical Theory of Context-Free Languages," Mc Graw-Hill Book Company, New York, 1966.

[4] S. Ginsburg and S. A. Greibach: "Abstract families of languages", System Development Corporation Report, TM-738/031/00, 17 April 1966.

[5] S. Ginsburg and E. H. Spanier: "Finite-turn pushdown automata," SIAM Journal on Control, Vol. 4, 1966, p. 429-453.

[6] S. Ginsburg and E. H. Spanier: "Control sets on grammars," to appear in Mathematical Systems Theory.

[7] S. Greibach: "A note on Undecidable Properties of Formal Languages," SDC Report, TM-738/038/00, 28 August 1967.

[8] S. Greibach and J. Hopcroft: "Independence of AFL Operations," SDC Report, TM-738/034/00, 12 July 1967.

[9] M. K. Yntema: "Inclusion Relations Among Families of Context-Free Languages," Information and Control, Vol. 10, 1967, p. 572-597.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) System Development Corporation 2500 Colorado Avenue Santa Monica, California 90406	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

DERIVATION-BOUNDED LANGUAGES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific. Interim

5. AUTHOR(S) (First name, middle initial, last name)

Seymour Ginsburg
Edwin H. Spanier

6. REPORT DATE 8 January 1968	7a. TOTAL NO. OF PAGES 38	7b. NO. OF REFS 9
8a. CONTRACT OR GRANT NO. F1962867C0008 AFOSR-1203-67 b. Project, Task, Work Unit Nos. 5632-05-01 c. DoD Element 61445OF d. Dod Subelement 681305	9a. ORIGINATOR'S REPORT NUMBER(S) TM-738/041/00 Scientific Report No. 14 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFCRL-68-0080	

10. DISTRIBUTION STATEMENT

1-Distribution of this document is unlimited. It may be released to the Clearinghouse, Department of Commerce, for sale to the general public.

11. SUPPLEMENTARY NOTES Hq. AFOSR, OAR (SRIR) United States Air Force Arlington, Virginia	12. SPONSORING MILITARY ACTIVITY Air Force Cambridge Research Laboratories (CRB) L. G. Hanscom Field Bedford, Massachusetts 01730

13. ABSTRACT

A derivation is a phrase-structure grammar is said to be k-bounded if each word in the derivation contains at most k occurrences of nonterminals. A set L is said to be derivation bounded if there exists a phrase-structure grammar G and a positive inter k such that L is the set of words in the language generated by G which have some k-bounded derivation. The main result is that every derivation-bounded set is a context-free language. Various characterizations of the derivation-bounded languages are then given. For example, the derivation-bounded languages coincide with the standard matching-choice sets discussed by Yntema. They also coincide with the smallest family of sets' containing the linear context free languages and closed under arbitrary substitution.

DD FORM 1 NOV 65 1473

Unclassified
Security Classification

Unclassified
Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Phase structure grammar						
Derivation bounded languages						
Contex-free languages						

Unclassified
Security Classification