1974

Abstract Families of Context-Free Grammars

David Workman

Report Number:
74-114

ABSTRACT FAMILIES OF CONTEXT-FREE GRAMMARS

David Workman
Computer Science Department
Purdue University
West Lafayette, Indiana 47907

March 1974

CSD-TR-114

Abstract.

An <u>abstract family of grammars</u> (AFG) may be deinfed as a class of grammars for which the corresponding class of languages forms an abstract family of languages (AFL) as defined by Ginsburg and Greibach. The derivation bounded grammars of Ginsburg and Spanier is an example of an AFG which is properly included in the class of all context-free grammars (also AFG). The main result is that there exists two distinct infinite hierarchies of AFG which exhaust the derivation bounded AFG such that the AFL associated with the kth member of one of these AFG hierarchies is properly included in the AFL associated with the k+1st member of that same hierarchy. Each hierarchy is shown to be strongly incomparable to the other; that is, the first member of each generates some language not generated by a fixed but arbitrary member of the other. We designate these hierarchies as the hierarchies of <u>left</u> and <u>right dominant grammars</u> (<u>languages</u>).

1. Introduction.

In [6] the notion of Abstract Family of Languages (AFL) was introduced to describe language classes which were closed under certain types of transformations. In most of the literature on AFL theory, specifically [6a, b] and [9], AFLs are generally characterized by some generating class of languages or family of acceptors. From a practical point of view, the theory gives very little explicit information concerning the nature of the underlying class of grammars that is associated with a given AFL. The obvious exceptions to this statement are the classes of right (left) linear grammars, the context-free, context-sensitive and general phrase structure grammars. However, this set of examples is by no means exhaustive. It is our purpose here to describe two distinct hierarchies of "abstract families of grammars" (AFG) which exhaust the class of all derivation bounded grammars studied by Ginsburg and Spanier [7]. By "abstract family of grammars" we shall mean any class of grammars for which the corresponding class of languages forms an AFL. An AFG is a useful concept only if there is some decision procedure for identifying members of the family -- a property which is not enjoyed by AFLs. One of our results is the specification of such a decision procedure for the class of grammars we have undertaken to study.

The technique we employ involves defining certain relations on the nonterminal alphabet of context-free grammars. By requiring that these relations be irreflexive we are able to isolate the class of all derivation bounded grammars. As pointed out in [7], this class of grammars defines an abstract family of languages properly included in the context-free. By virtue of the irreflexive property of our relations, which we have chosen to call the "generalized left and right dominant relations", we are able to associate a pair of nonnegative integers $\ell deg(G)$ and $rdeg(G)$ with every reduced derivation bounded grammar, G. These integers represent the "degree of left and right dominance", respectively, of G. For each integer $k \geq 0$ we define $\mathcal{G}_\ell(k)$ ($\mathcal{G}_r(k)$) to be the class of all derivation bounded grammars, G, for which $\ell deg(G) \leq k$ ($rdeg(G) \leq k$). Our main results state that for each $k \geq 0$, the classes $\mathcal{G}_\ell(k)$ and $\mathcal{G}_r(k)$ generate full AFLs. Furthermore, it is shown that the class of languages, $\mathcal{L}_\ell(k)(\mathcal{L}_r(k))$ associated with the grammar class $\mathcal{G}_\ell(k)(\mathcal{G}_r(k))$ is properly included in the class of next higher degree. Although the scope of our investigation has been limited to context-free grammars, we feel that perhaps the techniques employed here may have extensions which isolate classes of AFG which include context-sensitive or general phrase structure grammars.

The paper is divided into five other sections. In section 2 we present the basic notation and terminology used throughout the remaining sections. In addition, section 2 also presents results from other sources which are referred to in the sequel.

In section 3 we introduce the class of strictly linear languages which are fundamental to our characterization of the classes $\mathcal{L}_\ell(k)$ ($\mathcal{L}_r(k)$) presented in section 5.

Section 4 introduces the generalized left and right dominance relations referred to above. These relations are denoted Δ_ℓ and Δ_r, respectively. It is in this section that we also define the notion of "degree" of left and right dominance which allows us to describe the grammar hierarchies, $\mathcal{G}_\ell(k)$ and $\mathcal{G}_r(k)$, $k \geq 0$. The three major results of this section are theorems 4.4, 4.8 and 4.9. Theorem 4.4 establishes the equivalence of the derivation bounded (nonexpansive) grammars to the class of context-free grammars for which $\Delta_\ell(\Delta_r)$ is irreflexive. Theorem 4.8 places another interesting class of grammars, the nonterminal bounded grammars [2], within the hierarchy of left and right dominant grammars. We conclude this section with theorem 4.9 which gives an effective procedure for computing $\ell deg(G)$ $(rdeg(G))$ for an arbitrary reduced context-free grammar, G.

In section 5 we give a characterization of the language classes $\mathscr{A}_\ell(k)$ and $\mathscr{A}_r(k)$ in terms of substitutions applied to strictly linear languages. The class of substitutions we allow are restricted to having their range sets lie in certain language classes which are determined by the domain alphabet. To obtain the characterizations in a relatively straight forward manner it was necessary to introduce new relations (ρ_k and λ_k) which refine the classes $\mathscr{G}_\ell(k)$ and $\mathscr{G}_r(k)$ into yet another hierarchy of subclasses. The characterization of $\mathscr{A}_\ell(k)(\mathscr{A}_r(k))$ is expressed in terms of the subclasses of languages determined by the refinement of $\mathscr{G}_\ell(k)(\mathscr{G}_r(k))$ imposed by the relation $\rho_k(\lambda_k)$.

Section 6 contains most of the major results of this paper. It is shown that $\mathscr{A}_\ell(k)(\mathscr{A}_r(k))$ forms a full AFL and that for each $k \geq 0$, $\mathscr{A}_\ell(k) \subsetneq \mathscr{A}_\ell(k+1)$ $(\mathscr{A}_r(k) \subsetneq \mathscr{A}_r(k+1))$. Theorem 6.5 is a somewhat surprising result in that it is shown that $\mathscr{A}_\ell(0) - \mathscr{A}_r(k) \neq \Phi$ for each $k \geq 0$ and similarly $\mathscr{A}_r(0) - \mathscr{A}_\ell(k) \neq \Phi$ for each $k \geq 0$.

II. Notation, Definitions and Background results.

For the most part, our notational conventions and basic definitions follow those commonly found in the literature concerning language theory. Any background material not explicitly presented in this section can be found in Ginsburg [5] or Hopcroft and Ullman [11].

Definition 2.1. A context-free grammar is a four-tuple, $G = (V, T, P, \alpha)$, where V (nonterminals), T (terminals) and P (productions) are finite non-empty sets. The start symbol, α, belongs to V. Elements of V will usually be denoted by small Greek letters, while elements of T will usually be denoted by small letters early in the English alphabet.

Definition 2.2. Let $G = (V, T, P, \alpha)$ be a context-free grammar and let $p : \beta \to w$ denote an element of P. If $w \in T^*$, p is said to be a terminating production. If $w \in T^*VT^*$ (VT^*, T^*V), p is said to be linear (left-linear, right-linear). If all productions of G are linear or terminating, then G is said to be a linear grammar. The language generated by G will be denoted by $L(G)$.

Notation. Let $(p : \beta \to w) \in P$. If $u, v \in (V \cup T)^*$, then we write $u \overset{p}{\underset{G}{\Longrightarrow}} v$ whenever $u = u_1 \beta u_2$ and $v = u_1 w u_2$. If $\pi = p_1 p_2 \cdots p_n$ with $p_i \in P$, $1 \leq i \leq n$, then we write $u \overset{\pi}{\underset{G}{\Longrightarrow}} v$ if and only if there exists words $z_i \in (V \cup T)^*$, $0 \leq i \leq n$ such that $u = z_o$, $v = z_n$ and $z_{i-1} \overset{p_i}{\underset{G}{\Longrightarrow}} z_i$, $1 \leq i \leq n$. We write $u \overset{+}{\underset{G}{\Longrightarrow}} v$ if there exists π such that $u \overset{\pi}{\underset{G}{\Longrightarrow}} v$.

Furthermore, $u \overset{*}{\underset{G}{\Longrightarrow}} v$ if $u = v$ or $u \overset{+}{\underset{G}{\Longrightarrow}} v$. The sequence π in the above context is called a derivation of v from u in G. The words

z_i, $1 \leq i \leq n$, will be called u-sentential forms or, more simply, sentential forms if u is understood. In case p_i in π is always applied to the left-most (right-most) nonterminal of z_{i-1} we call π a left-most (right-most) derivation and write $u \overset{\pi}{\underset{lm}{\Longrightarrow}} v$ $(u \overset{\pi}{\underset{rm}{\Longrightarrow}} v)$.

If S is a set, then $|S|$ denotes the number of elements in S. If $x \in (V \cup T)^*$, then $||x||$ denotes the length of x. "ε" denotes the string of length zero. If $s \in (V \cup T)$, then $||x||_s$ represents the number of occurances of s in x.

We define $||x||_S = \sum_{s \in S} ||x||_s$.

Definition 2.3. Let $G = (V, T, P, \alpha)$ be a context-free grammar.

G is said to be reduced if for every $\beta \in V$, there exists π_1 and π_2 such that $\beta \overset{\pi_2}{\underset{G}{\Longrightarrow}} x \in T^*$.

The class of nonterminal bounded grammars and their corresponding languages have received considerable attention in the literature; e.g., Banerji [2], Fleck [4], Ginsburg and Spanier [7], Gruska [12] and Moriya [11] have studied a number of different and interesting properties of these grammars. Ginsburg and Spanier [7] were the first to study the more general, but related class of derviation bounded grammars and languages. This latter class of languages seems to be a "natural" subclass of context-free languages in the sense that they form a full AFL, a result also established in [7].

The next definition describes the aforementioned grammars.

Definition 2.4. Let $G = (V, T, P, \alpha)$ be a context-free grammar.

1. G is said to be <u>nonterminal bounded</u> if and only if there exists a fixed $k \geq 0$ such that for every derivation π in G, $\alpha \overset{\pi}{\underset{G}{\Longrightarrow}} w \in (V \cup T)^*$ implies $||w||_V \leq k$.

2. G is said to be <u>derivation bounded</u> if and only if there exists $k \geq 0$ such that for every $x \in L(G)$ <u>there exists</u> a derivation π of x which has the following property: $\alpha \overset{\pi_1}{\underset{G}{\Longrightarrow}} w \overset{\pi_2}{\underset{G}{\Longrightarrow}} x$ implies $||w||_V \leq k$, for all $\pi_1\pi_2 = \pi$.

3. G is said to be <u>nonexpansive</u> if and only if for every $\beta \in V$, $\beta \overset{+}{\underset{G}{\Longrightarrow}} w \in (V \cup T)^*$ implies $||w||_\beta \leq 1$.

The following theorem due to Ginsburg and Spanier [7] characterize the derivation bounded grammars and the languages they generate.

Theorem 2.5. Let $L \subseteq T^*$. The following statements are equivalent.

(1) L is generated by some derivation bounded grammar.

(2) L is generated by some nonexpansive grammar.

(3) L belongs to the smallest family of languages containing all linear languages and closed under abitrary substitution of sets in the family for letters.

One of our major results of this paper concerns the existence of hierarchies of grammars which generate full ALFs of derivation bounded languages. The concept of full AFL is presented in our next definition due to Ginsburg and Greibach [6].

Definition 2.6. Given an infinite set of symbols, Γ, an abstract family of languages (AFL) is a family $\mathcal{L}$ of subsets of Γ^* such that,

(1) For each $L \in \mathcal{L}$ there is a finite set $T \subseteq \Gamma$ such that $L \subseteq T^*$.

(2) There exists some nonempty $L \in \mathcal{L}$.

(3) $\mathcal{L}$ is closed under the operations, finite union, concatenation, $+$, inverse-homomorphism, ε-free homomorphism and intersection with regular sets.

(4) L is said to be full if it is closed under arbitrary homomorphism.

The following theorem due to Greibach and Hopcroft [9] will be useful in section 6. The original statement of this theorem is a stronger result than we shall need, we have therefore taken the liberty to present a weaker version which is more suitable for results presented in the sequel.

Theorem 2.7. If $\mathcal{L}$ is a family of languages closed under union and intersection with a regular set, regular substitution and homomorphism†, then $\mathcal{L}$ is also closed under inverse homomorphism.

†: The theorem as originally stated in [9] required closure only under a restricted type of regular substitution and required only that $\mathcal{L}$ be closed under ε-free homomorphism.

3. Strictly Linear Grammars and Languages.

In this section we introduce the strictly linear languages. This class of languages is a proper subclass of the class of all linear languages. Their distinguishing property is that every string z in a strictly linear language has the form xy, where x and y are strings over disjoint alphabets. Furthermore, the set of all x's (y's) is a regular set. An example of such a language is $\{a^n b^n \mid n \geq 0\}$. The importance of the strictly linear languages rests in the fact that they provide the basis for a characterization of the left and right dominant languages of degree k introduced in section 4 and representing the main object of study in this paper.

Proposition 3.4 is a simple but useful result which states that every linear language is the homomorphic image of some strictly linear language. Lemma 3.5 describes closure properties of the strictly linear languages under regular substitution.

Another fundamental concept developed in this section is the notion of "subgrammar". A subgrammar of a given context-free grammar is the grammar obtained by reducing the original relative to one of its nonterminals. Subgrammars become useful when one attempts to isolate and describe local properties of a given grammar. The language generated by a subgrammar can be described, under appropriate conditions, in terms of a substitution applied to a corresponding "restricted subgrammar". In a restricted subgrammar, a set of nonterminals are treated as terminal symbols. Lemma 3.7 is the last result of this section and provides a characterization of subgrammars in terms of a substitution applied to restricted subgrammars. This lemma is a valuable tool in proving key results of section 4.

Definition 3.1. Let $G = (V, T, P, \alpha)$ be a context-free grammar. G is said to be linear over (T_ℓ, T_r) biased left if and only if

(1) G is a linear grammar,

(2) $T = T_\ell \cup T_r$, and

(3) $P \subseteq V \times (T_\ell^* V T_r^* \cup T_\ell^*)$.

G is said to be biased right if (3) is replaced by,

(3') $P \subseteq V \times (T_\ell^* V T_r^* \cup T_r^*)$.

If in addition to (1), (2) and (3) or (3'), G satisfies (4), then G is said to be strictly linear over (T_ℓ, T_r) biased left (right), where

(4) $T_\ell \cap T_r = \Phi$.

A language, L, is said to be (strictly) linear over (T_ℓ, T_r) biased left (right), if there is a so-named grammar, G, such that $L = L(G)$.

If G satisfies (1), (2) and either (3) or (3'), then we simply say that G is linear over (T_ℓ, T_r); similarly, if G satisfies (1), (2), (4) and either (3) or (3') we say G is strictly linear over (T_ℓ, T_r).

In subsequent sections we will need special notation for representing a set of abstract symbols disjoint and in one-to-one correspondence with a given set. In addition, a special homomorphism will often be required to identify members of the abstract set with corresponding members of the original. These notational conventions are given formal status by the next definition.

Definition 3.2. Let S by any set, then $\overline{S} = \{\overline{s} \mid s \in S\}$ denotes a set of abstract symbols disjoint from S. In addition, the homomorphism $\overline{h}: (S \cup \overline{S})^* \to S^*$ defined by $\overline{h}(\overline{s}) = \overline{h}(s) = s$, for all $s \in S$, will henceforth be designated as the unmarking homomorphism on S .

The following definition points out that the class of linear grammars are in one-to-one correspondence with the strictly linear grammars of left (right) bias.

Definition 3.3. Let $G = (V, T, P, \alpha)$ be a linear grammar. The <u>strict image of G biased left</u> is the grammar $\overline{G}_\ell = (V, \Sigma, \overline{P}_\ell, \alpha)$, strictly linear over (Σ_ℓ, Σ_r) biased left, where

(i) $\Sigma_\ell \subseteq T$ is the smallest alphabet such that

$P \subseteq V \times (\Sigma_\ell^* V T^* \cup \Sigma_\ell^*)$;

(ii) $\Sigma_r = \overline{T}_r \subseteq \overline{T}$, where $T_r \subseteq T$ is the smallest alphabet such that $P \subseteq V \times (T^* V T_r^* \cup T^*)$;

(iii) $\overline{P}_\ell = \{(\beta \to u) \in P \mid u \in T^*\} \cup$

$\{\beta \to u\beta'\overline{v} \mid (\beta \to u\beta'v) \in P,\ \beta' \in V$ and

$\overline{v} = \overline{h}^{-1}(v) \cap \Sigma_r^*\}$. ($\overline{h}$ is the unmarking homomorphism on T).

The <u>strict image of G biased right</u> is the grammar $\overline{G}_r = (V, \Sigma, \overline{P}_r, \alpha)$, strictly linear over (Σ_ℓ, Σ_r) biased right, where

(i) $\Sigma_\ell = \overline{T}_\ell \subseteq \overline{T}$, where $T_\ell \subseteq T$ is the smallest alphabet such that, $P \subseteq V \times (T_\ell^* V T^* \cup T^*)$;

(ii) $\Sigma_r \subseteq T$ is the smallest alphabet such that $P \subseteq V \times (T^* V \Sigma_r^* \cup \Sigma_r^*)$;

(iii) $\overline{P}_r = \{(\beta \to v) \in P \mid v \in T^*\}$

$\{\beta \to \overline{u}\beta'v \mid (\beta \to u\beta'v) \in P,\ \beta' \in V$ and $\overline{u} = \overline{h}^{-1}(u) \cap \Sigma_\ell^*\}$.

Finally, $L(\overline{G}_\ell)$ $(L(\overline{G}_r))$ is called the strict image of $L(G)$ biased left (right).

The next proposition is a simple consequence of the definitions above and therefore no proof will be given. It emphasizes the fact that every linear language is a homorphic copy of its strict image.

Proposition 3.4. Let $G = (V, T, P, \alpha)$ be a linear grammar. Then $L(G) = \overline{h}(L(\overline{G}_\ell)) = \overline{h}(L(\overline{G}_r))$, where $\overline{h}$ is the unmarking homomorphism on T and $\overline{G}_\ell(\overline{G}_r)$ is the strict image of G biased left (right) .

Lemma 3.5. Let $G = (V, T, P, \alpha)$ be strictly linear over (T_ℓ, T_r) . For each $a \in T_\ell$ let $R_a \subseteq \Sigma_\ell^*$ be a regular set; similarly, for each $b \in T_r$ let $R_b \subseteq \Sigma_r^*$ be a regular set.

Then $\tau(L(G))$ is linear over (Σ_ℓ, Σ_r) with the same bias as $L(G)$, where τ is the substitution defined by $\tau(c) = R_c$ for all $c \in T$. If $\Sigma_r \cap \Sigma_\ell = \Phi$, then $\tau(L(G))$ is strictly linear.

Proof. We construct a grammar $G' = (V', \Sigma, P', \alpha)$ which is linear over (Σ_ℓ, Σ_r) and having the same bias as G such that $\tau(L(G)) = L(G')$. P' and V' are described as follows. For each $a \in T_\ell$ let G_a be a right-linear grammar generating R_a and similarly, let G_b be a left-linear grammar generating R_b for each $b \in T_r$. We shall assume that the nonterminal sets of all such grammars are pair-wise disjoint and disjoint from V. Let $p_1, p_2, \ldots, p_k$ be some ordering of the productions of P. If $c \in T = T_\ell \cup T_r$, then we call (c, i, j) an occurence of c if and only if c appears in the right-part of p_i and p_i has the form,

$p_i : \beta \to ucv$, where $||uc|| = j$ if $c \in T_\ell$ or $||cv|| = j$ if $c \in T_r$. Clearly if (c, i, j) and (c', i', j') are two occurances of $c, c' \in T$, then $(c, i, j) \neq (c', i', j')$. For each occurance (c, i, j) of $c \in T$ let G_c^{ij} be a unique copy of G_c obtained by renaming the nonterminal symbols; that is, if $\gamma \in V_c$ (the nonterminal set of G_c), then (γ^{ij}) will be the corresponding nonterminal of V_c^{ij} in G_c^{ij}. Let $G_c^{ij} = (V_c^{ij}, \Sigma_c, P_c^{ij}, \gamma_c^{ij})$, where $\Sigma_c = \Sigma_\ell$ if $c \in T_\ell$ and $\Sigma_c = \Sigma_r$ if $c \in T_r$. Clearly $L(G_c^{ij}) = L(G_c) = R_c$ for all i and j and furthermore all nonterminal sets, V_c^{ij}, are pair-wise disjoint and disjoint from V.

The property we desire for G' is the power of "simulating" a single production, p, of G by using only left or right linear productions which generate words in R_c for each occurrance of c introduced by production p. We describe the productions of G' that are constructed for each type of production p in G.

(a) if $p_i \in P$ is of the form $(\beta \to \varepsilon)$ or $(\beta \to \beta')$, where $\beta' \in V$, then add p_i to P'.

(b) If $p_i \in P$ is a terminating production of the form $(\beta \to c_1 c_2 \ldots c_k)$, $k \geq 1$, then we consider two cases.

<u>Case k = 1</u>. For this case add $\beta \to \gamma_{c_1}^{i,1}$ to P', where $\gamma_{c_1}^{i,1}$ is the start symbol of $G_{c_1}^{i,1}$. In addition, add all productions of $P_{c_1}^{i,1}$ to P'.

<u>Case k > 1</u>. For this case we identify two subcases which are associated with the bias of G.

<u>Left bias</u>: Add $\beta \to \gamma_{c_1}^{i,1}$ to P'. For each $j < k$ add all productions of $P_{c_j}^{i,j}$ to P' where the terminating productions, $(\delta \to w) \in P_{c_j}^{i,j}$, are replaced by $\delta \to w\gamma_{c_{j+1}}^{i,j+1}$. Finally, add all productions of $P_{c_k}^{i,k}$ to P'.

<u>Right bias</u>: Let the right-part of p_i be written as $c_k c_{k-1} \ldots c_1$. Follow the same construction given for left-bias except that w and $\gamma^{i,j+1}_{c_{j+1}}$ should be reversed.

(c) If $p_i \in P$ is of the form $\beta \to c_1 c_2 \ldots c_k \beta'$ or $\beta \to \beta' c_k c_{k-1} \ldots c_1$, $k \geq 1$, where $\beta' \in V$, then follow the construction given in (b) with the change that if $(\delta \to w)$ is a terminating production of $P^{i,k}_{c_k}$, replace it by $\delta \to w\beta'$ if p_i is right-linear and by $\delta \to \beta' w$ if p_i is left-linear.

(d) If $(p_i : \beta \to a_1 a_2 \ldots a_r \beta' b_s \ldots b_1)$ P, where $r, s \geq 1$ and $\beta' \in V$, then let β'_i be a unique abstract symbol not already defined. Let $p^1_i : \beta \to a_1 a_2 \ldots a_r \beta'_i$ and $p^2_i : \beta'_i \to \beta' b_s \ldots b_1$ be formed from p_i. Add to P' the productions constructed from p^1_i and p^2_i according to (c) above.

Finally, let

$$V' = V \cup \{\beta'_i \mid \beta'_i \text{ is defined by (d) above}\} \cup \Big(\bigcup_{i,j,c} V^{i,j}_c \Big) .$$

It is not difficult to show that $(p_i : \beta \to u\beta' v) \in P$, $uv \in T^*$, $\beta' \in V$, if and only if $\beta \overset{*}{\underset{G'}{\Longrightarrow}} x\beta' y$, where $x \in \tau(u)$ and $y \in \tau(v)$. And similarly, $(p_i : \beta \to w) \in P$, $w \in T^*_\ell(T^*_r)$ if and only if $\beta \overset{*}{\underset{G'}{\Longrightarrow}} x \in \tau(w) \subseteq \Sigma^*_\ell(\Sigma^*_r)$, where $\beta \in V' \cap V$. Therefore it follows that $\alpha \overset{*}{\underset{G}{\Longrightarrow}} x \in L(G)$ if and only if $\alpha \overset{*}{\underset{G'}{\Longrightarrow}} y \in \tau(x)$ and that the bias of G' agrees with the bias of G.

Definition 3.6. Let $G = (V, T, P, \alpha)$ be a context-free grammar. The subgrammar of G relative to $\beta \in V$, denoted $G(\beta)$, is the grammar $G(\beta) = (V(\beta), T, P(\beta), \beta)$ obtained by reducing (V, T, P, β). For every subset $U \subseteq V$ and $\beta \in V - U$ define $G(\beta, U)$ to be the subgrammar of G relative to β restricted on U obtained by reducing $(V - U, T \cup U, P, \beta)$.

It should be noted that if G is reduced and α is the start symbol of G, then $G = G(\alpha) = G(\alpha, \Phi)$. The notion of a subgrammar is useful in identifying the nonterminals and productions involved in derivations originating from a fixed nonterminal. A subgrammar restricted on a set, U, of nonterminals is a means of describing all sentential forms derivable in the original grammar from some fixed nonterminal where members of U are treated as terminals; that is, members of U cannot be re-written once they are introduced in a sentential form of some derivation. The next lemma explores a useful property of certain types of restricted subgrammars.

Lemma 3.7. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar and let $G(\beta) = (V(\beta), T, P(\beta), \beta)$ be the subgrammar of G relative to $\beta \in V$. If U is any subset of $V - \{\beta\}$ such that for all $\gamma \in U$, $\gamma \overset{+}{\underset{G}{\Longrightarrow}} w$ implies $w \in (T \cup U)^*$, then $L(G(\beta)) = \sigma(L(G(\beta,U)))$, where σ is a substitution defined by,

$\sigma(t) = t$ for all $t \in T$ and

$\sigma(\gamma) = L(G(\gamma))$ for all $\gamma \in U$.

Furthermore, if $G(\beta, U) = (V', T \cup U, P', \beta)$, then

$V' = V(\beta) - U$ and

$$P' = P(\beta) - \left(\bigcup_{\gamma \in U} P(\gamma) \right).$$

Proof. If $U = \Phi$, then $G(\beta,U) = G(\beta)$, σ becomes the identity homomorphism and the conclusions of the lemma follow trivially. Assume, therefore, that $U \neq \Phi$. We now establish an important property of G.

(A) For every $\beta \in V - U$ and derivation π such that $\beta \underset{G}{\overset{\pi}{\Longrightarrow}} w \in (T \cup V)^*$ there exists a permutation π' of π such that $\beta \underset{G}{\overset{\pi'}{\Longrightarrow}} w$ and such that $\pi' = \pi_1'\pi_2'$, where $\pi_1 \neq \epsilon$ rewrites only elements of $V - U$ and π_2', if non-null, rewrites only elements of U. To establish (A) let π be any derivation from $\beta \in V - U$. If π rewrites only elements of $V - U$, then $\pi = \pi' = \pi_1'$ $(\pi_2' = \varepsilon)$ and the result is immediate. If $\pi = \pi_{11}\pi_{12} \cdots \pi_{k1}\pi_{k2}$ for some $k \geqslant 1$ where π_{i1}, $1 \leq i \leq k$, represents a sequence of productions which rewrite elements of $V - U$ and π_{i2}, $1 \leq i \leq k$, represents a sequence of productions rewriting elements of U. Furthermore, for $k = 1$, $\pi_{k2} = \pi_{12} \neq \varepsilon$, and if $k > 1$, then for $1 \leq i \leq k-1$, $\pi_{i2} \neq \varepsilon$. That π must begin with a sequence π_{11} follows from the fact that $\beta \in V - U$. We now show that π_{i2} can be interchanged with $\pi_{i+1,1}$ to obtain an equivalent derivation and consequently reducing the value of "k" for the resulting sequence.

If $k = 1$ initially, then π is already in the desired form and we are finished. Assume that $k > 1$ and consider the sequence $\pi_{11}\pi_{12}\pi_{21}$

and let $\pi_{21} = p_1 p_2 \cdots p_r$ for some $r \geq 1$. Since π_{12} rewrites only elements of U and since π_{12} cannot introduce elements of $V - U$ into the sentential form (an assumption of the lemma), then the non-terminal rewritten by p_1 must have been introduced by π_{11}. We may therefore permute p_1 and π_{12} obtaining $\pi_{11} p_1 \pi_{12} p_2 p_3 \cdots p_r$. If $r = 1$ we have succeeded in permuting π_{12} and π_{21}, otherwise we can apply the same argument to the sequence $\pi'_{11} \pi_{12} \pi'_{21}$, where $\pi'_{11} = \pi_{11} p_1$ and $\pi'_{21} = p_2 p_3 \cdots p_r$. Thus it follows that the sequence $\pi_{11} \pi_{21} \pi_{12}$ is equivalent to the sequence $\pi_{11} \pi_{12} \pi_{21}$. By permuting the left-most pair, π_{i2} and $\pi_{i+1,1}$, we have reduced the number of such paired sequences. In this way the original sequence π may be modified to produce an equivalent derivation π' of the desired form.

Returning now to the main proof we establish that $V' = V(\beta) - U$ and that $P' = P(\beta) - (\bigcup_{\gamma \in U} P(\gamma))$. Since G is reduced it follows that for every $\beta \in V$ there exists π such that $\beta \overset{\pi}{\underset{G}{\Longrightarrow}} \in T^*$. Thus $\beta' \in V(\beta)$ if and only if $\beta' = \beta$ or there exists π such that $\beta \overset{\pi}{\underset{G}{\Longrightarrow}} u\beta' v$, where $uv \in (T \cup V)^*$. As a consequence of this we have that $\beta \overset{\pi}{\underset{G(\beta,U)}{\Longrightarrow}} w$ implies $\beta \overset{\pi}{\underset{G(\beta)}{\Longrightarrow}} w$. Thus $V' \subseteq V(\beta) - U$ and $P' \subseteq P(\beta) - (\bigcup_{\gamma \in U} P(\gamma))$. Now suppose $\beta' \in V(\beta) - U$. Then there exists π such that $\beta \overset{\pi}{\underset{G}{\Longrightarrow}} w_1 \beta' w_2$ for some $w_1 w_2 \in (T \cup V)^*$.

By (A) $\beta \xrightarrow[G]{\pi'} w_1\beta' w_2$, where $\pi' = \pi_1'\pi_2'$ and π_1' rewrites elements of $V - U$ and π_2' rewrites elements of U. As argued before, if $\beta' \in V-U$ then β' must be introduced by π_1'. Thus $\beta \xrightarrow[G(\beta,U)]{\pi_1'} w_1'\beta' w_2'$ for some $w_1'w_2'$ implying $\beta' \in V'$; we conclude $V' = V(\beta) - U$. From this equality and the assumption that G is reduced it also follows that $P' = P(\beta) - (\bigcup_{\gamma \in U} P(\gamma))$.

If $\beta \xrightarrow[G(\beta,U)]{\pi} w \in L(G(\beta,U))$, then $\beta \xrightarrow[G(\beta)]{\pi} w$. Now if $w \in T^*$, then $\sigma(w) = w \in L(G(\beta))$. If $w = y_o\gamma_1 y_i \ldots \gamma_n y_n$, where $y_o y_i \ldots y_n \in T^*$ and $\gamma_i \in U \cap V(\beta)$, $1 \leq i \leq n$, then $\sigma(w) = y_o x_1 y_1 \ldots x_n y_n$ where $x_i \in L(G(\gamma_i))$. But $\gamma_i \xrightarrow[G(\gamma_i)]{+} x_i$ implies $\gamma_i \xrightarrow[G(\beta)]{+} x_i$ and therefore $w \xrightarrow[G(\beta)]{+} \sigma(w)$. It follows that $\sigma(L(G(\beta,U))) \subseteq L(G(\beta))$.

Now suppose $\beta \xrightarrow[G(\beta)]{\pi} x \in T^*$. Then by (A) $\beta \xrightarrow[G]{\pi_1'} w \xrightarrow[G]{\pi_2'} x$, where π_1' rewrites elements of $V - U$ and π_2' rewrites elements of U. Since $w \in (T \cup U)^*$, then $w \in L(G(\beta,U))$. Now if $\pi_2' = \varepsilon$, then $x = w \in L(G(\beta,U)) \cap T^* \subseteq \sigma(L(G(\beta,U)))$. If $w = y_o\gamma_i y_i \ldots \gamma_n y_n$, where $y_o y_1 \ldots y_n \in T^*$ and $\gamma_i \in U$, $1 \leq i \leq n$, then $w \xrightarrow[G(\beta)]{\pi_2'} x \in T^*$ implies $\gamma_i \xrightarrow[G(\beta)]{+} x_i \in T^*$. But $\gamma_i \xrightarrow[G(\beta)]{+} x_i$ implies $\gamma_i \xrightarrow[G(\gamma_i)]{+} x_i$, thus $x \in \sigma(w) \subseteq \sigma(L(G(\beta,U)))$. This establishes the reverse inclusion and hence the relation $L(G(\beta)) = \sigma(L(G(\beta,U)))$.

4. Right and Left Dominant Grammars and Languages of Degree k.

Banerji [2] introduced a "dominance" relation on the nonterminal set of a context-free grammar. The class of grammars for which this relation is irreflexive corresponds precisely to the class of nonterminal bounded grammars [2] which have been treated in a variety of contexts by other authors; e.g., Fleck [3, 4], Ginsburg and Spanier [8] and Moriya [12]. In this section we introduce the "generalized left and right dominance relations", denoted Δ_ℓ and Δ_r, respectively. These relations are defined on the nonterminal set of a context-free grammar and are based upon a type of self-embedding exhibited by nonterminals. One of the principal results of this section is theorem 4.4 which essentially states that the class of derivation bounded grammars [7] corresponds precisely to the class of context-free grammars for which Δ_ℓ and Δ_r are irreflexive. In this fashion Δ_ℓ and Δ_r represent generalizations of Banerji's dominance relation by virtue of characterizing a much larger class of grammars and languages.

For any set, S, and any relation R on that set we define the "degree" of an element, $s \in S$, with respect to the relation, R, denoted deg(s, R). By choosing $R = \Delta_\ell$ or Δ_r and letting S represent the nonterminal set of some grammar we are able to classify all derivation bounded grammars according to their "degree of generalized left (right) dominance." For each $k \geq 0$ we denote the class of all reduced context-free grammars of "left-degree" k or less by $\mathscr{G}_\ell(k)$. The corresponding class of languages is denoted $\mathscr{L}_\ell(k)$. We call this class of languages the "Left Dominant Languages of Degree k". In a similar fashion we define $\mathscr{G}_r(k)$ and $\mathscr{L}_r(k)$.

Theorem 4.8, presented at the end of this section, gives a quantitative measure of the complexity of the class of nonterminal bounded grammars relative to the class of all derivation bounded grammars. In this result we show that G is nonterminal bounded if and only if G belongs to $\mathcal{G}_\ell(0) \cap \mathcal{G}_r(0)$.

We end this section by presenting an algorithm for computing the least k such that $G \in \mathcal{G}_\ell(k)$, where G is an arbitrary reduced context-free grammar. The algorithm also determines if such a k exists.

Definition 4.1. Let S be a non-empty set and let R be a relation on S. For each $s \in S$ define

$C(s) = \{k \mid$ there exists a sequence $s_0, s_1, \ldots, s_k$ of elements in S such that $s = s_0$ and $(s_{i-1}, s_i) \in R$ for $1 \leq i \leq k\}$.

The <u>degree of s under R</u>, denoted deg(s,R), is defined by,

$$\begin{aligned} \deg(s,R) &= \infty, \text{ if } C(s) \text{ is infinite} \\ &= \text{Max } C(s), \text{ if } 0 < |C(s)| < \infty \text{ and} \\ &= 0, \text{ if } C(s) = \Phi. \end{aligned}$$

It is obvious that if S is a finite set, then R is irreflexive if and only if $\deg(s, R) < \infty$ for all $s \in S$. The next lemma describes some general properties of deg(s, R) where R is defined on the nonterminal set of a context-free grammar and satisfies certain conditions with respect to derivations. This lemma will apply to the generalized dominance relations Δ_ℓ and Δ_r introduced in definition 4.3. Another class of relations satisfying the conditions of this lemma is introduced in section 5.

Lemma 4.2. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar and let R be a relation on V satisfying,

(i) $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ and $(\beta', \beta'') \in R$ implies $(\beta, \beta'') \in R$, where $uv \in (V \cup T)^*$.

(ii) $(\beta, \beta') \in R$ implies $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ for some $uv \in (V \cup T)^*$.

Then,

(A) R is transitive.

(B) For $\beta, \beta' \in V$, $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$, $uv \in (V \cup T)^*$, implies $\deg(\beta, R) \geq \deg(\beta', R)$.

(C) $\deg(\beta, R) \leq \deg(\alpha, R)$ for all $\beta \in V$; if R is irreflexive, then $\deg(\alpha, R) < |V|$.

(D) If R is irreflexive, then $(\beta, \beta') \in R$ implies $\deg(\beta, R) > \deg(\beta', R)$.

(E) If R is irreflexive, then $\deg(\beta, R) > 0$ implies there exists $\beta' \in V$ such that $\deg(\beta', R) = \deg(\beta, R) - 1$.

(F) $\deg(\beta, R) > (\deg(\beta', R)$ implies $\beta' \overset{+}{\underset{G}{\not\Longrightarrow}} u\beta v$ for any $uv \in (V \cup T)^*$.

Proof.

(A): Let (β_1, β_2) , $(\beta_2, \beta_3) \in R$. Property (ii) implies $\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta_2 v$.

This together with property (i) implies $(\beta_1, \beta_3) \in R$, thus R is transitive.

(B): Let $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$. If $\deg(\beta', R) = 0$, then (B) is immediate.

Assume, therefore, that $\deg(\beta', R) \neq 0$. Then there exists a chain $(\beta', \beta_1), (\beta_1, \beta_2), \ldots, (\beta_{k-1}, \beta_k)$ in R, where $k \geq 1$. By property (i) it follows that $(\beta, \beta_1) \in R$ and thus $(\beta, \beta_1), (\beta_1, \beta_2), \ldots, (\beta_{k-1}, \beta_k)$ is a chain in R initiated by β. Since for each such chain initiated by β' there is a corresponding chain of equal length initiated by β, then it follows that $\deg(\beta, R) \geq \deg(\beta', R)$.

(C): Since G is reduced, then $\alpha \overset{+}{\underset{G}{\Longrightarrow}} u\beta v$ for all $\beta \neq \alpha$ in V . Thus by (B), $\deg(\alpha, R) \geq \deg(\beta, R)$ for all $\beta \in V$. Let (β_1, β_2), $(\beta_2, \beta_3), \ldots, (\beta_{k-1}, \beta_k) \in R$, where $k > |V|$. Then there exists $1 \leq i < j \leq k$ such that $\beta_i = \beta_j$. By transitivity of R we obtain $(\beta_i, \beta_i) \in R$. Thus $k \geq |V|$ if and only if R is irreflexive. It follows that if R is irreflexive, then $\deg(\alpha, R) \leq |V| - 1$.

(D): Let $(\beta, \beta') \in R$. Property (ii) and (B) imply $\deg(\beta, R) \geq \deg(\beta', R)$. If R is irreflexive, then by (C) $\deg(\beta', R) < |V|$. If $\deg(\beta', R) = 0$, then $(\beta, \beta') \in R$ implies $\deg(\beta, R) \geq 1$ and (D) holds immediately. Suppose $\deg(\beta', R) = k > 0$ and let (β', β_1), (β_1', β_2), ..., $(\beta_{k-1}, \beta_k) \in R$. Then since $(\beta, \beta') \in R$ we can form the chain (β, β'), (β', β_1), ..., (β_{k-1}, β_k). This implies $\deg(\beta, R) \geq k + 1$ $> \deg(\beta', R) = k$.

(E): Supose R is irreflexive and suppose $k = \deg(\beta, R) > 0$. Let (β, β_1), (β_1, β_2), ..., (β_{k-1}, β_k) be a maximal chain in R initiated by β. The existence of such a chain implies $\deg(\beta_1, R) \geq k - 1$. (D) implies $\deg(\beta_1, R) < \deg(\beta, R)$. We therefore conclude that $\deg(\beta_1, R) = k - 1$.

(F): This is the contrapositive of (B).

The relations Δ_ℓ and Δ_r are called the "generalized left and right dominance relations", respectively. Our choice of the tags "left" and "right" for these relations was made for a reason that is not at all clear from the definition. In Workman [13] it is shown that for reduced context-free G, $\deg(\alpha, \Delta_\ell(G)) = 0$ if and only if the set of left-most derivations for G is regular (α denotes the start symbol of G); similarly $\deg(\alpha, \Delta_r(G)) = 0$ if and only if the set of right-most derivations of G is regular. The choice of notation and terminology here is based on these characterizations in terms of the one-sided derivation sets. It should be pointed out that in [13] the designators "left" and "right" are reversed from their use here.

Definition 4.3. Let $G = (V, T, P, \alpha)$ be a context-free grammar. Define the relations $\Delta_\ell(G)$ and $\Delta_r(G)$ on V as follows:

$(\beta_1, \beta_2) \in \Delta_\ell(G)$ (alternatively, $\Delta_r(G)$) if and only if at least one of the following conditions hold in G.

(1) $\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta_1 v\beta_2 w$ for some $uvw \in (V \cup T)^*$

(alternatively, $\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta_2 v\beta_1 w$).

(2) there exists $\beta' \in V$ such that $\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$

for some $uv \in (V \cup T)^*$ and $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta' y\beta_2 z$

for some $xyz \in (V \cup T)^*$ (alternatively,

$\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta_2 y\beta' z$).

Example. We illustrate definition 4.3 by determining the relations Δ_ℓ and Δ_r for the following grammar, G. Let $G = (V, T, P, \alpha)$, where

$V = \{\alpha, \beta_1, \beta_2, \beta_3, \beta_4\}$,

$T = \{a, b\}$

$P = \{1: \alpha \to \beta_1 \alpha \beta_2$

$2: \beta_1 \to \beta_2 \beta_1$

$3: \beta_1 \to \beta_3 \beta_4$

$4: \beta_2 \to \beta_2 \beta_4$

$5: \beta_2 \to \beta_3$

$6: \beta_3 \to a$

$7: \beta_4 \to b$

$8: \alpha \to \varepsilon\}$

$\Delta_\ell = \{(\alpha, \beta_2), (\alpha, \beta_3), (\alpha, \beta_4), (\beta_2, \beta_4)\}$

$\Delta_r = \{(\alpha, \beta_1), (\alpha, \beta_2), (\alpha, \beta_3), (\alpha, \beta_4), (\beta_1, \beta_2), (\beta_1, \beta_3), (\beta_1, \beta_4)\}$.

Note that Δ_ℓ and Δ_r are irreflexive and transitive.
Note also that G is nonexpansive.

Theorem 4.4 gives a characterization of the derivation bounded (nonexpansive) grammars in terms of the relations Δ_ℓ and Δ_r .

Theorem 4.4. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar. The following are equivalent.

(1) G is nonexpansive,

(2) $\Delta_\ell(G)$ is irreflexive,

(3) $\Delta_r(G)$ is irreflexive.

Proof. We show equivalence of (1) and (2) by showing that G is not nonexpansive if and only if $\Delta_\ell(G)$ is not irreflexive. The proof of equivalence of (1) and (3) is similar and will not be given.

If G is not nonexpansive, then there exists $\beta \in V$ such that $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta v\beta w$ for some $uvw \in (V \cup T)^*$. By definition of $\Delta_\ell(G)$ it follows that $(\beta,\beta) \in \Delta_\ell(G)$ and hence $\Delta_\ell(G)$ is not irreflexive. Conversely, suppose $(\beta,\beta) \in \Delta_\ell(G)$. Then either $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta v\beta w$ or $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ and $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta' y\beta z$, where $uvwxyz \in (V \quad T)^*$. In the former case it is immediate that G is not nonexpansive. In the latter case we may obtain $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta' yu\beta' vz$ which also implies G is not nonexpansive.

Lemma 4.5. $\Delta_\ell(\Delta_r)$ satisfy properties (i) and (ii) of lemma 4.2. Thus if $\Delta_\ell(\Delta_r)$ are irreflexive for some grammar, G, the conclusions of lemma 3.9 hold for $\Delta_\ell(G)$ $(\Delta_r(G))$.

Proof. A proof will be given for Δ_ℓ; the proof for Δ_r is similar and will not be presented. Property (ii) of lemma 4.2 is immediate from the definition of $\Delta_\ell(G)$. To show property (i) suppose $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ for some $uv \in (V \cup T)^*$ and suppose $(\beta',\beta'') \in \Delta_\ell(G)$. Then either $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta' y\beta'' z$ or $\beta' \overset{+}{\underset{G}{\Longrightarrow}} u'\gamma v'$ and $\gamma \overset{+}{\underset{G}{\Longrightarrow}} x'\gamma y'\beta'' z'$. In the former case $(\beta,\beta'') \in \Delta_\ell(G)$ by (2) of definition 4.3. In the latter case we obtain $\beta \overset{+}{\underset{G}{\Longrightarrow}} uu'\gamma v' v$, which together with $\gamma \overset{+}{\underset{G}{\Longrightarrow}} x'\gamma y'\beta'' z'$ also implies $(\beta,\beta'') \in \Delta_\ell(G)$. Thus (i) of lemma 4.2 holds for $\Delta_\ell(G)$.

If G is a reduced context-free grammar for which $\Delta_\ell(G)$ (and hence $\Delta_r(G)$) is irreflexive, then by virtue of lemma 4.2 we can assign to G a unique pair of nonnegative integers $\deg(\alpha, \Delta_\ell(G))$ and $\deg(\alpha, \Delta_r(G))$, where α is the start symbol of G. These integers, called the "left degree" and "right degree" of G, respectively, induce natural hierarchies of grammar classes within the class of all derivation bounded grammars. The next definition formalizes these ideas and introduces the grammar classes, $\mathscr{G}_\ell(k)$, $\mathscr{G}_r(k)$, and their corresponding language classes, $\mathscr{L}_\ell(k)$ and $\mathscr{L}_r(k)$. We shall refer to the class $\mathscr{G}_\ell(k)$ ($\mathscr{L}_\ell(k)$) as the class of "left dominant grammars (languages) of degree k". We similarly describe $\mathscr{G}_r(k)$ ($\mathscr{L}_r(k)$).

Definition 4.6. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar. The <u>left-degree of G</u> <u>(right-degree of G)</u>, denoted <u>ℓdeg(G)</u> <u>(rdeg(G))</u>, is defined by,

$$\deg(G) = \deg(\alpha,\Delta_\ell) \quad (\mathrm{rdeg}(G) = \deg(\alpha,\Delta_r)).$$

Furthermore define,

$$\mathscr{G}_\ell(k) = \{G \mid G \text{ is a reduced context-free grammar such that } \ell\mathrm{deg}(G) \leq k\},$$

$$\mathscr{G}_r(k) = \{G \mid G \text{ is a reduced context-free grammar such that } \mathrm{rdeg}(G) \leq k\},$$

$$\mathscr{L}_\ell(k) = \{L(G) \mid G \in \mathscr{G}_\ell(k)\},$$

$$\mathscr{L}_r(k) = \{L(G) \mid G \in \mathscr{G}_r(k)\},$$

$$\mathscr{G}_\ell = \{G \mid \text{ there exists } k<\infty \text{ such that } G \in \mathscr{G}_\ell(k)\},$$

$$\mathscr{G}_r = \{G \mid \text{ there exists } k<\infty \text{ such that } G \in \mathscr{G}_r(k)\},$$

$$\mathscr{L}_\ell = \{L(G) \mid G \in \mathscr{G}_\ell\},$$

$$\mathscr{L}_r = \{L(G) \mid G \in \mathscr{G}_r\},$$

Theorem 4.7. Let $\mathscr{G}$ be the class of all reduced, nonexpansive context-free grammars and let $\mathscr{L}$ be the corresponding class of languages. Then,

(1) $\mathscr{G}_\ell = \mathscr{G}_r = \mathscr{G}$.

(2) $\mathscr{L}_\ell = \mathscr{L}_r = \mathscr{L}$.

Proof. This follows directly from the fact that $\deg(\alpha, \Delta_\ell(G)) < \infty$ ($\deg(\alpha, \Delta_r(G)) < \infty$) if and only if $\Delta_\ell(G)$ ($\Delta_r(G)$) is irreflexive and theorem 4.4.

Theorem 4.7 simply states the fact that the left (right) grammars of finite degree exhaust the class of all reduced derivation bounded grammars. The following result places the nonterminal bounded grammars of Banerji [2] within the hierarchies $\mathscr{G}_\ell$ and $\mathscr{G}_r$.

Theorem 4.8. If G is a reduced context-free grammar, then G is nonterminal bounded if and only if $G \in \mathscr{G}_\ell(0) \cap \mathscr{G}_r(0)$.

Proof. The nonterminal bounded context-free grammars were characterized in Banerji [2] as those grammars for which the "dominance" relation, $\gg$, is irreflexive. This relation is defined on the nonterminal set of $G = (V, T, P, \alpha)$ as follows:

$$\beta_1 \gg \beta_2 \text{ if and only if } \beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta_2 v, \text{ where } uv \in (V \cup T)^* - T^*.$$

What we shall demonstrate is that $\gg$ is irreflexive if and only if $\deg(\alpha, \Delta_\ell(G)) = \deg(\alpha, \Delta_r(G)) = 0$.

Suppose $\gg$ is not irreflexive, then $\beta \gg \beta$ for some $\beta \in V$. This implies that $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta v$, where $uv \in (V \cup T)^* - T^*$. Thus either

$u = x\beta'y$ or $v = x\beta'y$ for some $\beta' \in V$. In the former case, $(\beta, \beta') \in \Delta_r(G)$. In the latter case $(\beta, \beta') \in \Delta_\ell(G)$. It follows from lemma 3.9 that

$\deg(\alpha, \Delta_r(G)) \geq \deg(\beta, \Delta_r(G)) \geq 1$ or $\deg(\alpha, \Delta_\ell(G)) \geq \deg(\beta, \Delta_\ell(G)) \geq 1$.

Now suppose $\deg(\alpha, \Delta_\ell(G)) \neq 0$ or $\deg(\alpha, \Delta_r(G)) \neq 0$. In the former case we have that $(\alpha, \beta) \in \Delta_\ell(G)$ for some $\beta \in V$. Therefore either $\alpha \overset{+}{\underset{G}{\Longrightarrow}} u\alpha v\beta w$, implying $\alpha \gg \alpha$, or $\alpha \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ and $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta' y\beta z$, implying $\beta' \gg \beta'$. Thus $\gg$ is not irreflexive. The argument is similar if $\deg(\alpha, \Delta_r(G)) \neq 0$. This completes the proof.

Corollary. $\mathscr{A}_\ell(0) \cap \mathscr{A}_r(0)$ contains the class of all nonterminal bounded languages.

Theorem 4.9. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar. There is an effective procedure for computing $\deg(\alpha, \Delta_\ell(G))$ and $\deg(\alpha, \Delta_r(G))$.

Proof. For each $\beta \in V$ define $D(\beta) = \{\beta' \in V \mid \beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$ for some $uv \in (V \cup T)^*\}$. It is easily shown that there is an effective procedure for determining $D(\beta)$.

The algorithm described below computes $\deg(\alpha, \Delta_\ell(G))$. The procedure for computing $\deg(\alpha, \Delta_r(G))$ is analogous and will not be given. To this end let $\beta_1, \beta_2, \ldots, \beta_n$ be some enumeration of V and let $\{p_1, p_2, \ldots, p_r\} = P'$ be the set of all productions of P for which the right-part of p_j, $1 \leq j \leq r$, contains at least two occurrences of nonterminal symbols.

Step 0. If $P' = \Phi$, then G is a reduced linear grammar and by theorem 4.8, $G \in \mathcal{G}_\ell(0)$. Thus $\deg(\alpha, \Delta_\ell(G)) = 0$. If $P' \neq \Phi$, then continue to step 1.

Step 1. For $i = 1, 2, \ldots, n$ compute $Q(\beta_i) = \{\beta' \in V \mid \beta_i \overset{*}{\underset{G}{\Longrightarrow}} u\beta_i u\beta' w,$ $uvw \in (V \cup T)^*\}$. It is easily seen that

$\beta_i \overset{*}{\underset{G}{\Longrightarrow}} u\beta_i v\beta' w$ if and only if there exists $\delta \in D(\beta_i)$ such that

$(p_j : \delta \to w_1 \gamma w_2 \gamma' w_3) \in P'$, where $w_1 w_2 w_3$ belong to $(V \cup T)^*$ and $\beta_i \in D(\gamma)$

and $\beta' \in D(\gamma') \cup \{\gamma'\}$. Thus we obtain the following procedure for determining $Q(\beta_i)$.

1a. Set $j = 1$.

1b. Let $(p_j : \delta \to x_0 \gamma_{j1} x_1 \cdots \gamma_{jm_j} x_{m_j}) \in P'$, where
$x_0 x_j \cdots x_{m_j} \in T^*$ and $\gamma_{js} \in V$, $1 \leq s \leq m_j$.

1c. If $\delta \in D(\beta_i)$, then continue, else go to 1e.

1d. For $s = 1, 2, \ldots, m_j - 1$ set

$Q(\beta_i) = Q(\beta_i) \cup D(\gamma_{j,s+1}) \cup \{\gamma_{j,s+1}\}$ if and only if

there exists $t \leq s$ such that $\beta_i \in D(\gamma_{jt})$.

1e. Increment j. If $j \leq r$, then go to 1b, else go to 1a with the next value of i.

Step 2. Since $\Delta_\ell(G)$ is not irreflexive if and only if there exists $\beta \in V$ such that $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta v\beta w$, then $\Delta_\ell(G)$ is not irreflexive if and only if $\beta \in Q(\beta)$ for some $\beta \in V$. If this is the case, then halt with $\deg(\alpha, \Delta_\ell(G)) = \infty$, otherwise continue.

Step 3. For each i, $1 \leq i \leq n$, determine $R(\beta_i) = \{\beta' \in V \mid (\beta_i, \beta') \in \Delta_\ell(G)\}$. By definition of $\Delta_\ell(G)$, $R(\beta_i) = Q(\beta_i) \cup (\bigcup_{\gamma \in D(\beta_i)} Q(\gamma))$.

Upon entry to this step it is known that $\Delta_\ell(G)$ is irreflexive. Therefore since $\Delta_\ell(G)$ is also transitive (lemma 4.2 A), then it follows that for some $\beta \in V$, $R(\beta) = \Phi$. Thus $R(\beta) = \Phi$ if and only if $\deg(\beta, \Delta_\ell(G)) = 0$.

Step 4. Set $S_0 = \{\beta \in V \mid R(\beta) = \Phi\}$. $S_0 = \{\beta \in V \mid \deg(\beta, \Delta_\ell(G)) = 0\}$ Set $k = 0$ and continue.

Step 5. If $\alpha \in S_k$, then halt with $\deg(\alpha, \Delta_\ell(G)) = k$. Otherwise set $k = k + 1$ and continue.

Step 6. Set $S_k = \{\beta \in V - \bigcup_{j=0}^{k-1} S_j \mid R(\beta) \subseteq \bigcup_{j=0}^{k-1} S_j\}$. Go to step 5.

If for some $k > 0$, $S_k = \Phi$, then either $V = \bigcup_{j=0}^{k-1} S_j$ or for all $\beta \in V - \bigcup_{j=0}^{k-1} S_j$ it holds that $R(\beta) \cap (V - \bigcup_{j=0}^{k-1} S_j) \neq \Phi$. If the latter case is true it follows from transitivity of $\Delta_\ell(G)$, that $(\beta, \beta) \in \Delta_\ell(G)$ for some β in $V - \bigcup_{j=0}^{k-1} S_j$. This is in contradiction to the fact that $\Delta_\ell(G)$ is known to be irreflexive at this point of the computation.

Suppose that $\alpha \in S_k$ for some k and $S_{k+1} \neq \Phi$. Since G is reduced it follows that $D(\alpha) = V$ and hence $R(\beta) \subseteq R(\alpha)$ for all $\beta \in V$.

But if $\beta \in S_{k+1}$ and $\alpha \in S_k$, then we have that

$R(\beta) \not\subseteq \bigcup_{j=0}^{k-1} S_j$ while $R(\alpha) \subseteq \bigcup_{j=0}^{k-1} S_j$. This implies $R(\beta) \not\subseteq R(\alpha)$, a contradiction. This implies that $\alpha \in S_k$, where k is the least integer for which $S_{k+1} = \Phi$. The loop defined by steps 5 and 6 must therefore terminate with α assigned to the last non-void set S_k, computed in step 5.

Finally, by a simple inductive argument it can be shown that $\beta \in S_k$ if and only if $\deg(\beta, \Delta_\ell(G)) = k$. Thus the procedure eventually halts having determined $\deg(\alpha, \Delta_\ell(G)) = \ell\deg(G)$.

5. A Characterization of the Right and Left Dominant Languages.

The major results of this section are theorems 5.5 and 5.6. They present a characterization of the classes $\mathscr{A}_\ell(k)$ and $\mathscr{A}_r(k)$, respectively, in terms of special types of substitution applied to the class of strictly linear languages.

To establish the characterizations we introduce, for each $k \geq 0$, a relation ρ_k defined on the nonterminal set of grammars for which Δ_ℓ is irreflexive. Lemma 5.2 establishes that ρ_k is irreflexive and satisfies lemma 4.2. As a consequence of the properties of ρ_k, we are able to decompose $\mathscr{G}_\ell(k)$ into a hierarchy of grammar classes, $\mathscr{G}_\ell(k, j)$, $j \geq 0$. Our characterization in theorem 5.5 is based on this decomposition. In a similar manner, relations λ_k, $k \geq 0$, are defined to obtain an analogous decomposition of the grammars in $\mathscr{G}_r(k)$, $k \geq 0$.

Lemma 5.3 is a technical result which is used primarily to simplify the proof of theorem 5.5. Definition 5.4 introduces the substitution mechanism employed in the characterization theorems.

Definition 5.1. Let $G = (V, T, P, \alpha) \in \mathscr{G}$ (see theorem 4.7). For each $i \geq 0$ define

$V_\ell^{(i)} = \{\beta \in V \mid \deg(\beta, \Delta_\ell) = i\}$ and

$V_r^{(i)} = \{\beta \in V \mid \deg(\beta, \Delta_r) = i\}$. For each $i \geq 0$ define the relations $\rho_i(G)$ and $\lambda_i(G)$ on V as follows:

$(\beta_1,\beta_2)\in\rho_i(G)$ if and only if,

$\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta_2 v\beta' w$, where β_2, $\beta'\in V_\ell^{(i)}$ and

$uvw\in(V\cup T)^*$;

similarly,

$(\beta_1,\beta_2)\in\lambda_i(G)$ if and only if,

$\beta_1 \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v\beta_2 w$, where $\beta',\beta_2\in V_r^{(i)}$.

Lemma 5.2. Let $G\in\mathcal{G}$. Therefore each $i\geq 0$ $\rho_i(G)$ and $\lambda_i(G)$ are irreflexive and satisfy properties (i) and (ii) of lemma 4.2.

Proof. We will prove these properties for $\rho_i(G)$, $i\geq 0$; the proof for $\lambda_i(G)$ is similar and will therefore be omitted.

Since $G\in\mathcal{G}$, then $\Delta_\ell(G)$ is irreflexive by theorem 3.12. If $\rho_i(G)$ is not irreflexive, then for some $\beta\in V$ it must be the case that $(\beta,\beta)\in\rho_i(G)$. This implies that $\beta\in V_\ell^{(i)}$ and there exists $\beta'\in V_\ell^{(i)}$ such that $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta v\beta' w$ for some $uvw\in(V\cup T)^*$. But by definition of $\Delta_\ell(G)$ it follows that $(\beta,\beta')\in\Delta_\ell$ implying by lemma 4.2 that $\deg(\beta',\Delta_\ell)<\deg(\beta,\Delta_\ell)$. This contradicts the fact that $\beta,\beta'\in V_\ell^{(i)}$ which implies $\deg(\beta,\Delta_\ell)=\deg(\beta',\Delta_\ell)=i$. Thus $\rho_i(G)$ must be irreflexive.

By definition of $\rho_i(G)$ it follows at once that property (ii) of lemma 4.2 is satisfied. To show property (i) suppose that $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ for some $uv \in (V \cup T)^*$ and suppose that $(\beta', \beta'') \in \rho_i(G)$. Then $\beta' \overset{+}{\underset{G}{\Longrightarrow}} x\beta'' y\gamma z$, where $\beta'', \gamma \in V_\ell^{(i)}$ and $xyz \in (V \cup T)^*$. But then we have that $\beta \overset{+}{\underset{G}{\Longrightarrow}} ux\beta'' y\gamma zv$ and it follows that $(\beta, \beta'') \in \rho_i(G)$. This concludes the proof.

Lemma 5.3. Let $G = (V, T, P, \alpha) \in \mathcal{G}$ and let $Z_{ij} = \{\beta \in V | \deg(\beta, \Delta_\ell(G)) = i$ and $\deg(\beta, \rho_i(G)) = j\}$. If $k = \ell\deg(G) = \deg(\alpha, \Delta_\ell(G))$, then

(A) For each i, $0 \leq i \leq k$, there exists n_i, $0 \leq n_i < |V|$, such that $V = \bigcup_{i=0}^{k} (\bigcup_{j=0}^{n_i} Z_{ij})$ where $Z_{ij} \neq \Phi$ if and only if $0 \leq i \leq k$ and $0 \leq j \leq n_i$ and $Z_{ij} \cap Z_{rs} = \Phi$ if $i \neq r$ or $j \neq s$.

(B) For all $\beta \in Z_{ij}$, $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ implies $\beta' \in Z_{rs}$, where either $0 \leq r < i$ and $0 \leq s \leq n_r$ or $r = i$ and $0 \leq s \leq j$ $(uv \in (V \cup T)^*)$.

(C) For all $\beta \in Z_{ij}$, $0 \leq i \leq k$, $0 \leq j \leq n_i$, the grammar $G(\beta, U_{ij})$ is linear over $(T \cup U_{ij}, T \cup U_i')$ biased left, where

$U_i' = \Phi$ if $i = 0$,

$U_i' = \bigcup_{q<i} (\bigcup_{j \leq n_q} Z_{qj})$ if $i > 0$,

$U_{ij} = U_i'$ if $j = 0$ and

$U_{ij} = U_i' \cup (\bigcup_{q<j} Z_{iq})$ if $j > 0$.

(D) For all $\beta \in Z_{ij}$, $0 \leq i \leq k$ and $0 \leq j \leq n_i$, $L(G(\beta)) = \sigma(L(G(\beta,U_{ij})))$, where σ is the substitution defined by $\sigma(a) = a$ for all $a \in T$ and $\sigma(\gamma) = L(G(\gamma))$ for all $\gamma \in U_{ij}$. Furthermore, if $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$ for all $\beta' \in Z_{ij}$ and some $uv \in (V \cup T)^*$, then $G(\beta,U_{ij}) = (Z_{ij}, T \cup U_{ij}, P_{ij}, \beta)$, where $P_{ij} = \{(\beta \to w) \in P \mid \beta \in Z_{ij}\}$.

Proof of (A). Since $G \in \mathscr{G}$, $\Delta_\ell(G)$ is irreflexive. Thus by lemma 4.2C $0 \leq \deg(\beta, \Delta_\ell(G)) \leq \deg(\alpha, \Delta_\ell(G)) = k < |V|$ for all $\beta \in V$. It follows that $Z_{ij} = \Phi$ for all $i > k$. Lemma 4.2E guarantees that $V_\ell^{(i)} \neq \Phi$ for each $i \leq k$ (see definition 5.1). By lemma 5.2 $\rho_i(G)$ is irreflexive for each $i \geq 0$ and by lemma 4.2C $0 \leq \deg(\beta, \rho_i(G)) \leq \deg(\alpha, \rho_i(G)) < |V|$ for all $\beta \in V$. Thus for each i, $0 \leq i \leq k$, $n_i = \max\{\deg(\beta, \rho_i(G)) \mid \beta \in V_\ell^{(i)}\}$ exists and $Z_{ij} = \Phi$ for all $j > n_i$. What remains to be shown is that $Z_{ij} \neq \Phi$ for $0 \leq j \leq n_i$. Clearly $Z_{in_i} \neq \Phi$. Suppose $\beta \in Z_{ij}$ for some $j > 0$. From the proof of lemma 4.2E it follows that there exists $\beta' \in V$ such that $(\beta, \beta') \in \rho_i(G)$ and $\deg(\beta', \rho_i(G)) = j - 1$. By definition of $\rho_i(G)$ it follows that $\beta' \in V_\ell^{(i)}$ and thus $Z_{ij-1} \neq \Phi$. It follows that $Z_{ij} \neq \Phi$ for $0 \leq j \leq n_i$.

Finally, since $\deg(., \Delta_\ell(G))$ and $\deg(., \rho_i(G))$ are functions, it follows that $Z_{ij} \cap Z_{rs} = \Phi$ whenever $i \neq r$ or $j \neq s$.

Proof of (B). If $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' v$ for some $uv \in (V \cup T)^*$, then by lemma 4.2(B,F) it follows that $\deg(\beta', \Delta_\ell(G)) \leq \deg(\beta, \Delta_\ell(G))$ and $\deg(\beta', \rho_i(G)) \leq \deg(\beta, \rho_i(G))$ for all $i \geq 0$. The result follows immediately from these relations and the definition of Z_{ij}.

Proof of (C). Suppose $\beta \in Z_{ij}$ for some i and j and let $(\beta \to w) \in P$. What must be shown is that,

(1) $w \in (U_{ij} \cup Z_{ij} \cup T)^*$ and

(2) if $w = u\beta' v$, where $\beta' \in Z_{ij}$ then

$u \in (U_{ij} \cup T)^*$ and $v \in (U_i' \cup T)^*$.

Since $(\beta \to w) \in P$, then $\beta \overset{+}{\underset{G}{\Longrightarrow}} w$. Therefore if $w = u\beta' v$, $\beta' \in V$, it follows from (B) above that $\beta' \in Z_{ij} \cup U_{ij}$. Thus $w \in (U_{ij} \cup Z_{ij} \cup T)^*$. In the remainder of the proof we drop the "(G)" when referring to $\Delta_\ell(G)$ and $\rho_i(G)$.

Now suppose $(\beta \to u\beta' v) \in P$, where $\beta' \in Z_{ij} \subseteq V_\ell^{(i)}$. Either $v \in T^*$ or $v = x\beta'' y$, where $xy \in (V \cup T)^*$ and $\beta'' \in U_{ij} \cup Z_{ij}$. Assume the latter case and suppose $\beta'' \in V_\ell^{(i)}$. Since $\beta', \beta'' \in V_\ell^{(i)}$ and since $\beta \overset{+}{\underset{G}{\Longrightarrow}} u\beta' x\beta'' y$, then by definition of ρ_i it follows that $(\beta, \beta') \in \rho_i$. By lemma 4.2D, $j = \deg(\beta', \rho_i) < \deg(\beta, \rho_i) = j$, a contradiction. Thus $\beta'' \in U_{ij} - V_\ell^{(i)} = U_i'$ and $0 \leq \deg(\beta'', \Delta_\ell) < \deg(\beta, \Delta_\ell) = i$. But this is possible only if $i > 0$. Thus if $i = 0$ we must conclude that $v \in T^* = (\Phi \cup T)^* = (U_0' \cup T)^*$. In either case it follows that $v \in (U_i' \cup T)^*$ for all $i \geq 0$.

Consider u. Again, either $u \in T^*$ or $u = x\beta''y$, where $xy \in (V \cup T)^*$ and $\beta'' \in Z_{ij} \cup U_{ij}$. If $\beta'' \in V_\ell^{(i)}$, then $\beta \overset{+}{\underset{G}{\Longrightarrow}} x\beta''y\beta'v$ and it follows that $(\beta,\beta'') \in \rho_i$ which implies by lemma 4.2D that $\deg(\beta'',\rho_i) < \deg(\beta,\rho_i) = j$. This is possible only if $j > 0$. Thus if $j > 0$, and $\beta'' \in V_\ell^{(i)}$, then $\beta'' \in Z_{iq}$ for some $q < j$. If $j = 0$, then $\beta'' \notin V_q^{(i)}$ and by (B) above it follows that $\deg(\beta'',\Delta_\ell) < i$. But this is possible only if $i > 0$. Therefore, if $i > 0$ and $j = 0$, then $\beta'' \in U_i'$. Finally, if $i = 0$ and $j = 0$, then β'' cannot exist and we conclude $u \in T^*$. In all cases $u \in (U_{ij} \cup T)^*$.

Finally, note that if $\beta \to w$ is a terminating production of $G(\beta, U_{ij})$, then w contains no elements of Z_{ij}. Thus from (1) above, $w \in (T \cup U_{ij})^*$ and $G(\beta, U_{ij})$ is biased left over $(T \cup U_{ij}, T \cup U_i')$.

Proof of (D). Clearly $U_{ij} \subseteq V - Z_{ij}$. Furthermore, if $\gamma \in U_{ij}$ and $\gamma \overset{+}{\underset{G}{\Longrightarrow}} u\gamma'v$ for some $\gamma' \in V$ and $uv \in (V \cup T)^*$, then by (B) above it follows that $\gamma' \in U_{ij}$. Thus by lemma 3.7 the result follows when U is taken to be U_{ij}.

Definition 5.4. Let $\mathscr{L}$ denote the class of all strictly linear languages. Let $\mathscr{A}$ and $\mathscr{B}$ represent language classes. Define

$(\mathscr{A},\mathscr{B}) = \{L \mid L = \tau(L')$, where $L' \in \mathscr{L}$ is strictly linear over (Σ_ℓ, Σ_r) for some such pair and τ is a substitution defined by, $\tau(a) \in \mathscr{A}$ for all $a \in \Sigma_\ell$ and $\tau(b) \in \mathscr{B}$ for all $b \in \Sigma_r$ $\cdot\}$.

Theorem 5.5. Let $\mathscr{R}$ be the class of all regular sets.

1. Let $\mathscr{C}_0^{(0)} = \mathscr{L}(\mathscr{R}, \mathscr{R})$ and define $\mathscr{C}_{i+1}^{(0)} = \mathscr{L}(\mathscr{C}_i^{(0)}, \mathscr{R})$ for $i \geq 0$. Then $L \in \mathscr{U}_\ell(0)$ if and only if there exists $j \geq 0$ such that $L \in \mathscr{C}_j^{(0)}$.

2. For $k > 0$ define $\mathscr{C}_0^{(k)} = \mathscr{L}(\mathscr{U}_\ell(k-1), \mathscr{U}_\ell(k-1))$ and $\mathscr{C}_{i+1}^{(k)} = \mathscr{L}(\mathscr{C}_i^{(k)}, \mathscr{U}_\ell(k-1))$, $i \geq 0$. Then $L \in \mathscr{U}_\ell(k)$ if and only if there exists $j \geq 0$ such that $L \in \mathscr{C}_j^{(k)}$.

3. Define $\mathscr{G}_\ell(i, j) = \{G \in \mathscr{G} \mid \deg(\alpha, \Delta_\ell(G)) \leq i$ and $\deg(\alpha, \rho_i(G)) \leq j$, where α is the start symbol of $G\}$, where $\mathscr{G}$ is the class of all reduced non-expansive grammars. Then $L \in \mathscr{C}_j^{(i)}$ if and only if $L = L(G)$ for some $G \in \mathscr{G}_\ell(i, j)$, $j, i \geq 0$.

Proof. The proof will consist of first showing that (1) and (2) are equivalent to (3) and then demonstrating (3). Suppose (3) holds. Let $L \in \mathscr{L}_j^{(i)}$, then $L = L(G)$ for some $G = (V, T, P, \alpha) \in \mathscr{G}_\ell(i, j)$. This implies $\deg(\alpha, \Delta_\ell(G)) \leq i$ and thus $G \in \mathscr{G}_\ell(i)$. It follows by definition of $\mathscr{A}_\ell(i)$ that $L(G) \in \mathscr{A}_\ell(i)$. Thus $\mathscr{L}_j^{(i)} \subseteq \mathscr{A}_\ell(i)$ for every $j \geq 0$. Now let $L \in \mathscr{A}_\ell(i)$. Then there exists $G \in \mathscr{G}_\ell(i)$ such that $L = L(G)$. Since $G \in \mathscr{G}_\ell(i) \subseteq \mathscr{G}$, then $\Delta_\ell(G)$ and $\rho_i(G)$ are irreflexive and thus by lemma 4.2C, $\deg(\alpha, \rho_i(G)) < |V|$ implying $G \in \mathscr{G}_\ell(i, |V| - 1)$. By (3) it follows that $L(G) \in \mathscr{L}_{|V|-1}^{(i)}$. Therefore $L \in \mathscr{A}_\ell(i)$ if and only if there exists j such that $L \in \mathscr{L}_j^{(i)}$. The proof will be complete if (3) can be established.

($\Leftarrow$): Let $G = (V, T, P, \alpha) \in \mathscr{G}_\ell(i, 0)$ and let $k = \ell\deg(G) = \deg(\alpha, \Delta_\ell(G)) \leq i$. We show that $L(G) \in \mathscr{L}_0^{(i)}$. If $0 \leq k < i$, then $G \in \mathscr{G}_\ell(i - 1)$ and $L(G) \in \mathscr{A}_\ell(i - 1)$. Let $L' = \{a\}$. Clearly L' is strictly linear over $(\{a\}, \Phi)$. If we choose the substitution, τ, such that $\tau(a) = L(G)$, then clearly $\tau(L') = \tau(a) = L(G) \in \mathscr{L}_0^{(i)}$. Therefore suppose $k = i$. Since G is reduced, then $L(G) = L(G(\alpha))$ and by lemma 5.3D, $L(G(\alpha)) = \sigma(L(G(\alpha, U_{i,0}))) = \sigma(L(G(\alpha, U_i')))$, where

$G(\alpha, U_i')$ is linear over $(T \cup U_i', \overline{T \cup U_i'})$ biased left and σ is the substitution defined by, $\sigma(a) = \{a\}$ for all $a \in T$ and $\sigma(\gamma) = L(G(\gamma))$ for all $\gamma \in U_i'$. Let $\overline{G}(\alpha, U_i')_\ell$ be the strict image of $G(\alpha, U_i')$ constructed as in definition 3.3. $\overline{G}(\alpha, U_i')_\ell$ is strictly linear over (Σ_ℓ, Σ_r) biased left, where $\Sigma_\ell \subseteq T \cup U_i'$ and $\Sigma_r \subseteq \overline{T \cup U_i'}$. By proposition 3.4, $L(G(\alpha, U_i')) = \overline{h}(L(\overline{G}(\alpha, U_i')_\ell))$, where $\overline{h}$ is the unmarking homomorphism on $T \cup U_i'$. Define the substitution $\tau = \sigma\overline{h}$. Clearly $\tau(L(\overline{G}(\alpha, U_i')_\ell)) = (\overline{h}(L(\overline{G}(\alpha, U_i')_\ell)) = \sigma(L(G(\alpha, U_i'))) = L(G)$. From lemma 5.3(A, C) it follows that $U_i' \neq \Phi$ if and only if $i > 0$. For all $a \in T$ and $\overline{a} \in \overline{T}$, $\tau(a) = \tau(\overline{a}) = \{a\}$. For all $\gamma \in U_i'$ and $\overline{\gamma} \in \overline{U}_i'$, $\tau(\gamma) = \tau(\overline{\gamma}) = L(G(\gamma))$. By definition of U_i', if $\gamma \in U_i'$, then $\deg(\gamma, \Delta_\ell(G)) \leq i - 1$. This implies $\ell\deg(G(\gamma)) \leq i - 1$ and therefore $G(\gamma) \in \mathscr{G}_\ell(i - 1)$ implying $L(G(\gamma)) \in \mathscr{U}_\ell(i - 1)$. By definition of $\mathscr{U}_\ell(k)$ it follows that $\mathscr{U}_\ell(0) \subseteq \mathscr{U}_\ell(k)$ for all $k \geq 0$. By the corollary to theorem 4.8 it follows that the singleton sets $\tau(a) = \tau(\overline{a}) = \{a\}$, which are regular, belong to $\mathscr{U}_\ell(i - 1)$ as well. Thus if $i > 0$, then $\tau(L(\overline{G}(\alpha, U_i')_\ell)) = L(G) \in \mathscr{L}_0^{(i)}$. If $i = 0$, then $U_i' = \Phi$ and τ is a regular substitution implying that $\tau(L(\overline{G}(\alpha, U_i')_\ell)) = L(G) \in \mathscr{L}_0^{(0)}$. Thus $G \in \mathscr{G}_\ell(i, 0)$ implies $L(G) \in \mathscr{L}_0^{(i)}$, for all $i \geq 0$.

Now assume that $G \in \mathscr{G}_\ell(i, j)$ implies $L(G) \in \mathscr{L}_j^{(i)}$. We show that $G \in \mathscr{G}_\ell(i, j+1)$ implies $L(G) \in \mathscr{L}_{j+1}^{(i)}$. By applying substitutions to singleton sets it clearly follows that $\mathscr{L}_j^{(i)} \subseteq \mathscr{L}_{j+1}^{(i)}$ for all $j \geq 0$. Thus if $G = (V, T, P, \alpha)$ and $\deg(\alpha, \Delta_\ell(G)) < i$ or if $\deg(\alpha, \rho_i(G)) < j + 1$, then $G \in \mathscr{G}_\ell(i, j)$ and by our previous remark together with the induction hypothesis it follows that $L(G) \in \mathscr{L}_{j+1}^{(i)}$. Assume therefore that $\ell\deg(G) = i$ and $\deg(\alpha, \rho_i(G)) = j + 1$. By lemma 5.3D and an argument similar to that given above it follows that $L(G) = \tau(L(\bar{G}(\alpha, U_{i,j+1})_\ell))$, where $\tau = \sigma\bar{h}$ as before and $\bar{G}(\alpha, U_{i,j+1})_\ell$ is strictly linear over (Σ_ℓ, Σ_r) biased left such that $\Sigma_\ell \subseteq T \cup U_{i,j+1}$ and $\Sigma_r \subseteq \overline{T \cup U_i'}$.

Suppose $i = 0$. Then $U_i' = \Phi$ and $U_{i,j+1} = \{\beta \in V \mid \deg(\beta, \rho_0(G)) \leq j\}$. Consider $\tau(c)$ for $c \in \Sigma_\ell$. If $c \in T$, then $\tau(c) = \{c\}$ is regular and clearly belongs to $\mathscr{L}_0^{(0)} \subseteq \mathscr{L}_j^{(0)}$. If $c = \gamma \in U_{0,j+1}$, then $\tau(\gamma) = L(G(\gamma))$. By definition of $U_{0,j+1}$ it follows that $G(\gamma) \in \mathscr{G}_\ell(0, j)$ and thus $L(G(\gamma)) \in \mathscr{L}_j^{(0)}$ by the induction hypothesis. For all $\bar{a} \in \Sigma_r = \bar{T}$, $\tau(\bar{a}) = \{a\}$ is regular. Thus by definition of $\mathscr{L}_{j+1}^{(0)}$ it follows that $L(G) = \tau(L(\bar{G}(\alpha, U_{0,j+1})_\ell)) \in \mathscr{L}_{j+1}^{(0)}$. If $i > 0$, then it follows that $\mathscr{L}_j^{(i)} \supseteq \mathscr{L}_0^{(i)} \supseteq \mathscr{A}_\ell(i-1) \supseteq \mathscr{A}_\ell(0) \supseteq \mathscr{R}$; the last inclusion follows from the corollary to theorem 4.8. For $\gamma \in U_{i,j+1}$ it follows that $G(\gamma) \in \mathcal{G}_\ell(i, j)$ and by the induction hypothesis $\tau(\gamma) = L(G(\gamma)) \in \mathcal{L}_j^{(0)}$. For $\gamma \in U_i'$ it follows that $G(\gamma) \in \mathcal{G}_\ell(i - 1)$ and hence $\tau(\bar{\gamma}) = L(G(\gamma)) \in \mathcal{M}_\ell(i - 1)$. Therefore $\tau(a) \in \mathcal{L}_j^{(i)}$ for all $a \in \Sigma_\ell$ and $\tau(a) \in \mathcal{M}_\ell(i - 1)$ for all $a \in \Sigma_r$. By definition of $\mathcal{L}_{j+1}^{(i)}$ it follows that $\tau(L(\bar{G}(\alpha, U_{i,j+1})_\ell)) = L(G) \in \mathcal{L}_{j+1}^{(i)}$.

($\Longrightarrow$): We conclude the proof with a demonstration that $L \in \mathcal{L}_j^{(i)}$ implies $L = L(G)$ for some $G \in \mathcal{G}_\ell(i, j)$. This will be shown by induction on j for each $i \geq 0$.

Let $L \in \mathcal{L}_0^{(i)}$, then there exists a strictly linear language, L', over (Σ_ℓ, Σ_r) such that $L = (L')$, where $\tau(a) \in \mathcal{M}_\ell(i - 1)$ for all $a \in \Sigma = \Sigma_\ell \cup \Sigma_r$ if $i > 0$ and $\tau(a) \in \mathcal{R}$, $a \in \Sigma$, if $i = 0$. We may assume without loss of generality that Σ_ℓ and Σ_r are the smallest such sets.

Let $G' = (V', \Sigma, P', \alpha)$ be a reduced grammar generating L'. For $i > 0$ and for all $a \in \Sigma$ let $G_a = (V_a, T_a, P_a, \gamma_a) \in \mathcal{G}_\ell(i - 1)$ be a grammar generating $\tau(a)$. We may assume that the sets V_a, $a \in \Sigma$, and V' are pairwise disjoint.

Consider the case when $i = 0$. Since $\tau(a) \in \mathscr{R}$ for all $a \in \Sigma$, then by lemma 3.5, $\tau(L')$ is linear and is therefore generated by some reduced linear grammar, $G = (V, T, P, \alpha)$. By theorem 4.8 $G \in \mathscr{G}_\ell(0)$. If β, $\beta' \in V$ and $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$, then $uv \in T^*$. It follows from the definition of $\rho_0(G)$ that $\deg(\beta, \rho_0(G)) = 0$ for all $\beta \in V$. Thus $G \in \mathscr{G}_\ell(0,0)$.

Suppose $i > 0$ and let $G = (V, T, P, \alpha)$, where

$$V = V' \cup (\bigcup_{a \in \Sigma} V_a),$$

$$T = \bigcup_{a \in \Sigma} T_a \quad \text{and}$$

$$P = \{\beta \to \eta(w) \mid (\beta \to w) \in P'\} \cup (\bigcup_{a \in \Sigma} P_a),$$

where η is the homomorphism defined on $\Sigma \cup V'$ by $\eta(\beta) = \beta$ for all $\beta \in V'$ and $\eta(a) = \gamma_a$ for all $a \in \Sigma$. Since every $a \in \Sigma$ appears in some production of P', since G' and G_a are reduced for each $a \in \Sigma$, then G is also reduced. Furthermore it is clear that $L(G) = \tau(L')$. We now show that $G \in \mathscr{G}_\ell(i, 0)$.

Let $\beta \in V$. If $\beta \in V_a$ for some $a \in \Sigma$, then $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$, for some $\beta' \in V$, implies $\beta' \in V_a$ and thus $\deg(\beta, \Delta_\ell(G)) = \deg(\beta, \Delta_\ell(G_a)) \leq \deg(\gamma_a, \Delta_\ell(G_a)) \leq i - 1$. The last inequalities follow from the fact that $G_a \in \mathscr{G}_\ell(i - 1)$ and lemma 4.2. Since the productions of P are linear in elements of V' and since $\beta \in V - V'$ cannot introduce elements of V' into a derivation, it follows that $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta v\beta' w$, $\beta \in V'$ and $\beta' \in V$, implies $\beta' \in V_a$ for some $a \in \Sigma$. Thus if $(\beta, \beta') \in \Delta_\ell(G)$, where $\beta \in V'$, then $\beta' \in V_a$. From this we conclude that $\deg(\beta, \Delta_\ell(G)) \leq i$. Since every sentential form of G contains at most one occurance of $\beta \in V'$, then it follows from the definition of $\rho_i(G)$, that $\deg(\beta, \rho_i(G)) = 0$ for all $\beta \in V$. Thus $G \in \mathscr{G}_\ell(i, 0)$.

Continuing with the general case, suppose that $L \in \mathscr{L}_j^{(i)}$ implies $L = L(G)$ for some G in $\mathscr{G}_\ell(i, j)$. We show that $L \in \mathscr{L}_{j+1}^{(i)}$ implies $L = L(G)$ for some $G \in \mathscr{G}_\ell(i, j + 1)$.

Let $L = \tau(L')$, where $L' = L(G')$ is strictly linear over (Σ_ℓ, Σ_r) and $\tau(a) \in \mathscr{L}_j^{(i)}$ for $a \in \Sigma_\ell$, $\tau(b) \in \mathscr{R}$, $b \in \Sigma_r$, if $i = 0$ and $\tau(b) \in \mathscr{U}_\ell(i - 1)$ if $i > 0$. Let G' and G_a, $a \in \Sigma$, be those described earlier except that by the induction hypothesis we will assume $G_a \in \mathscr{G}_\ell(i, j)$.

Consider the case $i = 0$. Choose $\bar{\Sigma}_\ell$ to be an abstract set of symbols in one-to-one correspondence with Σ_ℓ such that $\bar{\Sigma}_\ell \cap (\bigcup_{b \in \Sigma_r} T_b) = \phi$.

Let $\sigma_1(a) = \bar{a} \in \bar{\Sigma}_\ell$ for all $a \in \Sigma_\ell$ and let $\sigma_1(b) = \tau(b)$ for all $b \in \Sigma_r$. By lemma 3.5, $\sigma_1(L') = L_1$ is strictly linear over $(\bar{\Sigma}_\ell, T_r)$, where $T_r = \bigcup_{b \in \Sigma_r} T_b$. Now let $G_1 = (V_1, \bar{\Sigma}_\ell \cup T_r, P_1, \alpha)$ be a reduced strictly linear grammar generating L_1. Define $G = (V, T, P, \alpha)$, where

$$V = V_1 \cup (\bigcup_{a \in \Sigma_\ell} V_a),$$

$$T = T_r \cup (\bigcup_{a \in \Sigma_\ell} T_a),$$

$$P = \{\beta \rightarrow \eta(w) \mid (\beta \rightarrow w) \quad P_1\} \cup (\bigcup_{a \in \Sigma_\ell} P_a),$$

where η is the homomorphism defined on $\bar{\Sigma} \cup T_r \cup V_1$ by $\eta(c) = c$ for all $c \in T_r \cup V_1$ and $\eta(\bar{a}) = \gamma_a$, where γ_a is the start symbol of G_a defined earlier. Since each $\bar{a} \in \bar{\Sigma}_\ell$ appears in some production of P_1 and since G_1 and G_a are reduced for each $a \in \Sigma_\ell$, then clearly G is reduced. Furthermore, it is easy to see that $L(G) = \tau(L')$. What remains to be shown is that $G \in \mathcal{G}_\ell(0, j+1)$.

For all $\beta \in V - V_1$, $\beta \in V_a$ for some $a \in \Sigma_\ell$ and therefore from the fact that $G_a \in \mathcal{G}_\ell(0, j)$, lemma 4.2 and the fact that $\beta \overset{*}{\underset{G}{\Longrightarrow}} w$ implies $\beta \overset{*}{\underset{G_a}{\Longrightarrow}} w$, it follows that $\deg(\beta, \Delta_\ell(G)) = \deg(\beta, \Delta_\ell(G_a)) = 0$ and

$\deg(\beta, \rho_0(G)) = \deg(\beta, \rho_0(G_a)) \leq j$. If $\beta, \beta' \in V_1$, then $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$ implies $v \in T_r^*$, thus $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta v$ implies $v \in T_r^*$ and hence $\deg(\beta, \Delta_\ell(G)) = 0$. In addition, the string $u \in (T \cup (V - V_1))^*$. Thus if $\beta \in V_1$ and $(\beta, \beta') \in \rho_0(G)$, then $\beta' \in V_a$ for some $a \in \Sigma_\ell$. From this we can conclude $\deg(\beta, \rho_0(G)) \leq j + 1$ for all $\beta \in V_1$. It follows, therefore, that $G \in \mathscr{G}_\ell(0, j + 1)$.

For $i > 0$ the argument is similar. In this case we construct $G = (V, T, P, \alpha)$ directly from G' :

$$V = V' \cup (\bigcup_{a \in \Sigma} V_a) , \quad \Sigma = \Sigma_\ell \cup \Sigma_r ,$$

$$T = \bigcup_{a \in \Sigma} T_a ,$$

$$P = \{\beta \to \eta(w) \mid (\beta \to w) \in P'\} \cup (\bigcup_{a \in \Sigma} P_a),$$

where $\eta(a) = \gamma_a$ for all $a \in \Sigma$ and $\eta(\beta) = \beta$ for all $\beta \in V'$. It should be noted that $G_a \in \mathscr{G}_\ell(i - 1)$ for all $a \in \Sigma_r$ and $G_a \in \mathscr{G}_\ell(i, j)$ for all $a \in \Sigma_\ell$. The latter holding as a result of the induction hypothesis.

Again it is easily seen that $\tau(L') = L(G)$ and that G is reduced by virtue of the properties ascribed to G_a, G', Σ_ℓ and Σ_r . By arguments presented in the case $i = 0$, it follows that $\deg(\beta, \Delta_\ell(G)) \leq i$ for all $\beta \in V - V'$. For $\beta \in V_a$, $a \in \Sigma_\ell$, it is easily shown that $\deg(\beta, \rho_i(G)) \leq j$. For $\beta \in V_b$, $b \in \Sigma_r$, $\deg(\beta, \Delta_\ell(G)) = \deg(\beta, \Delta_\ell(G_b)) \leq i - 1$. Thus if

$\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta_1 v\beta_2 w$, where $\beta_1, \beta_2 \in V$, then $\beta_1, \beta_2 \in V_b$. This implies that for all $\beta \in \bigcup_{b \in \Sigma_r} V_b$, β can introduce only nonterminals, β', such that $\deg(\beta', \Delta_\ell(G)) \leq i - 1$. Therefore it follows that $\deg(\beta, \rho_i(G)) = 0$, for all $\beta \in \bigcup_{b \in \Sigma_r} V_b$.

Consider $\beta, \beta' \in V'$. If $\beta \overset{*}{\underset{G}{\Longrightarrow}} u\beta' v$, then $v \in (T \cup (\bigcup_{b \in \Sigma_r} V_b))^*$ and $u \in (T \cup (\bigcup_{a \in \Sigma_\ell} V_a)^*$. It follows that $(\beta, \gamma) \in \Delta_\ell(G)$ implies $\gamma \in \bigcup_{b \in \Sigma_r} V_b$ and thus $\deg(\beta, \Delta_\ell(G)) \leq i$. Since at most one occurrence of $\beta \in V'$ can appear in any sentential form of G, then it follows that for $\beta \in V'$, $(\beta, \gamma) \in \rho_i(G)$ implies $\gamma \notin V'$. By previous argument it follows that $\deg(\gamma, \rho_i(G)) \leq j$. Thus $\deg(\beta, \rho_i(G)) \leq j + 1$ for all $\beta \in V'$ and hence also true for all $\beta \in V$. We conclude that $G \in \mathscr{G}_\ell(i, j + 1)$ and thus completing the proof.

Theorem 5.6. Let $\mathscr{R}$ be the class of regular sets. Then,

1. Let $\mathscr{D}_0^{(0)} = \mathscr{L}(\mathscr{R}, \mathscr{R})$. For $j \geq 0$ let $\mathscr{D}_{j+1}^{(0)} = \mathscr{L}(\mathscr{R}, \mathscr{D}_j^{(0)})$. Then $L \in \mathscr{A}_r(0)$ if and only if there exists $j \geq 0$ such that $L \in \mathscr{D}_j^{(0)}$.

2. For each $k > 0$ let $\mathscr{D}_0^{(k)} = \mathscr{L}(\mathscr{A}_r(k - 1), \mathscr{A}_r(k - 1))$.

For $j \geq 0$ let $\mathscr{D}_{j+1}^{(k)} = \mathscr{L}(\mathscr{A}_r(k-1), \mathscr{D}_j^{(k)})$. Then

$L \in \mathscr{A}_r(k)$ if and only if there exists a $j \geq 0$ such that

$L \in \mathscr{D}_j^{(k)}$.

3. Define $\mathscr{G}_r(i, j) = \{G \in \mathscr{G} \mid \deg(\alpha, \Delta_r(G)) \leq i$ and $\deg(\alpha, \rho_i(G)) \leq j\}$.

Then

$L \in \mathscr{D}_j^{(i)}$ if and only if $L = L(G)$ for some $G \in \mathscr{G}_r(i, j)$.

Proof. The analog to lemma 5.3 holds, where Z_{ij} is defined by replacing Δ_ℓ by Δ_r and ρ_i by λ_i. "k" is then defined to be rdeg(G). Part (c) of lemma 5.3 must be altered to read "$G(\beta, U_{ij})$ is linear over $(T \cup U_i', T \cup U_{ij})$ biased right". The proof then follows as given after replacing $V_\ell^{(i)}$ by $V_r^{(i)}$, ρ_i by λ_i, Δ_ℓ by Δ_r and interchanging "$T \cup U_{ij}$" and "$T \cup U_i'$" whenever they appear related by context, e.g., proof of (c) condition (2) and all references to the pair "$(T \cup U_{ij}, T \cup U_i')$". The proof of this theorem then follows that of theorem 5.5 with similar modifications.

6. AFL Properties of the Left and Right Dominant Languages.

Our major results are presented in this section. The first of these is theorem 6.1 which states that $\mathscr{A}_\ell(k)$ is a full AFL for each $k \geq 0$. Theorem 6.3 establishes that the hierarchy, $\mathscr{A}_\ell(0) \subsetneq \mathscr{A}_\ell(1) \ldots$, is nontrivial by showing that each inclusion is proper.

Theorem 6.4 is especially important in that it describes the relationship between the class $\mathscr{A}_\ell(k)$ and its counterpart, $\mathscr{A}_r(k)$. It is shown that $L \in \mathscr{A}_\ell(k)$ if and only if L^R ("R" is the reversal operator) belongs to $\mathscr{A}_r(k)$. Important corollaries to this theorem establish that $\mathscr{A}_r(k)$ is a full AFL for each $k \geq 0$, and that the right dominant languages form a nontrivial hierarchy just as do the left dominant languages.

Theorem 6.5 demonstrates that the two hierarchies are incomparable in a very strong sense, i.e., $\mathscr{A}_\ell(0)$ contains languages that do not belong to $\mathscr{A}_r(k)$ for each $k \geq 0$ and similarly, $\mathscr{A}_r(0) - \mathscr{A}_\ell(k) \neq \Phi$ for each $k \geq 0$.

Theorem 6.6, an immediate consequence of theorem 6.1 and corollary 1 to theorem 6.4, states that $\mathscr{A}_\ell(i) \cap \mathscr{A}_r(j)$ is a full AFL for each i and $j \geq 0$.

Theorem 6.1. $\mathscr{A}_\ell(k)$ is a full AFL for each $k \geq 0$.

Proof. The general approach will be to show that if θ represents an AFL operation and $L = L(G)$ for some $G \in \mathscr{G}_\ell(k)$, then there exists $G' \in \mathscr{G}_\ell(k)$ such that $L(G') = \theta(L)$. This is to say that AFL operations do not increase the "complexity" of the grammar required to describe their effect, "complexity" being measured by the index k as determined by the relation, Δ_ℓ.

(i) $\mathscr{A}_\ell(k)$ is closed under arbitrary homomorphism. Let $L = L(G)$, where $G = (V, T, P, \alpha) \in \mathscr{G}_\ell(k)$. Let $h: T^* \to \Sigma^*$ be an arbitrary homomorphism. Let $h': (T \cup V)^* \to (\Sigma \cup V)^*$ be an extension of h such that $h'(a) = h(a)$ for $a \in T$ and $h'(\beta) = \beta$ for $\beta \in V$. Finally, let $G' = (V, \Sigma, P', \alpha)$ be a grammar constructed from G by replacing $(\beta \to w) \in P$ by $(\beta \to h'(w))$ to form P'. It should be clear that $h(L) = L(G')$ and furthermore that $(\beta_1, \beta_2) \in \Delta_\ell(G)$ if and only if $(\beta_1, \beta_2) \in \Delta_\ell(G')$. Thus $\ell\deg(G') = \ell\deg(G)$ and it follows that $h(L) \in \mathscr{A}_\ell(k)$.

(ii) $\mathscr{A}_\ell(k)$ is closed under $\cup$, $\cdot$, $*$.

Let $G_1 = (\{\alpha\}, \{a_1\}, \{\alpha \to \varepsilon, \alpha \to a_1\alpha\}, \alpha)$.

Let $G_2 = (\{\alpha\}, \{a_1, a_2\}, \{\alpha \to a_1a_2\}, \alpha)$.

Let $G_3 = (\{\alpha\}, \{a_1, a_2\}, \{\alpha \to a_1, \alpha \to a_2\}, \alpha)$.

It is clear that $L(G_1) = a_1^*$, $L(G_2) = a_1a_2$ and $L(G_3) = \{a_1\} \cup \{a_2\}$ are strictly linear over $(\{a_1, a_2\}, \Phi)$. If $L_1, L_2 \in \mathscr{A}_\ell(k)$, then by theorem 5.5, there exists j_1 and j_2 such that $L_1 \in \mathscr{L}_{j_1}^{(k)}$ and $L_2 \in \mathscr{L}_{j_2}^{(k)}$. Let $j = \text{Max}\{j_1, j_2\}$, then $L_1, L_2 \in \mathscr{L}_j^{(k)}$. This follows from the fact that $\mathscr{G}_{i+1}^{(k)} \supseteq \mathscr{G}_i^{(k)}$ for all i, $k \geq 0$. By definition of $\mathscr{G}_{j+1}^{(k)}$ and theorem 5.5, it follows that $\tau(L(G_1)) = L_1^*$, $\tau(L(G_2)) = L_1L_2$ and $\tau(L(G_3)) = L_1 \cup L_2$ all belong to $\mathscr{G}_{j+1}^{(k)}$ and hence to $\mathscr{A}_\ell(k)$.

(iii) If R is an arbitrary regular set and $L \in \mathscr{A}_\ell(k)$, then $R \cap L \in \mathscr{A}_\ell(k)$.

Let $L = L(G)$, where $G = (V, T, P, \alpha) \in \mathscr{G}_\ell(k)$. Let $R \subseteq T^*$ be regular and let $A = (Q, T, \delta, q_0, F)$ be a minimal-state, deterministic, finite state acceptor for R, where Q denotes the set of states, δ denotes the transition function, $q_0 \in Q$ denotes the initial state and $F \subseteq Q$ denotes the set of final states (assume $F \neq \Phi$). Now for each $f \in F$ let $R_f = \{x \in T^* \mid \delta(q_0, x)\} = f$. It clearly follows that $L \cap R = \bigcup_{f \in F}(R_f \cap L)$. Since $\mathscr{A}_\ell(k)$ is closed under union, then the result will follow once it can be shown that $R_f \cap L \in \mathscr{A}_\ell(k)$.

We now describe the construction of a grammar G_f such that for each $f \in F$, $L(G_f) = R_f \cap L$ and $G_f \in \mathscr{G}_\ell(k)$. G_f will be the grammar $(V_f, T, P_f, (q_0, \alpha, f))$ obtained by reducing the grammar $(Q \times V \times Q, T, P_f', (q_0, \alpha, f))$, where P_f' consists of all productions of the form:

(1) $(q_1, \beta, q_2) \to u$, if $(\beta \to u) \in P$, $u \in T^*$ and

$\delta(q_1, u) = q_2$, where $q_1, q_2 \in Q$;

(2) $(q_1, \beta, q_2) \to u_0(s_{11}, \gamma_1, s_{12})u_1 \ldots (s_{n1}, \gamma_n, s_{n2})u_n$, if

$(\beta \to u_0\gamma_1u_1 \ldots \gamma_nu_n) \in P$, where $u_0u_1 \ldots u_n \in T^*$,

$\gamma_i \in V$, $1 \leq i \leq n$, and the states $q_1 = s_{02}, s_{i1}, s_{i2}, 1 \leq i \leq n$,

$q_2 = s_{n+1,1}$ satisfy the conditions that,

$\delta(s_{i-1,2}, u_{i-1}) = s_{i1}$ for $1 \leq i \leq n+1$ and for

$1 \leq i \leq n$ there exists $x_i \in T^*$ such that $\delta(s_{i,1}, x_i) = s_{i,2}$.

If $\not{p} \in P_f$ is a production generated from $p \in P$, then we call p the "parent" of $\not{p}$. In a similar fashion we call β the parent of (q_1,β,q_2) for all $q_1,q_2 \in Q$ such that $(q_1,\beta,q_2) \in V_f$. For convenience we let $\psi: P_f^* \to P^*$ be a homomorphism such that $\psi(\not{p})$ is the parent of $\not{p}$ for each $\not{p} \in P_f$.

By induction on the lengths of derivations the following generalizations of (1) and (2) may be obtained:

(1*) for all $(q_1,\beta,q_2) \in V_f$ it follows that $(q_1,\beta,q_2) \overset{\pi}{\underset{G_f}{\Longrightarrow}} x \in T^*$ if and only if $\beta \overset{\psi(\pi)}{\underset{G}{\Longrightarrow}} x$ and $\delta(q_1,x) = q_2$.

(2*) for all $(q_1,\beta,q_2) \in V_f$ it follows that

$(q_1,\beta,q_2) \overset{\pi}{\underset{G_f}{\Longrightarrow}} u_0(s_{11},\gamma_1,s_{12})u_1 \ldots (s_{n1},\gamma_n,s_{n2})u_n$, for some $n \geq 1$, if and only if $\beta \overset{\psi(\pi)}{\underset{G}{\Longrightarrow}} u_0\gamma_1u_1 \ldots \gamma_nu_n$ and $\delta(q_1,u_0) = s_{11}$, there exists $x_i \in T^*$ such that $\delta(s_{i1},x_i) = s_{i2}$ for $1 \leq i \leq n$ and finally, $\delta(s_{n2},u_n) = q_2$.

From (1*) it follows that $(q_o,\alpha,f) \overset{*}{\underset{G_f}{\Longrightarrow}} x \in T^*$ if and only if $\alpha \overset{*}{\underset{G}{\Longrightarrow}} x$ and $\delta(q_o,x) = f$. Thus we have $L(G_f) = R_f \cap L$. Now suppose $(z,z') \in \Delta_\ell(G_f)$, where $z = (q_1,\beta,q_2)$ and $z' = (q_1',\beta',q_2')$. By definition of Δ_ℓ it follows that either $z \overset{\pi}{\underset{G_f}{\Longrightarrow}} uzvz'w$ or $z \overset{\pi'}{\underset{G_f}{\Longrightarrow}} uz_1v$ and $z_1 \overset{\pi''}{\underset{G_f}{\Longrightarrow}} xz_1yz'w$. By (2*) above we have either

$\beta \overset{\psi(\pi)}{\underset{G}{\Longrightarrow}} u'\beta v'\beta'w'$ or $\beta \overset{\psi(\pi')}{\underset{G}{\Longrightarrow}} u'\beta_1v$ and $\beta_1 \overset{\psi(\pi'')}{\underset{G}{\Longrightarrow}} x'\beta_1y'\beta'w'$. In either case it follows that $(\beta,\beta') \in \Delta_\ell(G)$. Thus if $(z_1,z_2),(z_2,z_3) \ldots (z_i,z_{i+1})$ is a chain in $\Delta_\ell(G_f)$, then $(\beta_1,\beta_2),(\beta_2,\beta_3), \ldots (\beta_i,\beta_{i+1})$ is a chain in $\Delta_\ell(G)$, where β_j is the parent of z_j, $1 \leq j \leq i+1$. It follows from this that $\ell\deg(G_f) = \deg((q_o,\alpha,f),\Delta_\ell(G_f)) \leq \deg(\alpha,\Delta_\ell(G)) = \ell\deg(G)$. Thus $G_f \in G_\ell(k)$ and $L \cap R_f \in \mathscr{L}_\ell(k)$.

(iv) $\mathscr{L}_\ell(k)$ is closed under regular substitution.

The proof will be by industion on k. To show $\mathscr{L}_\ell(0)$ is closed under regular substitution we show that $\mathscr{L}_0^{(0)}$ is closed under regular substitution and then show that $\mathscr{L}_j^{(0)}$ closed under regular substitution implies $\mathscr{L}_{j+1}^{(0)}$ is as well. Then by theorem 5.5 it follows that $\mathscr{A}_\ell(0)$ is closed under regular substitution. Let $L \in \mathscr{L}_0^{(0)}$ and let σ be a regular substitution defined on Σ, where $L \subseteq \Sigma^*$. Since $\mathscr{L}_0^{(0)} = \mathscr{L}(\mathscr{R}, \mathscr{R})$, then $L = \tau(L')$, where L' is strictly linear over (T_ℓ, T_r) and τ is a regular substitution. Let σ' be the substitution on $T = T_\ell \cup T_r$ defined by, $\sigma'(a) = \sigma(\tau(a))$ for all $a \in T$. Since the regular sets are closed under regular substitution, then it follows that σ' is regular. Thus $\sigma'(L') = \sigma(L) \in \mathscr{L}(\mathscr{R}, \mathscr{R}) = \mathscr{L}_0^{(0)}$.

Assume that $\mathscr{L}_j^{(0)}$ is closed under regular substitution and let $L \in \mathscr{L}_{j+1}^{(0)}$. $L = \tau(L')$, where L' is strictly linear over (T_ℓ, T_r) and τ is a substitution such that $\tau(a) \in \mathscr{L}_j^{(0)}$ for all $a \in T_\ell$ and $\tau(b)$ is regular for all $b \in T_r$. Let σ be an arbitrary regular substitution. Define $\sigma'(a) = \sigma(\tau(a))$ for all $a \in T = T_\ell \cup T_r$. Since $\tau(a) \in \mathscr{L}_j^{(0)}$ and $\mathscr{L}_j^{(0)}$ is closed under regular substitution by induction, then $\sigma'(a) \in \mathscr{L}_j^{(0)}$ for all $a \in T_\ell$. Furthermore, $\sigma'(a)$ is regular for all $a \in T_r$. Clearly $\sigma'(L') = \sigma(L) \in \mathscr{L}(\mathscr{L}_j^{(0)}, \mathscr{R}) = \mathscr{L}_{j+1}^{(0)}$. Thus $\mathscr{A}_\ell(0)$ is closed under regular substitution.

Assume that $\mathscr{A}_\ell(j)$ is closed under regular substitution for all $j \leq i$. We show that $\mathscr{A}_\ell(i + 1)$ is also closed under regular substitution. By essentially the same arguments as given for $\mathscr{A}_\ell(0)$ it can be shown that $\mathscr{C}_0^{(i+1)} = \mathscr{L}(\mathscr{A}_\ell(i), \mathscr{A}_\ell(i))$ is closed under regular substitution by virtue of closure for $\mathscr{A}_\ell(i)$ from the induction hypothesis. By a similar inductive argument to that given previously, it follows easily that $\mathscr{C}_j^{(i+1)}$ closed under regular substitution implies $\mathscr{C}_{j+1}^{(i+1)}$ is closed under regular substitution. Thus it follows that $\mathscr{A}_\ell(i + 1)$ is closed under regular substitution.

(v) $\mathscr{A}_\ell(k)$ is closed under inverse homomorphism for each $k \geq 0$. From the definition of $\mathscr{A}_\ell(k)$ it clearly follows that $\mathscr{A}_\ell(0) \subseteq \mathscr{A}_\ell(k)$ for $k \geq 0$. Thus $\mathscr{A}_\ell(k)$ contains all regular sets by the corollary to theorem 4.8. Since $\mathscr{A}_\ell(k)$ is closed under union, intersection with regular sets, regular substitution and arbitrary homomorphisms, then by theorem 2.7 $\mathscr{A}_\ell(k)$ is closed under inverse homomorphism and thus forms a full AFL.

Lemma 6.2. Let $G = (V, T, P, \alpha)$ be a reduced context-free grammar. There exists a reduced grammar $G' = (V', T, P', \alpha)$ such that,

(i) $L(G') = L(G)$,

(ii) α does not appear in the right-part of any production of P',

(iii) $P' \cap (V' \times V') = \Phi$, $P' \cap ((V' - \{\alpha\}) \times \{\varepsilon\}) = \Phi$ and $(\alpha \to \varepsilon) \in P$ if and only if $\varepsilon \in L(G)$,

(iv) $\deg(\beta, \Delta_\ell(G')) \leq \deg(\beta, \Delta_\ell(G))$ for all $\beta \in V' \subseteq V$.

Proof. The grammar G' is obtained by first applying the construction given in theorem 4.11 of Hopcroft and Ullman [11] (pages 62-63) and then applying the construction given in theorem 4.4 of Hopcroft and Ullman [11] (page 50). It can be easily verified by examining these constructions that if G is reduced, then G' will be also. Furthermore, it follows that $V' \subseteq V$ and that $\beta \overset{+}{\underset{G'}{\Longrightarrow}} w \in (V' \cup T)^*$ implies $\beta \overset{+}{\underset{G}{\Longrightarrow}} w$. Thus, if $(\beta, \beta') \in \Delta_\ell(G')$, then certainly $(\beta, \beta') \in \Delta_\ell(G)$. Hence $\deg(\beta, \Delta_\ell(G')) \leq \deg(\beta, \Delta_\ell(G))$ for all $\beta \in V'$.

Theorem 6.3. The language $L_k \in \mathscr{U}_\ell(k) - \mathscr{U}_\ell(k-1)$ for all $k \geq 1$, where L_k is defined as follows:

1. $L_0 = \{a_0^n \, b_0 \, (c_0 d_0 e_0)^n \mid n \geq 1\}$

2. For $k > 0$ define $L_k' = \{a_k^n \, b_k \, (c_k d_k e_k)^n \mid n \geq 1\}$ and let

   τ_k be the substitution defined by

   $\tau_k(a_k) = a_k$, $\tau_k(b_k) = b_k$, $\tau_k(c_k) = c_k$,

   $\tau_k(e_k) = e_k$ and

$\tau_k(d_k) = L_{k-1}$, then

$L_k = \tau_k(L_k')$.

3. $\sigma_i = \sigma_j'$ if and only if $\sigma = \sigma'$ and $i = j$,

where $\sigma, \sigma' \in \{a, b, c, d, e\}$.

Proof. We first establish that $L_k \in \mathscr{A}_\ell(k)$ for each $k \geq 0$. L_0 is clearly linear and hence by the corollary to 4.8, $L_0 \in \mathscr{A}_\ell(0)$. Assume that $L_i \in \mathscr{A}_\ell(i)$ for each i, $0 \leq i \leq k$. We show that $L_{k+1} \in \mathscr{A}_\ell(k+1)$. Clearly L_{k+1}' is strictly linear over $(\{a_{k+1}, b_{k+1}\}, \{c_{k+1}, d_{k+1}, e_{k+1}\})$. Since $L_k \in \mathscr{A}_\ell(k) = \mathscr{A}_\ell((k+1) - 1)$ by hypothesis, then by theorem 5.5 it follows that $\tau_k(L_k') = L_{k+1} \in \mathscr{A}_\ell(k+1)$.

Next it must be established that for all $k \geq 1$, $L_k \neq L(G)$ for any $G \in \mathscr{G}_\ell(k-1)$. This will be done by showing that if G is any reduced grammar generating L_k, then G has at least one nonterminal β, such that $\deg(\beta, \Delta_\ell(G)) \geq k$. To this end let G by any reduced grammar generating L_k, $k > 0$. By lemma 6.2 G has an equivalent grammar $G' = (V', T, P', \alpha)$ which contains no erasing rules $(\varepsilon \notin L_k)$ and no productions of the form $\beta \to \gamma$, where $\gamma \in V'$. Furthermore, $\deg(\beta, \Delta_\ell(G')) \leq \deg(\beta, \Delta_\ell(G))$ for all $\beta \in V'$. Thus if we can establish that $\deg(\beta, \Delta_\ell(G')) \geq k$ for some $\beta \in V'$, then it will certainly hold that $\deg(\beta, \Delta_\ell(G)) \geq k$ for some $\beta \in V$.

To be able to conveniently represent elements of L_k, $k \geq 0$, we shall define $X_k(n) \subseteq L_k$ as follows:

$$X_0(n) = \{y \in L_0 \mid y = a_0^i\, b_0(c_0 d_0 e_0)^i,\ \ i > n\},$$

for k > 1, let

$$X_k(n) = \{y \in L_k \mid y = a_k^i\, b_k\, (c_k z_1 e_k)\ (c_k z_2 e_k)\ \ldots\ (c_k z_i e_k)$$

$$\text{such that } i > n \text{ and } z_j \in X_{k-1}(n),\ \ 1 \leq j \leq i\}.$$

Let M be the least upper bound on the length of the right-parts of productions of G'. Consider all possible left-most derivations, π, having the following properties:

(i) $\alpha \overset{\pi}{\underset{lm}{\Longrightarrow}} w \in T^*$

(ii) If π_1 (possibly null), π_2 and π_3 are any substrings of π such that $\pi = \pi_1\pi_2\pi_3$ and

$$\alpha \overset{\pi_1}{\underset{lm}{\Longrightarrow}} u\beta_1 v \overset{\pi_2}{\underset{lm}{\Longrightarrow}} ux\beta_2 yv \overset{\pi_3}{\underset{lm}{\Longrightarrow}} w, \text{ where } \beta_1, \beta_2 \in V', \text{ then}$$

$\beta_1 \neq \beta_2$.

For such derivations it follows that $||w|| \leq M^\nu$, where $\nu = |V'|$.

Thus if $w \in L(G')$ and $||w|| > M^{\nu}$, then any left-most derivation, π, of w must be of the form, $\pi = \pi_1\pi_2\pi_3$ (π_1 possibly null), where

$$\alpha \underset{lm}{\overset{\pi_1}{\Longrightarrow}} u\beta v \underset{lm}{\overset{\pi_2}{\Longrightarrow}} ux\beta yv \underset{lm}{\overset{\pi_3}{\Longrightarrow}} w.$$

If, in addition, β is the first such nonterminal which derives itself in a left-most derivation, then it also follows that $||ux\beta yv|| \leq M^{\nu}$.

We now consider a left-most derivation $\pi = \pi_1\pi_2\pi_3$ of any string $w \in X_k(n)$, where $n > 2 \cdot M^{\nu}$. Furthermore, assume $\beta_1 \in V'$ is the first nonterminal for which $\alpha \underset{lm}{\overset{\pi_1}{\Longrightarrow}} u_1\beta_1 v_1 \underset{lm}{\overset{\pi_2}{\Longrightarrow}} u_1x_1\beta_1y_1v_1 \underset{lm}{\overset{\pi_3}{\Longrightarrow}} w$. What will be shown is that $\deg(\beta_1, \Delta_{\ell}(G')) \geq k$. To demonstrate this we show that the string $y_1 \in (V' \cup T)^* - T^*$ and must contain $\beta_1' \in V'$ such that $\deg(\beta_1', \Delta_{\ell}(G')) \geq k-1$. We proceed by showing first that $y_1 \notin T^*$.

Case 1. $y_1 \neq \varepsilon$. If $y_1 = \varepsilon$, then since G' contains no erasing rules and no rules of the form $\beta \rightarrow \gamma$, where $\gamma \in V'$, then $x_1 \in T$. From $||u_1x_1\,{}_1y_1v_1|| \leq M^{\nu}$ and the assumption that $w \in X_k(n)$ it follows that $x_1 = a_k^r$ for some $r \geq 1$. By iterating the derivation π_2 it would be possible to produce an unbalance between the number of a_k's appearing in a terminal string and the number of c_k's produced by π_3. This is in contradiction to the form of strings in L_k. Thus $y_1 \neq \varepsilon$.

Case 2. y_1 cannot contain b_k. By iterating π_2 it would be possible to introduce more than one b_k into a terminal string if y_1 contained a b_k, it follows that b_k cannot occur in y_1.

Case 2 implies $y_1 = a_k^i$ for some $i > 0$ or y_1 is a subword of $c_k z_1 e_k c_k z_2 e_k \ldots c_k z_r e_k$, where $r > n > M^\nu$ and $z_i \in X_{k-1}(n)$, $1 \leq i \leq r$. The former case is not possible by an argument similar to that given for case 1. Therefore consider the second possibility. Only four subcases need by considered based on the form of strings of L_k and the constraint that $||y_1|| < n$. Before discussing the possible subcases we note the following properties of strings in L_k :

(i) $||w||_{a_j} = ||w||_{c_j} = ||w||_{e_j}$ for $0 \leq j \leq k$.

(ii) $||w||_{a_j} = ||w||_{b_{j-1}}$ for $1 \leq j \leq k$.

In addition, since $||u_1 x_1\ _1 y_1 v_1|| < n$, then it follows that $x_1 = a_k^i$ for some i, $0 \leq i < n$.

Subcase 1. y_1 is not a subword of $c_k a_{k-1}^i$. This follows because iteration of π_2 would result in violation of $||w||_{c_k} = ||w||_{e_k}$ or $||w||_{a_{k-1}} = ||w||_{c_{k-1}}$.

Subcase 2. y_1 is not a subword of $a_j^r b_j c_j a_{j-1}^s$ for $1 \le j < k$. If $k = 1$ this case does not apply. For $k > 1$, iteration of π_2 would produce one of the following invalid contexts in a terminal string: $a_{j-1}a_j$, $a_{j-1}b_j$, $a_{j-1}c_j$, $c_j a_j$, $b_j a_j$, $c_j b_j$, $b_j b_j$ or $c_j c_j$. If $y_1 = a_j^i$, $i > 0$, for any j, then relations (i) and (ii) would be violated by iterating π_2.

Subcase 3. y_1 is not a subword of $a_0^i b_0 (c_0 d_0 e_0)^j$. y_1 cannot contain b_0, else iteration of π_2 would destroy relation (ii). In all other cases, relation (i) would be violated by iterating π_2.

Subcase 4. y_1 is not a subword of $(c_0 d_0 e_0)^i e_1 e_2 \ldots e_j c_j a_{j-1}^r$, for $1 \le j \le k$. In this case, iteration of π_2 would produce the following invalid contexts in terminal strings: $a_{j-1}c_j$, $a_{j-1}e_q$ $(0 \le q \le j)$, $a_{j-1}d_0$, $a_{j-1}c_0$ or $c_j e_j$. If y_1 does not contain a_{j-1}, then all other cases would result in violation of relation (i) by iterating π_2.

This completes the demonstration that $y_1 \notin T^*$. Thus $y_1 = u_1' \beta_2' v_1'$, where $u_1' \in T^*$, $v_1' \in (V' \cup T)^*$ and $\beta_2' \in V'$. By definition of Δ_ℓ it follows that $(\beta_1, \beta_2') \in \Delta_\ell(G')$.

By appealing to relations (i) and (ii) and the context properties of terminals appearing in strings of L_k, it can be shown that $x_1 = a_k^r$ and $y_1 \overset{*}{\underset{G'}{\Longrightarrow}} c_k z_{11} e_k c_k z_{12} e_k \cdots c_k z_{1r} e_k$ for some $r \geq 1$. Here $z_{1j} \in X_{k-1}(n)$, $1 \leq j \leq r$.

Employing the above argument repeatedly we may establish the following relations for each j, $1 \leq j \leq k$. It should be noted that $\pi_2^{(1)}$ is to be identified with π_2 defined earlier.

1. $\beta_j \overset{\pi_2^{(j)}}{\underset{lm}{\Longrightarrow}} x_j \beta_j\ y_j$, where $||x_j \beta_j\ y_j|| \leq M^{\nu}$ and

   $x_j = a_{k-j+1}^{r_j}$, $y_j \notin T^*$,

2. $y_j = u_j' \beta_{j+1}' v_j' \overset{*}{\underset{G'}{\Longrightarrow}} (c_{k-j+1} z_{j1} e_{k-j+1}) \cdots (c_{k-j+1} z_{jr_j} e_{k+j-1})$,

   where $u'_j \in T^*$, $v_j' \in (V' \cup T)^*$ and $z_{ji} \in X_{k-j}(n)$,

3. $u_j' \beta_{j+1}' v_j' \overset{*}{\underset{lm}{\Longrightarrow}} j_j' u_{j+1} \beta_{j+1} v_{j+1} v_j'$ such that $u_j' u_{j+1} \in T^*$ and

   $||u_j' u_{j+1} \beta_{j+1} v_{j+1} v_j'|| \leq M^{\nu}$.

From 1. and 3. it follows that

$||u_j' u_{j+1} x_{j+1} \beta_{j+1} y_{j+1} v_{j+1} v_j'|| \leq 2 \cdot M^{\nu}$ and thus from 2. it follows that $u_j' u_{j+1} x_{j+1}$ is of the form $c_{k-j+1} a_{k-j}^{i}$. This condition allows the argument to be applied repeatedly for each j. Relations 1., 2. and

3. imply that $(\beta_j, \beta_{j+1}) \in \Delta_\ell(G')$ for each j, $1 \leq j \leq k$. Thus $\deg(\beta_1, \Delta_\ell(G')) \geq k$ and we conclude G' and therefore G cannot belong to $\mathcal{G}_\ell(k-1)$.

Theorem 6.4. $L \in \mathcal{A}_\ell(k)$ if and only if Reverse $(L) \in \mathcal{A}_r(k)$ for all $k \geq 0$.

Proof. If $L \in \mathcal{A}_\ell(k)$, then there exists $G = (V, T, P, \alpha) \in \mathcal{G}_\ell(k)$ such that $L = L(G)$. Let $G' = (V, T, P', \alpha)$, where $P' = \{\beta \to$ Reverse (w) $(\beta \to w) \in P\}$. It is easily shown that $\beta \overset{\pi}{\underset{G}{\Longrightarrow}} x \in (V \cup T)^*$ if and only if $\beta \overset{\pi}{\underset{G'}{\Longrightarrow}}$ Reverse (x). From this it follows that $L(G') =$ Reverse (L) and furthermore that for all $\beta \in V$,

(i) $\deg(\beta, \Delta_\ell(G)) = \deg(\beta, \Delta_r(G'))$,

(ii) $\deg(\beta, \rho_i(G)) = \deg(\beta, \lambda_i(G'))$ for all $i \geq 0$.

The converse follows in a similar fashion.

Corollary. $\mathcal{A}_r(k)$ is a full AFL for each $k \geq 0$.

Proof. The result follows from theorem 6.4, the following relations and the fact that the regular sets are closed under reversal.

1. $h(L) =$ Reverse $(h^R($Reverse $(L)))$, h an arbitrary homomorphism $(h^R(a) =$ Reverse $(h(a)))$.

2. $L^* =$ Reverse ([Reverse $(L)]^*$.

3. $L_1 \cup L_2 =$ Reverse (Reverse $(L_1) \cup$ Reverse $(L_2))$.

4. $L \cap R =$ Reverse (Reverse $(L) \cap$ Reverse $(R))$, where R is a regular set.

5. $\tau(L) =$ Reverse $(\tau^R($Reverse $(L)))$, where τ is a regular substitution $(\tau^R(a) =$ Reverse $(\tau(a)))$.

6. Closure under h^{-1} follows from 3., 4. and 5. and theorem 2.7.

Corollary 3.27. Reverse $(L_k) \in \mathscr{A}_r(k) - \mathscr{A}_r(k-1)$ for all $k > 0$.

Theorem 6.5.

(i) $L_{k+1} \in \mathscr{A}_r(0) - \mathscr{A}_\ell(k)$ for all $k \geq 0$.

(ii) Reverse $(L_{k+1}) \in \mathscr{A}_\ell(0) - \mathscr{A}_r(k)$ for all $k \geq 0$.

L_k is defined as in theorem 6.3.

Proof. It can easily be verified that the grammar,

$G_{k+1} = (V_{k+1}, \Sigma_{k+1}, P_{k+1}, \alpha_k) \in \mathscr{G}_r(0) \cap \mathscr{G}_\ell(k+1)$ and $L(G_{k+1}) = L_{k+1}$

for all $k \geq 0$. We define G_k inductively as follows:

$G_0 = (V_0, \Sigma_0, P_0, \alpha_0)$,

$V_0 = \{\alpha_0\}$

$\Sigma_0 = \{a_0, b_0, c_0, d_0, e_0\}$

$P_0 = \{\alpha_0 \to a_0\alpha_0c_0d_0e_0,\ \ \alpha_0 \to a_0b_0c_0d_0e_0\}$.

For $k > 0$, define

$V_k = V_{k-1} \cup \{\alpha_k\}$

$\Sigma_k = \Sigma_{k-1} \cup \{a_k, b_k, c_k, e_k\}$

$P_k = P_{k-1} \cup \{\alpha_k \to a_k\alpha_k\ c_k\alpha_{k-1}e_k,\ \alpha_k \to a_kb_kc_k\alpha_{k-1}e_k\}$.

Part (ii) is proved by defining G_k' to be obtained from G_k be reversing the right-parts of all productions. It then follows that $G_{k+1}' \in \mathcal{G}_\ell(0) \cap \mathcal{G}_\ell(k+1)$ and $L(G_{k+1}') = $ <u>Reverse</u> $(L(G_{k+1}))$ for each $k \geq 0$.

It is worthy of note that $G_{k+1} \in \mathcal{G}_r(0, k+1)$ and that $G_{k+1}' \in \mathcal{G}_\ell(0, k+1)$, where $\mathcal{G}_\ell(i, j)$ and $\mathcal{G}_r(i, j)$ are defined in theorems 5.5 and 5.6, respectively.

Theorem 6.6. For each $i \geq 0$ and each $j \geq 0$, $\mathcal{A}_\ell(i) \cap \mathcal{A}_r(j)$ is a full AFL properly included in $\mathcal{A}_\ell(i+1) \cap \mathcal{A}_r(j)$ and $\mathcal{A}_\ell(i) \cap \mathcal{A}_r(j+1)$.

Proof. That $\mathcal{A}_\ell(i) \cap \mathcal{A}_r(j)$ is a full AFL follows easily from the fact that $\mathcal{A}_\ell(i)$ and $\mathcal{A}_r(j)$ are full AFL for each i and $j \geq 0$. Since $L_{i+1} \in \mathcal{A}_r(0) \subseteq \mathcal{A}_r(j)$ and since $L_{i+1} \in \mathcal{A}_\ell(i+1) - \mathcal{A}_\ell(i)$ for each $i \geq 0$, then $L_{i+1} \in \mathcal{A}_\ell(i+1) \cap \mathcal{A}_r(j) - \mathcal{A}_\ell(i) \cap \mathcal{A}_\ell(j)$. In a similar fashion <u>Reverse</u> $(L_{j+1}) \in \mathcal{A}_\ell(i) \cap \mathcal{A}_r(j+1) - \mathcal{A}_\ell(i) \cap \mathcal{A}_\ell(j)$.

A particularly interesting class of languages is the class $\mathscr{A}_\ell(0)\cap\mathscr{A}_r(0)$. L is a member of this class if and only if there exists grammars G and G' such that $L = L(G) = L(G')$ and $G\in\mathscr{G}_\ell(0)$ and $G'\in\mathscr{G}_r(0)$; that is, L is generated by some left dominant grammar of degree 0 and also by some right dominant grammar of degree 0. There is a striking analogy that can be drawn between the regular sets which are generated by some left as well as right <u>linear</u> grammar and the sets in $\mathscr{A}_\ell(0)\cap\mathscr{A}_r(0)$ which are generated by some left as well as right dominant grammar of degree -0. Because of this analogy we choose to call $\mathscr{A}_\ell(0)\cap\mathscr{A}_r(0)$ the class of "regularly dominant" languages. The analogy can be extended to the entire class of derivation bounded languages in that these languages are precisely those which are generated by some left as well as right dominant grammar of finite degree.

A final comment. The class of regularly dominant languages form a full AFL and contain the nonterminal bounded languages by the corollary to theorem 4.8. We conjecture that this is the smallest such full AFL. Another interesting problem would be to characterize the subclass of $\mathscr{G}_\ell(0)\cup\mathscr{G}_r(0)$ which generates those and only those languages of $\mathscr{A}_\ell(0)\cap\mathscr{A}_r(0)$.

1. Altman, E. and Banerji, R. (1965), "Some Problems of Finite Representability," Information and Control 8, 251-263.

2. Banerji, R.B. (1963), "Phrase Structure Languages, Finite Machines, and Channel Capacity," Information and Control 6, 153-162.

3. Fleck, A.C. (1969), "On the Complexity of Grammars," Tech. Report No. 4, The University of Iowa.

4. Fleck, A.C. (1971), "On the Combinatorial Complexity of Context-free Grammars," Proc. IFIP Congress 71, Booklet TA-2, 15-17.

5. Ginsburg, S. (1966), The Mathematical Theory of Context-free Languages, McGraw-Hill, New York, N.Y.

6. - Ginsburg, S. and Greibach, S. (1970), "Principal AFL," Journal of Computer and System Sciences 4, 308-338.

- Ginsburg, S. and Greibach, S. (1969), "Abstract Families of Languages," Memoirs of AMS, #87, 1-32.

7. Ginsburg, S. and Spanier, E.H. (1966), "Derivation Bounded Languages," J. Comput. Sys. Sci. 2, 228-250.

8. Ginsburg, S. and Spanier, E.H. (1966), "Finite-turn Push Down Automata," SIAM J. Control 4, 3, 429-453.

9. Greibach, S. and Hopcroft, J.E. (1969), "Independence of AFL Operations," Memoirs of AMS, #87, 33-40.

10. Gruska, J. (1970), "A Characterization of Context-free Languages," J. Comput. Sys. Sci. 5, #4, 353-364.

11. Hopcroft, J.E. and Ullman, J.D. (1969), Formal Languages and Their Relation to Automata, Addison-Wesley Publ. Co., Reading, Mass.

12. Moriya, E. (1973), "Associate Languages and Derivational Complexity of Formal Grammars and Languages," Information and Control 22, 139-162.

13. Workman, D. (1973), "The Structural Complexity of Context-free Grammars and Languages," Tech. Report No. 79-09, The University of Iowa, Iowa City, Iowa.