1974

A Measure of Structural Complexity for Context-free Grammars

D. A. Workman

Report Number:

74-129

A MEASURE OF STRUCTURAL COMPLEXITY FOR CONTEXT-FREE GRAMMARS

D. A. Workman
Computer Sciences Department
Purdue University
West Lafayette, Indiana 47907

CSD-TR 129

Title: "A Measure of Structural Complexity for Context-free Grammars"

Submitted by: D. A. Workman
Department of Computer Science
Purdue University
West Lafayette, Indiana 47907

Abstract: Fleck [2] suggested that the structural properties of grammars could be measured by the relative complexity of their derivation languages. Moriya [6] ranked context-free grammars within an infinite hierarchy by appealing to this measure. We present a similar measure based on left(right)-most derivations. The structural complexity of a context-free grammar is given by its "left(right)-degree" which is strongly related to the nature of its left(right)-most derivation language. Our measure appears to be "natural" in the sense that each structural complexity class defines a full AFL. The measure is applied to obtain the relative complexity of a grammar and an equivalent grammar in Greibach form.

The function μ_π applied to w specifies the maximum number of nonterminals that could appear in a sentential form at some step of π when treated as a derivation from w.

Definition 3. Let $G = (V,\Sigma,P,\alpha)$. L(G) denotes the language generated by G, while $\mathcal{D}(G)$, $\mathcal{D}_\ell(G)$ and $\mathcal{D}_r(G)$ are respectively, the languages of all derivations, left-most derivations and right-most derivations of G.

$$\mathcal{D}(G) = \{\pi \mid \alpha \xRightarrow[G]{\pi} x \in L(G)\},$$

$$\mathcal{D}_\ell(G) = \{\pi \mid \alpha \xRightarrow[G,\ell m]{\pi} x \in L(G)\},$$

$$\mathcal{D}_r(G) = \{\pi \mid \alpha \xRightarrow[G,rm]{\pi} x \in L(G)\}.$$

Definition 4. Let $G = (V,\Sigma,P,\alpha)$ be context-free. For each integer $k \geq 0$ the languages of k-bounded derivations, left-most derivations and right-most derivations are denoted, respectively, by $\mathcal{D}^{(k)}(G)$, $\mathcal{D}_\ell^{(k)}(G)$ and $\mathcal{D}_r^{(k)}(G)$ and are defined as follows:

$$\mathcal{D}^{(k)}(G) = \{\pi \in \mathcal{D}(G) \mid \mu_\pi(\alpha) \leq k\},$$

$$\mathcal{D}_\ell^{(k)}(G) = \{\pi \in \mathcal{D}_\ell(G) \mid \ell\mu_\pi(\alpha) \leq k\}, \text{ and}$$

$$\mathcal{D}_r^{(k)}(G) = \{\pi \in \mathcal{D}_r(G) \mid r\mu_\pi(\alpha) \leq k .$$

Definition 5. Let $G = (V,\Sigma,P,\alpha)$. G and L(G) are said to be:

(a) nonterminal bounded (ntb) if $\mathcal{D}(G) = \mathcal{D}^{(k)}(G)$ for some $k \geq 0$; the notions of left-most and right-most nonterminal bounded (lntb,rntb) grammars are analogously defined.

(b) derivation bounded (db) if there exists $k \geq 0$ such that for every $x \in L(G)$ there exists $\pi_x \in \mathcal{D}^{(k)}(G)$ such that $\alpha \xRightarrow[G]{\pi_x} x$.

Theorem 1. G is ntb(ℓntb,rntb) if and only if $\mathcal{D}(G)(\mathcal{D}_\ell(G), \mathcal{D}_r(G))$ is regular.

Proof. The result for ntb grammars can be found in Banerji [1] and Fleck [2]. The result for lntb and rntb grammars is given in Moriya [6].

Definition 6. Let $G = (V,\Sigma,P,\alpha)$ be context-free. The relations $\Delta_\ell(G)$ and $\Delta_r(G)$ are defined on V as follows:

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

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

(2) $\beta_1 \overset{+}{\underset{G}{\Rightarrow}} u\beta_1' v$ and $\beta_1' \overset{+}{\underset{G}{\Rightarrow}} x\beta_1' y\beta_2 z$ for some $\beta_1' \in V$ and $uvxyz \in (V \cup \Sigma)^*$.

$\Delta_r(G)$ is defined by replacing $u\beta_1 v\beta_2 w$ and $x\beta_1' y\beta_2 z$ by their reverses.

Definition 7. Let $\mathcal{G}_\ell$, $\mathcal{G}_r$ and $\mathcal{G}$ denote respectively, the class of all reduced context-free grammars for which $\Delta_\ell(G)$ is irreflexive, $\Delta_r(G)$ is irreflexive and G is derivation bounded.

Theorem 2. $\mathcal{G}_\ell = \mathcal{G}_r = \mathcal{G}$.

Proof. A context-free grammar is said to be nonexpansive if for every nonterminal, β, $\beta \overset{+}{\underset{G}{\Rightarrow}} w$ implies $||w||_\beta \leq 1$. Ginsburg and Spanier [4] have shown that G is db if and only if G is nonexpansive. It follows easily from definition 6 that G is nonexpansive if and only if $\Delta_\ell(G)$ and $\Delta_r(G)$ are irreflexive.

Definition 8. Let S be a set and R a relation on S. For each $s \in S$ define,

$C(s,R) = \{k \mid$ there exists a sequence $s = s_0, s_1, \ldots, s_k$, $k \geq 1$, elements of S (not necessarily distinct) such that (s_{i-1}, s_i) R for $1 \leq i \leq k\}$,

$\deg(s,R) = \infty$, if $C(s,R)$ is infinite;

$= \text{Max } C(s,R)$, if $C(s,R)$ is finite and nonempty;

$= 0$ if $C(s,R)$ is empty.

Lemma 1. Let $G = (V,\Sigma,P,\alpha)$ be a reduced context-free grammar, then

(1) $\Delta_\ell(G)$ and $\Delta_r(G)$ are transitive.

(2) $G \in \mathcal{G}_\ell(\mathcal{G}_r)$ if and only if $\deg(\alpha,\Delta_\ell(G))$ $(\deg(\alpha,\Delta_r(G)))$ is finite.

(3) $\deg(\alpha,\Delta_\ell(G)) \geq \deg(\beta,\Delta_\ell(G))$ for all $\beta \in V$ (similarly for $\Delta_r(G)$).

(4) If $G \in \mathcal{G}_\ell(\mathcal{G}_r)$, then $(\beta,\beta') \in \Delta_\ell(G)(\Delta_r(G))$ implies $\deg(\beta,\Delta_\ell(G)) > \deg(\beta',\Delta_\ell(G))$ (similarly for $\Delta_r(G)$).

Proof. (1), (2) and (3) are immediate from definitions 6 and 8. (4) follows directly from the irreflexive and transitive properties of $\Delta_\ell(G)$ and $\Delta_r(G)$.

Definition 9. Let $G = (V,\Sigma,P,\alpha)$ be a reduced context-free grammar. The left-degree and right-degree of G, denoted $\ell deg(G)$ and $rdeg(G)$, respectively, are defined by,

$$\ell deg(G) = deg(\alpha,\Delta_\ell(G))$$
$$rdeg(G) = deg(\alpha,\Delta_r(G)).$$

For $k < \infty$ we define,

$$\mathscr{G}_\ell(k) = \{G\in\mathscr{G}_\ell \mid \ell deg(G) \le k\},$$
$$\mathscr{G}_r(k) = \{G\in\mathscr{G}_r \mid rdeg(G) \le k\},$$
$$\mathscr{A}_\ell(k) = \{L(G) \mid G\in\mathscr{G}_\ell(k)\},$$
$$\mathscr{A}_r(k) = \{L(G) \mid G\in\mathscr{G}_r(k)\}.$$

The classes $\mathscr{G}_\ell(k)(\mathscr{G}_r(k))$ and $\mathscr{A}_\ell(k)(\mathscr{A}_r(k))$ are called respectively, the left (right) dominant grammars and languages of degree - k.

Theorem 3. Let $\mathscr{L}$ denote the class of all linear languages.

(1) $\mathscr{L} \subsetneq \mathscr{A}_\ell(k) \subsetneq \mathscr{A}_\ell(k+1)$ for each $k \ge 0$.

(2) $\mathscr{A}(k)$ is a full AFL (Abstract Family of Languages) closed under regular substitution for each $k \ge 0$.

(3) Let $L_0 = \{a_0^n b_0(c_0 d_0 e_0)^n \mid n \ge 1\}$ and for each $k > 0$ let $L_k' = \{a_k^n b_k(c_k d_k e_k)^n \mid n \ge 1\}$ and $L_k = \tau(L_k')$, where τ is the substitution defined by $\tau(\sigma) = \{\sigma\}$ for all $\sigma\in\{a_k,b_k,c_k,e_k\}$ and $\tau(d_k) = L_{k-1}$. Then $L_{k+1}\in\mathscr{A}_\ell(k+1) - \mathscr{A}_\ell(k)$ for each $k \geq 0$.

(4) $L\in\mathscr{A}_\ell(k)$ if and only if $\underline{Rev}(L)\in\mathscr{A}_r(k)$ for each $k \ge 0$. ($\underline{Rev}$ is the reversal operator).

Proof. These results have been established in Workman [8].

It should be noted that part (4) of theorem 3 implies that (1) and (2) hold for $\mathscr{A}_r(k)$, $k \ge 0$, as well. Part (3) holds for $\mathscr{A}_r(k)$ if L_k is replaced by $Rev(L_k)$.

The grammar classes, $\mathscr{G}_\ell(k)$ and $\mathscr{G}_r(k)$, may be further refined by defining relations ρ_k and λ_k. ρ_k decomposes $\mathscr{G}_\ell(k)$ into an infinite hierarchy of grammars, while $\mathscr{G}_r(k)$ is analogously decomposed by λ_k. These relations are important in the characterization of $\mathscr{A}_\ell(k)$ (and $\mathscr{A}_r(k)$) as presented in theorem 4.

Definition 10. Let $G = (V,\Sigma,P,\alpha) \in \mathscr{G}$. For each $k \geq 0$ let $\rho_k(\lambda_k)$ be the relation defined on V by $(\beta_1,\beta_2) \in \rho_k$ if and only if there exists $\beta_2' \in V$ such that $\deg(\beta_2,\Delta_\ell(G)) = \deg(\beta_2',\Delta_\ell(G)) = k$ and $\beta_1 \overset{+}{\underset{G}{\Rightarrow}} u\beta_2 v\beta_2' w$ for some $uvw \in (V \cup \Sigma)^*$. Similarly, λ_k is defined by replacing Δ_ℓ by Δ_r and $u\beta_2 v\beta_2' w$ by its reverse.

It follows easily from the definitions of $\Delta_\ell(G)$ and $\Delta_r(G)$ that if Δ_ℓ and Δ_r are irreflexive, then ρ_k and λ_k are also irreflexive; furthermore ρ_k and λ_k are transitive for each k. By applying definition 8 with $R = \rho_k$ or λ_k, lemma 1 can be demonstrated for ρ_k and λ_k. From these facts we are motivated to make the following definition.

Definition 11. $\mathscr{G}_\ell(i,j) = \{G \in \mathscr{G}_\ell(i) \mid \deg(\alpha,\rho_i(G)) \leq j\}$, $\mathscr{G}_r(i,j) = \{G \in \mathscr{G}_r(i) \mid \deg(\alpha,\lambda_i(G)) \leq j\}$, where α is the start symbol of G. The corresponding classes of languages are denoted $\mathscr{L}_\ell(i,j)$ and $\mathscr{L}_r(i,j)$, respectively.

The main results of this paper rely heavily on the properties outlined in theorem 3 together with an important characterization of the left(right) dominant languages of degree - k. Our characterization is presented in theorem 4 and is based on the class of left(right) strictly linear languages introduced in the next definition.

Definition 12. A context-free grammar $G = (V,\Sigma,P,\alpha)$ is said to be <u>left-strictly linear over</u> (Σ_ℓ,Σ_r) if

(1) G is linear,

(2) $P \subseteq V \times (\Sigma_\ell^* V \Sigma_r^* \cup \Sigma_\ell^*)$, where $\Sigma = \Sigma_\ell \cup \Sigma_r$ and $\Sigma_\ell \cap \Sigma_r = \Phi$.

Similarly, G is said to be <u>right-strictly linear</u> if instead of (2), (2') holds.

(2') $P \subseteq V \times (\Sigma_\ell^* V \Sigma_r^* \cup \Sigma_r^*)$.

A <u>language</u> is left(right) strictly linear over (Σ_ℓ,Σ_r) is generated by a so-named grammar.

Definition 13. Let $G = (V,\Sigma,P,\alpha)$ be a linear grammar. The <u>left-strict image</u> of G is the grammar $\overline{G} = (V,\Sigma',P',\alpha)$, left-strictly linear over $(\Sigma,\overline{\Sigma})$, where $(\beta \to u\beta'\overline{v}) \in P'$, $\beta' \in V$, if and only if $(\beta \to u\beta' v) \in P$ and $\overline{v}$ is the string v with each symbol $\sigma \in \Sigma$ replaced by its counterpart, $\overline{\sigma} \in \overline{\Sigma}$. The <u>right-strict image of G</u> is analogously defined by replacing u by $\overline{u}$ (instead of v by $\overline{v}$).

It is obvious that $L(G) = h(L(\overline{G}))$, where h is the homomorphism defined by $h(\sigma) = h(\overline{\sigma}) = \sigma$.

Definition 14. Let $G = (V,\Sigma,P,\alpha)$ be context-free. The <u>subgrammar of G relative to $\beta \in V$</u> is the grammar $G(\beta)$ obtained by reducing (V,Σ,P,α). For $U \subseteq V - \{\beta\}$ the subgrammar of G relative to β <u>restricted on U</u> is the grammar $G(\beta,U)$ obtained by reducing $(V-U,\Sigma \cup U,P,\beta)$.

In a subgrammar restricted on U, the nonterminals of U are treated as terminals in that they are not rewritten when introduced into a sentential form derivable in the restricted subgrammar.

Lemma 2. Let $G = (V,\Sigma,P,\alpha) \in \mathscr{G}_\ell$. Let $\ell\deg(G) = k$ and $\deg(\alpha,\rho_k(G)) = n$. Furthermore, define $V_{ij} = \{\beta \in V \mid \deg(\beta,\Delta_\ell(G)) = i \text{ and } \deg(\beta,\rho_i(G)) = j\}$. Then for each $\beta \in V_{ij}$, the restricted subgrammar $G(\beta,U_{ij})$ is linear and its left-strict image, $\bar{G}(\beta,U_{ij})$ is strictly linear over (Γ_ℓ,Γ_r), where $\Gamma_\ell = U_{ij} \cup \Sigma$ and $\Gamma_r = \overline{(U_i' \cup \Sigma)}$. The sets U_{ij} and U_i' are defined by,

$$U_i' = \phi, \text{ if } i = 0$$

$$= \bigcup_{q<i} (\bigcup_j V_{qj}), \text{ if } i > 0.$$

$$U_{ij} = U_i', \text{ if } j = 0$$

$$= U_i' \cup (\bigcup_{q<j} V_{iq}) \text{ if } j > 0.$$

Finally,

$L(G) = \tau(L(\bar{G}(\alpha,U_{kn})))$, where τ is the substitution defined by $\tau(\sigma) = \tau(\bar{\sigma}) = \sigma$ for all $\sigma \in \Sigma_\ell$ and $\bar{\sigma} \in \Sigma_r$. $\tau(\gamma) = \tau(\bar{\gamma}) = L(G(\gamma))$ for all $\gamma \in \Gamma_\ell \cap V$ and $\bar{\gamma} \in \Gamma_r \cap \bar{V}$.

Proof. The proof is given in Workman [8]; it is based on properties of the relations $\Delta_\ell(G)$ and $\rho_i(G)$, $i \geq 0$. It can be readily established that if $\beta \in V_{ij}$, then $\beta \overset{+}{\underset{G}{\Longrightarrow}} w$ implies $||w||_{V_{ij}} \leq 1$. Furthermore, for all $\beta \in V_{ij}$ and for all $\gamma \in U_{ij}$, $\gamma \overset{+}{\underset{G}{\Longrightarrow}} w$ implies $||w||_\beta = 0$.

Definition 15. Let $\mathscr{L}_\ell$ denote the class of all left-strictly linear languages and let $\mathscr{A}$ and $\mathscr{B}$ be classes of languages. Define,

$\mathscr{L}_\ell(\mathscr{A},\mathscr{B}) = \{L \mid L' \in \mathscr{L}_\ell$, left-strictly linear over (Σ_ℓ,Σ_r) and a substitution, τ, such that $L = \tau(L')$, where $\tau(a) \in \mathscr{A}$ for all $a \in \Sigma_\ell$ and $\tau(b) \in \mathscr{B}$ for all $b \in \Sigma_r\}$.

The remaining results are stated with respect to $\mathscr{U}_\ell(k)$, $\mathscr{G}_\ell(k)$ and $\mathscr{D}_\ell(k)$. Dual results hold with respect to $\mathscr{U}_r(k)$, $\mathscr{G}_r(k)$ and $\mathscr{D}_r(k)$. In the statement of the next theorem, the dual is obtained by replacing $\mathscr{L}_\ell$ by $\mathscr{L}_r$ and $\mathscr{L}_\ell(\mathscr{A},\mathscr{B})$ by $\mathscr{L}_r(\mathscr{B}',\mathscr{A}')$, where $\mathscr{B}'$ and $\mathscr{A}'$ represent the appropriate replacements for $\mathscr{B}$ and $\mathscr{A}$, respectively.

Theorem 4. Let $\mathscr{R}$ be the class of all regular sets.

(1) $L \in \mathscr{A}_\ell(k)$ if and only if there exists $j \geq 0$ such that $L \in \mathscr{A}_\ell(k,j)$.

(2) $\mathscr{A}_\ell(0,0) = \mathscr{L}_\ell(\mathscr{R},\mathscr{R})$.

$\mathscr{A}_\ell(0,j) = \mathscr{L}_\ell(\mathscr{A}_\ell(0,j),\mathscr{R})$, for all $j \geq 0$.

(3) For $k > 0$, $\mathscr{A}_\ell(k,0) = \mathscr{L}_\ell(\mathscr{A}_\ell(k-1), \mathscr{A}_\ell(k-1))$ and

For $j \geq 0$, $\mathscr{A}_\ell(k,j+1) = \mathscr{L}_\ell(\mathscr{A}_\ell(k,j), \mathscr{A}_\ell(k-1))$.

Theorem 5. $G \in \mathscr{G}_\ell(0)$ if and only if $\mathscr{D}_\ell(G)$ is regular.

Proof. Walljasper [7] has established that $G \in \mathscr{G}_\ell(0)$ if and only if G is lntb. It follows by theorem 1 that $G \in \mathscr{G}_\ell(0)$ if and only if $\mathscr{D}_\ell(G)$ is regular.

One of the main results of this paper is theorem 9 which generalizes the preceding theorem by establishing necessary and sufficient conditions on $\mathscr{D}_\ell(G)$ to guarantee that $G \in \mathscr{G}_\ell(k)$, for arbitrary k. Theorems 6, 7 and 8 provide a proof of theorem 9.

Theorem 6. $G \in \mathscr{G}_\ell(k)$ implies $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(k-1)$ for all $k > 0$.

Proof. We proceed by induction on k. Initially let $G = (V,\Sigma,P,\alpha) \in \mathscr{G}_\ell$ such that $\ell\deg(G) = k > 0$. Let $\overline{P}$ denote the set of production "labels" for P. Define $G_0 = (V,\overline{P},P_0,\alpha)$ to be the grammar obtained from G by replacing $(p{:}\beta \to w) \in P$ by $(p{:}\beta \to pw')$, where w' is obtained from w by deleting all elements of Σ. It is easily established that $L(G_0) = \mathscr{D}_\ell(G_0) = \mathscr{D}_\ell(G)$; furthermore, $\deg(\beta,\Delta_\ell(G)) = \deg(\beta,\Delta_\ell(G_0))$ and $\deg(\beta,\rho_i(G)) = \deg(\beta,\rho_i(G_0))$ for all $i \geq 0$ and $\beta \in V$.

<u>Case k = 1</u>. We establish that $G \in \mathscr{G}_\ell(1,j)$ implies $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(0,j)$ by induction on j. Then by theorem 4 part (1) the result follows for k = 1.

Let j = 0. By lemma 2 and our initial remarks it follows that $\mathscr{D}_\ell(G) = L(G_0) = \tau(L(\overline{G}_0(\alpha,U_{1,0})))$, where $\overline{G}_0(\alpha,U_{1,0})$ is left-strictly linear over $(U_{1,0} \cup \overline{P},\overline{U_1'}) = (U_1' \cup \overline{P},\overline{U_1'})$ and τ is the substitution defined by $\tau(p) = p$ for all $p \in \overline{P}$ and $\tau(\gamma) = \tau(\overline{\gamma}) = L(G(\gamma))$ for all $\gamma \in U_1'$ and $\overline{\gamma} \in \overline{U_1'}$. But $\gamma \in U_1'$ implies $G_0(\gamma) \in \mathscr{G}_\ell(0)$ and by the nature of G_0, $\mathscr{D}_\ell(G_0(\gamma)) = L(G_0(\gamma))$. Thus by theorem 5, $\mathscr{D}_\ell(G_0(\gamma))$ is regular for all $\gamma \in U_1'$. Thus by theorem 4 part (2), $\mathscr{D}_\ell(G) = \tau(L(\overline{G}_0(\alpha,U_{1,0}))) \in \mathscr{A}_\ell(0,0) \subseteq \mathscr{A}_\ell(0)$.

Now suppose $G \in \mathscr{G}_\ell(1,j)$ implies $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(0,j)$ for all $j \leq i$. Let $G \in \mathscr{G}_\ell(1,i+1)$. Then $G_0 \in \mathscr{G}_\ell(1,i+1)$ and by lemma 2 it follows that $\mathscr{D}_\ell(G) = L(G_0) = \mathscr{D}_\ell(G_0) = \tau(L(\overline{G}_0(\alpha, U_{1,i+1})))$, where $\overline{G}_0(\alpha, U_{1,i+1})$ is left-strictly linear over $(\overline{P} \cup U_{1,i+1}, \overline{U_1'})$ and $\tau(p) = p$, $p \in \overline{P}$, $\tau(\gamma) = L(G_0(\gamma))$ for $\gamma \in U_{1,i+1}$ and $\gamma(\overline{\beta}) = L(G_0(\beta))$ for $\overline{\beta} \in \overline{U_1'}$. But $\gamma \in U_{1,i+1}$ implies $G_0(\gamma) \in \mathscr{G}_\ell(1,i)$ and since $L(G_0(\gamma)) = \mathscr{D}_\ell(G_0(\gamma))$, then by the induction hypothesis it follows that $\tau(\gamma) \in \mathscr{A}_\ell(0,i)$ for all $\gamma \in U_{1,i+1}$. By a similar argument it follows that $\tau(\overline{\beta})$ is regular for all $\overline{\beta} \in \overline{U_1'}$. Thus by theorem 4, part (2), we have that $L(G_0) = \mathscr{D}_\ell(G) \in \mathscr{A}_\ell(0,i+1)$.

<u>Case $k > 1$</u>. Assume that $G \in \mathscr{G}_\ell(k)$ implies $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(k-1)$ for all k, $1 \leq k \leq i$. By induction on j and arguments similar to those given for the case $k = 1$ (for the general case we appeal to theorem 4, part (3)), it follows that $G \in \mathscr{G}_\ell(i+1,j)$ implies $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(i,j) \subseteq \mathscr{A}_\ell(i)$. In this fashion the theorem is proved.

Out next theorem establishes a partial converse to theorem 6.

Theorem 7. Let G be an arbitrary reduced context-free grammar. If $\Delta_\ell(G)$ is not irreflexive then $\mathscr{D}_\ell(G) \notin \mathscr{A}_\ell(k)$ for any $k \geq 0$.

Proof. Since $\mathscr{A}_\ell(k)$ is a full AFL for each $k \geq 0$, then $\mathscr{A}_\ell(k)$ is closed under sequential transducer maps (Ginsburg and Greibach [3]). The proof consists of showing that if $\Delta_\ell(G)$ is not irreflexive, then for each $k \geq 0$ there exists a sequential transducer map, T_{k+1}, such that $T_{k+1}(\mathscr{D}_\ell(G)) = L_{k+1}$, where $L_{k+1} \in \mathscr{A}_\ell(k+1) - \mathscr{A}_\ell(k)$ is the language defined in theorem 3. Thus $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(k)$ for all $k \geq 0$. A precise definition of T_{k+1} is given in Workman [9].

We now strengthen theorems 6 and 7 by establishing the precise relationship between the left-degree of G and that of $\mathscr{D}_\ell(G)$.

Theorem 8. For all $k \geq 1$, if $G \in \mathscr{G}_\ell(k+1) - \mathscr{G}_\ell(k)$, then $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(k) - \mathscr{A}_\ell(k-1)$. If $G \in \mathscr{G}_\ell(1) - \mathscr{G}_\ell(0)$, then $\mathscr{D}_\ell(G) \in \mathscr{A}_\ell(0) - \mathscr{R}$, where $\mathscr{R}$ is the class of regular sets.

Proof. The proof follows the same approach used for theorem 7. That is, if $\ell\deg(G) = k+1$, then there exists a sequential transducer map, T_k, such that $T_k(\mathscr{D}_\ell(G)) = L_k$, where $L_k \in \mathscr{A}_\ell(k) - \mathscr{A}_\ell(k-1)$ is defined as in theorem 3. For $k = 0$, L_0 is a linear language which is nonregular.

The results of theorems 5, 6, 7 and 8 are summarized in the next theorem.

Theorem 9. $G \in \mathcal{G}_\ell(k+1) - \mathcal{G}_\ell(k)$ if and only if $\mathcal{D}_\ell(G) \in \mathcal{U}_\ell(k) - \mathcal{U}_\ell(k-1)$ for all $k \geq 1$. $G \in \mathcal{G}_\ell(1) - \mathcal{G}_\ell(0)$ if and only if $\mathcal{D}_\ell(G) \in \mathcal{U}_\ell(0) - \mathcal{R}$.

The preceding theorem suggests that the hierarchy of AFLs, $\mathcal{U}_\ell(k)$, $k \geq 0$, may be "generated" from the regular sets as described in our next theorem.

Theorem 10. Let $\mathcal{B}_0 = \{L \mid L = L(G)$ for some G such that $\mathcal{D}_\ell(G)$ is regular$\}$. For $i \geq 0$ define $\mathcal{B}_{i+1} = \{L \mid L = L(G)$ for some G such that $\mathcal{D}_\ell(G) \in \mathcal{B}_i\}$. Then $\mathcal{B}_k = \mathcal{U}_\ell(k)$ for all $k \geq 0$.

If we define the "structural complexity" of a context-free grammar, G, to be the least k such that $G \in \mathcal{G}_\ell(k)$ or to be infinite if such a k does not exist, then it may be useful to know the degree to which various grammar transformations may change the structural complexity. Our final result establishes the relative structural complexity between a derivation bounded grammar and an equivalent grammar in left-Greibach form.

Theorem 11. Let G be a left-Greibach grammar in $\mathcal{G}_\ell(k+1)$, then $L(G) \in \mathcal{U}_\ell(k)$. Conversely, if $L \in \mathcal{U}_\ell(k)$, then $L = L(G)$ for some left-Greibach grammar, $G \in \mathcal{G}_\ell(k+1)$.

Proof. Suppose $G = (V,\Sigma,P,\alpha) \in \mathcal{G}_\ell(k+1)$, $k \geq 0$, is in left-Grebiach form. If $p \in P$, then p is of the form $\beta \to \sigma w$, where $\sigma \in \Sigma$ and $w \in V^*$. It is clear, therefore, that $L(G) = h(L(G_0)) = h(\mathcal{D}_\ell(G))$, where G_0 is defined in the proof of theorem 6 and h is a homomorphism. By theorems 4 and 6 it follows that $L(G) \in \mathcal{U}_\ell(k)$.

The converse is proved by showing that every linear grammar, G (G is linear if and only if $G \in \mathcal{G}_\ell(0,0)$), has a left-Greibach form in $\mathcal{G}_\ell(1,0)$. By an inductive construction based on that for linear grammars, it can be shown that for all $j \geq 0$, $G \in \mathcal{G}_\ell(0,j)$ implies there exists an equivalent left-Greibach grammar, $G' \in \mathcal{G}_\ell(1,j)$. The same inductive arguments are used to establish the result for $G \in \mathcal{G}_\ell(i,j)$ to obtain a $G' \in \mathcal{G}_\ell(i+1,j)$.

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

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

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

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

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

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

7. Walljasper, S. (1974), "Left-derivation Bounded Languages", J. Comput. Sys. Sci. 8, 1-7.

8. Workman, D. (1974), "Abstract Families of Context-free Grammars", Tech. Report No. CSD-TR-114, Purdue University, Lafayette, Indiana. (Submitted for publication to JCSS)

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