SIAM J. COMPUT.
Vol. 6, No. 1, March 1977

DERIVATION COMPLEXITY IN CONTEXT-FREE GRAMMAR FORMS*

SEYMOUR GINSBURG AND NANCY LYNCH†

Abstract. Let F be an arbitrary context-free grammar form and $\mathcal{G}(F)$ the family of grammars defined by F. For each grammar G in $\mathcal{G}(F)$, the derivation complexity function Φ_G, on the language of G, is defined for each word x as the number of steps in a minimal G-derivation of x. It is shown that derivations may always be speeded up by any constant factor n, in the sense that for each positive integer n, an equivalent grammar G' in $\mathcal{G}(F)$ can be found so that $\Phi_{G'}(x) \leqq |x|/n$ for all large words x, $|x|$ denoting the length of x.

Key words. complexity theory, grammar complexity, grammar forms

Introduction. In [2] the notion of a (context-free) grammar form was introduced, to model the situation where all grammars structurally close to a master grammar are being considered. The research on grammar forms to date [2], [3], [5] has been concerned with grammatical, structural, and language-theoretic problems. The present paper initiates the study of complexity-theoretic questions. Specifically, the derivation complexity function Φ_G, defined to be the minimal number of steps in a derivation of x, is examined with respect to all grammars defined by a grammar form F. It is trivial that Φ_G is at least linear and almost trivial that Φ_G is, in fact, linear. Our main result asserts that derivations may be speeded up by any constant factor n, in the sense that for each positive integer n, an equivalent grammar G' defined by F can be found so that $\Phi_{G'}(x) \leqq |x|/n$ for all large words x, $|x|$ denoting the length of x.

The basic question underlying this work is whether among the different grammar forms yielding the same family of languages, there are some which are more efficient than others. The results of this paper show that, if length of derivation is the only criterion, there is no difference among grammar forms. As will be seen, the cost of the speedup is a large increase in the size (e.g., number of productions) of the grammars used. It remains to study the resulting trade-offs.

The notion of derivation complexity was originally defined by Gladkii [6] and has been extensively studied by Book [1] for arbitrary phrase-structure grammars. Some of the results in [1] have a speedup flavor similar to ours, but the grammars in [1] accomplishing the speedup have structure very different from those of the original grammars. By carrying out our constructions within the framework of grammar forms, we preserve structure while speeding up derivations.

The paper is divided into three sections and an Appendix. Section 1 reviews grammar form concepts, defines the derivation complexity function, and determines a lower bound for it. Section 2 is concerned with proving Proposition 2.4, a special case of the main theorem. The main result itself, Theorem 3.2, is established in § 3. The proof involves first showing (Lemma 3.1) that the original grammar form may be assumed to have certain additional properties. An induction argument on the number of variables in the grammar form is then presented,

* Received by the editors November 27, 1974, and in revised form December 8, 1975.

† Computer Science Department, University of Southern California, Los Angeles, California 90007. This work was supported in part by a Guggenheim Fellowship and in part by National Science Foundation Grant GJ 42306.

with the case for one variable being exactly the situation handled in § 2. The Appendix is devoted to proving a technical combinatorial lemma, needed in § 2 to verify the main theorem for each grammar form defining the family of all context-free languages.

1. Preliminaries. In this section we first review the principal ideas relating to context-free grammar forms. Then we introduce the formalism for treating derivation complexity in grammar forms.

DEFINITION. A (*context-free*) *grammar form* is a 6-tuple $F = (V, \Sigma, \mathscr{V}, \mathscr{S}, \mathscr{P}, \sigma)$, where

(i) V is an infinite set of abstract symbols,

(ii) Σ is an infinite subset of V such that $V-\Sigma$ is infinite, and

(iii) $G_F = (\mathscr{V}, \mathscr{S}, \mathscr{P}, \sigma)$, called the *form grammar* (*of* F) is a context-free grammar[1] with $\mathscr{S} \subseteq \Sigma$ and $(\mathscr{V}-\mathscr{S}) \subseteq (V-\Sigma)$.

The reader is referred to [2] for motivation and further details about grammar forms.

Throughout, V and Σ are assumed to be fixed infinite sets satisfying conditions (i) and (ii) above. All context-free grammar forms considered here are with respect to this V and Σ. Also, the adjective "context-free" is usually omitted from the expression "context-free grammar form."

The purpose of a grammar form is to specify a family of grammars, each "structurally close" to the form grammar. This is accomplished by the notion of:

DEFINITION. An *interpretation* of a grammar form $F = (V, \Sigma, \mathscr{V}, \mathscr{S}, \mathscr{P}, \sigma)$ is a 5-tuple $I = (\mu, V_I, \Sigma_I, P_I, S_I)$, where

1. μ is a substitution on $\mathscr{V}^*$ such that $\mu(a)$ is a finite subset of Σ^*, $\mu(\xi)$ is a finite subset of $V-\Sigma$ for each ξ in $\mathscr{V}-\mathscr{S}$, and $\mu(\xi) \cap \mu(\eta) = \varnothing$ for each ξ and η, $\xi \neq \eta$, in $\mathscr{V}-\mathscr{S}$;

2. P_I is a subset of $\mu(\mathscr{P}) = \bigcup_{\pi \text{ in } \mathscr{P}} \mu(\pi)$, where $\mu(\alpha \to \beta) = \{u \to v / u \text{ in } \mu(\alpha), v \text{ in } \mu(\beta)\}$;

3. S_I is in $\mu(\sigma)$; and

4. $\Sigma_I(V_I)$ contains the set of all symbols in $\Sigma(V)$ which occur in P_I (together with S_I).

$G_I = (V_I, \Sigma_I, P_I, S_I)$ is called the *grammar* of I.

An interpretation is usually exhibited by indicating S_I, P_I, and (implicitly or explicitly) μ. The sets V_I and Σ_I are ordinarily not stated explicitly.

A grammar form determines a family of grammars and a family of languages as follows:

DEFINITION. For each grammar form F, $\mathscr{G}(F) = \{G_I / I \text{ an interpretation of } F\}$ is called the *family of grammars of* F and $\mathscr{L}(F) = \{L(G_I) | G_I \text{ in } \mathscr{G}(F)\}$ the *grammatical family of* F.

In this paper we are interested in studying derivation complexity in a grammar form, i.e., the derivation complexity of the grammars in $\mathscr{G}(F)$. To do this we consider the following:

[1] We assume the reader is familiar with the basic notions pertaining to context-free grammars. Here $\mathscr{V}$ is the total alphabet, $\mathscr{S}$ is the terminal alphabet, $\mathscr{P}$ is the set of productions, and σ is the start variable.

Notation. For every context-free grammar $G = (V_1, \Sigma_1, P, S)$ let Φ_G be the function on $L(G)$ in which $\Phi_G(x)$ is the minimum number of steps among all G-derivations of x,[2] for each x in $L(G)$.

Thus Φ_G is the minimum derivation function, in the sense that $\Phi_G(x)$ is the minimum number of steps necessary to derive x.

Notation. For every context-free grammar G let $\phi_G = \min\{\Phi_G(x) | x \text{ in } L(G), x \neq \varepsilon\}$ if G is not vacuous,[3] and $\phi_G = \infty$ otherwise.

Thus ϕ_G is the fewest number of steps needed to derive at least one non-ε word in $L(G)$.

The restriction in our definition of ϕ_G to non-ε words is needed because of the construction used in Lemma 2.1 to make finite patches on grammars.

From Lemma 2.1 of [2] we immediately get:

LEMMA 1.1. *For each grammar form F and each grammar G in $\mathcal{G}(F)$, $\phi_G \geqq \phi_{G_F}$.*

Using the above lemma we now obtain a lower bound for Φ_G.

PROPOSITION 1.2. *Let F be a grammar form and G in $\mathcal{G}(F)$. Then there exists a positive integer n so that $\Phi_G(x) \geqq \max\{\phi_{G_F}, |x|/n\}$ for[4] all $x \neq \varepsilon$ in $L(G)$.*

Proof. Let n be the largest number of terminal symbols on the right side of any production of G. Then for each $x \neq \varepsilon$ in $L(G)$, at least $|x|/n$ steps are needed to derive x and

$$\Phi_G(x) \geqq \phi_G, \quad \text{by definition,}$$
$$\geqq \phi_{G_F}, \quad \text{by Lemma 1.1.}$$

Hence the result.

Note that the right-hand expression in the conclusion of Proposition 1.2 decreases as n increases. The question arises whether the lower bound on the right-hand side is attainable as n gets larger.

DEFINITION. A grammar form F is *minimal* if for each L in $\mathcal{L}(F)$ and each positive integer n, there exists a grammar G in $\mathcal{G}(F)$ such that $L(G) = L$ and $\Phi_G(x) \leqq \max\{\phi_{G_F}, |x|/n\}$ for all x in L.

The purpose of the present paper is to prove that every grammar form is minimal. In other words, for each grammar form F and each language L in $\mathcal{L}(F)$ a grammar G in $\mathcal{G}(F)$ can be found which speeds up the derivation as much as desired. Thus, if derivation complexity is the only criterion being considered, there is no reason to select one grammar form instead of another.

2. Minimal forms for specific grammatical families. In this section we show that all grammar forms defining the finite languages, the regular sets, the linear languages, and the context-free languages are minimal. Using this result we then prove in the next section that all grammar forms are minimal.

To establish the result about all grammar forms for the above four grammatical families we need two lemmas. The first states that if a grammar form F has an

[2] By a *G-derivation of x, with n steps*, is meant a derivation $S = w_0 \underset{G}{\Rightarrow} w_1 \underset{G}{\Rightarrow} \cdots \underset{G}{\Rightarrow} w_n = x$.

[3] A grammar G is said to be *vacuous* if $L(G) = \varnothing$ or $L(G) = \{\varepsilon\}$. A grammar form F is said to be *vacuous* if G_F is vacuous.

[4] For each word x, $|x|$ denotes its length.

interpretation I which derives all long words x in $L(G_I)$ in at most $|x|/n$ steps, then F has another interpretation defining $L(G_I)$ which derives all possible words x in $L(G_I)$ in at most $|x|/n$ steps. In other words, a "finite patch" may be made for short words.

LEMMA 2.1. *Let F be a nonvacuous form and n a positive integer. Suppose there exists a positive integer k and a grammar G in $\mathcal{G}(F)$ such that $\Phi_G(x) \leqq \max\{k, |x|/n\}$ for all x in $L(G)$. Then there exists a grammar G' in $\mathcal{G}(F)$ such that $L(G') = L(G)$ and $\Phi_{G'}(x) \leqq \max\{\phi_{G_F}, |x|/n\}$ for all x in $L(G')$.*

Proof. The argument consists of adding to G productions that generate, in ϕ_{G_F} steps, the finite number of words x in $L(G)$ for which $|x|/n < k$.

Since $F = (V, \Sigma, \mathcal{V}, \mathcal{S}, \mathcal{P}, \sigma)$ is not vacuous, there is a ϕ_{G_F}-step derivation

$$\sigma = z_0 \underset{G_F}{\Longrightarrow} \cdots \underset{G_F}{\Longrightarrow} z_{\phi_{G_F}} = z \tag{1}$$

of a non-ε word z in $L(G_F)$. For each i, $1 \leqq i \leqq \phi_{G_F}$, let

$$\beta_i \to w_i \tag{2}$$

be the production used in $z_{i-1} \underset{G_F}{\Rightarrow} z_i$. Since $z \neq \varepsilon$, there exists some j such that z_j contains a symbol of $\mathcal{S}$, say $w_j = w_{j1} a_j w_{j2}$, where a_j is in $\mathcal{S}$ and w_{j1}, w_{j2} are in $\mathcal{V}^*$. Note that for every i and every occurrence of a variable on the right side of the ith production in (2), there is a unique integer l such that $\beta_l \to w_l$ is the production applied to that variable in the derivation (1).

Now let $G = (V_1, \Sigma_1, P_1, S)$ and x be an arbitrary word in $L(G)$, with $|x|/n < k$. For each i, $2 \leqq i \leqq \phi_{G_F}$, let $A_{i,x}$ be a new variable (in $V - \Sigma$). Let $A_{1,x} = S$. For each x, consider the set of new productions

$$\{A_{i,x} \to v_{i,x} \mid 1 \leqq i \leqq \phi_{G_F},\ i \neq j\} \cup \{A_{j,x} \to v_{j1x} x v_{j2x}\}, \tag{3}$$

where each v_{ix}, v_{j1x}, v_{j2x} is obtained from the corresponding word w_i, w_{j1}, w_{j2} by deleting all symbols in $\mathcal{S}$ and replacing each variable by the appropriate variable $A_{l,x}$. Clearly the rules in (3) derive just the word x and no production $A_{i,x} \to v_{i,x}$ can be applied to any variable $A_{i',y}$ unless $i = i' = 1$.

Let $G' = (V_2, \Sigma_1, P_2, S)$, where

$$V_2 = V_1 \cup \{A_{i,x} \mid x \text{ in } L(G), \frac{|x|}{n} < k,\ 2 \leqq i \leqq \phi_{G_F}\}.$$

Obviously G' is in $\mathcal{G}(F)$, $L(G') = L(G)$, and G' has a ϕ_F-step derivation for all x in $L(G) = L(G')$. Thus G' satisfies the conclusion of the lemma.

Remark. An alternative formulation of Lemma 2.1 is the following: Let F be a nonvacuous form and n a positive integer. Suppose there exists a positive integer r such that $\Phi_G(x) \leqq |x|/n$ for all x in $L(G)$ having $|x| \geqq r$. Then there exists a grammar G' in $\mathcal{G}(F)$ so that $L(G') = L(G)$ and $\Phi_{G'}(x) \leqq \max\{\phi_{G_F}, |x|/n\}$ for all x in $L(G')$.

COROLLARY 2.2. *Let $F = (V, \Sigma, \mathcal{V}, \mathcal{S}, \mathcal{P}, \sigma)$ and $F' = (V, \Sigma, \mathcal{V}', \mathcal{S}', \mathcal{P}', \sigma)$ be equivalent grammar forms,[5] with $\mathcal{P} \subseteq \mathcal{P}'$. If F is minimal, then so is F'.*

[5] Grammar forms F and F' are said to be *equivalent* if $\mathcal{L}(F) = \mathcal{L}(F')$.

The second lemma states that minimal grammar forms are not affected by either adding or removing "redundant" productions.

LEMMA 2.3. *Let F be a nonvacuous grammar form and $\beta \overset{*}{\underset{G_F}{\Rightarrow}} w$ a derivation, with β a variable. Let F' be the grammar form obtained by adding to F the production $\beta \to w$. Then F is minimal if and only if F' is minimal.*

Proof. Suppose F is minimal. By Proposition 2.1 of [2], $\mathscr{L}(F') = \mathscr{L}(F)$. Hence F' is minimal by Corollary 2.2.

Now assume that F' is minimal. Let L be in $\mathscr{L}(F) = \mathscr{L}(F')$ and n be a positive integer. Since $\beta \overset{*}{\underset{G_F}{\Rightarrow}} w$, $\beta \overset{k}{\underset{G_F}{\Rightarrow}} w$ for[6] some nonnegative integer k. Suppose $k \geqq 1$. Then there exists $G' = (V_1, \Sigma_1, P', S)$ in $\mathscr{G}(F')$ such that $L(G') = L$ and $\Phi_{G'}(x) \leqq \max\{\phi_{G_{F'}}, |x|/(kn)\}$ for all x in L. Also, there exists a sequence $\pi_1, \cdots, \pi_k$ of productions in F which realize $\beta \overset{k}{\Rightarrow} w$. Let $G = (V_2, \Sigma_1, P, S)$ in $\mathscr{G}(F)$ be obtained from G' as follows. Each production $A \to y$ in G' which comes from the production $\beta \to w$ in F' is replaced by a sequence of k productions, corresponding to $\pi_1, \cdots, \pi_k$, which realizes $A \overset{*}{\Rightarrow} y$. As in the proof of Lemma 2.1, new intermediate variables are introduced in the sequence in such a way that the new sequence can only derive $A \overset{*}{\Rightarrow} y$. Let V_2 be V_1 together with all the new variables introduced. Clearly $L(G) = L(G')$. Since every production of G' requires at most k productions of G to simulate it, $\Phi_G(x) \leqq \max\{k\phi_{G_{F'}}, k|x|/(kn)\}$ for all x in L. By Lemma 2.1, there exists $\bar{G}$ in $\mathscr{G}(F)$ such that $L(\bar{G}) = L$ and $\Phi_{\bar{G}}(x) \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for all x in L.

Suppose $k = 0$. Then there exists $G' = (V_1, \Sigma_1, P', S)$ in $\mathscr{G}(F')$ such that $L(G') = L$ and $\Phi_{G'}(x) \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for all x in L. Let $G = (V_2, \Sigma_1, P, S)$, where

$$P = (P' - \{A \to B \text{ in } P' | A \text{ and } B \text{ in } \mu(\beta)\}) \cup \{A \to y | A \text{ in } \mu(\beta),\ y \text{ in } V^* - \mu(\beta), \text{ there exist } l \geqq 1 \text{ and } A_1, \cdots, A_l \text{ in } \mu(\beta) \text{ such that } A \to A_1, A_i \to A_{i+1}, A_l \to y, 1 \leqq i \leqq l-1, \text{ are in } P'\}.$$

Clearly G is in $\mathscr{G}(F)$ and $L(G) = L$. Since each derivation δ' in G' of a word x in L has an obvious corresponding derivation δ in G of x, with at most the same number of steps as in δ', $\Phi_G(x) \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for each x in L. By Lemma 2.1, there exists $\bar{G}$ in $\mathscr{G}(F)$ such that $L(\bar{G}) = L$ and $\Phi_{\bar{G}}(x) \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for all x in L.

Thus, for any value of k, there exists $\bar{G}$ in $\mathscr{G}(F)$ so that $L(\bar{G}) = L$ and $\Phi_{\bar{G}}(x) \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for all x in L. Hence F is minimal.

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

PROPOSITION 2.4. *In each of the following cases, $F = (V, \Sigma, \mathscr{V}, \mathscr{S}, \mathscr{P}, \sigma)$ is minimal:*

(a) *$\mathscr{L}(F)$ is the family of all finite sets.*
(b) *$\mathscr{L}(F)$ is the family of all regular sets.*
(c) *$\mathscr{L}(F)$ is the family of all linear languages.*
(d) *$\mathscr{L}(F)$ is the family of all context-free languages.*

[6] By $u \overset{k}{\Rightarrow} v$ is meant that there exist $u_1, \cdots, u_{k-1}$ such that $u_0 = u \Rightarrow u_1 \Rightarrow \cdots \Rightarrow u_k = v$.

Proof. (a) Since $\mathscr{L}(F)$ is the family of all finite sets, by Theorem 2.1 of [2] there is a word w in[7] $\mathscr{S}^+$ such that $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} w$. Let F' be the grammar form obtained from F by adding the production $\sigma \to w$. By Lemma 2.3 it suffices to show that F' is minimal.

Let $L = \{x_1, \cdots, x_k\}$ be any finite set and let n be an arbitrary positive integer. If $k = 0$ there is nothing to prove. Suppose $k \geqq 1$. Let $G = (\{S\} \cup \Sigma_1, \Sigma_1, P, S)$ be the grammar where Σ_1 is the set of all symbols appearing in any of the x_i, $1 \leqq i \leqq k$, and $P = \{S \to x_i | 1 \leqq i \leqq k\}$. Obviously G is in $\mathscr{G}(F')$, $L(G) = L$, and $\Phi_G(x) = 1 \leqq \max\{\phi_{G_{F'}}, |x|/n\}$ for all x in L. Thus F' is minimal.

(b) Since $\mathscr{L}(F)$ is the class of all regular sets, $L(G_F)$ is an infinite set by Theorem 2.1 of [2]. By [9], there exist x_1, x_2, x_3, x_4, x_5 in $\mathscr{S}^*$ and β in $\mathscr{V} - \mathscr{S}$ such that x_5 is in $\mathscr{S}^+$, x_3x_4 is in $\mathscr{S}^+$, $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} x_1\beta x_2$, $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_3\beta x_4$, and $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_5$. By Lemma 2.3, there is no loss of generality in assuming that $\sigma \to x_1\beta x_2$, $\beta \to x_3\beta x_4$, and $\beta \to x_5$ are in $\mathscr{P}$. By symmetry, there is no loss in assuming $x_3 \neq \varepsilon$.

Let L be any regular set. Then $L = L(G)$ for some right-linear grammar $G = (V_1, \Sigma_1, P, S)$ in which $A \to wB$ in P, A and B variables, implies w is in Σ_1^+. Let n be an arbitrary positive integer. Let S' be a new variable,[8]

$$P_2 = \{S' \to S\} \cup \{A \to wB | A, B \text{ in } V_1 - \Sigma_1, A \overset{\leqq n+1}{\underset{G}{\longrightarrow}} wB\}$$

$$\cup \{A \to w | A \text{ in } V_1 - \Sigma_1, w \text{ in } \Sigma_1^*, A \overset{\leqq n+1}{\underset{G}{\Longrightarrow}} w\},$$

and $G' = (\{S'\} \cup V_1, \Sigma_1, P_2, S')$. Clearly $L(G') = L(G) = L$. Also G' is in $\mathscr{G}(F)$. [For one can construct an interpretation (μ, G') of F for which $S' \to S$ is in $\mu(\sigma \to x_1\beta x_2)$, $A \to wB$ is in $\mu(\beta \to x_3\beta x_4)$ for every production $A \to wB$, A and B variables, in P_2, and $A \to w$ is in $\mu(\beta \to x_5)$ for every production $A \to w$, w in Σ_1^*, in P_2.] Consider any word x in L. Obviously there exists a derivation in G' of x so that except, perhaps, for the first and last productions, each production deposits at least $n+1$ terminals. Thus $\Phi_{G'}(x) \leqq 2 + |x|/(n+1)$. For x sufficiently large, $2 + |x|/(n+1) < |x|/n$. Hence, $\Phi_{G'}(x) \leqq |x|/n$ for all large x. By Lemma 2.1, F is minimal.

(c) Since $\mathscr{L}(F)$ is the family of all linear languages, by Theorem 2.4 of [2] there exist x_1, x_2, x_3, x_4, x_5 in $\mathscr{S}^*$ and β in $\mathscr{V} - \mathscr{S}$ such that x_3, x_4, x_5 are in $\mathscr{S}^+$ and $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} x_1\beta x_2$, $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_3\beta x_4$, and $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_5$. By Lemma 2.3, we may assume that $\sigma \to x_1\beta x_2$, $\beta \to x_3\beta x_4$, and $\beta \to x_5$ are in $\mathscr{P}$.

Now let L be any linear language. Then $L = L(G)$ for some linear grammar $G = (V_1, \Sigma_1, P, S)$ such that $A \to uBv$, A and B in $\mathscr{V} - \mathscr{S}$, implies uv in Σ_1^+. Let n be

---

[7] For each set E of words, $E^+ = \cup_{i=1}^{\infty} E^i$.

[8] By $u \overset{\leqq k}{\underset{G}{\Longrightarrow}} v$ is meant that there exists a derivation of v from u in at most k (possibly 0) steps.

an arbitrary positive integer. Let S' be a new variable,

$$P_2 = \{S' \to S\} \cup \{A \to uBv | A, B \text{ in } V_1 - \Sigma_1, A \overset{\leqq n+1}{\underset{G}{\Longrightarrow}} uBv\}$$

$$\cup \{A \to w | A \text{ in } V_1 - \Sigma_1, w \text{ in } \Sigma_1^*, A \overset{\leqq n+1}{\underset{G}{\Longrightarrow}} w\},$$

and $G' = (\{S'\} \cup V_1, \Sigma_1, P_2, S')$. The remainder of the argument is as in part (b).

(d) As often happens in proofs about grammar forms, the case where $\mathscr{L}(F)$ is the family of all context-free languages is proved very differently from the other cases, although the statement of the result is similar. Here we cannot simply compose productions as in (b) and (c), since the resulting productions would not necessarily be in the required form. Instead, we introduce productions where possible which simulate within the required form the result of composing sequences of productions. We then use a combinatorial lemma to show that the simulating productions are sufficient for the task at hand.

Since $\mathscr{L}(F)$ is the family of all context-free languages, by Theorem 2.2 of [2] there exist x_1, x_2, x_3, x_4, x_5 in $\mathscr{S}^*$, x_6 in $\mathscr{S}^+$, and β in $\mathscr{V} - \mathscr{S}$ such that $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} x_1\beta x_2$, $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_3\beta x_4\beta x_5$, and $\beta \overset{*}{\underset{G_F}{\Longrightarrow}} x_6$. By Lemma 2.3, we may assume that $\sigma \to x_1\beta x_2$, $\beta \to x_3\beta x_4\beta x_5$, and $\beta \to x_6$ are in $\mathscr{P}$. Intuitively this means that it suffices to prove the results for $G_F = (\{\sigma, a\}, \{a\}, \{\sigma \to \sigma\sigma, \sigma \to a\}, \sigma)$.

Now let L be any context-free language. Then there exists a grammar $G = (V_1, \Sigma_1, P_1, S)$ such that $L = L(G)$ and each production of P_1 is of the type $A \to BC$ or $A \to w$, where A, B, C are variables, w is in Σ_1^*, neither B nor C is S, and $A = S$ if $w = \varepsilon$. (Thus S never appears on the right side of any production and S is the only variable which derives ε.) Intuitively, we shall construct a grammar G' as follows. We consider a derivation of a word w in G, represented by a derivation tree. We put into G' productions which simulate the effects of G-productions used near the ends of branches. (See part 4 below.) Thus if the derivation tree of w is very wide, then these productions yield the needed speed-up. On the other hand, the derivation tree of w may be very narrow, with very little internal branching. In this case, the new productions do not speed up the derivation sufficiently. To obtain the speed-up here, we put into G' productions which "condense" long internal paths having little branching. (See part 6 below.)

Proceeding more formally, note that

1. if $A \overset{s}{\underset{G}{\Longrightarrow}} w$, where A is in $V_1 - \Sigma_1$ and w is in $\Sigma_1^*(V_1 - \Sigma_1)\Sigma_1^*$, then $s \leqq 2|w| - 2$, and

2. if $A \overset{s}{\underset{G}{\Longrightarrow}} w$, where A is in $V_1 - \Sigma_1$ and w is in Σ_1^+, then $s \leqq 2|w| - 1$.

Let n be an arbitrary positive integer. Let $G' = (V_2, \Sigma_1, P_2, S')$, where S' is a new variable, V_2 consists of the symbols of V_1 together with all variables in P_2, and P_2 is defined as follows:

3. $S' \to S$ is in P_2.

4. $A \to w$ is in P_2 if $A \overset{\leqq 112n}{\underset{G}{\Longrightarrow}} w$, where A is in $V_1 - \Sigma_1$ and w is in Σ_1^*.

5. $A \to BC$ if A, B, C are in $V_1 - \Sigma_1$ and $A \to BC$ is in P_1.

6. For each derivation $\delta\colon A \underset{G}{\Rightarrow} \cdots \underset{G}{\Rightarrow} vBw$ of at most $112n$ steps, with A, B in $V_1 - \Sigma_1$ and v, w in Σ_1^*, let D_δ, E_δ, and H_δ be new variables. Let $A \to D_\delta E_\delta$, $E_\delta \to BH_\delta$, $D_\delta \to v$, and $H_\delta \to w$ be in P_2.

The indexing of the variables in part 6 is done, as in the proof of Lemma 2.1, to keep the new variables distinct, so that each new production can be used only as part of the simulation of the derivation of G for what it was intended.

From parts 2 and 4, we get

7. $A \to w$ is in P_2 if $A \overset{*}{\underset{G}{\Longrightarrow}} w$, with A in $V_1 - \Sigma_1$, w in Σ_1^*, and $|w| \leqq 56n$.

It is easily seen that G' is in $\mathcal{G}(F)$. Since $P_1 \subseteq P_2$ and new productions (except $S' \to S$) in P_2 are only used to simulate productions in P_1, it follows that $L(G') = L(G) = L$. To complete the argument, it remains to show that G' has sufficiently short derivations of all words in L.

Let x be an arbitrary word in L. For each G'-derivation of x, there is associated in the obvious way [4] a G'-derivation tree[9] $T(x)$. Let $\bar{T}(x)$ be the tree obtained from $T(x)$ by deleting the root, all leaf nodes, and all edges incident to these nodes. Note that $\bar{T}(x)$ is a binary tree.[10] Consider the set of all such derivation trees $\bar{T}(x)$ for x. Let $\bar{T}_0(x)$ be one such tree with the fewest number of nodes, and let $T_0(x)$ be a tree giving rise to $\bar{T}_0(x)$. Then $\bar{T}_0(x)$ has the following three properties:

8. No two consecutive[11] internal nodes of $\bar{T}_0(x)$ have both their node names in $V_2 - V_1$.

9. Each internal node of $\bar{T}_0(x)$ with at least one daughter an internal node generates a subtree of $T_0(x)$ whose terminal word is of length $\geqq 56n$.

10. Let $n_1, \cdots, n_6$ be six consecutive internal nodes of $\bar{T}_0(x)$ and for each i, $2 \leqq i \leqq 6$, let m_i be the other daughter of n_{i-1}. Suppose that each m_i is a leaf in $\bar{T}_0(x)$ and $B_i \to w_i$ in $T_0(x)$, B_i the node name of m_i. Then $|w_2 \cdots w_6| \geqq 56n$.

For consider part 8. Let n_1 and n_2 be consecutive internal nodes of $\bar{T}_0(x)$. Let A and B be the node names of n_1 and n_2 respectively. Suppose A is in $V_2 - V_1$. Then by part 6, there is a δ such that A is E_δ and B is H_δ. By construction, $B = H_\delta$ cannot be an internal node of $\bar{T}_0(x)$.

Consider part 9. Suppose it is false. Then there exist consecutive internal nodes n_1 and n_2 of $\bar{T}_0(x)$ such that n_1 (thus n_2) generates a subtree of $T_0(x)$ whose terminal word w_1 (w_2) is of length smaller than $56n$. Let A and B be the node names of n_1 and n_2 respectively. By part 8, one of the variables, say A, is in V_1. Then $A \overset{*}{\underset{G}{\Longrightarrow}} w_1$. By part 7 $A \to w_1$ is in P_2. Replacing the subtree in $T_0(x)$ realizing

[9] Trees are viewed with the root node at the top. A node leading downward to another node is called an *internal* node. Otherwise, the node is called a *leaf*. If nodes n_1 and n_2 are jointed by an edge, with n_2 below n_1, then n_2 is called a *daughter* of n_1, and n_1 the *father* of n_2.

[10] A tree is called *binary* if each internal node has exactly two daughters.

[11] Nodes $n_1, \cdots, n_r$ are *consecutive* if each n_{i+1} is a daughter of n_i, $i \geqq 2$.

$A \overset{*}{\underset{G'}{\Rightarrow}} w_1$ by the subtree representing $A \to w_1$ gives rise to a G'-derivation tree $T_1(x)$, deriving x, with the property that $\bar{T}_1(x)$ has fewer nodes than $\bar{T}_0(x)$. (The two daughter nodes of n_1 in $\bar{T}_0(x)$ are no longer present.) This is a contradiction. A similar contradiction arises if B is in V_1. Hence part 9 holds.

Consider part 10. Suppose it is false. Let A_i, $1 \leqq i \leqq 6$, and B_j, $2 \leqq j \leqq 6$, be the node names of n_i and m_j, respectively. By part 8, either A_1 or A_2, say A_2, and either A_5 or A_6, say A_5, is in V_1. [An analogous argument holds if any of the other three possibilities occurs.] Since m_3, n_3 are daughters of n_2; m_4, n_4 are daughters of n_3; and m_5, n_5 are daughters of n_4, we have as productions in P_2, $A_2 \to A_3B_3$ or $A_2 \to B_3A_3$, $A_3 \to A_4B_4$ or $A_3 \to B_4A_4$, and $A_4 \to A_5B_5$ or $A_4 \to B_5A_5$. Suppose $A_2 \to B_3A_3$, $A_3 \to A_4B_4$, and $A_4 \to B_5A_5$ are the productions in P_2 realizing the above daughter relations. [An analogous argument holds if one of the other combinations occur.] Thus

$$A_2 \underset{G'}{\Rightarrow} B_3A_3 \underset{G'}{\Rightarrow} B_3A_4B_4 \underset{G'}{\Rightarrow} B_3B_5A_5B_4 \overset{*}{\underset{G'}{\Rightarrow}} w_3w_5A_5w_4.$$

Since A_2 and A_5 are in V_1, $A_2 \overset{*}{\underset{G}{\Rightarrow}} w_3w_5A_5w_4$. By assumption, $|w_2 \cdots w_6| < 56n$. Thus $|w_3w_5w_4| < 56n$. By part 1, there exists a derivation $\delta : A_2 \underset{G}{\Rightarrow} \cdots \underset{G}{\Rightarrow} w_3w_5A_5w_4$, of at most $112n$ steps. Replacing in T_0 the subgraph realizing $A_2 \to B_3A_3, B_3 \to w_3, A_3 \to A_4B_4, B_4 \to w_4, A_4 \to B_5A_5, B_5 \to w_5$ by the graph realizing $A_2 \to D_\delta E_\delta, E_\delta \to A_5H_\delta, D_\delta \to w_3w_5, H_\delta \to w_4$ gives rise to a G' derivation tree $T_1(x)$ for x. Then $\bar{T}_1(x)$ has two nodes fewer than $\bar{T}_0(x)$. This contradicts the minimality of $\bar{T}_0(x)$. Hence part 10 holds.

Using the above symbolism for trees, let r be a positive integer such that $\bar{T}_0(x)$ has at least two internal nodes for each x in L, $|x| \geqq r$. Clearly r exists. By Lemma 2.1, it suffices to show that $\Phi_{G'}(x) \leqq |x|/n$ for each word x in L, $|x| \geqq r$.

Let x be any word in L, with $|x| \geqq r$. Consider the following result, whose proof is in the Appendix.

LEMMA 2.5. *Let T be a binary tree with at least two internal nodes and exactly l leaf nodes. Suppose there exists a positive integer k and a weight function ω which assigns a nonnegative integer to every leaf node in such a way that the following two properties hold*:

(a) *For each internal node n_0 which has at least one daughter an internal node, $\sum_{m \text{ in } Q(n_0)} \omega(m) \geqq k$, where $Q(n_0)$ is the set of all leaf nodes in the subtree generated by n_0.*

(b) *If $n_1, \cdots, n_6$ are six arbitrary consecutive internal nodes and $m_2, \cdots, m_7$ are leaf nodes such that each m_i is a daughter of n_{i-1}, then $\sum_{i=2}^{7} \omega(m_i) \geqq k$.*

Then

$$\sum_{m \text{ a leaf}} \omega(m) \geqq \frac{kl}{28}.$$

In Lemma 2.5, let $k = 56n$, let $T = \bar{T}_0(x)$, and for each leaf node m in $\bar{T}_0(x)$ let $\omega(m) = |w|$, where A is the node name of m and $A \to w$ is the production in P_2 realizing the subtree in $T_0(x)$ generated by m. By parts 9 and 10, (a) and (b) of

Lemma 2.5 are satisfied. (There is some redundancy in (b).) By the conclusion of Lemma 2.5, $\sum_{m \text{ a leaf}} \omega(m) \geqq (56n/28)l = 2nl$, where l is the number of leaf nodes in $\bar{T}_0(x)$. Now the number of steps in any derivation realizing $T_0(x)$ is 1 (for $S' \to S$) plus the number of internal nodes of $\bar{T}_0(x)$ plus l. Since $\bar{T}_0(x)$ is a binary tree, it is easily seen that l is 1 plus the number of internal nodes. Hence the number of steps in any derivation realizing $T_0(x)$ is $2l$. But $|x| = \sum_{m \text{ a leaf}} \omega(m)$. Therefore $\Phi_{G'}(x) = 2l \leqq |x|/n$, and the proof of Proposition 2.4 is complete.

3. Minimality of arbitrary grammar forms. In the previous section we proved that all the grammar forms for some special types of grammatical families were minimal. In the present section we establish the result for all grammar forms for all grammatical families.

The argument for the main result is as follows. In § 3 of [2] a procedure was given for converting an arbitrary grammar form into an equivalent, completely reduced, sequential one. This procedure is exploited here to show that the transformation cannot convert a nonminimal grammar form into a minimal one. It is then proved that the resulting grammar form is always minimal.

LEMMA 3.1. *For every grammar form F there exists an equivalent, completely reduced,*[12] *sequential grammar form*[13] *F' such that F is minimal if F' is.*

Proof. If F is vacuous then $L(G_F) = \varnothing$ or $L(G_F) = \{\varepsilon\}$. Thus $\mathscr{L}(F) = \{\varnothing\}$ or $\mathscr{L}(F) = \{\varnothing, \{\varepsilon\}\}$. In the former case, let F' be a form with no productions, and in the latter let F' be a grammar form with the single production $\sigma \to \varepsilon$. Clearly F and F' are both minimal, and F' satisfies the conclusion of the lemma.

Suppose that F is not vacuous. For the remainder of this proof, we assume the reader is familiar with the contents of § 3 of [2]. We follow the transformation procedure given there, noting that each step of the procedure cannot change a nonminimal grammar form into a minimal one. There are five parts to consider.

(a) By the proof of Lemma 3.1 of [2] a reduced, equivalent grammar form F_a is obtained from F. Since F_a is constructed by deleting the useless productions of F, i.e., those productions involved in no derivation of a terminal word, only useless productions of each G in $\mathscr{G}(F)$ are deleted. Thus F is minimal if F_a is.

(b) By Lemma 3.2 of [2], an equivalent, reduced, noncyclic grammar form F_b is obtained from F_a. Assume F_b is minimal. Let F'' be the grammar form obtained by adding to F_b all productions $\beta \to \gamma$, β and γ variables, for which $\beta \overset{*}{\underset{G_{F_a}}{\Longrightarrow}} \gamma$. Then $\mathscr{L}(F_b) = \mathscr{L}(F_a) = \mathscr{L}(F'')$. Since each production in F_b is in F'', it follows from Corollary 2.2 that F'' is minimal. By repeated use of Lemma 2.3, F_a is minimal.

(c) By Lemma 3.3 of [2] an equivalent reduced grammar form F_c containing no production of the kind $\xi \to \eta$, ξ and η variables, is obtained from F_b. Suppose F_c is minimal. Repeating the argument in (b) above, with F_c replaced by F_a and F_a by F_b, it is easily seen that F_b is minimal.

---

[12] A grammar form $F = (V, \Sigma, \mathscr{V}, \mathscr{S}, \mathscr{P}, \sigma)$ is said to be *completely reduced* if G_F is reduced, there are no variables α and β in $\mathscr{V} - \mathscr{S}$ such that $\alpha \to \beta$ is in $\mathscr{P}$, and for each variable α in $\mathscr{V} - (\mathscr{S} \cup \{\sigma\})$ there exist x and y in $\mathscr{S}^*$, $xy \neq \varepsilon$, such that $\alpha \to x\alpha y$ is in $\mathscr{P}$.

[13] A context-free grammar (V_1, Σ_1, P, S) is *sequential* if the variables can be ordered $S = A_1, \cdots, A_r$ such that if $A_i \to uA_jv$ is any production in P, then $j \geqq i$. A grammar form F is *sequential* if G_F is sequential.

(d) By Lemma 3.4 of [2], an equivalent, completely reduced grammar form F_d is obtained from F_c. Assume F_d is minimal. If every variable of F_c is partially self-embedding, then F_c is minimal by Lemma 2.3. Suppose that F_c has exactly $k>0$ variables which are not partially self-embedding. The procedure in (β) of Lemma 3.4 of [2] shows how to obtain from F_c an equivalent, reduced grammar form F_c' having exactly $k-1$ variables which are not partially self-embedding. This procedure is iterated k times until F_d is obtained. To show that F_c is minimal, it therefore suffices to prove that F_c is minimal provided that F_c' is minimal.

Assume F_c' is minimal. Let L be in $\mathscr{L}(F_c)=\mathscr{L}(F_c')$ and n be an arbitrary positive integer. Let l be 1 plus the maximum number of times any single variable appears on the right of any single production of F_c. Then there exists a grammar G' in $\mathscr{G}(F_c')$ such that $L(G')=L$ and $\Phi_{G'}(x)\leqq\max\{\phi_{G_{F'}}, |x|/(ln)\}$ for all x in L. By the method of construction of F_c' from F_c, there exists a grammar G in $\mathscr{G}(F_c)$ such that $L(G)=L$ and the following holds: For every G'-derivation δ' of a word x in $L(G')$ there is a G-derivation δ of x in which each step of δ' is simulated by at most l steps of δ. Then $\Phi_G(x)\leqq\max\{l\phi_{G_{F_c'}}, |x|/n\}$ for all x in L. The minimality of F_c follows from Lemma 2.1.

(e) By Theorem 3.1 of [2], an equivalent, completely reduced, sequential grammar form F' is obtained from F_d. Assume F' is minimal. Let F_e' be a grammar form obtained by adding to F_d one production of the kind $\beta\to v\gamma w$, v and w in $\mathscr{S}^*$, for all variables β and γ, $\beta\neq\gamma$, in F_d such that $\beta \overset{*}{\underset{G_{F_d}}{\Longrightarrow}} v'\gamma w'$ for some v' and w' in $\mathscr{S}^*$. By Lemma 2.3, it suffices to show that F_c' is minimal.

Let L be in $\mathscr{L}(F_e')=\mathscr{L}(F_d)=\mathscr{L}(F')$ and n be an arbitrary positive integer. Let k be 2 plus the maximum number of variables on the right side of any production in F'. Then there exists a grammar G' in $\mathscr{G}(F')$ such that $L(G)=L$ and $\Phi_{G'}(x)\leqq\max\{\phi_{G_{F'}}, |x|/(kn)\}$. In an obvious way there exists a grammar G_e' in $\mathscr{G}(F_e')$ such that $L(G_e')=L(G')$ and for each G'-derivation δ' of a word x in L there corresponds a G_e'-derivation δ_e' of x in which each step of δ' is simulated by at most k steps in δ_e'. (Specifically, each production p, corresponding to a production in F', may be replaced by one production corresponding to a production in F_d, plus one production corresponding to a production of the form $\beta\to v\gamma w$ for every variable in p.) Thus $L(G_e')=L$ and $\Phi_{G_e'}(x)\leqq\max\{k\phi_{G_{F'}}, |x|/n\}$ for all x in L. This implies the minimality of F_e'.

Combining (a)–(e), we obtain our result.

THEOREM 3.2. *Every grammar form is minimal.*

Proof. Clearly each vacuous grammar form is minimal. Consider nonvacuous grammar forms. By Lemma 3.1, we may restrict our attention to completely reduced, sequential grammar forms. The proof will be by induction on the number k of variables in the grammar form.

Suppose F is a completely reduced, sequential grammar form with just one variable. By Lemma 5.1 of [2], $\mathscr{L}(F)$ is either the family of finite, regular, linear, or context-free languages. By Proposition 2.4, F is minimal.

Now assume the theorem is true for all completely reduced, sequential grammar forms with at most k variables. Let $F=(V,\Sigma,\mathscr{V},\mathscr{S},\mathscr{P},\sigma)$ be a completely reduced, sequential grammar form with $k+1$ variables. Thus, the variables in $\mathscr{V}$ can be arranged into a sequence $\sigma=\alpha_0,\cdots,\alpha_k$ so that for each

production $\alpha_i \to u\alpha_j v$ in $\mathscr{P}$, $i \leqq j$. If $\mathscr{L}(F)$ is the family of context-free languages, then by Proposition 2.4 we are through. Thus assume $\mathscr{L}(F)$ is not the family of all context-free languages. We may assume similarly that F is nontrivial[14] (since otherwise F is either vacuous or generates the family of all finite sets.) By Lemma 2.3 we may assume the following:

1. For each i, $0 \leqq i \leqq k$, there exists v_i in $\mathscr{S}^+$ such that $\alpha_i \to v_i$ is in $\mathscr{P}$.

2. If there exist w_1, w_2 in $(\mathscr{V}-\{\sigma\})^+$ such that $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} w_1\sigma w_2$, then there exist x_1, x_2 in $\mathscr{S}^+$ such that $\sigma \to x_1\sigma x_2$ is in $\mathscr{P}$.

3. If there exist w in $(\mathscr{V}-\{\sigma\})^+$ such that $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} w\sigma$ $(\sigma \to \sigma w)$ then there exists x in $\mathscr{S}^+$ such that $\sigma \to x\sigma$ $(\sigma \to \sigma x)$ is in $\mathscr{P}$.

Intuitively, we proceed as follows. Consider the collection of "component" grammar forms arising from F by treating each variable of F except σ, in turn, as the start variable. Each such component form has at most k variables and thus, by induction, is minimal. Given any interpretation grammar G of F, the speed-up is accomplished by a grammar G' constructed as follows: Consider the collection of "components" of G, i.e., the parts of G which correspond to respective component forms of F. By the minimality of the component forms, each component of G may be sped up. The productions accomplishing this are then placed into G'. (See part 6 below.) In addition, productions which speed up short derivations are also placed into G'. (See part 4 below.) Similarly, productions which speed up the portion of G not part of any component of G (i.e., the portion involving variables corresponding to σ) are placed into G'. (See part 5 below.)

More formally, let $F_i = (V, \Sigma, \mathscr{V}_i, \mathscr{S}, \mathscr{P}_i, \alpha_i)$ be the grammar form in which $\mathscr{V}_i = \mathscr{S} \cup \{\alpha_j | \alpha_i \overset{*}{\underset{G_F}{\Longrightarrow}} w_1\alpha_j w_2$ for some w_1, w_2 in $\mathscr{V}^*\}$ and $\mathscr{P}_i$ be the set of all productions in F involving only variables in $\mathscr{V}_i$. Then F_i is nontrivial, completely reduced, sequential, and has at most k variables. For each i, $\phi_{G_{F_i}} = 1$ by assumption 1 above.

Let L be in $\mathscr{L}(F)$ and n be an arbitrary positive integer. There exists an interpretation (μ, G) of F such that $L(G) = L$. Let $G = (V_1, \Sigma_1, P, S)$. We shall modify G to obtain a new grammar G' obtaining the speedup of L by constant n. Let A be an arbitrary variable in $V_1 - (\Sigma_1 \cup \mu(\sigma))$. Then A is in $\mu(\alpha_i)$ for some $i \geqq 1$. Let $G'_A = (V'_A, \Sigma_1, P'_A, A)$, where $P'_A = P \cap \mu(\mathscr{P}_i)$ and V'_A is $\{A\} \cup \Sigma_1$ together with all the symbols appearing in productions of P'_A. Obviously G'_A is in $\mathscr{G}(F_i)$. By induction, there exists a grammar $G_A = (V_A, \Sigma_1, P_A, A')$ in $\mathscr{G}(F_i)$ such that $L(G_A) = L(G'_A)$ and $\Phi_{G_A}(x) \leqq \max\{1, |x|/(2(n+1))\} = \max\{\phi_{G_{F_i}}, |x|/(2(n+1))\}$ for all x in $L(G_A)$. There is no loss of generality in assuming that $A' = A$, A does not occur on the right side of any production in P_A, and each symbol in $V_A - (\Sigma_1 \cup \{A\})$ is a new symbol in $V - \Sigma$.

Let s be the number of elements in $\mu(\sigma)$ and r the maximum number of variables (not necessarily distinct) on the right side of any production in $\mathscr{P}$. Let $G' = (V'_1, \Sigma_1, P', S)$, where V'_1 is $\Sigma_1 \cup \{S\}$ together with all symbols in P', and P' consists of P and the following productions:

---

[14] A grammar form F is said to be *nontrivial* if $L(G_F)$ is infinite.

4. Foir every variable A in $V_1-\Sigma_1$ and w in Σ_1^*, with $|w|\leqq 2(n+1)(r+2)$ and $A \overset{*}{\underset{G}{\Longrightarrow}} w$, let $A \to w$ be in P'.

[Part 4 speeds up the G-derivation of short terminal words from variables.]

5. Suppose A, B are in $\mu(\sigma)$ and $A \overset{\leqq(n+1)s}{\underset{G}{\Longrightarrow}} w_1Bw_2$ for some w_1, w_2 in V_1^*. Let $B_1, \cdots, B_q$ be the variables (not necessarily distinct) of G appearing in w_1w_2 in order from left to right, and for each i suppose $B_i \overset{*}{\underset{G}{\Longrightarrow}} u_i$, where u_i is in Σ_1^* and $|u_i|\leqq 2(n+1)(r+2)$. Let $\bar{w}_1$ and $\bar{w}_2$ be the words obtained by replacing each B_i by u_i in w_1 and w_2 respectively. Then $A \to \bar{w}_1B\bar{w}_2$ is in P'.

[Part 5 speeds up G-derivations of the form $A \overset{*}{\underset{G}{\Longrightarrow}} w_1Bw_2 \overset{*}{\underset{G}{\Longrightarrow}} \bar{w}_1B\bar{w}_2$, A and B corresponding to σ, in which $A \overset{*}{\underset{G}{\Longrightarrow}} w_1Bw_2$ does not have too many steps and each variable of w_1w_2 has a G-derivation of a short terminal word.]

6. For each A in $V_1-(\Sigma_1 \cup \mu(\sigma))$ let each production of P_A be in P'.

[Part 6 speeds up G-derivations of long words from variables not corresponding to σ.]

It is readily seen that G' is in $\mathcal{G}(F)$ and that $L(G')=L$. To show that F is minimal, by Lemma 2.1 it suffices to show that $\Phi_{G'}(x)\leqq|x|/n$ for all words x in L such that $|x|/(n+1)+4+2r\leqq|x|/n$. [For the number of words x in L such that $|x|/(n+1)+4+2r>|x|/n$ is finite.] Thus consider any word x in L such that $|x|/(n+1)+4+2r\leqq|x|/n$. There exists a minimal G-derivation δ of x so that productions (possibly none) of the type $A \to vBw$, A, B in $\mu(\sigma)$ and v, w in $(V_1-\mu(\sigma))^*$ are applied first; then exactly one production of the type $A \to v$, A in $\mu(\sigma)$ and v in $(V_1-\mu(\sigma))^*$, and finally productions involving only variables in $V_1-(\Sigma_1 \cup \mu(\sigma))$. Three cases arise.[15]

(α) There are $(n+1)s$ consecutive productions $p_1, \cdots, p_{(n+1)s}$ in δ with a variable in $\mu(\sigma)$ on both sides such that for each variable A in $V-\mu(\sigma)$ generated by each of the p_i, the subword of x generated by A has length at most $2(n+1)(r+2)$. Then by part 5, there is a production π in G' which simulates the sequence $p_1, \cdots, p_{(n+1)s}$ as well as the derivations into subwords of x by every variable not in $\mu(\sigma)$ produced by each of the p_i. Since δ is minimal, each time a variable in $\mu(\sigma)$ is repeated, either a terminal symbol or a variable deriving (in δ) a terminal symbol is produced. Since there are only s distinct variables in $\mu(\sigma)$ and the production π simulates the effect of all the productions alluded to above, π deposits at least $n+1$ symbols.

(β) Fewer than $(n+1)s$ consecutive productions as in (α) occur, followed by a production p: $A \to vBw$ such that A, B are in $\mu(\sigma)$, v, w are in $(V_1-\mu(\sigma))^*$, and $|u|>2(n+1)(r+2)$ for some subword u of x generated by some variable in vw. As in (α), there is a production π of G' which simulates the sequence of productions preceding p, plus the derivation of variables not in $\mu(\sigma)$ into subwords of x. Let

[15] We implicitly use the fact that since $\mathcal{L}(F)$ is not the family of all context-free languages, it follows from Theorem 2.2 of [2] that there are no words u_1, u_2, u_3 in $\mathcal{V}^*$ such that $\sigma \overset{*}{\underset{G_F}{\Longrightarrow}} u_1\sigma u_2\sigma u_3$.

$B_1, \cdots, B_q$ be the occurrences of the variables in vw, in order, and let $x_1, \cdots, x_q$ be the corresponding subwords of x which the B_i generate. Note that $q \leqq r$. For each i such that $|x_i| \leqq 2(n+1)(r+2)$, there is a production $B_i \to x_i$ in P', by part 4. For each i such that $|x_i| > 2(n+1)(r+2)$, there is a sequence of at most $|x_i|/(2(n+1))$ productions in P' which converts B_i into x_i by part 6. Thus, the simulation of the sequence of productions, plus p, plus the derivation into subwords of x of all generated variables not in $\mu(\sigma)$, requires at most 1 (for the initial sequence of productions) $+1$ (for p) $+r$ (for expanding all B_i such that $|x_i| \leqq 2(n+1)(r+2)$) $+\sum_i^q |x_i|/(2(n+1))$ (for i such that $|x_i| > 2(n+1)(r+2)$) productions of G', i.e., at most $2+r+\sum_{i=1}^q |x_i|/(2(n+1))$ productions. Since $\sum_{i=1}^q |x_i| > 2(n+1)(r+2)$, the number of productions is at most $\sum_{i=1}^q |x_i|/(n+1)$. Since at least $\sum_{i=1}^q |x_i|$ terminals are deposited by these productions, it follows that at least $n+1$ terminal symbols are deposited for each production of G' used (although each production may not itself deposit $n+1$ terminals.)

(γ) Fewer than $(n+1)s$ consecutive productions as in (α) occur, followed by a production $p: A \to v$, where A is in $\mu(\sigma)$ and v is in $(V_1 - \mu(\sigma))^*$. As in (β), one production of G' simulates the sequence of productions preceding p, plus the derivation of the variables not in $\mu(\sigma)$ into subwords of x. Let $B_1, \cdots, B_q$ be the sequence of occurrences of variables, in order, and $x_1, \cdots, x_q$ the corresponding subwords of x. As in (β), the total number of productions needed to simulate the initial productions, plus p, plus the derivation of all generated variables into subwords of x, is at most $2+r+\sum_{i=1}^q |x_i|/(2(n+1))$. If $\sum_{i=1}^q |x_i| > 2(n+1)(r+2)$, then as in case ($\beta$), $n+1$ terminals are deposited for every production of G' used.

We now apply (α) and (β) to δ in the obvious way until (γ) arises. The number of applications of productions in G' is at most $\sum_{i=1}^q |x_i|/(n+1)$, where the x_i are the subwords of x derived from the variables not in $\mu(\sigma)$. Suppose that $\sum_{i=1}^q |x_i| > 2(n+1)(r+2)$ for the x_i arising in (γ). Then $\Phi_{G'}(x) \leqq |x|/(n+1) < |x|/n$. Suppose that $\sum_{i=1}^q |x_i| \leqq 2(n+1)(r+2)$ for the x_i arising in (γ). Then

$$\begin{aligned}
\Phi_{G'}(x) &\leqq \frac{|x|}{n+1} \quad \text{(from } (\alpha) \text{ and } (\beta)\text{)} \\
&\quad + 2 + r + \frac{\sum_{i=1}^q |x_i|}{2(n+1)} \quad \text{(for the } x_i \text{ arising in } (\gamma)\text{)} \\
&\leqq \frac{|x|}{n+1} + 2 + r + 2 + r \\
&= \frac{|x|}{n+1} + 4 + 2r \\
&\leqq \frac{|x|}{n}, \quad \text{by hypothesis on } x.
\end{aligned}$$

Hence the result.

The basic question we are interested in is whether some grammar forms are "more efficient" than others, either for representing particular languages or for their entire language families. By our main result, this question has a negative answer if our measure of complexity is derivation length. (That is, each grammar

form has the power of expressing each language in its grammatical family as efficiently as liked.) Of course, derivation length is not the only criterion for judging the efficiency of a grammar. Other possibilities are "size" measures (e.g., total number of symbols needed to represent the grammar or number of productions in the grammar), as studied, for example, in [7], [8]. The reader will note that the cost of the speedup using our construction is a large increase in the size of the grammar: if $S(n)$ is the size of the grammar constructed to accomplish speedup of a language by constant n, then $S(n)$ can be roughly equal to $S(1)^{kn}$, where k is a constant depending on the form and language. It remains to study comparative efficiency of forms with respect to size measures, and to examine trade-offs between the two types of measures.

Appendix. We now establish Lemma 2.5. Suppose there are at least $l/7$ leaf nodes m with the property that

(A.1) for some leaf $m' \neq m$, m and m' are daughters of the same father.

Then there are at least $l/14$ pairs of distinct leaf nodes, the two nodes in each pair having a common father. Thus there are at least $l/14$ such fathers, and since T has at least two internal nodes, $l/28$ fathers of such fathers. Each such father of a father is an internal node with at least one daughter an internal node. By (a) of the hypothesis, the sum of the weights below each such father of a father is at least k. Thus $\sum_{m \text{ a leaf}} \omega(m) > k(l/28)$.

Suppose there are $e < l/7$ leaf nodes m satisfying (A.1). Call an internal node both of whose daughters are internal nodes a *branch* node. Let i be the number of internal nodes and b the number of *branch* nodes. Since T is a binary tree, it is readily seen that $l = i + 1$ and $e/2 = b + 1$. Now remove all branch nodes and their incident edges from the tree T, obtaining a graph with g connected components, $\Gamma_1, \cdots, \Gamma_g$. Clearly

$$g \leqq 2b + 1 = e - 1 < \frac{l}{7}.$$

Let n be the number of original internal nodes in all the g components. Then

$$n = i - b = l - \frac{e}{2} > \frac{13}{14} l.$$

Also observe that each component is one of the following two types:

(A.2) For some $r \geqq 1$, the nodes are $\{n_i, m_i | 1 \leqq i \leqq r\} \cup \{m_r'\}$, where for each i, $1 \leqq i \leqq r$, m_i is a daughter of n_i, and for each i, $1 \leqq i \leqq r-1$, n_{i+1} is a daughter of n_i. Also, m_r' is a daughter of n_r. In addition, $n_1, \cdots, n_r$ are internal nodes of T but not branch nodes, and each m_i, m_r' is a leaf node of T.

(A.3) For some $r \geqq 1$, the nodes are $\{n_i, m_i | 1 \leqq i \leqq r\}$, where for each i, $1 \leqq i \leqq r$, m_i is a daughter of n_i, and for each i, $1 \leqq i \leqq r-1$, n_{i+1} is a daughter of n_i. In addition, $n_1, \cdots, n_r$ are internal nodes of T but not branch nodes, and each m_i is a leaf node of T.

Define a 6-chain as a 6-tuple $(n_1, \cdots, n_6)$ in which each n_i is an internal, nonbranch node of T, and n_j is a daughter of n_{j-1} for all $j \geqq 2$. All 6 nodes of a 6-chain are in some common component since no n_i is a branch node. For each i,

$1 \leqq i \leqq g$, let a_i be the number of original internal nodes in Γ_i. Then there is a set of [16]

$$\begin{aligned}\sum_{i=1}^{g} \lfloor a_i/6 \rfloor &\geqq \frac{1}{6}\sum_{1}^{g} a_i - \frac{5}{6}g \\ &= \frac{1}{6}(n-5g) \\ &> \frac{1}{6}\left(\frac{13}{14}l - \frac{5}{7}l\right) \\ &= l/28\end{aligned}$$

pairwise disjoint 6-chains (i.e., 6-chains having no elements in common). Since every internal node not a branch node has a daughter which is a leaf, it follows from (b) of the hypothesis that the sum of the weights of the leaf nodes which are daughters of nodes in a given 6-chain is at least k. Since there are at least $l/28$ disjoint 6-chains, the sum of the weights of the leaf nodes in T is at least $kl/28$, completing the proof of Lemma 2.5.

REFERENCES

[1] R. V. BOOK, *Time-bounded grammars and their languages*, J. Comput. System Sci., 5 (1971), pp. 397–429.

[2] A. B. CREMERS AND S. GINSBURG, *Context-free grammar forms*, Ibid., 11 (1975), pp. 86–117.

[3] A. B. CREMERS, S. GINSBURG AND E. H. SPANNER, *The structure of context-free grammatical families*, submitted.

[4] S. GINSBURG, *The Mathematical Theory of Context-Free Languages*, McGraw-Hill, New York, 1966.

[5] S. GINSBURG AND E. H. SPANIER, *Substitution of context-free grammar forms*, Acta Math., 5 (1975), pp. 377–386.

[6] A. GLADKII, *On the complexity of derivations in phase-structure grammars*, Algebra i Logika Sem., 3 (1964), pp. 29–44.

[7] J. GRUSKA, *On the size of context-free grammars*, Kybernetica (Prague), 8 (1972), pp. 213–218.

[8] A. R. MEYER AND M. J. FISCHER, *Economy of description by automata, grammars and formal systems*, Twelfth Annual Symp. on Switching and Automata Theory, 1971, pp. 188–191.

[9] S. SCHEINBERG, *Note on the Boolean properties of context-free languages*, Information and Control, 3 (1960), pp. 372–375.

[16] For each number t, $\lfloor t \rfloor$ is the largest integer less than or equal to t.