INFORMATION AND CONTROL 17, 92–108 (1970)

Reduction of Context-Free Grammars*

KENICHI TANIGUCHI AND TADAO KASAMI

Department of Control Engineering, Faculty of Engineering Science, Osaka University, Toyonaka, Japan

Received October 24, 1969; revised December 31, 1969

This paper is concerned with the following problem: Given a context-free grammar G, find a context-free grammar with the fewest nonterminal symbols or with the fewest rules that is equivalent to G. A reduction procedure is presented for finding such a reduced context-free grammar that is structurally equivalent to a given G. On the other hand, it is proved that there is no finite procedure for finding such a reduced context-free grammar that is weakly equivalent to a given G.

I. INTRODUCTION

This paper is concerned with the following problem: Given a context-free grammar (CFG) G, find a CFG with the fewest nonterminal symbols (NTS's) or with the fewest rules that is equivalent to G. Reduction of grammars has some practical significances. For example, the recognition and parsing algorithm for the language is less-time consuming, if a simplified grammar is used. Also, a simplified grammar often displays a predominant characteristic of the language.

Two CFG's are said to be weakly equivalent if they generate the same language. In Section IV, it is proved that there is no finite procedure for finding a CFG with the fewest NTS's or with the fewest rules that is weakly equivalent to a given G.

Two CFG's are said to be structurally equivalent if they not only generate the same sentences but they structure these sentences in the same manner. It is decidable whether two arbitrary CFG's are structurally equivalent (Fujii [1] and Paull [2]). Consequently, it is possible in principle to find

* This paper is based on the authors' previous report, Simplification of context-free grammars, A68-32, papers of Tech. Group on Automaton, IECE, Japan (in Japanese), September 1968.

a CFG with the fewest NTS's or with the fewest rules that is structurally equivalent to a given G. The problem is to find a procedure which yields the result with a reasonable amount of work. In Section III, some properties of such a reduced CFG that is structurally equivalent to a given G are investigated and a reduction procedure is presented, which is more efficient than the only existing procedure, namely the exhaustive search. In the case where each rule of a given CFG is of one of the two forms $X \to aY$ or $X \to a$ where X and Y are NTS's and a is a terminal symbol, the problem of finding a CFG with the fewest NTS's that is structurally equivalent to a given CFG reduces to the problem of finding a minimum state nondeterministic finite automaton that is equivalent to a given automaton. The reduction procedure presented here is a generalization of Kameda's procedure for reducing a nondeterministic finite automaton (Kameda [3]).

II. Preliminaries

The basic definitions and notations of the theory of context-free grammars and languages used in this paper are as in Ginsburg [4], unless stated otherwise.

A context-free grammar (briefly CFG) is a 4-tuple $G = (V_N, V_T, P, S)$, where V_N is the set of nonterminal symbols (briefly NTS's), V_T is the set of terminal symbols, P is the set of rules and $S \in V_N$ is the initial symbol of G.

For an NTS X of a CFG $G = (V_N, V_T, P, S)$, let

$$L(X; G) = \{w \mid X \overset{*}{\underset{G}{\Rightarrow}} w, w \in V_T{}^*\}.$$

The language generated by the CFG G, $L(G)$, is the set $L(S; G)$.

In the following, a CFG with one or more initial symbols is sometimes considered. In such cases, a CFG is denoted as $G = (V_N, V_T, P, V_S)$, where V_S is the set of initial symbols. The language generated by such a CFG G is defined by

$$L(G) = \bigcup_{S \in V_S} L(S; G).$$

For two CFG's G and G', if $L(G) = L(G')$, i.e., they generate the same language, then they are said to be *weakly equivalent*.

For two CFG's G and G', we write $G < G'$ if for any sentence w in $L(G)$ and for any generation tree for w in G there exists some generation tree for w in G' such that those two generation trees differ only in labelling the nodes (that is, they are geometrically identical). If $G < G'$ and $G' < G$, then they are said to be *structurally equivalent*. The following equivalent definition of structural equivalence is more convenient to work with.

Let G be a CFG with rules $\{X_i \rightarrow w_i \mid i = 1 \text{ to } n\}$. Then $[G]$, the *parenthesized version* of G is the CFG with rules $\{X_i \rightarrow [w_i] \mid i = 1 \text{ to } n\}$ where "["and"]" are special brackets that are not terminal symbols of G. Two CFG's G and G' are *structurally equivalent* if $L([G]) = L([G'])$. (The notation $G < G'$ means $L([G]) \subseteq L([G'])$. The brackets used in constructing $[G]$ and $[G']$ are the same and are not in either G or G'.) Structurally equivalent CFG's not only generate the same sentences but they structure these sentences in the same manner.

An NTS X of a CFG $G = (V_N, V_T, P, V_S)$ is said to be *useless* if for each S in V_S there is no ψ_1 and ψ_2 in $(V_N \cup V_T)^*$ such that $S \overset{*}{\underset{G}{\Rightarrow}} \psi_1 X \psi_2$, or if there is no w in $V_T{}^*$ such that $X \overset{*}{\underset{G}{\Rightarrow}} w$. It is well-known that there is a simple procedure to determine the useless NTS's. Any useless NTS and every rule containing that NTS can be discarded from the CFG without changing the language of the CFG.

III. Reduction Within the Structural Equivalence

In this section, a procedure is presented for finding a CFG with the fewest NTS's or with the fewest rules among the CFG's structurally equivallent to a given CFG. As usual, such a reduced CFG as well as a given CFG has a single initial symbol. However, by a minor modification, the procedure given in this paper is also applicable to the case where one or more initial symbols are admitted.

1. *Reduced Backwards-Deterministic Grammar and C Matrix*

Let $G = (V_N, V_T, P, V_S)$ be a CFG and let $[G]$ be the parenthesized version of G. For each NTS X in V_N, let

$$\bar{C}(X; [G]) = \{(w_1, w_2) \mid S \overset{*}{\underset{[G]}{\Rightarrow}} w_1 X w_2, \quad S \in V_S, \quad w_1, w_2 \in (V_T \cup \{[,]\})^*\}.$$

For a subset M of V_N, let

$$C(M; [G]) = \bigcap_{X \in M} \bar{C}(X; [G]) - \bigcup_{Y \notin M} \bar{C}(Y; [G]).$$

For any (w_1, w_2) in $C(M; [G])$, there is an $S \in V_S$ such that $S \overset{*}{\underset{[G]}{\Rightarrow}} w_1 X w_2$ when and only when X is in M. A subset M of V_N such that $C(M;[G]) \neq \phi$ is called a *distinguishable* set of $[G]$. The following proposition follows directly from the definitions.

PROPOSITION 1. *If* $M \neq M'$ $(M, M' \subseteq V_N)$, *then*

$$C(M; [G]) \cap C(M'; [G]) = \phi.$$

A CFG $\bar{G}$ with no two rules having the same right side is called a *backwards-deterministic* grammar (BDG). Two NTS's X and Y of a BDG $\bar{G}$ are *equivalent* if there is no distinguishable set of $[\bar{G}]$ having one of these without the other. (This is an equivalence relation.) A BDG is *reduced* if no two distinct NTS's are equivalent and if it has no useless NTS's (McNaughton [5]).

The first step of the reduction procedure presented here is to transform a given CFG $G = (V_N, V_T, P, S)$ into a reduced BDG $\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$ that is structurally equivalent to G. This step is essentially the same as the procedures of Theorems 1 and 4 of McNaughton [5]. For convenience, the procedure is presented in Appendix 1. In general, more than one initial symbols appear in $\bar{G}$. The following proposition holds for a BDG [5].

PROPOSITION 2. *If* $X \neq Y$ $(X, Y \in \bar{V}_N)$, *then* $L(X; [\bar{G}]) \cap L(Y; [\bar{G}]) = \phi$.[1]

$\bar{G}$ is the CFG with the fewest NTS's and with the fewest rules among the BDG's structurally equivalent to the given G. However, $\bar{G}$ is not necessarily the CFG with the fewest NTS's or with the fewest rules among the general CFG's structurally equivalent to the given G.

Denote the set of the distinguishable sets of $[\bar{G}]$ by $\mathscr{D}$.[2] (The procedure for finding the distinguishable sets is shown in Appendix 1, which is essentially the same as the one shown in Theorem 3 of McNaughton [5]. Construct the *C matrix* as follows: There is a column corresponding to each NTS of $\bar{G}$ and there is a row corresponding to each distinguishable set in $\mathscr{D}$. For simplicity, the row corresponding to $D \in \mathscr{D}$ and the column corresponding to $X \in \bar{V}_N$ are called the row D and the column X, respectively. The intersection of the row D and the column X is 1 if $X \in D$ and 0 if $X \notin D$.

The following proposition follows immediately from Propositions 1 and 2 and the definition of the C matrix.

[1] We use X, Y,... for NTS's of $\bar{G}$.

[2] It can be shown that the number of the distinguishable sets of $[\bar{G}]$ is at most 2^n, where n is the number of the NTS's of the CFG $G = (V_N, V_T, P, S)$ from which the BDG $\bar{G}$ was obtained.

Proposition 3. *If the intersection of the row D and the column X is* 1, *then for each* $(w_1, w_2) \in C(D; [\bar{G}])$ *and for each* $w \in L(X; [\bar{G}])$, w_1ww_2 *is in* $L([\bar{G}])$. *Furthermore, if the intersection of the row D and the column X is* 0, *then for any* $(w_1, w_2) \in C(D; [\bar{G}])$ *and for any* $w \in L(X; [\bar{G}])$, w_1ww_2 *is not in* $L([\bar{G}])$.

2. *Some Properties of Structural Equivalence*

In this subsection some necessary conditions are given for an arbitrary CFG to be a CFG with the fewest NTS's or with the fewest rules that is structurally equivalent to the given CFG. The structurally equivalent CFG with the fewest NTS's or with the fewest rules has only to be searched for among the CFG's that satisfy those conditions. This reduces the number of candidates to be tested.

It is said that a CFG G has a *redundant* NTS if, for some two NTS's A_1 and A_2 of G, the CFG obtained from G by replacing all the A_1 and A_2 appearing in G by a new NTS B (B is an initial symbol if at least one of A_1 and A_2 is an initial symbol of G) is structurally equivalent to G.

Let a CFG $G^1 = (V_N^1, V_T, P^1, S^1)$ be an *arbitrary* CFG with the fewest NTS's or with the fewest rules that is structurally equivalent to the given CFG G. Of course, G^1 is structurally equivalent to the BDG $\bar{G}$ obtained in Subsection 1. Evidently, we can assume that G^1 has neither useless NTS's nor redundant NTS's.

Let σ be defined by

$$\sigma(A) = \{X \mid L(A; [G^1]) \cap L(X; [\bar{G}]) \neq \phi, X \in \bar{V}_N\}^3$$

for an NTS A of G^1. From Proposition 2, the definition of σ and the fact that an NTS of $\bar{G}$ is not useless, it follows that $\sigma(S^1) = \bar{V}_S$. Furthermore, the following proposition follows from the definition of σ and the fact that an NTS A of G^1 is not useless.

Proposition 4. $L(A; [G^1]) \subseteq \bigcup_{X \in \sigma(A)} L(X; [\bar{G}])$.

For an NTS A of $[G^1]$, let

$$\tilde{C}(A; [G^1]) = \{(w_1, w_2) \mid \text{ if } w \in L(A; [G^1]), \text{ then } w_1ww_2 \in L([G^1]) = L([\bar{G}])\}.$$

Proposition 5. $\tilde{C}(A; [G^1]) \subseteq \bigcap_{X \in \sigma(A)} \bar{C}(X; [\bar{G}])$.

[3] We use $A, B, \ldots$ for NTS's of G^1.

Proof. Suppose the proposition does not hold. Then there is some (w_1, w_2) such that

(1) for any $w \in L(A; [G^1])$, $w_1ww_2 \in L([G^1]) = L([\bar{G}])$, and

(2) there is some X in $\sigma(A)$ such that for any $S \in \bar{V}_S$, w_1Xw_2 is not derivable from S in $[\bar{G}]$.

By the definition of σ, if X is in $\sigma(A)$, then there is a string w in $L(A; [G^1] \cap L(X; [\bar{G}])$. By Proposition 2, this w is not derivable from any other NTS in $[\bar{G}]$ other than X. Hence, w_1ww_2 is not in $L([\bar{G}])$ by (2). This contradicts (1). Q.E.D.

Now, for the mapping σ defined above, the following proposition holds. The proof is given in Appendix 2.

PROPOSITION 6. *Let A and B be two distinct NTS's of G^1. Then, $\sigma(A) \neq \sigma(B)$.*

Let $V_N{}^2$ be the set of $\sigma(A)$'s where A is an NTS of G^1, and denote $\sigma(S^1) = \bar{V}_S$ in $V_N{}^2$ by S^2. Let $G^2 = (V_N{}^2, V_T, P^2, S^2)$ be the CFG obtained from G^1 by replacing each NTS A of G^1 by its corresponding NTS $\sigma(A)$. From Proposition 6, it follows immediately that G^2 is structurally equivalent to G^1.

Consider a CFG $\tilde{G} = (\tilde{V}_N, V_T, \tilde{P}, \tilde{S})$ such that each NTS of $\tilde{G}$ is represented by a subset of the set of NTS's of the BDG $\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$. The CFG $\tilde{G}$ is said to have the *property SP* if for each rule of $\tilde{G}$, $M \to \alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}$, where $\alpha_j \in V_T{}^*$ $(1 \leqslant j \leqslant n+1)$,[4] $M \in \tilde{V}_N$, $M \subseteq \bar{V}_N$, $M_i \in \tilde{V}_N$, and $M_i \subseteq \bar{V}_N$ $(1 \leqslant i \leqslant n)$, it holds that for any X_1 in M_1, X_2 in M_2,..., and X_n in M_n, there is a rule in $\bar{G}$ whose right side is $\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ and whose left side is an NTS in M.[5]

The following proposition is important.

PROPOSITION 7. *The CFG G^2 has the property SP.*

Proof. Suppose that G^2 does not have the property *SP*. Then, there exists a rule[6] $N \to \alpha_1 N_1 \alpha_2 N_2 \cdots \alpha_n N_n \alpha_{n+1}$ in G^2 such that for some $X_1 \in N_1$, $X_2 \in N_2$,..., and $X_n \in N_n$ (1) there is no rule in $\bar{G}$ whose right side is $\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$, or (2) there is a rule in $\bar{G}$ whose right side is

[4] In the following, if we use α_1, α_2,..., and α_{n+1}, then it is assumed that they are in $V_T{}^*$.

[5] There is at most one rule in $\bar{G}$ whose right side is $\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$, since $\bar{G}$ is a BDG.

[6] We use N, M, ... for NTS's of G^2.

$\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ and whose left side, X, is not in N. By the definition of σ, there is a string w_i in $L(X_i\,; [\bar{G}]) \cap L(N_i\,; [G^2])$ for each i, since X_i is in N_i. Thus, there exists a derivation $N \overset{*}{\Rightarrow} [\alpha_1 w_1 \alpha_2 w_2 \cdots \alpha_n w_n \alpha_{n+1}]$ in $[G^2]$. In the case of (1), there is no NTS in $[\bar{G}]$ which derives $[\alpha_1 w_1 \alpha_2 w_2 \cdots \alpha_n w_n \alpha_{n+1}]$. Therefore, $\bar{G}$ and G^2 cannot be structurally equivalent, a contradiction. In the case of (2), there is a derivation $X \overset{*}{\Rightarrow} [\alpha_1 w_1 \alpha_2 w_2 \cdots \alpha_n w_n \alpha_{n+1}]$ in $[\bar{G}]$. Hence, $L(X; [\bar{G}]) \cap L(N; [G^2]) \neq \phi$. By the definition of σ, this implies that X is in N, which contradicts the assumption that X is not in N. Q.E.D.

We say that the rule of $\bar{G}$, $X \rightarrow \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ is *contained* in the rule (of $\tilde{G}$) $M \rightarrow \alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}$, if $X \in M$, and $X_i \in M_i$ for each i.

PROPOSITION 8. *For each rule of $\bar{G}$, there is at least one rule of G^2 which contains it.*

Proof. Consider the terminal string generated by a derivation in $[\bar{G}]$ which uses the rule $X \rightarrow [\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}]$ in the first place. By the structural equivalence this string must be generated by a derivation in $[G^2]$ which uses a rule of the form $N \rightarrow [\alpha_1 N_1 \alpha_2 N_2 \cdots \alpha_n N_n \alpha_{n+1}]$ in the first place. Suppose that the rule of G^2, $N \rightarrow [\alpha_1 N_1 \alpha_2 N_2 \cdots \alpha_n N_n \alpha_{n+1}]$, was used in the first place to generate that terminal string. Then, $X \in N$ and $X_i \in N_i$ for each i; for, if Y is not in M, $L(Y; [\bar{G}]) \cap L(M; [G^2]) = \phi$. Q.E.D.

So far it has been shown that each NTS of G^1 corresponds to a subset of the set of NTS's of $\bar{G}$ (Proposition 6) and that for the rules Propositions 7 and 8 hold. In the following, the relation between the NTS's of G^1 and the C matrix is considered. In the synthesis we can obtain from the C matrix some informations about the NTS's that are necessary in order to be structurally equivalent to $\bar{G}$.

For a subset M of $\bar{V}_N$, the set of NTS's of $\bar{G}$, let $f(M)$ be a set of 1's in the C matrix that lie at the intersections of a set of rows D's such that $D \supseteq M$ and a set of columns X's such that $X \in M$. We call such a set of 1's in the C matrix a *grid*. An NTS A of G^1 corresponds to a grid $f(\sigma(A))$. A set of grids forms a *cover*, if every 1 in the C matrix is a point of at least one grid in the set.

PROPOSITION 9. *The set that consists of all the grids $f(N)$ where N is an NTS of G^2 forms a cover.*

Proof. Suppose that there exist row D and column X such that the intersection of the row D and the column X is 1 that is not a point of grid

$f(N)$ for any NTS N of G^2. Hence, this 1 is not a point of grid $f(\sigma(A))$ for any NTS A of G^1. Thus by the definition of grid it does not hold that $\sigma(A) \subseteq D$ and $X \in \sigma(A)$ for any NTS A of G^1. By Proposition 3, for each $(w_1, w_2) \in C(D; [\bar{G}])$ and $w \in L(X; [\bar{G}])$, w_1ww_2 is in $L([\bar{G}])$. We show that w_1ww_2 cannot be in $L([G^1])$, contradicting the assumption that G^1 and $\bar{G}$ are structurally equivalent. In $[G^1]$, w can be derived from only an NTS A such that $\sigma(A) \ni X$ by the definition of σ. Therefore, in order to prove the proposition, it suffices to show that for such an A, w_1Aw_2 cannot be derived (from S^1) in $[G^1]$, that is, $\bar{C}(A; [G^1]) \cap C(D; [\bar{G}]) = \phi$. Since $\bar{C}(A; [G^1]) \subseteq \tilde{C}(A; [G^1])$, it follows from Proposition 5 that

$$\bar{C}(A; [G^1]) \subseteq \bigcap_{Y \in \sigma(A)} \bar{C}(Y; [\bar{G}]).$$

By the definition of grid, the right side $\bigcap_{Y \in \sigma(A)} \bar{C}(Y; [\bar{G}])$ coincides with the union of $C(D'; [\bar{G}])$ for row D' such that $\sigma(A) \subseteq D'$. Therefore, from Proposition 1 and the fact that $\sigma(A) \subseteq D$ does not hold, it follows that $\bar{C}(A; [G^1] \cap C(D; [\bar{G}]) = \phi$. Q.E.D.

3. *Reduction Procedure*

In the previous subsection it was shown that, in order to find a CFG with the fewest NTS's or with the fewest rules that is structurally equivalent to the given CFG G, we have only to consider CFG's that are constructed as follows: Choose a cover in the C matrix, let each grid correspond to an NTS and form a rule when it will possess the property *SP*. In particular, the condition of possessing property *SP* is very usefull to reduce the number of candidates.

Next, it will be shown that we have only to choose those covers which consist only of prime grids. A grid $f(M)$, where $M \subseteq \bar{V}_N$, is called a ***prime grid***, if it is not properly contained in any other grid $f(M')$, where $M' \subseteq \bar{V}_N$ and $M \neq M'$. For each grid $f(M)$ there exists a unique prime grid that contains it.

For a subset M of $\bar{V}_N$, let $p(M)$ be the subset of $\bar{V}_N$ which consists of NTS X such that the prime grid that contains the grid $f(M)$ contains at least one point on the column X. The following proposition follows from the definitions.

PROPOSITION 10. *$Y \in p(M)$ if, and only if, $\bar{C}(Y; [\bar{G}]) \supseteq \bigcap_{X \in M} \bar{C}(X; [\bar{G}])$.*

Let $V_N{}^3$ be the set of $p(N)$'s where N is an NTS of G^2 and let $G^3 = (V_N{}^3, V_T, P^3, S^3)$ be the CFG obtained from G^2 by replacing each

NTS N of G^2 by its corresponding NTS $p(N)$. The initial symbol of G^3, S^3, is $\bar{V}_S$ which is the initial symbol of G^2; for, since $\bar{V}_S \in \mathscr{D}$, $p(\bar{V}_S) = \bar{V}_S$. The number of distinct NTS's of G^3 is less than or equal to that of G^2, and the number of distinct rules of G^3 is less than or equal to that of G^2.

It is clear that for each rule of $\bar{G}$ there is at least one rule of G^3 which contains it and that the set which consists of all the grids $f(M)$ where M is an NTS of G^3 forms a cover.

In order to show that we have only to consider those covers which consist only of prime grids, it suffices to prove the following proposition.

PROPOSITION 11. *The CFG G^3 is structurally equivalent to $\bar{G}$. Furthermore, G^3 has the property SP.*

Proof. First, we show that G^3 has the property SP. Let

$$N \to \alpha_1 N_1 \alpha_2 N_2 \cdots \alpha_n N_n \alpha_{n+1}, \tag{1}$$

where $\alpha_j \in V_T^*$ $(1 \leqslant j \leqslant n+1)$, $N \in V_N{}^2$ and $N_i \in V_N{}^2$ $(1 \leqslant i \leqslant n)$ be a rule of G^2 and let

$$p(N) \to \alpha_1 p(N_1)\, \alpha_2 p(N_2) \cdots \alpha_n p(N_n)\, \alpha_{n+1}, \tag{2}$$

where $\alpha_j \in V_T^*$ $(1 \leqslant j \leqslant n+1)$, $p(N) \in V_N{}^3$ and $p(N_i) \in V_N{}^3$ $(1 \leqslant i \leqslant n)$ be the corresponding rule of G^3. By Proposition 7, rule (1) possesses the property SP.

We show by induction that rule (2) possesses the property SP. It is clear that the property SP holds for $p(N) \to \alpha_1 N_1 \alpha_2 N_2 \cdots \alpha_n N_n \alpha_{n+1}$, since $p(N) \supseteq N$. Suppose that the property SP holds for

$$p(N) \to \alpha_1 p(N_1) \cdots \alpha_{i-1} p(N_{i-1})\, \alpha_i N_i \alpha_{i+1} N_{i+1} \cdots \alpha_n N_n \alpha_{n+1}.$$

Consider

$$p(N) \to \alpha_1 p(N_1) \cdots \alpha_{i-1} p(N_{i-1})\, \alpha_i p(N_i)\, \alpha_{i+1} N_{i+1} \cdots \alpha_n N_n \alpha_{n+1}.$$

By Proposition 10, if $Y_i \in p(N_i) - N_i$, then $\bar{C}(Y_i\,; [\bar{G}]) \supseteq \bigcap_{X_i \in N_i} \bar{C}(X_i\,; [\bar{G}])$. Therefore, if there is a rule

$$X \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_{i-1} X_{i-1} \alpha_i X_i \alpha_{i+1} X_{i+1} \cdots \alpha_n X_n \alpha_{n+1}$$

in $\bar{G}$, where $X \in p(N)$, $X_j \in p(N_j)$ $(1 \leqslant j \leqslant i-1)$, $X_i \in N_i$, and $X_k \in N_k$ $(i+1 \leqslant k \leqslant n)$, then there must be a rule

$$Y \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_{i-1} X_{i-1} \alpha_i Y_i \alpha_{i+1} X_{i+1} \cdots \alpha_n X_n \alpha_{n+1}$$

in $\bar{G}$ for each $Y_i \in p(N_i) - N_i$. Furthermore, it must hold that

$$\bar{C}(Y; [\bar{G}]) \supseteq \bigcap_{X \in N'} \bar{C}(X; [\bar{G}]),$$

where $N' = \{X \mid X \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1} \in \bar{P}, X_j \in p(N_j) \ (1 \leqslant j \leqslant i-1),$ and $X_k \in N_k \ (i \leqslant k \leqslant n)\}$.

Since $N' \subseteq p(N)$, $\bigcap_{X \in N'} \bar{C}(X; [\bar{G}]) \supseteq \bigcap_{X \in p(N)} \bar{C}(X; [\bar{G}])$. Hence, $\bar{C}(Y; [\bar{G}]) \supseteq \bigcap_{X \in p(N)} \bar{C}(X; [\bar{G}])$. Therefore, by Proposition 10, $Y \in p(p(N)) = p(N)$. The above discussion verifies that the property SP holds for

$$p(N) \to \alpha_1 p(N_1) \cdots \alpha_{i-1} p(N_{i-1}) \, \alpha_i p(N_i) \, \alpha_{i+1} N_{i+1} \cdots \alpha_n N_n \alpha_{n+1}.$$

So we have shown that G^3 has the property SP.

Next, we show that G^3 is structurally equivalent to $\bar{G}$. In general, $G^2 < G^3$; for there is the case where $p(N) = p(N')$ for two distinct NTS's N and N' of G^2. Since G^3 has the property SP, as shown above, and the initial symbol of G^3 is $\bar{V}_S$, it follows from the following proposition that $G^3 < \bar{G}$. Therefore, G^3 is structurally equivalent to $\bar{G}$, since G^2 and $\bar{G}$ are structurally equivalent.

Q.E.D.

PROPOSITION 12. *Suppose that each NTS of a CFG $\tilde{G} = (\tilde{V}_N, V_T, \tilde{P}, \tilde{S})$ is represented by a subset of the set of NTS's of the BDG $\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$. If $\tilde{G}$ has the property SP and $\tilde{S}$ is $\bar{V}_S$, then $\tilde{G} < \bar{G}$.*

Proof. We first prove that for each NTS M of $[\tilde{G}]$,

$$L(M; [\tilde{G}]) \subseteq \bigcup_{X \in M} L(X; [\bar{G}]) \tag{3}$$

by induction on depth of the derivation. Depth of the derivation is defined as follows. For simplicity, denote $V_T \cup \{[,]\}$ by V_T'. The derivation $M \underset{[\tilde{G}]}{\overset{*}{\Rightarrow}} [w] \in V_T'^*$ has depth 1 if it corresponds to a single use of the terminating rule $M \to [w]$ in $[\tilde{G}]$. The derivation

$$M \underset{[\tilde{G}]}{\Rightarrow} [\alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}] \underset{[\tilde{G}]}{\overset{*}{\Rightarrow}} [\alpha_1 [w_1] \alpha_2 [w_2] \cdots \alpha_n [w_n] \alpha_{n+1}] \in V_T'^*$$

(which uses the rule $M \to [\alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}]$ in the first place) where $M_i \underset{[\tilde{G}]}{\overset{*}{\Rightarrow}} [w_i] \in V_T'^*$ for each i has depth $h+1$ if for each i, depth of the derivation $M_i \underset{[\tilde{G}]}{\overset{*}{\Rightarrow}} [w_i]$ is less than or equal to h and for at least one i, the derivation $M_i \underset{[\tilde{G}]}{\overset{*}{\Rightarrow}} [w_i]$ has depth h.

If the derivation $M \overset{*}{\underset{[\tilde{G}]}{\Rightarrow}} [w] \in V'^*_T$ has depth 1, then the rule $M \to [w]$ is in $[\tilde{G}]$. From the assumption that $\tilde{G}$ has the property SP, it follows that for some X in M, $X \to w$ is in $\bar{P}$, that is, $X \overset{*}{\underset{[\bar{G}]}{\Rightarrow}} [w]$.

Assume true that, if the derivation $M' \overset{*}{\underset{[\tilde{G}]}{\Rightarrow}} [w] \in V'^*_T$ has depth less than or equal to h, then for some X' in M', $X' \overset{*}{\underset{[\bar{G}]}{\Rightarrow}} [w]$. Suppose that there is a derivation

$$M \underset{[\tilde{G}]}{\Rightarrow} [\alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}] \overset{*}{\underset{[\tilde{G}]}{\Rightarrow}} [\alpha_1[w_1]\alpha_2[w_2] \cdots \alpha_n[w_n]\alpha_{n+1}] \in V'^*_T$$

of depth $h + 1$, where $M_i \overset{*}{\underset{[\tilde{G}]}{\Rightarrow}} [w_i]$ for each i. Then, the derivation $M_i \overset{*}{\underset{[\tilde{G}]}{\Rightarrow}} [w_i]$ has depth less than or epual to h for each i. By the inductive hypothesis, $X_i \overset{*}{\underset{[\bar{G}]}{\Rightarrow}} [w_i]$ for some X_i in M_i. Since $\tilde{G}$ has the property SP and

$$M \to \alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}$$

is in $\tilde{P}$, there is a rule $X \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ in $\bar{P}$ for some X in M. Thus for some X in M,

$$X \underset{[\bar{G}]}{\Rightarrow} [\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}] \overset{*}{\underset{[\bar{G}]}{\Rightarrow}} [\alpha_1[w_1]\alpha_2[w_2] \cdots \alpha_n[w_n]\alpha_{n+1}].$$

We have shown that (3) holds.

In particular, $L(\tilde{S}; [\tilde{G}]) \subseteq \bigcup_{X \in \bar{V}_S} L(X; [\bar{G}])$, since $\tilde{S}$ is $\bar{V}_S$ by the assumption. This is equivalent to $\tilde{G} < \bar{G}$. Q.E.D.

Now we are ready to describe our reduction procedure of CFG's. In the following we give a procedure for finding a CFG with the fewest NTS's that is structurally equivalent to a given CFG.

Procedure for Finding a Structurally Equivalent CFG with the Fewest NTS's

(i) Transform a given CFG G into a reduced BDG $\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$ that is structurally equivalent to G and construct the C matrix.

(ii) Let I be the number such that no set of less than I prime grids forms a cover and that at least one set of I prime grids forms a cover. Set $i = I$.

(iii) Choose the cover of i prime grids in the C matrix (which has not been chosen until now) and go to step (v). If all the covers of i prime grids have been chosen, then go to step (iv).

(iv) Set $i = i + 1$ and go to step (iii).

(v) For each prime grid chosen, let a set of NTS X of $\bar{G}$ such that the prime grid contains at least one point on the column X of the C matrix be an NTS. Form the rules by using these NTS's when they possess the property SP. Let $\bar{V}_S$ be the initial symbol. Denote the CFG constructed here by G^0.

(vi) If either G^0 has a useless NTS or there is a rule of $\bar{G}$ which is not contained in any rule of G^0, then go to step (iii). Otherwise, test whether $G < G^0$ (or, $\bar{G} < G^0$) or not (for the test procedure, see Appendix 3). If $G < G^0$, the process terminates. G^0 is a CFG with the fewest NTS's that is structurally equivalent to G.[7] If not, go to step (iii).

Taking into account the condition of structural equivalence in the test procedure (in Appendix 3), if we choose NTS's one by one in the above procedure, the search procedure might be more efficient. A procedure for finding a structurally equivalent CFG with the fewest rules is similar to the above procedure. So the procedure is not presented here.

IV. Reduction Within the Weak Equivalence

In this section we prove the following theorem.

THEOREM. *There is no finite procedure for finding a CFG with the fewest NTS's or with the fewest rules that is weakly equivalent to a given CFG.*

Before describing the proof, we show several lemmas.

LEMMA 1. *It is not decidable whether a given CFG generates $\{a, b\}^*c$.*

Proof. Follows from the well-known result that it is not decidable whether a given CFG generates $\{a, b\}^*$.

LEMMA 2. *A CFG with the fewest rules that generates $\{a, b\}^*c$ has three rules, $S \to aS$, $S \to bS$ and $S \to c$ (S is a initial symbol). Any CFG with three rules other than the above three rules cannot generate $\{a, b\}^*c$.*

Proof. It is clear that the CFG with the above three rules generates $\{a, b\}^*c$. Suppose a CFG $G = (V_N, \{a, b, c\}, P, S)$ with three rules generates $\{a, b\}^*c$. In order to generate sentences c, ac and bc, it is necessary to use the rules whose right side are $\alpha c\beta$, $\alpha' a\beta'$ and $\alpha'' b\beta''$ respectively, where α, β, α' and α'' are in V_N^* and β' and β'' are in $\{c\} \cup V_N^*$. Thus three rules are necessary to generate $\{a, b\}^*c$. Since G has only three rules, G cannot have the

[7] By Proposition 12, it is guaranteed that $G^0 < G$.

rule whose right side is ϵ (empty string). Therefore, $\alpha = \beta = \alpha' = \alpha'' = \epsilon$ and β' and β'' are in $\{c\} \cup \{\epsilon\} \cup V_N$ and the rule $S \rightarrow c$ must be in P. If $\beta'(\beta'')$ is c, then G cannot generate the sentences other than $ac(bc)$ which contain $a(b)$. If $\beta'(\beta'')$ is ϵ or an NTS other than S, then the sentence $ac(bc)$ cannot be generated. Therefore, $\beta' = \beta'' = S$. It is easily verified that if the NTS of the left side of the rules is the one other than S, then G cannot generate $\{a, b\}^*c$. Q.E.D.

LEMMA 3. *The language $\{a, b\}^*c$ can be generated by a CFG with only one NTS (initial symbol). Each rule of a CFG with only one NTS S that generates $\{a, b\}^*c$ is of one of the two forms $S \rightarrow \alpha S$ or $S \rightarrow \alpha' c$ where α and α' are in $\{a, b\}^*c$.*

Proof. The first part is easily verified (the CFG in Lemma 2 is an example). Suppose a CFG G with only one NTS S generates $\{a, b\}^*c$. G does not have the rule $S \rightarrow \epsilon$, since ϵ is not in $\{a, b\}^*c$. The rule that does not contain S in the right side must be of the form $S \rightarrow \alpha' c$, where α' is in $\{a, b\}^*$, and G must have at least one such a rule. In fact, G has the rule $S \rightarrow c$ in order to generate the sentence c. Therefore, the rule that contains S in the right side must be of the form $S \rightarrow \alpha S$, where α is in $\{a, b\}^*$. Q.E.D.

Now we prove the theorem.

Proof of the Theorem. Suppose that there exists a finite procedure for finding a CFG with the fewest NTS's or with the fewest rules that is weakly equivalent to a given CFG G. We show that if such a reduction procedure exists, it is decidable whether a given CFG G generates $\{a, b\}^*c$, which contradicts Lemma 1.

Assume that such a reduction procedure exists. By the reduction procedure, find a CFG G' with the fewest NTS's or G'' with the fewest rules that is weakly equivalent to a given G. If G'' has more than three rules or less than three rules, then G'' does not generate $\{a, b\}^*c$ by Lemma 2. If G'' has three rules, then it is decidable by Lemma 2 whether G'' generates $\{a, b\}^*c$. If G' has two or more NTS's, then G' does not generate $\{a, b\}^*c$ by Lemma 3. If G' has only one NTS and G' has a rule of the form other than the one in Lemma 3, then G' does not generate $\{a, b\}^*c$ by Lemma 3. If G' has only one NTS and each rule of G' is of the form in Lemma 3, it is decidable whether G' generates $\{a, b\}^*c$, since G' is a right-linear CFG and $\{a, b\}^*c$ is a regular set and it is decidable whether a right-linear CFG generates a particular regular set (see Ginsburg [4]). Thus in any case, it is decidable whether a given G generates $\{a, b\}^*c$. This completes the proof.

APPENDIX 1: A Procedure for Transforming a CFG $G = (V_N, V_T, P, S)$ into a Reduced BDG $\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$ that is Structurally Equivalent to G

First, transform a CFG $G = (V_N, V_T, P, S)$ into a structurally equivalent BDG $\bar{\bar{G}} = (\bar{\bar{V}}_N, V_T, \bar{\bar{P}}, \bar{\bar{V}}_S)$, which is not necessarily reduced. The procedure is as follows.

(i) Set $\tilde{V}_N = \phi$ and $\tilde{P} = \phi$.

(ii) For each distinct α in V_T^* such that for some X, $X \rightarrow \alpha$ is a rule in P, form a rule $M_\alpha \rightarrow \alpha$, where $M_\alpha = \{X \mid X \rightarrow \alpha \text{ is in } P\}$. Let $M_\alpha \rightarrow \alpha$ be in $\tilde{P}$ and M_α in $\tilde{V}_N$. Then, go to step (iii).

(iii) Choose a string $\eta = \alpha_1 M_1 \alpha_2 M_2 \cdots \alpha_n M_n \alpha_{n+1}$, where $\alpha_j \in V_T^*$ and $M_i \in \tilde{V}_N$ for each j and i such that there is no rule in $\tilde{P}$ whose right sides is η and that for some $X_1 \in M_1$, $X_2 \in M_2$,..., and $X_n \in M_n$, there is a rule in P whose right side is $\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$. Then, go to step (iv). If there is no such an η, then go to step (v).

(iv) Form a rule $M \rightarrow \eta$, where $M = \{X \mid X \rightarrow \alpha_1 X_1 \alpha_2 X_2 \ldots \alpha_n X_n \alpha_{n+1} \in P$ where $X_i \in M_i$ for each $i\}$. Let $M \rightarrow \eta$ be in $\tilde{P}$, and if M is not in $\tilde{V}_N$, let M be in $\tilde{V}_N$. Then, go to step (iii).

(v) Let $\bar{\bar{V}}_N = \tilde{V}_N$, $\bar{\bar{P}} = \tilde{P}$ and $\bar{\bar{V}}_S = \{M \mid S \in M, M \in \tilde{V}_N\}$. A CFG $\bar{\bar{G}} = (\bar{\bar{V}}_N, V_T, \bar{\bar{P}}, \bar{\bar{V}}_S)$ is a BDG that is structurally equivalent to G. $\bar{\bar{G}}$ may have a useless NTS which is not derivable from any initial symbol. From $\bar{\bar{G}}$, discard useless NTS's and rules containing those NTS's, if any.

Next, find the distinguishable sets of $[\bar{\bar{G}}]$. The procedure is as follows.

(i) Set $\mathscr{D} = \phi$.

(ii) Let $\bar{\bar{V}}_S$ be in $\mathscr{D}$.

(iii) Select any element M in $\mathscr{D}$ that is not marked off and go to step (iv). If every element in $\mathscr{D}$ has been marked off, the procedure terminates. $\mathscr{D}$ is the set of all the distinguishable sets of the BDG $[\bar{\bar{G}}]$.

(iv) For each (ψ_1, ψ_2) such that for some X in M, $X \rightarrow \psi_1 Y \psi_2$ is a rule in $\bar{\bar{P}}$, where $Y \in \bar{\bar{V}}_N$, find $M' = \{Y \mid X \rightarrow \psi_1 Y \psi_2 \in \bar{\bar{P}}, X \in M, Y \in \bar{\bar{V}}_N\}$ and if M' is not in $\mathscr{D}$, let M' be in $\mathscr{D}$. Then, mark off M and go to step (iii).

Then, by using $\mathscr{D}$, classify the set of NTS's of $\bar{\bar{G}}$ by an equivalence relation of NTS's. (An equivalence class of NTS's is a nonempty set having all, and only all, the NTS's equivalent to some given NTS.) Let the equivalent classes be, without repetition, E_1, E_2,..., and E_m. A reduced BDG

$\bar{G} = (\bar{V}_N, V_T, \bar{P}, \bar{V}_S)$ that is structurally equivalent to $\bar{\bar{G}}$ (consequently, structurally equivalent to the given G) is constructed as follows: Let $\bar{V}_N = \{E_1, E_2, ..., E_m\}$ and $\bar{V}_S = \{E_i \mid E_i \cap \bar{\bar{V}}_S \neq \phi\}$. For every rule of $\bar{\bar{G}}$, $X \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ where $\alpha_j \in V_T{}^*$ $(1 \leqslant j \leqslant n+1)$ and $X_i \in \bar{\bar{V}}_N$ $(1 \leqslant i \leqslant n)$, take for $\bar{G}$ the rule $E_{i_0} \to \alpha_1 E_{i_1} \alpha_2 E_{i_2} \cdots \alpha_n E_{i_n} \alpha_{n+1}$, where $X \in E_{i_0}$, $X_1 \in E_{i_1}$, ..., and $X_n \in E_{i_n}$.

APPENDIX 2: The Proof of Proposition 6

Let $\hat{G}^1$ be the CFG obtained from G^1 by replacing all the A and B in G^1 by a new NTS C (C is an initial symbol if at least one of A and B is an initial symbol of G^1). Since G^1 has no redundant NTS, $\hat{G}^1$ and G^1 are not structurally equivalent. In order to prove the proposition, it suffices to show that $\hat{G}^1$ and G^1 are structurally equivalent under the assumption that $\sigma(A) = \sigma(B)$. Since $G^1 < \hat{G}^1$ and $\bar{G}$ is structurally equivalent to G^1, it suffices to show that $\hat{G}^1 < \bar{G}$ under the assumption that $\sigma(A) = \sigma(B)$.

For each $w \in L(A; [G^1])$ or $L(B; [G^1])$, let $T(w)$ be a generation tree of w in $[G^1]$. Let

$$\mathscr{T}(A; [G^1]) = \{T(w) \mid w \in L(A; [G^1])\}$$

and

$$\mathscr{T}(B; [G^1]) = \{T(w) \mid w \in L(B; [G^1])\}.$$

Let $\hat{t}$ be a generation tree for a sentence in $[\hat{G}^1]$. By the definition of $[\hat{G}^1]$, $\hat{t}$ can be obtained from a generation tree t in $[G^1]$ by a finite number of relabelling the node or replacing the subtree as follows. Choose a node V with node name A or B such that there is no node with node name A or B on the path from the node V to the root of the tree under consideration.

(1) Change the node name of the node V to C, or

(2) Delete the subtree generated by the node V (for the terminology, refer to Ginsburg [4]), place an appropriate $\Delta t_B \in \mathscr{T}(B; [G^1])$ (or $\Delta t_A \in \mathscr{T}(A; [G^1])$) with its root on the node V if the node name of V is A (or B), and change the node name of V to C. Repeat the above step until the names A and B disappear. $\hat{t}$ is the tree obtained in such a manner.

Suppose that $\sigma(A) = \sigma(B) = N$. From Proposition 5, it follows that

$$\tilde{C}(A; [G^1]) \subseteq \bigcap_{X \in N} \bar{C}(X; [\bar{G}]) \qquad \text{A-1)}$$

and

$$\tilde{C}(B; [G^1]) \subseteq \bigcap_{X \in N} \tilde{C}(X; [\bar{G}]). \tag{A-2}$$

From Proposition 4, it follows that

$$L(A; [G^1]) \subseteq \bigcup_{X \in N} L(X; [\bar{G}]) \tag{A-3}$$

and

$$L(B; [G^1]) \subseteq \bigcup_{X \in N} L(X; [\bar{G}]). \tag{A-4}$$

Since for each $(w_1, w_2) \in \bigcap_{X \in N} \tilde{C}(X; [\bar{G}])$ and $w \in \bigcup_{X \in N} L(X; [\bar{G}])$, $w_1 w w_2$ is in $L([\bar{G}])$, it follows from (A-1) and (A-4) that for each $(w_1, w_2) \in \tilde{C}(A; [G^1])$ and $w \in L(B; [G^1])$, $w_1 w w_2$ is in $L([\bar{G}])$. Similary, it follows from (A-2) and (A-3) that for each $(w_1, w_2) \in \tilde{C}(B; [G^1])$ and $w \in L(A; [G^1])$, $w_1 w w_2$ is in $L([\bar{G}])$. This means that if in the transformation of the generation tree stated above a tree t' generates a sentence in $L([\bar{G}])$, then the tree t'' obtained from t' by the transformation (1) or (2) generates a sentence in $L([\bar{G}])$, too. Obviously, t generates a sentence in $L([\bar{G}])$, since $\bar{G}$ and G^1 are structurally equivalent. Therefore, $\hat{t}$ generates a sentence in $L([\bar{G}])$. Thus $\hat{G}^1 < \bar{G}$ under the assumption that $\sigma(A) = \sigma(B)$, completing the proof of the proposition.

APPENDIX 3: A Test Procedure Whether $G < G'$ for two CFG's $G = (V_N, V_T, P, V_S)$ and $G' = (V_N', V_T, P', V_S')$

Assume that G has no useless NTS's.

(i) Set $\tilde{V}_N = \phi$ and $\tilde{P} = \phi$.

(ii) For each terminating rule $X \to \alpha$, $\alpha \in V_T{}^*$ in P, let $M_\alpha = \{X' \mid X' \to \alpha$ is in $P'\}$. If $M_\alpha \neq \phi$, then let $(X, M_\alpha) \to \alpha$ be in $\tilde{P}$ and (X, M_α) in $\tilde{V}_N$, and go to step (iii). If $M_\alpha = \phi$ for some such an α, then it does not hold that $G < G'$.

(iii) Choose a string

$$\eta = \alpha_1(X_1, M_1)\, \alpha_2(X_2, M_2) \cdots \alpha_n(X_n, M_n)\, \alpha_{n+1},$$

where $\alpha_j \in V_T{}^*$ and $(X_i, M_i) \in \tilde{V}_N$ for each j and i such that there is no rule in $\tilde{P}$ whose right side is η and that there is at least one rule in P whose right side is $\alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$. Then, go to step (iv). If there is no such an η, then go to step (v).

(iv) Let

$$M = \{X' \mid X' \to \alpha_1 X_1{}' \alpha_2 X_2{}' \cdots \alpha_n X_n{}' \alpha_{n+1} \in P', \text{ where } X_i{}' \in M_i \text{ for each } i\}.$$

If $M = \phi$, then it does not hold that $G < G'$. Otherwise, for each rule $X \to \alpha_1 X_1 \alpha_2 X_2 \cdots \alpha_n X_n \alpha_{n+1}$ in P, form a rule $(X, M) \to \eta$. Let $(X, M) \to \eta$ be in $\tilde{P}$, and if (X, M) is not in $\tilde{V}_N$, let (X, M) be in $\tilde{V}_N$. Then, go to step(iii).

(v) If for each (X, M) in $\tilde{V}_N$ where X is in V_S, $M \cap V_S{}' \neq \phi$, then $G < G'$. Otherwise, it does not hold that $G < G'$.

References

1. M. Fujii and T. Kasami, "Some Structural Equivalence Relations and Well Transformation of Context-Free Grammars," papers of Tech. Group on Automaton, I.E.C.E., Japan, 1968 (in Japanese).
2. M. C. Paull and S. H. Unger, Structural equivalence of context-free grammars, *J. Comput. System Sci.* 2 (1968), 427–463.
3. T. Kameda and P. Weiner, On the reduction of non-deterministic finite automata, submitted for publication to *IEEE Trans. Comput.*
4. S. Ginsburg, "The Mathematical Theory of Context-Free Languages," McGraw-Hill Book Co., New York, 1966.
5. R. McNaughton, Parenthesis grammars, *J. Ass. Comput. Mach.* 14 (1967), 490–500.