INFORMATION AND CONTROL 14, 217–231 (1969)

Tree Generating Regular Systems

WALTER S. BRAINERD

Department of Mathematics, Naval Postgraduate School, Monterey, California 93940

Trees are defined as mappings from tree structures (in the graph-theoretic sense) into sets of symbols.

Regular systems are defined in which the production rules are of the form $\phi \to \psi$, where ϕ and ψ are trees. An application of a rule involves replacing a subtree ϕ by the tree ψ.

The main result is that the sets of trees generated by regular systems are exactly those that are accepted by tree automata. This generalizes a theorem of Büchi, proved for strings.

1. INTRODUCTION

In recent years, mathematical models of machines which operate on strings of symbols have been studied extensively. These models include the Turing machine and its many variations, such as the linear bounded automaton, automaton with push down storage and the finite automaton. For each of these machines, there is a string generating system or grammar which produces exactly the set of strings recognized by the machine. In this manner, the context-sensitive and context-free grammars of Chomsky correspond to linear bounded automata and finite automata with a single push down storage, respectively, and the regular systems of Büchi correspond to finite automata.

Büchi has placed the study of finite automata within the framework of more traditional mathematics by considering an automaton to be an abstract algebra, in which the operations are unary, and an output relation. With this point of view, it is natural to consider the case where operators may have arbitrary finite index. Doner, Thatcher, and Wright have investigated the properties of these algebras, considered as automata which accept trees which are formed, using Polish notation, from a finite alphabet of symbols, each of which may be thought of as an operator of fixed index. In this paper, we investigate a class of systems which generate exactly the sets of trees accepted by the automata.

2. TREES

We wish to adopt a formalism used by Doner (1967) and Gorn (1965) to define trees. It is interesting to note that Thue (1910) used similar ideas in a paper written almost 60 years ago. This method of presentation allows precise and simple definition of subtree, replacement of a subtree, etc.

DEFINITION 2.1. Let N^+ be the set of positive integers. Let U be the free monoid generated by N^+. Let $\cdot$ be the operation and 0 the identity of U. The *depth* of $a \in U$ is denoted $d(a)$ and defined as follows: $d(0) = 0$, $d(a \cdot i) = d(a) + 1$, $i \in N^+$. $a \leqq b$ iff there exists $x \in U$ such that $a \cdot x = b$. a and b are *incomparable* iff $a \nleqq b$ and $b \nleqq a$. Figure 1 illustrates the partial ordering on U, called the universal tree domain by Gorn.

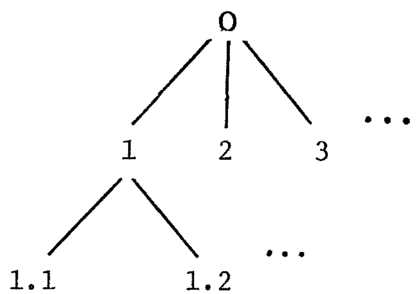


FIG. 1.

DEFINITION 2.2. D is a *tree domain* iff D is a finite subset of U satisfying (1) $b \in D$ and $a < b$ implies $a \in D$ and (2) $a \cdot j \in D$ and $i < j$ in N^+ implies $a \cdot i \in D$.

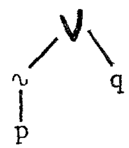
A tree (i.e., an ordered labelled tree or value tree) will be defined as a mapping from a tree domain (an ordered tree in the graph theoretic sense) into some set of symbols. However, only trees of a certain type will be considered; therefore, we first make the following definition:

DEFINITION 2.3. A *stratified alphabet*, Gorn (1965), is a pair $\langle A, \sigma \rangle$ where A is a finite set of symbols and $\sigma : A \to N$. Let $A_n = \sigma^{-1}(n)$.

DEFINITION 2.4. A *tree* over A (i.e., over $\langle A, \sigma \rangle$) is a function $\alpha : D \to A$ such that D is a tree domain and $\sigma[\alpha(a)] = \max \{i \mid a \cdot i \in D\}$. I.e., the stratification of a label at a must be equal to the number of branches in the tree domain at a. The domain of a tree is denoted $D(\alpha)$ or $D\alpha$.

Let A^T be the set of all trees over A. The *depth* of α, $d(\alpha) = \max \{d(a) \mid a \in D\alpha\}$.

EXAMPLE 2.5. Let $A = \{\vee, \sim, p, q\}$, $\sigma(\vee) = 2$, $\sigma(\sim) = 1$, $\sigma(p) = \sigma(q) = 0$. In this case A^T is the set of well formed formulas of propositional calculus involving one or two statement letters. The tree $\{(0, \vee), (1, \sim), (1.1, p), (2, q)\}$ would usually be written



For typographical reasons, trees will be represented using Polish postfix notation, *e.g.*, $p \sim q\vee$ for the tree above; however, extensive use will be made of the functional nature of trees in definitions and proofs.

It should be noted that trees are generalizations of strings. Let $A = \{\#, x_1, \cdots, x_k\}$ where $\sigma(\#) = 0$ and $\sigma(x_i) = 1$, $1 \leqq i \leqq n$. There is a natural correspondence between A^T and $(A - \{\#\})^*$ given by $\#\alpha \leftrightarrow \alpha$. The systems discussed in this paper are generalizations of Büchi's (1964) regular canonical systems, which generate strings.

DEFINITION 2.6. Let a, b, b' be members of U such that $b = a \cdot b'$. Then $b/a = b'$. b/a is not defined unless $a \leqq b$.

We have $b/0 = b$, $a/a = 0$ and $a \cdot (b/a) = (a \cdot b)/a = b$.

DEFINITION 2.7. Let $\alpha \in A^T$ and $a \in D\alpha$. $\alpha/a = \{(b, x) \mid (a \cdot b, x) \in \alpha\}$. α/a is the *subtree* of α at a and α/a *occurs* at a in α.

We have $\alpha/0 = \alpha$ and $\alpha \in A^T$, $a \in D\alpha$ implies $\alpha/a \in A^T$.

DEFINITION 2.8. Let $\alpha \in A^T$, $a \in U$. $a \cdot \alpha = \{(b, x) \mid (b/a, x) \in \alpha\}$.

We have $0 \cdot \alpha = \alpha$ and $a \cdot \alpha \in A^T$ *iff* $a = 0$. Using postfix notation, the tree $\bigcup_{i=1}^{n} i \cdot \alpha_i \cup \{(0, x)\}$ is represented by $\alpha_1\alpha_2 \cdots \alpha_n x$.

DEFINITION 2.9. Let $a \in D\alpha$, $\alpha, \beta \in A^T$.

$$\alpha(a \leftarrow \beta) = \{(b, x) \in \alpha \mid b \ngeqq a\} \cup a \cdot \alpha.$$

This is the result of replacing the subtree α/a at a by the tree β.

EXAMPLE 2.10. Let A and $\alpha = p \sim q\vee$ be as given in Example 2.5. Let $\beta = pq\vee$. Then $\alpha/1 = p\sim$ and $\alpha(1 \leftarrow \beta) = pq\vee q\vee$.

Note that the notation α/a allows one to reference any well-formed subformula in many formal systems.

The following results are immediate consequences of the above definitions. Proofs may be found in Brainerd (1967).

LEMMA 2.11. *If* $a \leqq b \in D\alpha$, *then* $\alpha/b = (\alpha/a)/(b/a)$.

LEMMA 2.12. *If* $a \leqq b$ *and* $\beta = \alpha(b \leftarrow \psi)$, *then* $\beta/a = (\alpha/a)(b/a \leftarrow \psi)$, i.e., $(\alpha(b \leftarrow \psi))/a = (\alpha/a)(b/a \leftarrow \psi)$.

LEMMA 2.13. $\gamma(b \leftarrow \alpha)/b \cdot a = \alpha/a$.

LEMMA 2.14. $\gamma(b \leftarrow \alpha(a \leftarrow \psi)) = \gamma(b \leftarrow a)(b \cdot a \leftarrow \psi)$.

LEMMA 2.15. *If* $\alpha, \beta, \gamma \in A^T$ *and if* a *and* b *are incomparable elements of* $D\gamma$, *then* $\gamma(a \leftarrow \alpha)(b \leftarrow \beta) = \gamma(b \leftarrow \beta)(a \leftarrow \alpha)$; i.e., *replacements at* a *and* b *may be made in either order with the same result.*

Lemma 2.15 is an example of the general principle "anything done to a subtree at a has no effect on the subtree at b, provided a and b are incomparable." Acceptance of this principle should make some of the rather complicated proofs somewhat easier to follow.

3. REGULAR SYSTEMS

In this section, a class of systems which generate trees is defined and investigated.

DEFINITION 3.1. A *regular system* over $\langle A, \sigma\rangle$ is a system $S = \langle B, \sigma', P, \Gamma\rangle$ satisfying

(a) $\langle B, \sigma'\rangle$ is a finite stratified alphabet such that $A \subseteq B$ and $\sigma' \mid A = \sigma$. The elements of A and B-A are called *terminal* and *nonterminal* symbols, respectively.

(b) P is a finite set of *production rules* of the form $\phi \rightarrow \psi$, where $\phi, \psi \in B^T$.

(c) $\Gamma \subseteq B^T$ is a finite set of *axioms.*

The following definition indicates how a regular system generates trees.

DEFINITION 3.2. $\alpha \xrightarrow{a} \beta$ is S iff there is a rule $\phi \rightarrow \psi$ in P such that $\alpha/a = \phi$ and $\beta = \alpha(a \leftarrow \psi)$. $\alpha \rightarrow \beta$ in S iff there exists $a \in D\alpha$ such that $\alpha \xrightarrow{a} \beta$. $\alpha \vdash \beta$ in S iff there exist $\alpha_0, \alpha_1, \cdots, \alpha_m$, $m \geqq 0$ such that $\alpha = \alpha_0 \rightarrow \alpha_1 \rightarrow \cdots \rightarrow \alpha_m = \beta$ in S. The sequence $\alpha_0, \alpha_1, \cdots, \alpha_m$ is called a *derivation* or *deduction* of β from α, and m is the *length* of the derivation.

Note, in particular, that α is a derivation of length zero of $\alpha \vdash \alpha$, for all $\alpha \in B^T$.

We want to think of a system over A as generating or producing trees in A^T, i.e., trees containing only terminal symbols.

DEFINITION 3.3. If S is a regular system over A, then $T(S) = \{\alpha \in A^T \mid \exists \gamma \in \Gamma \ni \gamma \vdash \alpha \text{ in } S\}$ is the set of trees generated by S. Σ is a *regular set* iff $\Sigma = T(S)$ for some regular system S. Systems S and S' are *equivalent* iff $T(S) = T(S')$.

EXAMPLE 3.4. Let $\langle A, \sigma\rangle$ be as in Example 2.5. Let $S = \langle B, \sigma', P, \Gamma\rangle$ where $B = A \cup \{X\}$, $\sigma'(X) = 0$, $P = \{X \to Xp\mathsf{V}, X \to p\}$ and $\Gamma = \{X\}$. It may be verified that $T(S) = \{p^{n+1}\mathsf{V}^n \mid n \geqq 0\}$.

Notice that if the trees of this example are expressed in postfix (or prefix) form, S becomes a context-free grammar, Chomsky (1959), Ginsburg (1966). However, the following example does not even correspond to a context-sensitive grammar, Chomsky (1959), Ginsburg (1966), because it contains rules in which the right hand side is shorter than the left hand side. A later theorem, however, shows that each regular system generates a set of trees which form a context-free set when written in postfix (or prefix) form.

EXAMPLE 3.5. Let $S = \langle A, \sigma, P, \Gamma\rangle$ where $\langle A, \sigma\rangle$ is as in example 2.5, $P = \{p \to p \sim q\mathsf{V}, p \sim \to p\sim\sim, p \sim\sim q\mathsf{V} \to pq\mathsf{V}\sim\}$ and $\Gamma = \{p\sim\}$. The following is a derivation of $p \sim q\mathsf{V}q\mathsf{V}\sim\sim$: $p \sim \to p \sim q\mathsf{V} \sim \to p \sim\sim q\mathsf{V} \sim \to pq\mathsf{V} \sim\sim \to p \sim q\mathsf{V}q\mathsf{V}\sim\sim$. This grammar has no nonterminal symbols. It will be shown (Theorem 4.8) that each regular set is generated by a system having no nonterminal symbols.

In the following three lemmas, let $S = \langle B, \sigma', P, \Gamma\rangle$ be a regular system over $\langle A, \sigma\rangle$

LEMMA 3.6. *If* $\alpha \xrightarrow{b} \beta$ *and* $a \leqq b$, *then* $\alpha/a \xrightarrow{b/a} \beta/a$.

Proof. $\alpha \xrightarrow{b} \beta$ implies there is a rule $\phi \to \psi$ in P such that $\alpha/b = \phi$ and $\beta = \alpha(b \leftarrow \psi)$. Hence $(\alpha/a)/(b/a) = \alpha/b = \psi$, by lemma 2.11 and $\beta/a = (\alpha/a)(b/a \leftarrow \psi)$, by Lemma 2.12, which implies $\alpha/a \xrightarrow{b/a} \beta/a$.

LEMMA 3.7. *If* $\alpha \to \beta$ *and* $b \in D\gamma$, *then* $\gamma(b \leftarrow \alpha) \to \gamma(b \leftarrow \beta)$.

Proof. If $\alpha \to \beta$, then there exist $a \in D\alpha$ and a rule $\phi \to \psi$ in P such that $\alpha/a = \phi$ and $\beta = \alpha(a \leftarrow \psi)$. Hence $\gamma(b \leftarrow \alpha)/b \cdot a = \alpha/a = \phi$ by Lemma 2.13 and $\gamma(b \leftarrow \beta)(b \cdot a \leftarrow \psi) = \gamma(b \leftarrow \alpha(a \leftarrow \psi))$ by Lemma 2.14. Thus $\gamma(b \leftarrow \alpha) \xrightarrow{b \cdot a} \gamma(b \leftarrow \beta)$.

COROLLARY 3.8. *If* $\alpha \vdash \beta$ *and* $b \in D\gamma$, *then* $\gamma(b \leftarrow \alpha) \vdash \gamma(b \leftarrow \beta)$.

LEMMA 3.9. *Suppose* $\alpha = \alpha_1\alpha_2 \cdots \alpha_n x, \beta = \beta_1\beta_2 \cdots \beta_n x, \sigma^{-1}(x) = n > 0$ *and* $\alpha = \gamma_0 \xrightarrow{a_1} \cdots \xrightarrow{a_m} \gamma_m = \beta$ *is a deduction of* $\alpha \vdash \beta$ *in which* $a_j \neq 0$, $1 \leqq j \leqq m$, i.e., *an entire tree is never replaced. Then* $\alpha_i \vdash \beta_i$, $1 \leqq i \leqq n$. *Conversely, if* $\alpha_i \vdash \beta_i$, $1 \leqq i \leqq n$, *then* $\alpha \vdash \beta$.

Proof. Assume $\alpha \vdash \beta$ as described above. The proof that $\alpha_i \vdash \beta_i$ is by induction on m, the length of the derivation.

(a) $m = 0$ implies $\alpha = \beta$ implies $\alpha_i = \beta_i$ implies $\alpha_i \vdash \beta_i$, for $1 \leqq i \leqq n$.

(b) Let $m > 0$ and assume the lemma holds for all derivations of length $m - 1$. Since $a_1 \neq 0$, $\gamma_1(0) = \gamma_0(0) = x$. Let $\gamma_1 = \alpha_1' \cdots \alpha_n' x$. Also, $a_1 \neq 0$ implies there exists $k \in N^+$ such that $a_1 = k \cdot a'$. (The first replacement is made in the kth branch of the tree.) Thus, for $i \neq k$, $\alpha_i' = \alpha_i$ and $\alpha_i \vdash \alpha_i'$. By Lemma 3.6, since $\gamma_0 \xrightarrow{a_1} \gamma_1$, $\gamma_0/k \xrightarrow{a_1/k} \gamma_1/k$, i.e., $\alpha_k' \xrightarrow{a_1} \alpha_k'$ and $\alpha_k \vdash \alpha_k'$. By the induction hypothesis, $\alpha_i' \vdash \beta_i$ for $1 \leqq i \leqq n$, thus $\alpha_i \vdash \beta_i$, since $\vdash$ is transitive.

For the converse, let $\gamma = y^n x$, where y is any member of B_0. Then $\alpha = \alpha_1 \cdots \alpha_n x = \gamma(1 \leftarrow \alpha_1) \cdots (n \leftarrow \alpha_n) \vdash \gamma(1 \leftarrow \beta_1) \cdots (n \leftarrow \beta_n) = \beta_1 \cdots \beta_n x$ by Corollary 3.8.

Our goal is to show that each regular set is accepted by a tree automaton. We proceed to show how, given a regular system, to construct equivalent systems, which are simpler in form.

LEMMA 3.10. *For every regular system* $S = \langle B, \sigma, P, \Gamma \rangle$, *one can effectively construct an equivalent system* $S' = \langle B', \sigma', P', \Gamma' \rangle$ *such that* Γ' *consists of a single nonterminal symbol*, i.e., *a tree of the form* $\{(0, Z)\}, Z \in B_0'$.

Proof. Let Z be any symbol not in B. Let $B' = B \cup \{Z\}$, $\sigma'(Z) = 0$, $P' = P \cup \{Z \to \gamma \mid \gamma \in \Gamma\}$ and $\Gamma' = \{Z\}$. Clearly $T(S') = T(S)$.

As a result of Lemma 3.10, it will usually be assumed that $\Gamma = \{Z\}$ and the system will be written $\langle B, \sigma, P, Z \rangle$.

DEFINITION 3.11. A system $\langle B, \sigma, P, Z \rangle$ is *simple* iff all rules of P are of the form $X_0 \to X_1 \cdots X_{\sigma(x)} x$, $X_0 \to X$, or $X_1 \cdots X_{\sigma(x)} x \to X_0$, where each $X_0, X_1, \cdots$ is a nonterminal symbol.

LEMMA 3.12. Given a system $S = \langle B, \sigma, P, Z \rangle$, one can effectively find a simple system $S' = \langle B', \sigma', P', Z \rangle$ which is equivalent to S.

Proof. Let $P = \{\phi_i \to \psi_i \mid 1 \leqq i \leqq r\}$. To construct S', introduce new symbols U_a^i and V_b^i for each $a \in D\phi_i$ and $b \in D\psi_i$. Let P^i consist of (1) rules that contract ϕ_i one level at a time, having the form $U^i_{a\cdot1} \cdots U^i_{a\cdot m}x \to U_a{}^i$, where $\phi_i(a) = x \in B_m$; (2) the rules $U_0{}^i \to V_0{}^i$; and 3) rules of the form $V_b{}^i \to V^i_{b\cdot1} \cdots V^i_{b\cdot n}y$ that expand $V_0{}^i$ to the tree ψ_i. In detail, let $B^i = \{U_a{}^i \mid a \in D\phi_i\} \cup \{V_b{}^i \mid b \in D\psi_i\}$ and $B' = \bigcup_{i=1}^r B' \cup B$. Let

$$P^i = \{U^i_{a\cdot1} \cdots U^i_{a\cdot m}x \to U_a{}^i \mid a \in D\phi_i, \phi_i(a) = x \in B_m\}$$
$$\cup \{U_0{}^i \to V_0{}^i\} \cup \{V_b{}^i \to V^i_{b\cdot1} \cdots V^i_{b\cdot n}y \mid b \in D\psi_i, \psi_i(b) = y \in B_n\}.$$

Let $P' = \bigcup_{i=1}^r P^i$.

The construction of S' is clearly effective. We must now show that $T(S') = T(S)$. First, note that $\phi_i \vdash \psi_i$ in S', for each rule $\phi_i \to \psi_i$ in P. The deduction is obtained by applying each rule of P^i in some appropriate order. Now suppose $\alpha \xrightarrow{a} \beta$ in S. Then for some rule $\phi_i \to \psi_i$ in S, $\alpha/a = \phi_i$ and $\beta = \alpha(a \leftarrow \psi)\alpha$. By the above argument $\alpha/a = \phi_i \vdash \psi_i = \beta/a$ in S', hence $\alpha = \alpha(a \leftarrow \alpha/a) \vdash \alpha(a \leftarrow \beta/a) = \alpha(a \leftarrow \psi_i) = \beta$ in S' by Corollary 3.8 and $T(S) \subseteq T(S')$.

For the converse, suppose $\alpha \in T(S')$, i.e., $Z \vdash \alpha \in A^T$ in S'. A deduction of $Z \vdash \alpha$ in S may be constructed as follows: Examine the deduction $Z \vdash \alpha$ in S'; each time a rule $U_0{}^i \to V_0{}^i$ is a applied at c, apply the rule $\phi_i \to \psi_i$ at c. The result will be derivation of $Z \vdash \alpha$ in S, because if the rule $U_0{}^i \to V_0{}^i$ can be applied at c, all contracting rules in P^i (i.e., those involving $U_a{}^i$, $a \in D\phi_i$) must have been applied previously at the corresponding addresses $c \cdot a$ and all expanding rules of P^i (i.e., those involving $V_b{}^i$, $b \in D\psi_i$) must be applied later at c.b, since all symbols $U_a{}^i$ and $V_b{}^i$ are nonterminal. This also means that if one rule of P^i is applied in a derivation, then all rules in P^i must be applied to eliminate the nonterminal symbols $U_a{}^i$ and $V_b{}^i$ and if all can be applied, then the single rule $\phi_i \to \psi_i$ in S is also applicable. It should be noted that in S', it is not necessary to apply all rules of some set P^i in sequence; other rules may be applied to independent subtrees (see Lemma 2.15).

LEMMA 3.13. *Given a simple system $S = \langle B, \sigma, P, Z\rangle$ over A and given $x \in A_n$, $X_0, X_1, \cdots, X_n \in B\text{-}A$, $X_0 \vdash X_1 \cdots X_n x$ in S in decidable.*

Proof. Suppose $X_0 = \alpha_0 \to \alpha_1 \to \cdots \to \alpha_m = X_1 \cdots X_n x$. Call $d = \max\{d(\alpha_i) \mid 0 \leqq i \leqq m\}$ the *depth* of the derivation, where $d(\alpha_i)$ is the depth of α_i (see Definitions 2.1 and 2.4). Let

$$M = \max\{n \mid \sigma^{-1}(n) \neq 0\}$$

and let K be the number of elements in B. It will be shown that if

$d > 1 + K^{M+2}$, then there is another derivation of $X_0 \vdash X_1 \cdots X_n x$ of length less than m. Suppose $d > 1 + K^{M+2}$ and let α_t be a tree in the derivation such that $d(\alpha_t) = d$. Let $k_1 \cdot k_2 \cdots k_d$, $k_j \in N^+$, $1 \leqq j \leqq d$ be an element of $D\alpha_t$ of maximal depth d and let $a_j = k_1 \cdots k_j$ so that $a_1 < a_2 < \cdots < a_d$. Let $\alpha_t(a_j) = u_j \in B$. For $1 \leqq j < d$, let $p_j = \min \{i \leqq t \mid i \leqq h < t \text{ implies } \alpha_h(a_j) = u_j\}$ and $q_j = \max \{i \geqq t \mid t < h < i \text{ implies } \alpha_h(a_j) = u_j\}$. Note that if $p_j \leqq h < q_j$ then $\alpha_h(a_j) = \alpha_t(a_j) = u_j$. If $\sigma(u_m) = n_j$, then $n_j > 0$, $1 \leqq j < d$. Since S is simple, each step $\alpha_{p_{j-1}} \to \alpha_{p_j}$, $1 \leqq j < d$, involves the application of a rule of the form $U_0^j \to U_1^j \cdots U_{n_j}^j u_j$ at a_j and each step $\alpha_{q_{j-1}} \to \alpha_{q_j}$, $1 \leqq j < d$ involves the application of a rule of the form $V_1^j \cdots V_{n_j}^j v_j \to V_0^j$. (Here the superscript refers to the depth at which the rule is applied.)

For each j, $1 \leqq j < d$, write the sequence $s_j = \langle u_j, U_1^j, \cdots, U_{n_j}^j, V_0^j \rangle$. There are at most K^{M+2} such sequences, so there exist integers p and q, such that $s_p = s_q$ and $1 \leqq p < q \leqq d - 1$, since $d - 1 > K^{M+2}$. A shorter deduction of $X_0 \vdash X_1 \cdots X_n x$ can be constructed by eliminating from the given deduction all steps in which a rule is applied at a, where $a > a_p$ and $a \not> a_q$, and replacing all steps involving an application of a rule at $a > a_q$ by an application of the same rule at $a_p \cdot (a/a_q)$.

If the resulting deduction still has depth $d > 1 + K^{M+2}$, this process can be repeated, yielding a sequence of deductions with strictly decreasing lengths. Thus, after a finite number of steps, a deduction of depth $d \leqq 1 + K^{M+2}$ is obtained. There are only finitely many trees over B of depth $d < 1 + K^{M+2}$. Let ν be the number of such trees. If the deduction $X_0 = \alpha_0 \to \cdots \to \alpha_m = X_1 \cdots X_n x$ has length $m > \nu$, then $\alpha_p = \alpha_q$ for some $0 \leqq p < q \leqq m$ and there is a shorter deduction $X_0 = \alpha_0 \to \cdots \to \alpha_p \to \alpha_{q+1} \to \cdots \to \alpha_m = X_1 \cdots X_n x$. We have established that if $X_0 \vdash X_1 \cdots X_n x$, there is a deduction of depth $d \leqq 1 + K^{M+2}$ and length $m < \nu$. Since there are only finitely many such deductions and it may be effectively decided when the list is complete, Lemma 3.13 is established.

Definition 3.14. A system $S = \langle B, \sigma, P, Z \rangle$ over A is *expansive* iff each rule in P is of the form $X_0 \to X_1 \cdots X_n x$ where $x \in A_n$ and $X_0, X_1, \cdots, X_n \in B\text{-}A$.

Lemma 3.15. *For each simple system $S = \langle B, \sigma, P, Z \rangle$, one can effectively construct an expansive system $S' = \langle B, \sigma, P', Z \rangle$ such that $T(S') = T(S)$.*

Proof. Let $P' = \{X_0 \to X_1 \cdots X_n x \mid X_0 \vdash X_1 \cdots X_n x \text{ in } S, x \in A_n, X_0, X_1, \cdots, X_n \in B\text{-}A\}$.

Each derivation in S' can obviously by duplicated in S, thus $T(S') \subseteq T(S)$.

To prove $T(S) \subseteq T(S')$, we first prove that if $X \vdash \alpha \in A^T$ in S, then $X \vdash \alpha$ in S', for any $X \in B\text{-}A$. The proof is by induction on the depth of α.

(a) $d(\alpha) = 0$ implies $\alpha = x \in A_0$ implies $X \to x \in P'$ implies $X \vdash x = \alpha$ in S'.

(b) Assume $d(\alpha) > 0$ and that if $d(\alpha') < d(\alpha)$ and $X \vdash \alpha'$ in S, then $X \vdash \alpha'$ in S'. Let $\alpha = \alpha_1\alpha_2 \cdots \alpha_n x$, $x \in A_n$, $n > 0$. Let $X = \beta_0 \xrightarrow{a_1} \beta_1 \xrightarrow{a_2} \cdots \xrightarrow{a_m} \beta_m = \alpha$ be a deduction of $X \vdash \alpha$ in S. Let $p = \max\{i \mid \beta_i(0) \neq \alpha(0)\} + 1$, so that $\beta_{p-1}(0) \neq \beta_p(0) = \beta_{p+1}(0) = \cdots = \alpha(0)$. Thus $\beta_{p-1} = X_0$ and $\beta_p = X_1 \cdots X_n x$ for some $X_0, X_1, \cdots, X_n$ in $B\text{-}A$, since S is simple. Since $X \vdash \beta_p = X_1 \cdots X_n x$ in S, $X \to X_1 \cdots X_n x$ is a rule in P' and so $X \vdash X_1 \cdots X_n x = \beta_p$ in S'.

Since $\beta_j(0) = x$ for $p \leqq j \leqq m$, $a_j \neq 0$ for $p < j$, hence $X_i \vdash \alpha_i$ in S, $1 \leqq i \leqq n$, by one part of Lemma 3.9. Hence $X_i \vdash \alpha_i$ in S' by the induction hypothesis, since $d(\alpha_i) < d(\alpha)$, $1 \leqq i \leqq n$. By the other part of Lemma 3.9, $\beta_p = X_1 \cdots X_n x \vdash \alpha_1 \cdots \alpha_n x = \alpha$ in S', hence $X \vdash \beta_p \vdash \alpha$ in S'.

If $\alpha \in T(S)$, then $\alpha \in A^T$ and $Z \vdash \alpha$ in S. Hence $Z \vdash \alpha$ in S' and $\alpha \in T(S')$.

P' can be effectively constructed from S by Lemma 3.13 and the proof of Lemma 3.15 is complete.

The results of Lemmas 3.10, 3.12, and 3.15 are summarized by the following theorem.

THEOREM 3.16. *For each regular system, one can effectively construct an equivalent expansive system* (i.e., *one which has a single nonterminal axiom and whose rules are all of the form* $X_0 \to X_1 \cdots X_n x$, $x \in A_n$ *and* $X_0, X_1, \cdots, X_n \in B\text{-}A$).

COROLLARY 3.17. If S is a regular system over A, then the set of strings in A^* obtained by writing the members of $T(S)$ in postfix (or prefix) form is a context-free language.

Proof. Observe that the expansive system guaranteed by Theorem 3.16 is a context-free grammar when the trees are written in postfix (or prefix) form.

This result is of special interest when one observes that the rules of a regular system, when written in postfix (or prefix) form are semi-Thue rules. The rules may be quite general in the sense that they do not even need to be context-sensitive rules, as was illustrated in Example 3.5. There is, instead, the important restriction that both sides of a rule must be trees expressed as strings. Although this may be a fairly natural restriction in some respects, it severely restricts the class of sets generated, because, although semi-Thue sets generate all recursively enumerable sets, Davis (1958), the following restatement of corollary 3.17 shows that if both sides of each rule of a semi-Thue system can be interpreted as trees, then only a context-free set is generated.

THEOREM 3.18. *Let* $S = \langle B, P, \Gamma \rangle$ *be a semi-Thue system over the alphabet* $A \subseteq B$. *Let* $T(S)$ *be the set of strings in* A^* *derivable in* S *from one of the axioms* Γ. *Suppose* $P = \{\phi_i \to \psi_i \mid 1 \leqq i \leqq r\}$. *If there is a function* $\sigma: B \to N$ *such that* ϕ_i, ψ_i, $1 \leqq i \leqq r$ *and the members of* Γ *are trees over* $\langle A, \sigma \rangle$ *written in postfix (or prefix) form,* $T(S)$ *is a context-free set.*

Another close connection between regular sets of trees and context-free sets is exhibited by the following result proved by Mezei and Wright (1965), who gave credit to D. Muller. The proof given here provides an alternate approach using regular systems.

DEFINITION 3.19. Define a mapping $h: A^T \to A_0^*$ as follows:

$$h(\alpha) = x \quad \text{if} \quad \alpha = x \in A_0$$

$$h(\alpha_1 \cdots \alpha_n x) = h(\alpha_1) \cdots h(\alpha_n) \quad \text{if} \quad x \in A_n, \qquad n > 0.$$

The function h forms a string in A_0^* obtained from a tree α by writing in order the images (labels) of all end points of α.

THEOREM 3.20. *If* Σ *is a regular set of trees, then* $h(\Sigma)$ *is a context-free set. Conversely, if* L *is a context-free set, then there is a regular set of trees* Σ, *such that* $L = h(\Sigma)$.

Proof. By Theorem 3.16, if Σ is a regular set, there is an expansive system $S = \langle B, \sigma, P, Z \rangle$ such that $\Sigma = T(S)$. Let

$$P' = \{X_0 \to X_1 \cdots X_n \mid x \in A_n, n > 0, X_0 \to X_1 \cdots X_n x \in P\}$$
$$\cup \{X \to x \mid X \to x \in P, x \in A_0\}.$$

Then if G is the context-free grammar $\langle B, P', Z \rangle$,

$$T(G) = h(T(S)) = h(\Sigma).$$

For the converse, suppose L is generated by the context-free grammar

$G = \langle B, P, Z\rangle$. It may be assumed that all rules of G are of the form $X_0 \to X_1X_2$ or $X_0 \to x$, where $X_0, X_1, X_2 \in B\text{-}A$ and $x \in A_0$, Chomsky (1959). Let $A' = A \cup \{+\}$ and $B' = B \cup \{+\}$, where $+ \notin B$. Let $\sigma(x) = 0$, $x \in B$ and $\sigma(+) = 2$. Let

$$P' = \{X_0 \to X_1X_2 + \mid X_0 \to X_1X_2 \in P\} \cup \{X_0 \to x \mid X_0 \to x \in P\}.$$

Let $S = \langle B', \sigma, P', Z\rangle$. Then if $\Sigma = T(S)$, $L = T(G) = h(\Sigma)$.

This section is concluded with the following theorem.

THEOREM 3.21. Let $S = \langle B, \sigma, P, \Gamma\rangle$ satisfy all conditions for a regular system except possibly the one that requires Γ to be finite. As before, define $T(S) = \{\alpha \in A^T \mid \exists\gamma \in \Gamma \ni \gamma \vdash \alpha \text{ in } S\}$. If Γ is a regular set in A^T, then so is $T(S)$.

Proof. If Γ is regular, then by theorem 3.16, $\Gamma = T(S')$ for some expansive system $S' = \langle B', \sigma', P', S'\rangle$ over A. By changing symbols in $B'\text{-}A$ if necessary, it may be assumed that $(B\text{-}A) \cap (B'\text{-}A) = \varnothing$, since $\Gamma \subseteq A^T$. Let $S'' = \langle B'', \sigma'', P'', Z\rangle$, where $B'' = B \cup B'$, $\sigma'' = \sigma \cup \sigma'$ and $P'' = P \cup P'$. (σ'' is a function because $(B\text{-}A) \cap (B'\text{-}A) = \varnothing$ and $\sigma \mid A = \sigma' \mid A$ by Definition 3.1(a).) We now show that $T(S') = T(S)$.

First suppose $\alpha \in T(S)$. Then there exists $\gamma \in \Gamma$ such that $\gamma \vdash \alpha$ in S, which implies that $\gamma \vdash \alpha$ in G'', since $P \subseteq P''$. On the other hand, $\gamma \in \Gamma$ implies $Z \vdash \gamma$ in S', hence $Z \vdash \gamma$ in G'', since $P' \subseteq P''$. Thus $Z \vdash \gamma \vdash \alpha$ in G'' and $\alpha \in T(S'')$.

Suppose $\alpha \in T(S'')$. To prove that $\alpha \in T(S)$, we show that any derivation $Z \vdash \alpha$ in S'' can be arranged so that all rules of P' which are used in the derivation are applied first, yielding an element $\gamma \in \Gamma$, after which all rules of P which are used in the derivation are applied. This will show that $\gamma \in \Gamma$ and $\gamma \vdash \alpha$ in S, hence $\alpha \in T(S)$.

Suppose in the deduction $Z \to \cdots \to \alpha_{k-1} \to \alpha_k \to \alpha_{k+1} \to \cdots \to \alpha$ in S'', $\alpha_{k-1} \xrightarrow{a_k} \alpha_k$ by an application of a rule $\phi \to \psi$ in P at a_k and $\alpha_k \xrightarrow{a_{k+1}} \alpha_k$ by an application of a rule in P' at a_{k+1}. Since all rules of P' are expansive $\alpha_k(a_{k+1}) = X_0$ for some $X_0 \in B\text{-}A$, which implies that $a_{k+1} \ngeq a_k$, because $a_{k+1} \geq a_k$ implies $X_0 = \alpha_k(a_{k+1}) = \psi(a_{k+1}/a_k) \in B$. On the other hand $a_k > a_{k+1}$ implies a_{k+1} is not an end point of α_k, which contradicts $\alpha_k(a_{k+1}) = X_0$, since $\sigma'(X_0) = 0$. Thus a_k and a_{k+1} are incomparrable and by Lemma 2.15, $Z \to \cdots \to \alpha_{k-1} \to \alpha_{k+1} \to \alpha_k \to \cdots \to \alpha$ is a derivation in S''. This process of interchanging steps in the derivation may be repeated until a derivation is obtained in which all rules of P'

used are applied before any rules of P are applied. This completes the proof by the comments made above.

4. THE EQUIVALENCE OF REGULAR SYSTEMS AND TREE AUTOMATA

The main purpose of this section is to show that the sets produced by regular systems are exactly those accepted by machines which are generalizations of ordinary finite automata.

DEFINITION 4.1. Let $\langle A, \sigma \rangle$ be a stratified alphabet, where $A = \{x_1 x_2, \cdots x_k\}$. A finite tree *automaton* over A is a system

$$M = \langle Q, t_1, \cdots, t_k, F \rangle$$

where

(a) Q is a finite set of *states*,

(b) For each i, $1 \leqq i \leqq k$, t_i is a relation on $Q^{\sigma(x_i)} \times Q$, and

(c) $F \subseteq Q$ is a set of *final* or *accept* states.

If each t_i is a function $t_i : Q^{\sigma(x_i)} \to Q$, then M is *deterministic*, otherwise M is *nondeterministic*, in which case we write $t_i(X_1, \cdots, X_n) \sim X_0$ iff $((X_1, \cdots, X_n), X_0) \in t_i$. If $\sigma(x_i) = 0$, we write $t_i \sim X$ iff $X \in t_i$.

We now indicate how each automaton accepts or rejects a member of A^T, and so defines a set of accepted trees.

Notation. If $x = x_i \in A$, then t_x means t_i.

DEFINITION 4.2. The *response relation* ρ of an automaton M is defined as follows:

(a) if $x \in A_0$, $\rho(x) \sim X$ iff $t_x \sim X$,

(b) if $x \in A_n$, $n > 0$, $\rho(\alpha_1 \cdots \alpha_n x) \sim X$ iff there exist $\exists\, X_1, \cdots, X_n \in Q$, $t_x(X_1, \cdots, X_n) \sim X$ and $\rho(\alpha_i) \sim X_i$, $1 \leqq i \leqq n$.

Note that if M is deterministic, ρ is a function $\rho: A^T \to Q$, characterized by the following:

(a) if $x \in A_0$, $\rho(x) = t_x$,

(b) if $x \in A_n$, $n > 0$, $\rho(\alpha_1 \cdots \alpha x_n) = t_x(\rho(\alpha_1), \cdots, \rho(\alpha_n))$.

DEFINITION 4.3. $T(M) = \{\alpha \in A^T \mid \exists\, X \in F \ni \rho(\alpha) \sim X\}$ is the set of trees accepted by M. M_1 and M_2 are *equivalent* iff $T(M_1) = T(M_2)$. A regular system S is equivalent to an automaton M iff $T(S) = T(M)$.

Note that if M is deterministic, $T(M) = \{\alpha \in A^T \mid \rho(\alpha) \in F\}$.

LEMMA 4.4. *For every expansive system $S = \langle B, \sigma, P, Z \rangle$ over A, one can effectively construct a nondeterministic automaton M, such that $T(M) = T(S)$.*

Proof. Let $M = \langle B - A, t_1, \cdots, t_k, \{Z\}\rangle$, where $t_x(X_1, \cdots, X_n) \sim X_0$ iff $X_0 \to X_1 \cdots X_n x$ is a rule of P. We first prove that $X \vdash \alpha$ in S iff $\rho(\alpha) \sim X$ in M by induction on the depth of α.

(a) $d(\alpha) = 0$ implies $\alpha = x \in A_0$, thus $X \vdash \alpha$ in S iff $X \to x$ in S, since S is expansive, iff $X \to x$ is a rule in P iff $t_x \sim X$, by definition of M, iff $\rho(x) \sim X$, by definition of ρ.

(b) Assume $\alpha = \alpha_1 \cdots \alpha_n x$, $x \in A_n$, $n > 0$ and that if $d(\alpha') < d(\alpha)$, then $X \vdash \alpha'$ in S iff $\rho(\alpha') \sim X$ in M. Then $X \vdash \alpha = \alpha_1 \cdots \alpha_n x$ in S iff $\exists X_1, \cdots, X_n \in B\text{-}A \ni X \to X_1 \cdots X_n x \vdash \alpha_1 \cdots \alpha_n x$ in S, since S is expansive, iff $X \to X_1 \cdots X_n x$ is a rule of P and $X_i \vdash \alpha_i$ in S, by Lemma 2.15, iff $t_x(X_1, \cdots, X_n) \sim X$, by the definition of M and $\rho(\alpha_i) \sim X_i$, by the induction hypothesis, iff $\rho(\alpha) \sim X$, by the definition of ρ.

Now suppose $\alpha \in A^T$, then $\alpha \in T(M)$ iff $\rho(\alpha) \sim a$ in M iff $Z \vdash \alpha$ in S iff $\alpha \in T(S)$, which proves the lemma.

Lemma 4.5. *For every nondeterministic automaton, one can effectively construct an equivalent deterministic automaton.*

Proof. This has already been shown by Doner (1967) and by Thatcher and Wright (1966). The proof is a direct adaptation of the subset construction used by Myhill for ordinary automata.

The results of Lemmas 3.10, 3.12, 3.15, 4.4, and 4.5 are summarized by the following theorem.

Theorem 4.6. *For every regular system, one can effectively find an equivalent deterministic automaton.*

It is likely that the application of the procedures given by Lemmas 3.10, 3.12, 3.5, 4.4, and 4.5 will yield an automaton with far more than the minimum number of states. An algorithm for constructing a minimal automaton is given in Brainerd (1968).

We now prove a result that is stronger than the converse of Theorem 4.6.

Lemma 4.7. *Let M be a deterministic automaton with q states. For each state $X \in Q$, if there is a tree α such that $\rho(\alpha) = X$, then there is a tree β such that $d(\beta) < q$ and $\rho(\beta) = X$.*

Proof. This result has been proved by Doner (1967) and by Thatcher and Wright (1966) in a manner analagous to the result for ordinary automata, Rabin and Scott (1959).

THEOREM 4.8. *Let* $M = \langle Q, f_1, \cdots, f_k, F\rangle$ *be a deterministic automaton over* $\langle A, \sigma\rangle$. *One can effectively construct a regular system* $S = \langle A, \sigma, P, \Gamma\rangle$, *such that* $T(S) = T(M)$. *That is, one can construct a system* S, *equivvlent to* M, *in which no nonterminal symbols are used.*

Proof. Suppose Q *contains* q *states. Let* $P = \{\phi \to \psi \mid d(\phi) < d(\psi) = q$ and $\rho(\phi) = \rho(\psi)$ in $M\}$. Let $\Gamma = \{\alpha \in T(M) \mid d(\alpha) < q\}$. P and Γ may be found effectively since there is only a finite number of trees of depth $d \leqq q$ and given M, $\rho(\phi)$ in M is computable.

To show $T(M) \subseteq T(S)$, we first prove the following result:

$$(^*)\quad \forall \alpha \in A^T,\quad \exists \beta \in A^T \ni d(\beta) < q,\quad \beta \vdash \alpha \text{ in } S, \quad\text{and}$$
$$\rho(\beta) = \rho(\alpha) \text{ in } M.$$

If $d(\alpha) < q$, pick $\beta = \alpha$. If $d(\alpha) \geqq q$, pick α' to be any subtree of α such that $d(\alpha') = q$. By Lemma 4.7, there is a β' such that $d(\beta') < d(\alpha') = q$ and $\rho(\beta') = \rho(\alpha')$, which means $\beta' \to \alpha'$ is a rule of P. Let β_1 be the result of replacing an occurrence of α' by β'. Then $\beta_1 \to \alpha$ in S, since $\beta' \to \alpha' \in P$. Also $\rho(\beta_1) = \rho(\alpha)$, since $\rho(\beta') = \rho(\alpha')$. If $d(\beta_1) \geqq q$, the process may be repeated, yielding $\beta_2, \beta_3, \cdots$, but since $d(\beta') < d(\alpha')$, the process must terminate after a finite number of steps with $d(\beta_j) < q$, $\beta = \beta_j \to \cdots \to \beta_1 \to \alpha$ in S and $\rho(\beta) = \rho(\beta_j) = \cdots = \rho(\beta_1) = \rho(\alpha)$, which proves $(^*)$.

Now suppose $\alpha \in T(M)$. Then $\exists \beta \in A^T \ni \beta \vdash \alpha$ in S, $\rho(\beta) = \rho(\alpha)$ and $d(\beta) < q$, by $(^*)$. Since $\rho(\alpha) \in F$, so is $\rho(\beta)$, which implies that $\beta \in \Gamma$, hence $\alpha \in T(S)$.

Conversely, if $\alpha \in T(S)$, then there is a tree $\beta \in \Gamma \ni \beta \vdash \alpha$ in S. $\rho(\alpha) = \rho(\beta) \in F$ by the definition of P and Γ, hence $\alpha \in T(M)$. This completes the proof of Theorem 4.8.

Theorems 4.6 and 4.8 yield the following result:

THEOREM 4.9. *The sets of trees generated by regular systems are exactly those accepted by finite tree automata.*

ACKNOWLEDGMENT

Discussions with J. R. Büchi, L. Landweber, W. Hosken and especially R. Korfhage were very helpful during the development of the results in this paper.

RECEIVED: October 21, 1968; revised January 6, 1969

REFERENCES

BRAINERD, W. S. (1967), Tree Generating Systems and Tree Automata, Doctoral Dissertation, Purdue University.

BRAINERD, W. S. (1968), The Minimalization of Tree Automata, *Inform. Control*, **13**, 484–491.

BÜCHI, J. R. (1964), Regular Canonical Systems, *Archiv. fur Mathematische Logik und Grundlagenforschung*, **6**, 91–111.

CHOMSKY, N. (1959), On Certain Formal Properties of Grammars, *Inform. Control*, **2**, 137–167.

DAVIS, M. (1958), "Computability and Unsolvability," McGraw-Hill, New York.

DONER, J. E. (1967), Tree Acceptors and Some of Their Applications, System Development Corporation, Scientific Report No. 8.

GINSBURG, S. (1966), "The Mathematical Theory of Context-Free Languages," McGraw-Hill, New York.

GORN, S. (1965), Explicit Definitions and Linguistic Dominoes, Presented at the Systems and Computer Science Conference, University of Western Ontario.

MEZEI, J. AND WRIGHT, J. B., (1965), "Generalized ALGOL-like Languages," IBM Research Paper RC-1528.

RABIN, M. O. AND SCOTT, D. (1959), Finite Automata and Their Decision Problems, *IBM J. Res. Develop.*, **3**, 114–125.

THATCHER, J. W. AND WRIGHT, J. B. (1966), "Generalized Finite Automata Theory with an Application to a Decision Problem of Second-order Logic," IBM Research Paper RC-1713.

THUE, A. (1910), Die Lösung eines Spezialfalles eines Generellen Logischen Problems, *Videnskabs-Selskabets Skrifter, I. Math.-Naturv. Klasse*, No. 8.