Reprinted from INFORMATION AND CONTROL, Volume 6, No. 3, September 1963


INFORMATION AND CONTROL 6, 246–264 (1963)

# On Context-Free Languages and Push-Down Automata

M. P. SCHÜTZENBERGER

*International Business Machines Corporation, Thomas J. Watson Research Center, Yorktown Heights, New York*

This note describes a special type of one-way, one-tape automata in the sense of Rabin and Scott that idealizes some of the elementary formal features used in the so-called "push-down store" programming techniques. It is verified that the sets of words accepted by these automata form a proper subset of the family of the *unambiguous context-free languages* of Chomsky's and that this property admits a weak converse.

## INTRODUCTION

This note is concerned with some relations observed by N. Chomsky and myself between *context-free languages* and what will be called *push-down automata*.[1]

Informally, a push-down automaton is a special type of *one-way, one-tape* automaton in the sense of Rabin and Scott in which the memory is a (potentially infinite) tape, used in a certain restricted manner. For each successive letter of the input word, the word stored on the tape is modified by deletion or adjunction at its right end. This is done under the control of a finite state device which can scan the word stored on the tape and carry some additional information. The input word is accepted iff after reading its last letter both the word written on the tape and the state of the finite part belong to a *finite* prescribed set.

This operation appears as an abstraction of some elementary features of the programming technique known as "*push-down store*" (Newell and Shaw, 1957).

I am indebted to C. C. Elgot for enlightening discussions which have lead to the clarification of many points and to a definition of push-down automata which may be less unrealistic than the ones I had previously considered.

[1] *Note added in proof:* Our definition does not coincide with the one introduced by Shepherdson and Sturgis (1963).

We recall that a context-free language on an alphabet $X$ is a subset $L$ of the set $F$ of all words in this alphabet that can be obtained by the following procedure, which is a special type of a Post production:

Let $\Xi = \{\xi_j\} (1 \leqq j \leqq n)$ be another set of letters and $H$ be the set of all words in the letters of $X \cup \Xi$. In Chomsky's terminology $\Xi$ is the *nonterminal* alphabet. A *grammar* is an assignment to each $\xi_j \in \Xi$ of a finite set $p_j$ of words of $H$ that does not contain the empty word $e$ or any word consisting of a single letter of $\Xi$. Both $X$ and $\Xi$ are finite.

Let $L_1$ be the least subset of $H$ that contains $\xi_1$ and every word $h'hh''$ $(h, h', h'' \in H)$ if it contains $h'\xi_j h''$ $(1 \leqq j \leqq n)$ and if $h \in p_j$. Then by definition $L = L_1 \cap F$ is the context-free language produced by the grammar $\{p_j\}$.

In the first part of the paper we verify the equivalence of this definition with another one which relates context-free languages with algebraic formal power series in noncommuting variables. The treatment given here is more elementary than that of Ginsburg and Rice (1962) and the notation introduced in this part is needed later on. Furthermore, except for trivial changes, the main bulk of the notation carries over to the still simpler case of the power series.

In the second part of the paper we verify that the set of words accepted by any push-down automaton (as defined here!) is a context-free language and, as an example, we consider the simplest nondegenerate type of such automata. The corresponding languages are called "*standard context-free languages.*"

In the fourth part we verify a weak converse of the first property. More explicitly, let $F$ and $F'$ be the set of all words in the letters of the alphabets $X$ and $X'$ (i.e., the free monoids generated by these sets). A homomorphism (of monoid) $\theta: F \rightarrow F'$ is any mapping from $F$ to $F'$ given by a mapping $\theta_1$ from $X$ to $F'$ and the rule that $\theta e = e$ ($e =$ the empty word) and that for all $f = x_{i_1} x_{i_2} \cdots x_{i_m} \in F$, $(m > 0)$, $\theta f = \theta_1 x_{i_1} \theta_1 x_{i_2} \cdots \theta_1 x_{i_m}$.

We verify that for any context-free language $L' \subset F'$ one can find a set $X$, a homomorphism $\theta: F \rightarrow F'$ and a standard context-free language $L \subset F$ such that $L' = \theta L$ $(= \{\theta f : f \in L'\})$.

A closer connection between push-down devices and context-free languages is to be found in (Chomsky, 1962).

## I. DEFINITIONS

Let $X$, $\Xi$, $F$, $H$ be as in the introduction. $\mathfrak{P}(H)$ is the collection of all subsets of $H$. We shall consider $\mathfrak{P}(H)$ as an algebraic system (in fact, a semiring) with respect to the usual operations of set union and set product (the intersection is not used).

We shall reserve the notation $\pi_m$ (resp. $\pi_F$) for the projection of $\mathfrak{P}(H)$ onto the union of the elements of $H$ of length $\leqq m$ (resp. onto $\mathfrak{P}(F)$). Thus, for $H' \in \mathfrak{P}(H)$, $\pi_0 H' = \phi$ iff $e \notin H'$ and we say then, as usual, that $H'$ is *quasi regular*.

For any $n$-tuple $q = (q_1, q_2, \cdots, q_n)$ of elements of $\mathfrak{P}(H)$, we denote by $\lambda_q$ the homomorphism (of semiring) of $\mathfrak{P}(H)$ induced by the substitution $\{\xi_j\} \rightarrow q_j$ for $1 \leqq j \leqq n$.

Now let $L$ be a (nonempty) context-free language with grammar $p = (p_1, p_2, \cdots, p_n)$ and nonterminal alphabet $\Xi$. We can assume that no $p_j$ is empty and we recall that by hypothesis (1) $p$ is quasi regular (i.e., $\pi_0 p_j = \phi$ for each $j$); (2) no $p_j$ contains a word consisting of a single nonterminal letter.

Let $\mathcal{P}$ (resp. $\mathcal{P}_F$) be the set of all quasi-regular $n$-tuples of subsets of $H$ (resp. of $F$). For any mapping $\phi$ of $\mathfrak{P}(H)$ and $r \in \mathcal{P}$, $\phi r$ denotes the $n$-tuple $(\phi r_1, \phi r_2, \cdots, \phi r_n)$.

Now let $q$ be any element of $\mathcal{P}$. The conditions (1) and (2) on $p$ imply the following identity or all $m \geqq 0$:

$$(*) \qquad \pi_{m+1}\lambda_q p = \pi_{m+1}\lambda_{\pi_m q} p.$$

Indeed, by the very definition of $\lambda$, for each $j$ and $m, m' \geqq 0$, $\lambda_{\pi_m q} p_j \subset \lambda_{\pi_{m+m'} q} p_j$ and so, with obvious notations $\pi_{m+1}\lambda_q p \supset \pi_{m+1}\lambda_{\pi_m q} p$.

On the other hand, any word $f$ of length $|f| \geqq m + 1$ belonging to some component of $q$ can only intervene in $\lambda_q p$ by being substituted for some $\xi \in \Xi$ which (by condition (2)) is a factor of a word of length at least two. Thus the word resulting from this substitution has length at least $|f| + 1 \geqq m + 2$, and it disappears after the application of of $\pi_{m+1}$.

Similar reasoning shows that $\pi_F \lambda_q p = \lambda_{\pi_F q} p$.

Let us consider the sequence $p(m)$ $(m \geqq 0)$ of elements of $\mathcal{P}$ defined inductively by $p(0) = (\phi)(=(\phi, \phi, \cdots, \phi))$, and, for each $m \geqq 0$, $p(m + 1) = \lambda_{p(m)} p$.

We prove that, for each $m', m'' \geqq 0$, $\pi_{m'} p(m') = \pi_{m'} p(m' + m'')$ and $p(m') \in \mathcal{P}_F$.

The relations are trivial for $m' \leqq 1$. Assume that they hold for $m' \leqq m$. Then: $\pi_{m+1} p(m + 1) = \pi_{m+1}\lambda_{p(m)} p$ (by definition) $= \pi_{m+1}\lambda_{\pi_m p(m)} p$ (by (*)) $= \pi_{m+1}\lambda_{\pi_m p(m+m'')} p$ (by the induction hypothesis) $= \pi_{m+1}\lambda_{p(m+m'')} p$

(by (*)) $= \pi_{m+1}p(m + m'' + 1)$ (by definition). This gives the first relation, and $p(m+1) \in \mathcal{P}_F$ follows by induction from $p(0) \in \mathcal{P}_F$.

Hence the limit for $m \to \infty$ of $p(m)$ is a well defined quasi-regular $n$-tuple of subsets of $F$ which we denote by $p(\infty)$ and which satisfies the identity $p(\infty) = \lambda_{p(\infty)}p$. Let us verify that any $p' \in \mathcal{P}$ which satisfies $p' = \lambda_{p'}p$ is equal to $p(\infty)$. Indeed, if $\pi_m p' = \pi_m p(\infty)$ for some $m \geqq 0$, the same relation holds for $m + 1$ since $\pi_{m+1}p' = \pi_{m+1}\lambda_{\pi_m p'}p = \pi_{m+1}\lambda_{\pi_m p(\infty)}p = \pi_{m+1}p(\infty)$. Since the hypothesis of quasi regularity on $p'$ implies $\pi_0 p' = \pi_0 p(\infty)$, the verification is completed.

It remains to show that $L = \pi_F L_1$ is equal to the first component of $p(\infty)$.

For this, let $\lambda_p{}^0 p = p$ and, inductively, $\lambda_p^{m+1}p = \lambda_p{}^m(\lambda_p p)$. The classical identity $\lambda_q\lambda_{q'} = \lambda_{\lambda_q q'}$ (valid for any $q$, $q' \in \mathcal{P}$) shows that, for all $m$, $\pi_{m+1}p(m + 1) = \pi_{m+1}\lambda_p{}^m p$. Thus $p(\infty)$ can also be defined as $\lim_{m\to\infty} \lambda_p{}^m p$.

Consider now $p' = p \cup (\xi)$ where $(\xi)$ is the $n$-tuple $(\{\xi_1\}, \{\xi_2\}, \cdots, \{\xi_n\})$ and where $\cup$ is to be performed component-wise. By definition the relation

$\pi_F\lambda_{p'}^{m'+1}p' = \pi_F\lambda_p^{m'+1}p$ is true for $m' = 0$; if it is true for $m' = m$, it is still true for $m' = m + 1$ because, setting $p''' = \lambda_{p'}^m p'$ and $p'' = \lambda_p{}^m p$,

$$\pi_F\lambda_{p'}^{m+1}p' = \pi_F(\lambda_{p'}^m(\lambda_{p'}p')) = \pi_F\lambda_{p'''}p' = \lambda_{\pi_F p'''}p$$

$$= \lambda_{\pi_F p''}p = \pi_F\lambda_{p''}p = \pi_F\lambda_p^{m+1}p.$$

Hence $\lim_{m\to\infty} \pi_F\lambda_{p'}^m p' = p(\infty)$. Since our original definition amounts to the definition of $L_1$ as the first component of $\lim_{m\to\infty} \lambda_{p'}^m p'$ the result is entirely verified.

For the sake of completeness we recall the proof of the following theorem which is needed later on. To simplify notation, $+$ and $\Sigma$ are used instead of $\cup$; $\mathfrak{R}(F)$ is the set of all *regular events* on $F$; for each $q \in \mathfrak{P}(F)$, $q^* = \Sigma\{q^n : n > 0\}$.

THEOREM (Chomsky and Miller, 1956). *If for all* $j$ $(1 \leqq j \leqq n)$

$$p_j = q_{j,0} + \Sigma\{\xi_{j'}q_{j,j'} : 1 \leqq j' \leqq n\}$$

*with* $q_{j,j'} \in \mathfrak{R}(F)$ *for all* $j$, $j'$ $(1 \leqq j \leqq n)$, $(0 \leqq j' \leqq n)$, *then every component of* $p(\infty)$ *belongs to* $\mathfrak{R}(F)$.

PROOF: The result follows by induction on $n$ from the unicity property of $p(\infty)$ mentioned above. Indeed for $n = 1$, $p' = q_{1,0}q_{1,1}^*$ satisfies $\lambda_{p'}p = p'$ and thus $p(\infty) = p' \in \mathfrak{R}(F)$. Assume now the result proved for $n < n'$ and let $(p_j')$ $(1 \leqq j \leqq n)$ be defined by

$$p_n' = q_{n,0}q_{n,n}^* + \Sigma\{\xi_j, q_{n,j'}q_{n,n}^*: 1 \leqq j' < n\}$$

and for $1 \leqq j < n$, $p_j'$ obtained by substituting in $p_j$ the right member of this last equation for $\xi_n$.

The hypothesis $p(\infty) = \lambda_{p(\infty)}p$ shows that $p'(\infty) = p(\infty)$. However, the grammar $(p_j')$ $(1 \leqq j < n)$ has only $n - 1$ nonterminal letters and the result follows from the induction hypothesis.

## II. PUSH-DOWN AUTOMATA

In all this section $X = \{x\}$ and $Y = \{y\}$ denote respectively the input alphabet and the internal alphabet (used for writing on the tape). The corresponding sets of words are $F$ and $G$. It can be proved that there would be no loss in generality in *card* $X$ = *card* $Y = 2$; it is not so for card $Y = 1$.

If $\mathfrak{K}$ is a finite automaton with input alphabet $Y$, we denote by $\chi g$ the state reached after reading the word $g \in G$, the initial state being fixed. A standard argument shows that there is no loss of generality (for our present purpose) in taking $\chi$ as a *finite homomorphism*—that is, as a mapping $\chi$ of $G$ onto a finite monoid $K$ such that for all $g$, $g'$, $g''$, $g'''$ the relation $\chi g = \chi g'$ implies $\chi g'' g g''' = \chi g'' g' g'''$.

For any $K' \in \mathfrak{P}(K)$ (the set of all subsets of $K$) and $g \in G$, $\rho_{K'}g$ denotes the longest left factor $g'$ of $g$ such that $g = g'g''$ with $\chi g'' \in K' \cup \{\chi g\}$.

With this notation, an elementary $\chi$-push-down mapping $\mu: G \to G$ is given by:

(1) a finite homomorphism $\chi: G \to K$,

(2) a mapping $\alpha: K \to G$;

(3) a mapping $\bar{\rho}: K \to \mathfrak{P}(K)$.

For each $g \in G$, $\mu g = \rho_{\bar{\rho}\chi g}(g \cdot \alpha\chi g)$. In more concrete manner, one may think of a device $\mathfrak{K}$ which performs the following cycle of operations:

(1′) $\mathfrak{K}$ reads $g$ and determines $\chi g \in K$;

(2′) $\mathfrak{K}$ writes the word $\alpha\chi g \in G$ to the right of the word $g$;

(3′) $\mathfrak{K}$ reads from right to left the word $g \cdot \alpha\chi g$ erasing successively each letter (eventually none) till it reaches a state belonging to the subset $\bar{\rho}\chi g$ of $K$ or until it has erased the whole word $g \cdot \alpha\chi g$.

The word left is $\mu g$. Thus, either $g$ is a left factor of $\mu g$ (in symbols $g \mid \mu g$) or $\mu g \mid g$; we shall refer to this situation by saying that $g$ and $\mu g$ are *comparable*.

Now let $S$ be a finite set in which a subset $\bar{S}$ and an element $s_\infty \in \bar{\mathcal{U}}$ have been distinguished. We denote by $U$ (resp. $\bar{U}$, resp. $\bar{U}_\infty$) the set of all pairs $(s, g)$ with $s \in S$ (resp. $s \in \bar{S}$, resp. $s = s_\infty$) and $g \in G$. We call *states* the elements of $U$. It will be understood that for $u, u' \in U$, the notation $u \mid u'$ (resp. $u \nmid u'$) means that $u = (s, g)$ $u' = (s', g')$ with $g \mid g'$ (resp. with *not* $g \mid g'$).

DEFINITION 1. A $\chi$-push-down mapping $\mu: U \to U$ is given by:

(1) A mapping $\sigma: (S, K) \to S$ such that for all $k \in K$, $\sigma(s_\infty, k) = s_\infty$;

(2) For each $s \in S$ an elementary $\chi$-push-down mapping $\mu_s$ on $G$ with the restriction that for all $u \in \bar{U}_\infty$, $\mu_{s_\infty} u = u$.

For each $u = (s, g) \in U$, $\mu u = (\sigma(s, \chi g), \mu_s g)$. For simplicity, we shall rather deal with the mapping $\mu^*$ defined as follows.

Let $\mu^1 u = \mu u$ and for each $i > 0$, $\mu^{i+1} u = \mu(\mu^i u)$. The largest positive $i$ (possibly infinite) such that, for all positive $i' < i$, $\mu^{i'} u \notin \bar{U}$ will be denoted by $j(u)$. Then

$$\mu^* u = \mu^{j(u)} u \quad \text{if} \quad j(u) < \infty;$$

$$= (s_\infty, g) \quad \text{if} \quad j(u) = \infty \quad \text{and} \quad u = (s, g).$$

Thus $\mu^* = \mu$ iff for all $u \in U$, $\mu u \in \bar{U}$.

DEFINITION 2. A *push-down automaton* $\mathfrak{A}$ is given by:

1) The finite alphabets $X$ and $Y$;

2) A finite homomorphism $\chi: G \to K$;

3) A finite set $S$ (with $\bar{S}$, $s_\infty$ and $U$ as above) and a $\chi$-push-down mapping $\mu: U \to U$;

4) A mapping $\beta: (\bar{S}, X) \to S$ with $\beta(s_\infty, x) = s_\infty$ identically;

5) An *initial state* $u_0 \in \bar{U} \backslash \bar{U}_\infty$ (i.e., $u_0 \in \bar{U}$ and $u_0 \notin \bar{U}_\infty$); a finite set $\bar{U}_{fin} \subset \bar{U} \backslash \bar{U}_\infty$ of *final states*.

For each $u = (s, g) \in U$ and $x \in X$, the "next state" $u \cdot x$ is $\mu^*(\beta(s, x), g)$. For each $f \in F$, $u \cdot f = u$ if $f$ is the empty word and $u \cdot f = (u \cdot f') \cdot x$ if $f = f'x$ $(f' \in F, x \in X)$. The set $Acc\ \mathfrak{A}$ of the words accepted by $\mathfrak{A}$ is $Acc\ \mathfrak{A} = \{f \in F: u_0 \cdot f \in \bar{U}_{fin}\}$.

Finally $\mathfrak{A}$ is *simple* iff $\bar{S} = S$, and then, clearly, $\mu = \mu^*$.

According to *Definition* 2 the cycle of the automaton that is initiated by each input letter consists of two successive operations: the mapping $\beta: (\bar{S}, X) \to S$ and the mapping $\mu^*: U \to \bar{U}$; further, any state $u \in \bar{U}_\infty$ is a *sink*, i.e. for all $u \in \bar{U}_\infty$ and $f \in F$, one has $u \cdot f = u$. Intuitively, one might think of a device $\mathfrak{A}_0$ acting in the following manner for each state $(s, g) \in \bar{U}$ and incoming input letter $x \in X$:

$\mathfrak{A}_0$ goes first to the state $u = (\beta(s, x), g)$ and it performs the push-down mapping $\mu_{\beta(s,x)}$ which brings it to $u' = (s', g')$, say. If $s'$ belongs to $\bar{S}$ the cycle initiated by $x$ is already completed and $\mathfrak{A}_0$ reads the next input letter (this is always the case when $\mathfrak{A}$ is simple). If $s'$ does not belong to $\bar{S}$, $\mathfrak{A}_0$ performs $\mu_{s'}$ and goes to $u'' = \mu_{s'}u'$. Again, if $u'' \in \bar{U}$ the cycle is completed; if it is not so, $\mathfrak{A}_0$ goes on performing a succession of push down mappings till, eventually, it reaches a state of $\bar{U}$. Clearly, in the general case, $\mathfrak{A}_0$ may never reach this subset and consequently it may happen that $\mathfrak{A}_0$ does not read the input word further than $x$. Our more formal definition of $\mathfrak{A}$ by the mapping $\mu^*$ is intended to obviate this minor notational difficulty.

PROPERTY 1. *For any push-down automaton $\mathfrak{A}$ there exists a simple push-down automaton $\mathfrak{A}'$ such that Acc $\mathfrak{A}$ = Acc$\mathfrak{A}'$.*

PROOF: The property amounts to the statement that for any given push-down mapping $\mu: U \to U$ there exists a finite set $S'$, a surjection (i.e., mapping onto) $\zeta: S' \to S$ and a push-down mapping $\mu: (S', G') \to (S', G')$ that have the following properties:

(i) $\mu'^* = \mu'$;

(ii) for all $g \in G$, $s' \in S'$, the relation $\mu'(s', g) = (s'', g')$ implies $\mu^*(\zeta s', g) = (\zeta s'', g')$.

To simplify notation we first verify that there is no loss in generality in assuming that $\mu$ satisfies the conditions (1), (2), and (3) below. For $u = (s, g) \in U$, we write $\chi u = (s, \chi g)$ and $\chi U = \{(s, k): s \in S, k \in K\}$; $\chi\bar{U} = \{(s, k): s \in \bar{S}, k \in K\}$.

CONDITION (1). There corresponds to each $(s, k) \in \chi U$ a subset $K(s, k)$ of $K$ (eventually empty) such that $g \in \chi^{-1}K(s, k)$ (i.e., $\chi g \in K(s, k)$) iff for all $g' \in G$ one has $\rho_{\bar{p}(s,k)}g'g\alpha(s, k) = g'$ (where $\bar{p}(s, k)$ denotes the subset of $K$ defined by the function $\bar{p}$ associated with the elementary push-down mapping $\mu_s$).

Indeed let the quotient monoid $K_1$ of $G$ and the epimorphism (= homomorphism onto) $\chi_1 : G \to K_1$ be defined by the condition that, for any $g, g' \in G$, $\chi_1 g = \chi_1 g'$ iff for all $K', K'' \subset K$ one has:

$$\rho_{K'}g \mid \rho_{K''}g \text{ is equivalent to } \rho_{K'}g' \mid \rho_{K''}g'.$$

Clearly $K_1$ is finite and there exists a homomorphism $\bar{\chi}: K_1 \to K$ such that $\bar{\chi}\chi_1 = \chi$. Hence, any of the $\chi$-push-down mappings entering in the definition of $\mathfrak{A}$ could have been defined as well as a $\chi_1$-push-down mapping. Thus we can assume that $\chi$ has been replaced by $\chi_1$ in the definition of $\mathfrak{A}$ and condition (1) is trivially satisfied.

CONDITION (2). For each $u \in U$, if $u \mid \mu u$ then $\mu u \in \bar{U}$.

Indeed let $j^+(u)$ denote the largest number $i$ (possibly infinite) such that, for all positive $i' < i$, one has $\mu^{i'}u \notin \bar{U}$ and $u \mid \mu^{i'}u$. By construction $j^+(u) \leqq j(u)$.

Let $\mu' : U \to U$ be defined by:

$$\mu' u = \mu^{j^+(u)} u \text{ if } j^+(u) \text{ if finite;}$$
$$= \mu^* u \text{ if } j^+(u) \text{ is infinite.}$$

Clearly $\mu'^* = \mu^*$ and it suffices to verify that $\mu'$ is a $\chi$-push-down mapping since, by construction, $\mu'$ satisfies Condition (2).

Consider $u = (s, g)$ such that $j^+(u) > 1$. Thus if $j' < j^+(u)$, the state $\mu^{j'}u$ has the form $(s', g\bar{g})$ for a certain $\bar{g} \in G$. Induction on $j'$ shows easily that if $\chi g = \chi g'$ then $\mu^{j'}(s, g') = (s', g'\bar{g})$ with the same $s'$ and the same $\bar{g}$. Hence $j^+(u)$ is a function of $\chi u$ only.

It follows that the mapping $\mu'' : U \to U$, defined for all $u \in U$ by:

$$\mu'' u = \mu u \text{ if } j^+(u) \text{ is finite;}$$
$$\mu'' u = \mu^* u \text{ if } j^+(u) \text{ is infinite,}$$

is a $\chi$-push-down mapping and, since $\mu''^* = \mu^*$, we can assume from now on that $\mu = \mu''$, i.e., that, for all $u \in U$, $j^+(u)$ is finite.

Then, as we have seen, for each $u = (s, g) \in U$, the state $\mu' u$ $(= \mu^{j^+(u)}u)$ has the form $(s', \rho_{K'} g\bar{g})$ where $s' \in S$, $K' \subset K$, and $\bar{g} \in G$ are functions of $\chi u \in \chi U$ only, and the verification is completed.

We point out that Condition (2) implies that, for each $(s, g) \in U$, the number $j(s, g) - 1$ is at most equal to the length $|g|$ of $g$. Indeed, if $\mu(s, g) \notin \bar{U}$, the state $(s', g') = \mu(s, g)$ is such that $g$ is not a left factor of $g'$ and, since $g$ and $g'$ are comparable, this implies $g' \mid g$ and $g' \neq g$, hence $|g'| < |g|$.

CONDITION (3). There exists a $\chi$-push-down mapping $\bar{\mu}$ such that $\mu^* = \mu\bar{\mu}$.

We assume (1) and (2). Let $U_1 = \{u : j(u) = 1\} = \{u : \mu u \in \bar{U}\} = \{u : \mu u = \mu^* u\}$. By the definition of a $\chi$-push-down mapping there exists a subset $\chi U_1$ of $\chi U$ such that $U_1 = \{u : \chi u \in \chi U_1\}$. When $u = (s, g) \notin U_1$, we have seen above that the state $(s', g') = \mu(s, g)$ is such that $g' \mid g$. Hence, introducing a new mapping $\bar{\mu} : U \to U_1$ by the rule that for all $u \in U$:

$$\bar{\mu} u = u \quad \text{if} \quad u \in U_1,$$
$$\bar{\mu} u = \mu^{j(u)-1} u \quad \text{if} \quad u \notin U_1,$$

we have identically $\mu^* = \mu\bar{\mu}$ and the verification of (3) amounts to the verification that $\bar{\mu}$ can be defined as a $\chi'$-push-down mapping for a suitable finite homomorphism $\chi'$.

For showing this, let $V$ denote the set of all quadruples $v = (s, k, s', k')$ with $s, s' \in S, k, k' \in K$. $V_1 \subset V$ is defined by the restriction $(s, k) \in \chi U_1$.

For $v = (s, k, s', k') \in V_1$ (resp. $\in V$) and $n > 0$, we define $\bar{G}_n(v)$ (resp. $G(v)$) as the set of all $g' \in G$ satisfying $k\chi g' = k'$ which are such that for some $g \in \chi^{-1}k$ one has $j(s', gg') = n + 1$ and $\mu^n(s', gg') = (s, g)$ (resp. $j(s, gg')$ arbitrary and $\mu(s', gg') = (s, g)$). Thus for $v \in V_1$, $\bar{G}_1(v) = G(v)$.

Because of the definition of $\bar{\mu}$ it is easily seen that if $g' \in G_n(v)$ (resp. $\in G(v)$) then for all $g \in \chi^{-1}k$ one has $j(s', gg') = n + 1$ and $\mu^n(s, gg') = (s, g)$ (resp. $\mu(s', gg') = (s, g)$).

For $v \in V_1$ $(v = (s, k, s', k'))$ let $\bar{G}(v)$ be the union of the sets $\bar{G}_n(v)$ over all positive $n$; also let $(v', v'') \in \omega(v)$ mean that there exist $(s'', k'') \in \chi\bar{U}$ $(= \{\chi u\colon u \in \bar{U}\})$ such that $v' = (s, k, s'', k'')$ (hence $v' \in V_1$) and $v'' = (s'', k'', s', k')$. Thus, by construction, $G(v) \subset \bar{G}(v)$ and, for $n > 0, g' \in \bar{G}_{n+1}(v)$ iff there exist $(v', v'') \in \omega(v)$ and a factorization $g' = g''g'''$ such that $g'' \in \bar{G}_n(v')$ and $g''' \in G(v'')$.

Introduce now a set $\{\xi_v\}$ of new letters indexed by the elements of $V_1$ and, for each $v \in V_1$, define

$$p_v = G(v) + \Sigma\{\xi_{v'}G(v'')\colon (v', v'') \in \omega(v)\}.$$

The system $(p) = (p_v)_{v \in V_1}$ defines a grammar and, by construction, $p_v(\infty) = \bar{G}(v)$ for each $v \in V_1$.

However, for arbitrary $v = (s, k, s', k')$, the condition that a word $g'$ belongs to $G(v)$ can be explicitly stated as:

$$s = \sigma(s', k'); \qquad k' = k\chi g'; \qquad \text{for all} \quad g \in \chi^{-1}k, \quad g = \rho_{\bar{p}(s,k)}gg'\alpha(s, k).$$

Thus, because of Condition (1), each of the sets $G(v)$ $(v \in V_1)$ is a regular event. By the theorem recalled at the end of the first section, it follows that each of the sets $\bar{G}(v)$ $(v \in V_1)$ is also a regular event. In other terms, there exists a homomorphism $\chi'$ of $G$ onto a finite monoid $K'$ and for each $v \in V_1$ a subset $\bar{K}'(v)$ of $K'$ such that $G(v) = \chi'^{-1}\bar{K}'(v)$.

Since $\chi'$ may be chosen so that $K'$ admits $K$ as a homomorphic image, we may argue as in the verification of Condition (1) that in fact $\chi = \chi'$. Under this hypothesis it is immediate that $\bar{\mu}$ can be defined as a $\chi$-push-down mapping and Condition (3) is verified.

The rest of the proof (i.e., that under the conditions (1), (2), and (3), the mapping $\mu^* = \mu\bar{\mu}$ can be defined as a push-down mapping) is trivial and it is omitted.

PROPERTY 2. *For any push-down automaton $\mathcal{A}$ the set Acc $\mathcal{A}$ is a context-free language.*

PROOF: By Property 1 we can assume that $\mathcal{A}$ is simple. Hence for each $s \in S$, $x \in X$, $g \in G$, the operation performed by $\mathcal{A}$ consists of a transition $s \rightarrow s \cdot x \in S$ and of an elementary push-down mapping $\mu_{s,x}$ on $G$.

For simplicity we shall speak of the states as if they consisted only of a word. Thus the *length* $|u|$ of $u = (s, g)$ is the length $|g|$ of $g$; $u$ is a *left factor* of $u' = (s' g')$ if $g$ is a left factor of $g'$ etc.

Let $f = x_{i_1}x_{i_2} \cdots x_{i_m}$ be an input word of length $|f| = m \geqq 2$. For any state $u$ we consider the $m - 1$ *intermediate states* $u_1 = u \cdot x_{i_1}$, $u_2 = u \cdot x_{i_1}x_{i_2}, \cdots, u_{m-1} = u \cdot x_{i_1}x_{i_2} \cdots x_{i_{m-1}}$ and we define min $(u; f)$ to be the minimum of their length. If $u_j$ $(1 \leqq j \leqq m - 1)$ is such that $|u_j| =$ min $(u \cdot f)$ and if further either $j = m - 1$ or $|u_{j'}| > |u_j|$ for $j < j' \leqq m - 1$, we call $u_j$ the *critical state* (of $f$ at $u$); $f' = x_{i_1}x_{i_2} \cdots x_{i_j}$ and $f'' = x_{i_{j+1}}x_{i_{j+2}} \cdots x_{i_m}$ are the *critical factors* (of $f$ at $u$). Clearly the critical state always exists (when $f \notin X$) and it is uniquely determined.

REMARK 1. *The critical state $u_j$ of $f$ at $u$ is a left factor of all the intermediate states and it is comparable with both $u$ and $u' = u \cdot f$.*

PROOF: The statement is trivial if $f$ has length two, i.e., if $f = x_{i_1}x_{i_2}$, because there is only one intermediate state, viz. $u_1 = u \cdot x_{i_1}$, which, by force, is the critical one. The fact that $u$ and $u_1$ and $u_1$ and $u' = u_1 \cdot x_{i_2}$ are comparable is a direct consequence of the definitions.

Assume now the property verified for all words of length $<m$ $(m > 2)$ and consider $f = \bar{f}x_{i_m}$ of length $m = |\bar{f}| + 1$. The intermediate states of $f$ at $u$ are those of $\bar{f}$ at $u$ plus the state $\bar{u} = u_{m-1} = u \cdot \bar{f}$. Hence we distinguish two cases:

(i) $|\bar{u}| >$ min $(u, \bar{f})$. Then, $u_j$ $(1 \leqq j \leqq m - 2)$ is at the same time the critical state of $f$ and $\bar{f}$.

(ii) $|\bar{u}| \leqq$ min $(u, \bar{f})$. Then $\bar{u}$ is the critical state of $f$. In case (i), because of the induction hypothesis we have only to prove that $u_j$ is a left factor of $\bar{u}$ comparable with $u' = u \cdot f$. The first statement follows directly from the hypothesis that $\bar{u}$ is comparable with $u_j$ and that $|u_j| < |\bar{u}|$. The second statement follows from the first, the fact that $u' = \bar{u} \cdot x_{i_m}$ and the remark that if $a$ is a left factor of $b$ and if $b$ is comparable with $c$ then in turn $a$ is comparable with $c$. In case (ii), the induction hypothesis and $|\bar{u}| \leqq |u_j|$ show that $\bar{u}$ is a left factor of $u_j$, hence a left factor of every intermediate state of $f$. By our last remark above it follows that $\bar{u}$ is comparable with $u$ and since $\bar{u}$ is comparable with $u' = u \cdot x_{i_m}$, the verification is completed.

Let us introduce the notation $C(u, u', n)$ for denoting the set (eventually empty) of all input words $f \neq e$ such that $u \cdot f = u'$ and either $|f| = 1$ or min $(u, f) \geqq n$. The symbols $+$ and $\Sigma$ denote disjoint union of sets. Finally $|a|$ is the maximum of the lengths of the words $\alpha(s, k)$ used in the definition of the push-down mappings $\mu$. Thus for any $u$ and $f \in F$ of length at least two, we have min $(u, f) \leqq |a| + |u|$ since this is an upper bound to the length of the first intermediate word. Our definition of the critical factors is summarized by the following relations:

REMARK 2. *For any triple* $(u, u', n)$

$$C(u, u', n) = X \cap C(u, u'n) + \Sigma\{C(u, u'', |u''|)C(u'', u', |u''| + 1) : u'' \in U, \; n \leqq |u''| \leqq |u| + |a|\}$$

*where* $\Sigma\{\ \}$ *is understood to be* $\phi$ *unless* $n \leqq |u| + |a|$.

PROOF: The left member is contained in the right member because any $f \in C(u, u', n)$ of length two or more has critical factors $f^*$ and $f''$ and a critical state $u''$ satisfying the condition indicated.

Conversely if $f' \in C(u, u'', |u''|)$ and $f'' \in C(u'', u, |u''| + 1)$ for some $u''$ such that $n \leqq |u''| \leqq |u| + a$, the product $f = f'f''$ belongs to $C(u, u', n)$ and has $u''$ as its critical state. Finally, the right member is a disjoint union of sets because of the unicity of the critical factorization.

Let $J_l$ (resp. $J_r$) be the set of all triples $(u, u', n)$ such that $C(u, u', n)$ is a nonempty set of words which can be left (resp. right) critical factors of other words. According to the definition and to Remark 1, $(u, u', n) \in J_l$ (resp. $\in J_r$) iff $C(u, u', n) \neq \phi$, $u$ and $u'$ are comparable, $n = |u'|$ (resp. $n = |u| + 1$).

For fixed $d > 0$ we consider the following subsets of $J_l \cup J_r$:

$$J_l^+ = \{(u, u', |u'|) : |u| \leqq |u'|\}$$

$$J_l^{-d} = \{(u, u', |u'|) : |u| = |u'| + d\}$$

$$J_r^- = \{(u, u', |u| + 1) : |u| \geqq |u'|\}$$

$$J_r^{+d} = \{(u, u', |u| + 1) : |u'| = |u| + d\}.$$

REMARK 3. *For each finite d there exists only a finite number of distinct sets* $C(u, u', n)$ *with* $(u, u', n) \in J(d) = J_l^+ \cup J_l^{-d} \cup J_r^- \cup J_r^{+d}$.

PROOF: For any quadruple of states of the form $u = (s, gg')$, $\bar{u} = (s, \bar{g}g')$, $u' = (s', gg'')$, $\bar{u}' = (s', \bar{g}g'')$ with $\chi g = \chi \bar{g}$, the definition of pushdown mappings and the hypothesis that $\chi$ is a homomorphism shows that

$$C(u, u', | g | + n') = C(\bar{u}, \bar{u}', | \bar{g} | + n')(n' \geqq 0).$$

This proves directly the statement for $J_l^{-d} \cup J_r^{+d}$. For $J_l^+$ we need to observe first that, if $| u | < | u' |$, $C(u, u', | u' |) \neq \phi$ only if $| u' | \leqq | u | + | a |$. For $J_r^-$ it suffices to check that (with $\chi_1$ as defined in the proof of Property (1)) the relations $\chi_1 q = \chi_1 \bar{q}$, $\chi_1 q' = \chi_1 \bar{q}'$ imply identically

$$C((s, gg'), (s', g), | gg' | + 1) = C((s, \bar{g}\bar{g}'), (s', \bar{g}), | \bar{g}\bar{g}' | + 1).$$

REMARK 4. *Each set* $C(u, u', n)$ *is a context-free language.*

PROOF: Let the triple $(u, u', n) = j_0$ be fixed. Let $d = \max(| u | - n, | u' | - n, | a |)$ and consider a minimal set $J^*$ of triples (containing $j_0$) such that any set $C(j)$ with $j \in \bar{J}(d) = J(0) \cup J(1) \cdots \cup J(d)$ is equal to one and only one set $C(j')$ with $j' \in J^*$. By Remark 3 we know that $J^*$ is finite. Furthermore, by construction, any critical (left or right) factor of a word from a set $C(j)$ with $j \in \bar{J}(d)$ or $j = j_0$ belongs itself to some set $C(j')$ with $j' \in \bar{J}(d)$—hence to some set $C(j'') = C(j')$ with $j'' \in J^*$. Hence, by Remark 2, there corresponds to each $j \in J^*$ an equation

$$C(j) = X_j + \Sigma\{C(j')C(j''): (j', j'') \in \omega(j)\}$$

where $X_j = X \cap C(j)$ and where $\omega(j)$ denote a finite set of pairs of triples $j', j'' \in J^*$.

We introduce a set $\Xi = \{\xi_j\}(j \in J^*)$ of new letters and, for each $j \in J^*$ we define the set $p_j$ as the union of $X_j$ and of the products $\xi_{j'}\xi_{j''}$ where $(j', j'') \in \omega(j)$. This reduces the problem to the proof of equivalence of the two definitions of a context-free language described in the first part of the paper.

Taking into account that the set of all context-free languages is closed under (finite) union, Property 2 is verified.

## III. EXAMPLE

Let us recall briefly the definition of the free group $\Gamma$ generated by a set $Y' = \{y_{i'}: 1 \leqq i' \leqq n\}$. Let $Y$ consist of $2n$ letters $y_i$ ($i = \pm i'$,

$1 \leqq i' \leqq n$) and say that a word of $G$ (the set of all words in the letters of $Y$) is *reduced* if it does not contain a pair of adjacent letters having opposite indices. Clearly to each $g \in G$ is associated a unique reduced word $\tau g$ obtained by successive cancellation of such pairs of adjacent letters. For instance,

$$\tau(y_1 y_{-2} y_2 y_{-1}) = \tau(y_1 y_{-1}) = e$$

and

$$\tau(y_1 y_2 y_{-1} y_{-2}) = y_1 y_2 y_{-1} y_{-2} .$$

The homomorphism $\gamma_0 : G \rightarrow \Gamma$ is defined by $\gamma_0 g = \gamma_0 g'$ iff $\tau g = \tau g'$. We shall identify $\Gamma$ with the set $\tau G$ of all reduced words of $G$ endowed with the (associative) multiplication $\tau(\tau g \cdot \tau g')(= \tau(gg'))$.

Thus, in particular, $\gamma_0 y_{-i} = (\gamma_0 y_i)^{-1}$ for all $y_i \in Y$. With this notation, any homomorphism $\gamma: F \rightarrow \Gamma$ is given by a homomorphism $\gamma': F \rightarrow G$ and the rule $\gamma f = \gamma f'$ iff $\gamma_0\gamma' f = \gamma_0\gamma' f'$ (i.e., iff $\tau\gamma' f = \tau\gamma' f'$).

We verify that the inverse image $\gamma^{-1}\Gamma'$ by $\gamma$ of any *finite* subset $\Gamma'$ of $\Gamma$ is a context-free language, by constructing a pushdown automaton $\mathfrak{A}$.

For this, consider a state $s_0$ and, for each word $\gamma' x_j \in \gamma X$, introduce $n_j$ new states $s_{j,1}, s_{j,2}, \cdots s_{j,n_j}$ where $n_j$ is the length of the word $\gamma' x_j$. Let $S$ be the union of $s_0$ and all the states $s_{j,k}$ and $\bar{S}$ consist of $s_0$ only. The state $s_\infty$ is not needed.

Now, in the notation of Definition 2 we define for each $x_j \in X$, $\beta(s_0, x_j) = s_{j,1}$ and to each state $s_{j,j'}$ we associate (1) a mapping $\sigma: S \rightarrow S$ defined by

$$\begin{aligned} \sigma s_{j,j'} &= s_{j,j'+1} \quad \text{if } j < n_j ; \\ &= s_0 \qquad\quad \text{if } j = n_j . \end{aligned}$$

(We do not need to introduce $K$ explicitly.) (2) A push-down mapping $\mu_i$ on $G$ where $i(-n \leqq i \leqq n)$ is determined by $y_i = y_{j,j'}$ (i.e., where $y_i$ is the $j'$th letter of $\gamma' x_j$) and where for each $y \in G$:

$$\begin{aligned} \mu_i g &= g' \quad \text{if} \quad g = g' y_{-i} \qquad\qquad (g' \in G); \\ &= g y_i \qquad\qquad\qquad\qquad\qquad \text{otherwise.} \end{aligned}$$

For instance, if $\gamma' x_1 = y_{-1} y_2$ the resulting transformation on $G$ is $\mu_2 \mu_{-1}$ and it has the following effect: $g \rightarrow g'$ if $g = g' y_1 y_{-2}$; $g \rightarrow g' y_2$ if $g = g' y_1$ and $g' \notin G y_{-2}$; $g \rightarrow g y_{-1} y_2$ in any other case.

Induction of the length of the input word $f$ shows that if the initial word stored in the memory is a reduced word $g \in \tau G$, the word stored after reading $f$ is $\tau(g\gamma' f)$. Thus, in particular, taking $g = e$ (the empty word), Property 2 shows that for any finite subset $\Gamma'$ of $\Gamma$ the set $\gamma^{-1}\Gamma' = \{f \in F: \gamma f \in \Gamma'\}$ is a context language.

Clearly, for any regular event $R$ on $F$, it is possible to add enough new states to the finite part of the automaton that it accepts a set of the form $R \cap \gamma^{-1}\Gamma'$.

Let us consider the special case where $X = \{x_i: i = \pm i', 1 \leqq i' \leqq n\}$ and $\gamma' x_i = y_i$ identically. Taking the empty word $e$ as initial word and the set $\{e\}$ as set of final word, it is clear that the set of words accepted by the automaton is $\gamma^{-1}e$, the kernel of the homomorphism $\gamma$ of $F$ onto $\Gamma$ that satisfies identically $(\gamma x_i)^{-1} = \gamma x_{-i}$. Thus we have $(\mathfrak{U}_n)$.

For all $f, f', f'' \in F$, any two of the following relations imply the third one: $f \in \gamma^{-1}e; f'f'' \in \gamma^{-1}e; f'ff'' \in \gamma^{-1}e$.

In fact, it can be proved that $\gamma^{-1}e$ can be defined abstractly as the least subset of $F$ that satisfies $(\mathfrak{U}_n)$ and that contains $e$ and every product $x_i x_{-i}$ $(1 \leqq i \leqq n)$.

We construct explicitly the grammar defining $\gamma^{-1}e$. For each $i$ $(i = \pm i', 1 \leqq i' \leqq n)$ let $D_i$ denote the set of all $f \in x_i F \cap \gamma^{-1}e$ which are such that $f = f'f''$; $f', f'' \in \gamma^{-1}e$ implies $f'$ or $f'' = e$. Intuitively, $f \in D_i$ iff the first letter of $f$ is $x_i$ and if the first return of the internal memory to the empty word occurs at the last letter of $f$. Thus $f \in D_i$ implies $f = x_i f' x_{-i}$ and either $f' = e$ or $f' = f_1 f_2 \cdots f_m$ where $f_1 \in D_{j_1}$, $f_2 \in D_{j_2} \cdots f_m \in D_{j_m}$ or in more concrete manner, where the words $f_k$ are characterized by the fact that at that end of their last letter (when reading $f$) the word stored in the memory is reduced to the letter $x_i$. Clearly this factorization is unique and our hypothesis that the memory is never empty before the end of $f$ implies that $j_1, j_2, \cdots, j_m \neq -i$. Conversely, if $D_i' = \Sigma\{D_j: j \neq -i\}$, any word of the form $x_i f' x_{-i}$ belongs to $D_i$ if $f'$ is a product of words from $D_i'$. Hence:

$$D_i = x_i x_{-i} + x_i D_i'^* x_{-i}$$

where $D_i'^*$ is defined by $D_i'^* = D_i' + D_i'^* D_i'$.

Introducing $2n$ new variables $\xi_i$ and $\xi_i'^*$ we deduce the $2n$ equations:

$$\xi_i = x_i x_{-i} + x_i \xi_i'^* x_{-i}\,; \qquad \xi_i'^* = (e + \xi_i'^*)\Sigma\{x_j x_{-j} + x_j \xi_j'^* x_{-j}: j \neq -i\}.$$

Finally, denoting by $D^*$ the set of the nonempty words of $\gamma^{-1}e$, we have $D^* = (e + D^*)\Sigma D_i$ from which we deduce the equation

$$\xi^* = (e + \xi^*)\Sigma\{x_i x_{-i} + x_i \xi_i'^* x_{-i}: i = \pm i', 1 \leqq i' \leqq n\}$$

in the new variable $\xi^*$ corresponding to $D^*$. The fact that $D^*$ is related to an algebraic system of equations goes back to Kesten (Kesten, 1959). When the internal alphabet consists of a single letter, $\gamma$ becomes a homomorphism into a cyclic group and the memory can be identified with an unbounded counter. The corresponding theory is due to Raney (Raney, 1960) and it relates to the enumeration of well formed formulas (free notation); an approach similar to the present one has been given in (Schützenberger, 1959). As a point of marginal interest it may be mentioned that the set $D^*$ (which we shall call a *Dyck set*) and, more generally, the standard context free languages defined below, can also be defined as the complement of the support of certain rational (non commutative) formal power series. This results instantly from the existence of isomorphic representations of the free group $\Gamma$ by integral finite dimensional matrices (cf. e.g. Sanov, 1957).

For further reference we introduce the following definition in which it is assumed that $X = \{x_{\pm i}\}$ as above.

DEFINITION. A *standard context-free language* is a set $L = D^* \cap R(X_1, V)$ where $D^*$ is a Dyck set and where the regular event $R(X_1, V)$ is given by a subset $X_1$ of $X$, a subset $V$ of $X^2$ and the relation $R(X_1, V) = X_1F \backslash FVF$ (= the set of all words that begin with a letter from $X_1$ and that have no factor belonging to $V$).

## IV. A WEAK CONVERSE PROPERTY

PROPERTY 3. *Each context-free language $L$ can be represented as the homomorphic image of some standard context-free language.*

PROOF: Let $L$ be produced by the grammar $(p_j)(1 \leqq j \leqq n)$, the notation being as in the first section of the paper.

Each word $h \in p_j$ has a unique factorization $h = f_1'\xi_{i_1}f_2'\xi_{i_2}f_3' \cdots f'_{\delta h}\xi_{i_{\delta h}}f'_{\delta h+1}$ where $f_1', f_2', \cdots, f'_{\delta h}, f'_{\delta h+1} \in F$, $\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_{\delta h}} \in \Xi$, $0 \leqq \delta h < d^*$ and where $d^* = 1 + \max\{\delta h : h \in p_j, 1 \leqq j \leqq n\}$. Eventually $h \in F$ in which case $\delta h = 0$ and $h = f_1'$.

Let us introduce a set $X'$ of new letters $x(j, h, d, \epsilon)$ indexed by quadruples with $1 \leqq j \leqq n$; $h \in p_j$; $0 \leqq d \leqq d^*$; $\epsilon = \pm 1$. For given $j$ and $h \in p_j$ the writing $q(h, d, d)$ denotes the word $x(j, h, d, +1)\xi_{i_d}x(j, h, d, -1)$ if $1 \leqq d \leqq \delta h$ and the word $x(j, h, d, +1)x(j, h, d, -1)$ if $\delta h < d \leqq d^*$; finally,

$$q(h, 0, 0) = x(j, h, 0, -1)q(h, 1, 1)q(h, 2, 2) \cdots q(h, d^*, d^*)x(j, h, 0, +1).$$

Let $F'$ be the free monoid generated by $\Xi \cup X'$ and define the homomorphism $\varphi: H' \to H$ in the obvious way: for all $\xi \in \Xi$, $\varphi\xi = \xi$; for

$1 \leqq j \leqq n$, $h \in p_j$, $1 \leqq d \leqq 1 + \delta h$, $\epsilon = +1$, $\varphi x(j, h, d, \epsilon) = f_j'$; $\varphi x(j, h, d, \epsilon) = e$ for all the other elements of $X'$.

Thus for all $h \in p_j$, we have $h = \varphi q(h, 0, 0)$. It follows that if the grammar $(p_j')(1 \leqq j \leqq n)$ is defined by $p_j' = \{\zeta q(h,0,0) : h \in p_j\}$, we have for all $m > 0$ the identity $\varphi p'(m) = p(m)$. Hence $L = \varphi p_1'(\infty)$ and, without loss of generality, we shall assume henceforth that $L$ *is* $p_1'(\infty)$ itself, that is, $X = X'$; $H = H'$, $(p_j) = (p_j')(1 \leqq j \leqq n)$. Under this assumption we shall write $x(h, d, \epsilon)$ instead of $x(j, h, d, \epsilon)$ since every word $h \in P = \bigcup\{p_j : 1 \leqq j \leqq n\}$ appears in one and only one set $p_j$ $(1 \leqq j \leqq n)$.

For $1 \leqq d \leqq d' \leqq n$ we define:

$$q(h, d, d') = q(h, d, d)q(h, d + 1, d + 1) \cdots q(h, d', d').$$

(Thus $q(h, d, d')$ is defined only when none or both of $d$ and $d'$ are 0.) Finally:

$$Q = \{q(h, d, d') : h \in P, \quad d = d' = 0 \quad \text{or} \quad 1 \leqq d \leqq d' \leqq d^*\}$$

$$T = \{\lambda_{p(\infty)} q(h, d, d') : q(h, d, d') \in Q\}.$$

We now define the standard (right) context-free language $D \cap R \cap H_j$ by:

(i) $D$ is the Dyck set $D = \{f \in F : f \neq e; \gamma f = e\}$ where the homomorphism $\gamma : H \to \Gamma$ is defined by:

$$\text{for all } \xi \in \Xi, \qquad \gamma\xi = e$$

$$\text{for all } x(h, d, \epsilon) \in X, \qquad (\gamma x(h, d, \epsilon))^{-1} = \gamma x(h, d, -\epsilon).$$

(ii) $R$ is the set of the words $f \in F$ such that each of their factors of length two belongs to $\bar{V} = \bar{V}' \cup \bar{V}'' \cup \bar{V}'''$ where:

$\bar{V}'$ is the set of all products $x(h, d, +1)x(h', 0, -1)$ and $x(h', 0, +1)$ $x(h, d, -1)$ for which $d$, $h$, and $h'$ are such that $q(h, d, d) = x(h, d, +1)\xi_{jd}x(h, d, -1)$ with $h' \in p_{jd}$.

$\bar{V}''$ is the set of all products $x(h, d, +1)$ $x(h, d, -1)$ with $h \in P$ and $\delta h < d \leqq d^*$.

$\bar{V}'''$ is the set of all words $x(h, d, -1)x(h, d', -1)$ with $h \in P$ and either $0 \leqq d < d' = d + 1 \leqq d^*$ or $d = d^*$ and $d' = 0$.

(iii) $H_j(1 \leqq j \leqq n)$ is the set of all words of $H$ whose first (left) letter has the form $x(h, 0, -1)$ with $h \in p_j$.

We shall use repeatedly the fact that if the $n$-tuple $a = (a_1, a_2, \cdots, a_n)$ of elements of $H$ is such that $a_1, a_2, \cdots, a_n \in \gamma^{-1}e$

(for short, if $a \subset (\gamma^{-1}e)$) then, for any $h \in H$, the two relations $\gamma h = e$ and $\lambda_a h \subset \gamma^{-1}e$ are equivalent.

Now we have:

(1) For each $j$ $(1 \leqq j \leqq n)$, $p_j(\infty) = T \cap H_j$.
This is a direct consequence of $Q \cap H_j = P$ and the definition of $T$.

(2) $T \subset D \cap R$.

It is clear that $p(0) \subset (D \cap R \cap H_j)$ (i.e., for each $j$, $p_j(0) \subset D \cap R \cap H_j$). Assume $p(n') \subset (D \cap R \cap H_j)$ proved for $n' \leqq n$. Since, trivially, $Q \subset \gamma^{-1}e$, it follows that $\lambda_{p(n)}q \subset D$ for each $q \in Q$; hence, in particular, $p_j(n+1) \subset D$ for each $j$ $(1 \leqq j \leqq n)$. Further, every factor of length two of a word of the set $\lambda_{p(n)}q (q \in Q)$ is a factor of a word contained in one of the sets $p_j(n)$ or belongs to $\bar{V}$. Hence $\lambda_{p(n)}q \subset R$ and, in particular, $p_j(n+1) \subset R$ for all $j$ $(1 \leqq j \leqq n)$. The fact that $p_j(n) \subset H_j$, identically is trivial and the result follows by induction on $n$.

Since (1) and (2) show $L \subset p_1(\infty) \subset D \cap R \cap H_1$ and $T \cap H_1 \subset L$, the verification of Property 3, i.e., of $L = D \cap R \cap H_1$, will follow from:

(3) $D \cap R \subset T$.

Let $f \in D \cap F$. It is trivial that $f \in T$ for $|f| \leqq 2$. Hence we can assume the result proved for all $f' \in F$ of length $<n$ and $|f| = n > 2$.

We consider the factorization $f = g'g''(g', g'' \in F)$ where $g'$ is defined as the shortest left factor of $f$ that belongs to $D$. Since $f \in D$, this implies $\gamma g'' = (\gamma g')^{-1}\gamma f = e$ and we distinguish two cases:

(i) $g'' \neq e$. Then $g'' \in D$; $|g'|, |g''| < n$, and, by the induction hypothesis, there exist two elements $q', q'' \in Q$ such that $g' \in \lambda_{p(\infty)}q'$ and $g'' \in \lambda_{p(\infty)}q''$. Let $x(h', d', \epsilon')$ and $x(h'', d'', \epsilon'')$ denote respectively the last (right) letter of $q'$ and the first (left) letter of $q''$. By the definition of $Q$, we have

$$1 \leqq d' \leqq d^* \text{ and } \epsilon' = 1 \text{ or } d' = 0 \text{ and } \epsilon' = +1; \quad \text{similarly:}$$

$$1 \leqq d'' \leqq d^* \text{ and } \epsilon'' = +1 \text{ or } d'' = 0 \text{ and } \epsilon'' = -1.$$

However, $g'g'' \in R$ implies $v = x(h', d', \epsilon')x(h'', d'', \epsilon'') \in \bar{V}$. Obviously $v \notin V' \cup V''$. Hence $v \in \bar{V}'''$ and, thus, $h = h' = h''$, $d'' = d' + 1$, $\epsilon' = -1$, $\epsilon'' = +1$. This means that $q'q'' = q(h, d_1, d_2) \in Q$ for some $d_1$, $d_2$. Since, now, $f \in \lambda_{p(\infty)}q(h, d_1, d_2)$, the result is verified in this case.

(ii) $g'' = e$, that is, $f = g'$.

Since $|f| = n > 0$, this implies $f = af'b$ with $a, b \in X$, $f' \in F$, $|f'| > 0$.

Because of the definition of $\gamma$, $\gamma f = e$ and the hypothesis that $f$ has no left factor in $D$ imply $\gamma f' = \gamma ab = e$. Thus $f' \in D \cap R$ and, by the induction hypothesis, $f' \in \lambda_{p(\infty)} q(h', d, d')$ for some $h' \in P$, $0 \leqq d \leqq d' \leqq d^*$.

It follows that $f' = x(h', d, \epsilon)f''x(h', d', \epsilon)$ either with $1 \leqq d < d' \leqq d^*$, $\epsilon = +1$, $\epsilon' = -1$ or with $d = d' = 0$, $\epsilon = -1$, $\epsilon' = +1$. Let $v = ax(h', d, \epsilon)$, $v' = x(h', d', \epsilon')b$. Since $f \in R$, both belong to $\bar{V}$. Clearly $v, v' \notin \bar{V}''$. Further, $v, v' \notin \bar{V}'''$ because, for instance, $v \in \bar{V}'''$ would imply $1 \leqq d < d' \leqq d^*$, $\epsilon = +1$, $\epsilon = -1$. $a = x(h', d - 1, -1)$, hence $b = x(h', d - 1, +1)$ (since $\gamma b = (\gamma a)^{-1}$) and, finally, $v' \in \bar{V}'''$ giving $d' = d - 2$, in contradiction of $d < d'$. Thus, $v, v' \in \bar{V}'$, that is, $f = x(h, d'', +1)f'x(h, d'', -1)$ and $f' = x(h', 0, -1)f''x(h', 0, +1)$, where $h$, $h'$ and $d''$ are such that $q(h, d'', d'') = x(h, d'', +1)\xi_j x \cdot (h, d'', -1)$ with $h' \in p_j$. Thus $f \in \lambda_{p(\infty)} q(h, d'', d'')$, concluding the verification of Property 3.

Remark. In the case of formal power series over a ring $A$, the sets $p_j$ $(1 \leqq j \leqq n)$ defining the grammar are replaced by the elements $\bar{p}_j = \Sigma\{a_{j,h}h : h \in p_j\}$ (with $a_{j,h} \in A$) of the free algebra over $A$ generated by $X \cup \Xi$. It is trivial that Property 3 and its proof remain valid provided that the homomorphism $\varphi$ is replaced by a homomorphism $\bar{\varphi}$ sending the large algebra (over the integers) of the free monoid generated by $X' \cup \Xi$ into the large algebra (over $A$) of the free monoid generated by $X \cup \Xi$. For this, it suffices to replace $\varphi$ by $\bar{\varphi}$ in all the definitions except for the conditions $\varphi x(j, h, 0, -1) = e$ $(1 \leqq j \leqq n, h \in p_j)$ which have to become $\bar{\varphi} x(j, h, 0, -1) = a_{j,h}$.

## References


Chomsky, N., and Miller, G. A., (1958), Finite state languages. *Information and Control* **1,** 91–112.

Chomsky, N., (1962), Context-free grammars and push-down storage. Quarterly Progress Report No. 65, Research Laboratory of Electronics, M.I.T.

Chomsky, N., and Schützenberger, M. P., (1962), The algebraic theory of context-free language, to appear in "Computer Programming and Formal Systems," P. Braffort and D. Hirschberg, eds. North Holland, Amsterdam.

Ginsburg, S., and Rice, H. G., (1961), Two families of languages related to Algol. *J. Assoc. Computing Mach.* **9,** 350–370.

Kesten, H., (1959), Symmetric random walks on groups. *Trans. Am. Math. Soc.* **92,** 336–354.

Newell, A., and Shaw, J. C., (1957), Programming the logic theory machine. *Proc. Western Joint Computer Conf.*, p. 230.

RABIN, M. O., AND SCOTT, D., (1959), Finite automata and their decision problems. *IBM J. Research Develop.* **3,** 115–125.

RANEY, G. N., (1960), Functional composition patterns and power-series reversion, *Trans. Am. Math. Soc.* **94,** 441–451.

REDEI, L., (1950), Die Anwendung des Schiefen Produktes in der Gruppentheorie. *J. reine u. angew. Math.* **188,** 201–227.

SANOV, I. N., (1947), A property of a representation of a free group. *Dokl. Akad. Nauk SSSR* **57,** 657–659.

SCHÜTZENBERGER, M. P., (1959), "Un problème de la théorie des automates." Séminaire Dubreil-Pisot, Institut H. Poincaré, Paris.

SHEPHERDSON, J. C., (1959), The reduction of two-way automata to one-way automata. *IBM J. Research Develop.* **3,** 198–200.

SHEPERDSON, J. C., AND STURGIS, H. E., (1963), Computability of recursive functions. *J. Assoc. Computing Mach.* **10,** 217–255.