JOURNAL OF COMPUTER AND SYSTEM SCIENCES 15, 185–208 (1977)

Theory of ω-Languages. II: A Study of Various Models of ω-Type Generation and Recognition

RINA S. COHEN

Department of Computer Science, Technion, Haifa, Israel

AND

ARIE Y. GOLD

Faculty of Mathematics, Technion, Haifa, Israel

Received September 25, 1975; revised September 14, 1976

ω-languages are sets consisting of ω-length strings; ω-automata are recognition devices for ω-languages. In a previous paper the basic notions of ω-grammars, ω-context-free languages (ω-CFL's), and ω-pushdown automata (ω-PDA's) were first defined and studied. In this paper various modes of ω-type generation are introduced and the effect of certain restrictions on the derivations in ω-grammars is investigated. Several distinct models of recognition in ω-PDA's are considered, giving rise to a hierarchy of subfamilies of the ω-CFL's. The relations among these subfamilies are established and characterizations for each family are derived. Non-leftmost derivations in ω-CFG's are studied and it is shown that leftmost generation in ω-CFG's is strictly more powerful than non-leftmost generation.

0. INTRODUCTION AND PRELIMINARIES

This paper constitutes the second part of [3]. In part I the notions of ω-grammars, ω-context-free languages (ω-CFL's), and ω-pushdown automata (ω-PDA's) were first introduced. Some fundamental results were presented and several characterizations of the family of ω-CFL's (CFL_ω) were derived.

In this paper the properties of ω-CFL's are studied, with particular emphasis on characterizing and comparing the various modes of generation in ω-CFG's and the modes of acceptance of ω-PDA's. In Section 1 some closure properties of the ω-CFL's are obtained. The use of control sets in ω-grammars is studied in Section 2 and certain known results from language theory concerning control sets and leftmost generation in ω-PSG's are generalized to ω-languages. In Section 3 the various types of i-acceptance by ω-PDA's are investigated and the corresponding families of ω-CFL's are characterized and shown to constitute a proper hierarchy within CFL_ω. Section 4 is devoted to the study of nonleftmost derivations in ω-CFG's. It is established that leftmost and nonleftmost generation in ω-CFG's are not equivalent; in fact, leftmost generation is

ISSN 0022-0000

more powerful and the ω-languages generated by ω-CFG's by non-leftmost derivations form a proper subclass of the ω-CFL's.

The reader of this paper is assumed to be familiar with Part I [3]. Therefore, most of the definitions concerning ω-grammars, ω-CFL's, and ω-automata of the various kinds will not be repeated here; we only briefly recall here a few definitions, which are particularly important for this paper, and also introduce some new notation.

DEFINITION 0.1. For any mapping $\psi: A \to B$, define $\text{In}(\psi) = \{b \mid b \in B, \text{card}(\psi^{-1}(b)) \geqslant \omega\}$ (card(D) denotes the cardinality of the set D).

Let $f: N \to S$ be a mapping from the set of natural numbers into a set S, and let $F \subseteq 2^S$. We say that mapping f is

1-*accepting* w.r.t. F	if	$(\exists H \in F)(\exists t)\, f(t) \in H$;
1′-*accepting* w.r.t. F	if	$(\exists H \in F)(\forall t)\, f(t) \in H$;
2-*accepting* w.r.t. F	if	$(\exists H \in F)\, \text{In}(f) \cap H \neq \varnothing$;
2′-*accepting* w.r.t. F	if	$(\exists H \in F)\, \text{In}(f) \subseteq H$;
3-*accepting* w.r.t. F	if	$\text{In}(f) \in F$.

DEFINITION 0.2. Let $M = (M_1, F)$ be an ω-PDA (ω-FSA). For $i = 1, 1', 2, 2', 3$ define $T_i(M) = \{\sigma \in \Sigma^\omega \mid$ there exists a run r of M on σ s.t. f_r is i-accepting w.r.t. $F\}$ where for each $j \geqslant 1, f_r(j)$ is the state entered in the jth step of the computation described by run r.

$T_i(M)$ ($i = 1, 1', 2, 2', 3$) is the *ω-language i-accepted by M*. For $i = 3$, i-acceptance is usually referred to as acceptance and the subscript 3 from $T_3(M)$ is omitted. An ω-language accepted by an ω-FSA (ω-PDA) is an ω-regular language (ω-CFL). CFL_ω denotes the class of ω-CFL's.

For $i = 1, 1', 2, 2'$, an ω-language which is i-accepted by some ω-FSA will be called an *Ai-ω-regular language.*

For $i = 1, 1', 2, 2'$, the class of ω-languages i-accepted by ω-PDA's will be denoted by Ai-PDL_ω.

The families of Ai-ω-regular languages were studied in [9, 11]; the families Ai-PDL_ω will be studied in Section 3 of this paper.

Notation 0.3. Let $G = (V_N, V_T, P, S)$ be a CFG. For $X \in V_N$, $P(X)$ will denote the set of all X-productions in P. For every $H \subseteq V_N$, let $P(H) = \bigcup_{X \in H} P(X)$ be the set of all productions of the variables in H.

Let d be the following (nonleft) derivation $d: \alpha_0 \Rightarrow_G \alpha_1 \Rightarrow_G \cdots \Rightarrow_G \alpha_l$. For each $1 \leqslant i \leqslant l$, let A_i be the variable rewritten at step i of the derivation; define $\text{Var}(d) = \{A \in V_N \mid A = A_i$ for some $1 \leqslant i \leqslant l\}$. Var$(d)$ is the set of all variables rewritten at least once during derivation d.

The above notation will also be used for finite derivations in ω-CFG's.

Recall that for an infinite derivation d in G

$$d: \alpha \underset{G}{\Rightarrow} \alpha_1 \underset{G}{\Rightarrow} \cdots \underset{G}{\Rightarrow} \alpha_i \underset{G}{\Rightarrow} \alpha_{i+1} \Rightarrow \cdots$$

the mappings $d_P: N \to P$ and $d_V: N \to V$ are defined by $d_P(i)$ = production used in the ith step of d; and $d_V(i)$ = variable rewritten in the ith step of d. Also define $\text{INV}(d) = \text{In}(d_V)$ and $\text{INP}(d) = \text{In}(d_P)$.

The following is a restatement of the ω-Kleene closure characterization theorem for ω-CFL's.

THEOREM 0.4 [3].

$$\text{CFL}_\omega = \omega\text{-KC(CF)} = \left\{ \bigcup_{i=1}^{k} U_i V_i^\omega \;\middle|\; U_i, V_i \textit{ are } \text{CFL's}, i = 1,..., k, k = 1, 2,... \right\}.$$

1. OPERATIONS ON ω-LANGUAGES

Nearly all operations studied in classical language theory can be redefined for ω-languages. However, here we restrict ourselves mainly to those operations, which are essential for obtaining the results of this paper.

DEFINITION 1.1. Let L_1, L_2 be ω-languages over Σ. Define the *quotient* of L_1 with respect to L_2 to be $L_1/L_2 = \{x \in \Sigma^* \mid \exists y \in L_2 \text{ s.t. } xy \in L_1\}$. (Note that the quotient of two ω-languages is a finite-string language.) For any ω-language L over Σ, define $\text{Init}(L)$ to be L/Σ^ω.

The next lemma follows directly from the ω-Kleene closure characterization theorem (Theorem 0.4) and from the closure properties of the context-free and regular languages.

LEMMA 1.2. *Let L be an* ω-CFL *(ω-regular language). Then* $\text{Init}(L)$ *is a* CFL *(regular language).*

As expected from the finite case, we have

PROPOSITION 1.3. (a) CFL_ω *is not closed under intersection and complementation*; (b) CFL_ω *is closed under intersection with ω-regular languages.*

Proof. (a) $L_0 = \{a^n b^n a^n \mid n \geq 1\} b^\omega$ is not an ω-CFL by Lemma 1.2 above. But $L_0 = L_1 \cap L_2$, where $L_1 = \{a^i b^i a^j \mid i, j \geq 1\} b^\omega$ and $L_2 = \{a^j b^i a^i \mid i, j \geq 1\} b^\omega$. L_1 and L_2 are ω-CFL's, hence the result follows.

(b) The proof follows the classical direct product construction of an ω-PDA and an ω-DFSA. ∎

DEFINITION 1.4. Let Σ, Δ be two finite alphabets. A substitution f is a mapping $f: \Sigma \to 2^{\Delta^*}$. f is extended to strings in Σ^* as (1) $f(\epsilon) = \epsilon$; (2) $f(xa) = f(x) f(a)$ for every $x \in \Sigma^*$, $a \in \Sigma$. For any $L \subseteq \Sigma^*$, define $f(L) = \bigcup_{x \in L} f(x)$. f is extended to strings in Σ^ω as follows. For $\sigma = \prod_{i=1}^{\infty} a_i \in \Sigma^\omega$, $a_i \in \Sigma\ \forall i \geq 1$, define $f(\sigma) = \{\prod_{i=1}^{\infty} b_i \mid b_i \in f(a_i)\}$. For each $L \subseteq \Sigma^\omega$, define $f(L) = \bigcup_{\sigma \in L} f(\sigma)$. Note that for ω-language L, $f(L) \subseteq \Delta^* \cup \Delta^\omega$. If for $L \subseteq \Sigma^\omega$, $f(L) \subseteq \Delta^\omega$, then we say that f is *ω-preserving on L*. f is called an *ω-preserving substitution* iff f is *ω-preserving* on all of Σ^ω, i.e., $f(\Sigma^\omega) \subseteq \Delta^\omega$.

Substitution f is said to be an *ϵ-free substitution* if for each $a \in \Sigma$, $\epsilon \notin f(a)$. f is said to be a *finite substitution* if $f(a)$ is a finite set for all a in Σ. In case $f(a)$ consists of a single word for each $a \in \Sigma$, f is a *homomorphism*. f is said to be a *context-free* (*regular*) *substitution* if $f(a)$ is a context-free (regular) language for all a in Σ.

As can be readily seen, a substitution f is ω-preserving iff it is ϵ-free.

EXAMPLE 1.5. Let $L = \{(ab)^\omega\}$ and define $f(a) = \epsilon$, $f(b) = b$. Then $f(L) = \{b^\omega\}$. Hence f is ω-preserving on L, though not ϵ-free.

Next, we consider GSM[1] mappings on ω-languages. Since every GSM can be viewed as an FSM which also emits a finite output string for each input symbol, we shall modify the notion of "run" for GSM to include also the sequence of output strings emitted by the machine.

DEFINITION 1.6. Let $S = (K, \Sigma, \Delta, \delta, q_0)$ be a (Σ, Δ)-GSM. Let $\sigma = \prod_{i=1}^{\infty} a_i \in \Sigma^\omega$, where $a_i \in \Sigma$ $\forall i \geqslant 1$. An infinite sequence $r = \{(q_i, x_i)\}_{i \geqslant 1}$, where $q_i \in K$ and $x_i \in \Delta^*$ $\forall i \geqslant 1$, is called a *run of S on σ* if $(q_1, x_1) = (q_0, \epsilon)$ and for each $i \geqslant 1$, $(q_{i+1}, x_{i+1}) \in \delta(q_i, a_i)$. Define $S(\sigma) = \{\sigma_1 \in \Delta^* \cup \Delta^\omega \mid$ there exists a run $r = \{(q_i, x_i)\}_{i \geqslant 1}$ of S on σ s.t. $\sigma_1 = \prod_{i=1}^{\infty} x_i\}$. For $L \subseteq \Sigma^\omega$, let $S(L) = \bigcup_{\sigma \in L} S(\sigma)$. For each ω-language L over Σ, GSM S is called *ω-preserving on L* if $S(L) \subseteq \Delta^\omega$. S is ω-preserving if $S(\Sigma^\omega) \subseteq \Delta^\omega$.

DEFINITION 1.7. A class $\mathscr{L}$ of ω-languages over Σ will be called *closed under substitution $h: \Sigma \to 2^{\Sigma^*}$* (closed under (Σ, Σ)-GSM mapping S) if for every $L \in \mathscr{L}$, $h(L) \cap \Sigma^\omega \in \mathscr{L}(S(L) \cap \Sigma^\omega \in \mathscr{L}\}$.

The next result follows from the ω-Kleene closure characterization (Theorem 0.4) and the closure properties of the CFL's and regular languages.

PROPOSITION 1.8. (a) *Let $L \subseteq \Sigma^\omega$ be an ω-*CFL *and $\tau: \Sigma \to 2^{\Delta^*}$ a context-free substitution. Then $\tau(L) \cap \Delta^*$ is a* CFL *and $\tau(L) \cap \Delta^\omega$ is an ω-*CFL.

(b) *Let $L \subseteq \Sigma^\omega$ be an ω-regular language and $\tau: \Sigma \to 2^{\Delta^*}$ a regular substitution. Then $\tau(L) \cap \Delta^*$ is a regular language and $\tau(L) \cap \Delta^\omega$ is an ω-regular language.*

Generalizing some well-known results concerning quotients from the classical theory (see Lemma 9.5 in [8]), we have

LEMMA 1.9. *Let $\mathscr{L}$ be a class of ω-languages closed under ϵ-free finite substitution and intersection with ω-regular languages. Let $\mathscr{D}$ be a family of finite-string languages closed under homomorphism and intersection with regular languages* s.t. $\{\text{Init}(L) \mid L \in \mathscr{L}\} \subseteq \mathscr{D}$; *then the quotient of any $L \in \mathscr{L}$ with respect to any ω-regular language belongs to $\mathscr{D}$.*

PROPOSITION 1.10. (a) *Let L be an ω-*CFL *and R an ω-regular language; then L/R is a*

[1] *A generalized sequential machine* over input alphabet Σ and output alphabet $\Delta((\Sigma, \Delta)$-GSM) is a 5-tupel $S = (K, \Sigma, \Delta, \delta, q_0)$, where K is a finite set of states, δ is a mapping from $K \times \Sigma$ to finite subsets of $K \times \Delta^*$ and q_0 is the initial state.

CFL. (b) *Let L be an ω-regular language and L_1 an arbitrary ω-language; then L/L_1 is a regular language.*

Proof. (a) follows from Lemmas 1.2, 1.9, and Propositions 1.3, 1.8. (b) is proved as in the finite-string case. ▮

PROPOSITION 1.11. *Let $\mathscr{L}$ be a class of ω-languages over alphabet Σ, closed under finite substitution and intersection with* A1′-*ω-regular languages. Then $\mathscr{L}$ is closed under* GSM *mapping.*

Proof. Let $S = (K, \Sigma, \Sigma, \delta, q_0)$ be a GSM. For every $a \in \Sigma$, let $D_a = \{[q, a, x, p] \mid q, p \in K, x \in \Delta^*, (p, x) \in \delta(q, a)\}$ and let f be the finite substitution $f(a) = D_a \; \forall a \in \Sigma$. Define the ω-language $R = \{\sigma \in (\bigcup_{a\in\Sigma} D_a)^\omega \mid \sigma = \prod_{i=1}^{\infty} [q_{i-1}, a_i, x_i, q_i], q_0$ is the initial state of S, $a_i \in \Sigma$ and $q_i \in K$ for $i = 1, 2,...\}$; clearly R is A1′-ω-regular. Let h be the homomorphism on $\bigcup_{a\in\Sigma} D_a$ defined as $\forall y = [q, a, x, p] \in \bigcup_{a\in\Sigma} D_a$, $h(y) = x$. Then for every $L \in \mathscr{L}$, $S(L) = h(f(L) \cap R)$ and $S(L) \cap \Sigma^\omega = h(f(L) \cap R) \cap \Sigma^\omega$; hence $\mathscr{L}$ is closed under GSM mapping. ▮

By Propositions 1.3, 1.8, and 1.11 we have

COROLLARY 1.12. (a) CFL_ω *and the class of ω-regular languages are each closed under* GSM *mapping*; (b) *For every ω*-CFL *(ω-regular language) $L \subseteq \Sigma^\omega$, and for every (Σ, Δ)*-GSM *S, $S(L) \cap \Delta^*$ is a* CFL *(regular language).*

2. CONTROL SETS

Control sets serve as the main tool in investigating certain variations and extensions of the definition of leftmost generation by ω-grammars. In this section certain invariance properties of ω-language families are derived; in particular, it is shown that leftmost derivations in ω-PSG's yield only ω-CFL's.

The representation of control sets follows [7].

DEFINITION 2.1. An *unrestricted* ω-PSG is an ω-PSG of the form[2] $(V_N, V_T, P, S, 2^P)$, i.e., all subsets of P are repetition sets. An *unrestricted* ω-CFG (ω-RLG) is an unrestricted ω-PSG in which the rules are context free (right linear).

The following lemma is a generalization of a result in [12].

LEMMA 2.2. *For every unrestricted ω*-PSG *G, $L_l(G)$[2] is an ω*-CFL.

Proof. Let $G = (V_N, V_T, P, S, 2^P)$ be an unrestricted ω-PSG. Construct an ω-PDA $M = (K, V_T, \Gamma, \delta, q_0, Z_0, F)$, where $Z_0 = S$ and $\Gamma = V_N \cup V_T$. If l is the

[2] Recall that, by our definitions (see Section 3 in [3]), the left-hand side of each rule in an ω-PSG is in V_N^+; furthermore, in a leftmost derivation, in each step the leftmost variable of the sentential form must be included in the rewritten substring.

maximal length of the left-hand sides of the rules of P, then $K = \{q_0\} \cup \{q_{[\alpha]} \mid \alpha \in \bigcup_{i=1}^{l} V_N{}^i\}$, where $V_N{}^i$ denotes the set of all words of length i over V_N. δ is defined as $\forall a \in V_T$, $\delta(q_0, a, a) = (q_0, \epsilon)$; $\forall A \in \Gamma$, $\delta(q_0, \epsilon, A) = (q_{[A]}, \epsilon)$ if $A\alpha \rightarrow \gamma \in P$ for some $\alpha \neq \epsilon$ and $\gamma \in V^*$; $\delta(q_0, \epsilon, A) = (q_0, \gamma)$ if $A \rightarrow \gamma \in P$; $\delta(q_{[\alpha]}, \epsilon, A) = (q_{[\alpha A]}, \epsilon)$ if $\alpha A \gamma_1 \rightarrow \gamma \in P$ for some $\gamma_1, \gamma \in V^*$, and $\delta(q_{[\alpha]}, \epsilon, A) = (q_0, \gamma A)$ if $\alpha \rightarrow \gamma \in P$. Define $F = \{D \subseteq K \mid q_0 \in D\}$. It can be easily verified that $T(M) = L_l(G)$. ∎

Definition 2.3. Given an unrestricted ω-PSG $G = (V_N, V_T, P, S, 2^P)$, let $\tilde{P}$ denote the set of labels of the productions in P. For $y \in V^*$, $\eta \in \tilde{P}^\omega$, and $\sigma \in \Sigma^\omega$, define $g_G(y, \eta) = \sigma$ if $\eta = \prod_{i=1}^{\infty} p_i$, $\forall i \geqslant 1$ $p_i \in \tilde{P}$, and there exists an infinite leftmost derivation

$$y = u_1\alpha_1 \Rightarrow u_1u_2\alpha_2 \Rightarrow \cdots \Rightarrow u_1 \cdots u_i\alpha_i \Rightarrow \cdots,$$

where $\forall i$, $u_i \in V_T{}^*$, $\alpha_i \in V_N V^*$, $\sigma = \prod_{i=1}^{\infty} u_i$, and step i of this leftmost derivation involves the production labeled p_i, i.e., $d_P(i) = p_i$. The function $g_G(y, \eta)$ will be undefined if no such σ exists.

For $U \subseteq V^*$ and $C \subseteq \tilde{P}^\omega$, let $g_G(U, C) = \{g_G(y, \eta) \mid y \in U, \eta \in C\}$. Note that $g_G(U, C)$ may be empty. For any set $C \subseteq \tilde{P}^\omega$, define $L_C(G) = g_G(\{S\}, C)$. $L_C(G)$ is called the *ω-language generated by G with control set C.*

Clearly, $L_C(G) \subseteq L_l(G)$.

Lemma 2.4. *Given an unrestricted ω-PSG G and a control set C, there exist an unrestricted ω-PSG $H(G)$ of the same type as G, a homomorphism h, and an ϵ-free regular substitution τ* s.t. $L_C(G) = h(L_l(H(G)) \cap \tau(C))$.

Proof. Let $G = (V_N, V_T, P, S, 2^P)$ be the unrestricted ω-PSG, then define the unrestricted ω-PSG $H(G) = (V_N, V_T', P', S, 2^{P'})$ where $V_T' = V_T \cup \tilde{P}$, $P' = \{u \rightarrow pv \mid p \in \tilde{P}$ is the label of $u \rightarrow v\}$. The elements of $\tilde{P}$ are added to V_T to serve as left brackets in $H(G)$. Define a homomorphism h: $V_T' \rightarrow V_T \cup \{\epsilon\}$ by $h(a) = a$ $\forall a \in V_T$ and $h(p) = \epsilon$ $\forall p \in \tilde{P}$. Also define $\tau(p) = V_T{}^*pV_T{}^*$ $\forall p \in \tilde{P}$. Then $L_C(G) = h(L_l(H(G)) \cap \tau(C))$. τ is an ϵ-free regular substitution; hence the lemma is proved. ∎

Theorem 2.5. *The class of ω-languages of the form $L_C(G)$, where C is an ω-regular language and G is an unrestricted ω-PSG, coincides with* CFL_ω.

Proof. For any C and G as above, by Lemma 2.4, $L_C(G) = h(L \cap L')$, where $L' = \tau(C)$, $L = L_l(H(G))$. By the closure properties, L' is an ω-regular language, L is an ω-CFL and hence $L_C(G)$ is also an ω-CFL.

To prove the other direction, let L be an ω-CFL generated by an ω-CFG with variable repetition sets $G = (V_N, V_T, P, S, F)$, and assume $F = \{F_1\}$, where $F_1 = \{A_i\}_{i=1}^{l}$. For each $i = 1, \ldots, l$, let the collection of labels of all A_i-rules be H_i, and let $D_i = \bigcup_{i \neq j} H_j$. The set of all ω-derivations in G in which F_1 is the set of all variables used infinitely often is the following ω-regular set $C(F_1) = \tilde{P}^*(H_1D_2{}^*H_2D_3{}^* \cdots H_lD_1{}^*)$ s.t. $L_{C(F_1)}((V_N, V_T, P, S, 2^P)) = L((V_N, V_T, P, S, \{F_1\}))$. The above proof can be extended to the case in which F includes more than one repetition set. ∎

COROLLARY 2.6. *For any ω-PSG $G = (V_N, V_T, P, S, F)$, $L_l(G)$ is an ω-CFL.*

Proof. An ω-regular control set C, representing the set of all ω-derivations in G for which the set of variables used infinitely often belongs to F, can be constructed similarly as in the proof of Theorem 2.5 above. Then we have $L_l(G) = L_C(G_1)$, where $G_1 = (V_N, V_T, P, S, 2^P)$, and by Theorem 2.5, $L_C(G_1)$ is an ω-CFL. ▌

Remark 2.7. As a special case of Corollary 2.6 we may deduce that for any ω-CFG G with production repetition sets, $L(G)$ is an ω-CFL; this justifies our choice of ω-CFG with variable repetition sets as our standard model of ω-CFG (see [3], Theorem 3.1.4).

The next corollary of Theorem 2.5 shows that we still obtain only ω-CFL's even if we change the criterion for ω-generation in ω-PSG's as follows.

COROLLARY 2.8. *Let $G = (V_N, V_T, P, S, F)$ be an ω-PSG, where $F \subseteq P$. Define $L_2(G) = \{\sigma \in V_T^\omega \mid$ there exists a leftmost derivation d: $S \Rightarrow^{\omega}_{(G)} \sigma, \mathrm{INP}(d) \cap F \neq \varnothing\}$. Then $L_2(G)$ is an ω-CFL.*

Following the construction in Lemma 2.4, we obtain the following

COROLLARY 2.9. *Let G be an unrestricted ω-RLG and C an ω-CFL. Then $L_C(G)$ is an ω-CFL.*

Utilizing Lemma 2.4 and the proof techniques in Theorem 2.5, we obtain

COROLLARY 2.10. *The class of ω-languages of the form $L_C(G)$, where G is an unrestricted ω-RLG and C is an ω-regular language, coincides with the class of ω-regular languages.*

3. i-ACCEPTANCE BY ω-PDA'S

In Part I of this paper it was established that the family of ω-languages (3-)accepted by ω-PDA's coincides with CFL_ω, and also coincides with the family of ω-languages 2-accepted by ω-PDA's ($\mathrm{A2\text{-}PDL}_\omega$). This means that 2-acceptance and 3-acceptance by ω-PDA's are equivalent (incidentally, this is no longer true w.r.t. deterministic ω-PDA's [4]). As will be shown below, for $i = 1, 1', 2'$, the families of ω-languages i-accepted by ω-PDA's ($\mathrm{A}i\text{-}\mathrm{PDL}_\omega$) constitute proper subfamilies of CFL_ω; in fact $\mathrm{A1'\text{-}PDL}_\omega \subsetneqq \mathrm{A1\text{-}PDL}_\omega = \mathrm{A2'} - \mathrm{PDL}_\omega \subsetneqq \mathrm{CFL}_\omega$. This section is devoted to the study of these families.

3.1. *Properties of the Families* $\mathrm{A}i\text{-}\mathrm{PDL}_\omega$

DEFINITION 3.1.1. Two ω-PDA's M and M' will be called i-equivalent (for $i = 1, 1', 2, 2'$) iff $T_i(M) = T_i(M')$.

THEOREM 3.1.2. $\mathrm{A1\text{-}PDL}_\omega = \mathrm{A2'\text{-}PDL}_\omega$.

Proof. It follows from the definition that A1-PDL$_\omega \subseteq$ A2'-PDL$_\omega$. Let L be 2'-accepted by an ω-PDA $M = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, where $F = \{F_i\}_{i=1}^l$. Clearly, we may assume w.l.o.g. that for $i \neq j$ $F_i \not\subseteq F_j$. Construct a U-ω-PDA $M_1 = (K_1, \Sigma, \Gamma, \delta_1, q_0, Z_0, \bar{F})$, where for $1 \leqslant i \leqslant l$, $\bar{F}_i = \{q^{(i)} \mid q \in F_i\}$, $\bar{F} = \bigcup_{i=1}^l \bar{F}_i$, and $K_1 = K \cup \bar{F}$. Whenever M enters some state $q \in F_i$, M_1 may choose to enter the corresponding state $q^{(i)}$ in $\bar{F}_i$ and continue to imitate M within $\bar{F}_i$, guessing that from now on M will stay in F_i. M_1 is blocked in case M later on enters a state outside F_i. ∎

PROPOSITION 3.1.3. *For each* $i = 1, 1', 2, 2'$, *every* ω-PDA *can be replaced by an i-equivalent* U-ω-PDA (*i.e.* ω-PDA *with a single designated set*).

Proof. For $i = 1'$ the construction of a 1'-equivalent U-ω-PDA is similar to that in the proof of Theorem 3.1.2 above. For $i = 1$ and $i = 2$ the result is obvious and for $i = 2'$ the result follows from Theorem 3.1.2. ∎

Utilizing Proposition 3.1.3 above, one can prove

LEMMA 3.1.4. A1'-PDL$_\omega \subseteq$ A1-PDL$_\omega$.

THEOREM 3.1.5. *The class of* ω*-regular languages is properly included in* A1'-PDL$_\omega$.

Proof. Let R be an ω-regular language; then R is 2-accepted by some U-ω-FSA $A = (K, \Sigma, \delta, q_0, F)$ [1, 13]. Construct a 1'-equivalent U-ω-PDA M with two special symbols Z, Z_1 among its pushdown symbols. By a sequence of ϵ-moves, M nondeterministically writes $Z^j Z_1$ on top of the pushdown store for some $j \geqslant 1$, guessing that A is about to enter a state in F within the next j steps. M then starts imitating A and is blocked if the guess turns out to be wrong; otherwise M proceeds to make a new guess j and the whole procedure is repeated.

To show that the inclusion is proper, let $\Sigma = \{a, b\}$ and consider the ω-language $L_a = \{\sigma \in \Sigma^\omega \mid \forall n \geqslant 1, \#_a(\sigma/n) \geqslant \#_b(\sigma/n)\}$, where for $x \in \Sigma^*$, $\#_c(x)$ denotes the number of occurrences of letter c in x. L_a is an example of a nonregular ω-language in A1'-PDL$_\omega$. ∎

THEOREM 3.1.6. A2'-PDL$_\omega$ *equals the class* $\mathscr{L}$ *of* ω*-languages of the form* $\bigcup_{i=1}^l L_i L_i'$, *where* $l \geqslant 1$ *and for each* $1 \leqslant i \leqslant l$, L_i is a CFL *and* $L_i' \in$ A1'-PDL$_\omega$.

Proof. Let L be a CFL, and $L' \in$ A1'-PDL$_\omega$. Then there exists a PDA A accepting L by empty store, and a U-ω-PDA B which 1'-accepts L'. Using our standard techniques, one can construct from A and B a new ω-PDA P which 2'-accepts LL'. Hence $\mathscr{L} \subseteq$ A2'-PDL$_\omega$.

Let $L \in$ A2'-PDL$_\omega$; by Theorem 3.1.2 L is 1-accepted by a U-ω-PDA $M = (K, \Sigma, \Gamma, \delta, q_0, Z_0, F)$. Let $B = \{(q, \gamma) \mid |\gamma| > 0$ and (q, γ) in the range of $\delta\}$. For each $(q, \gamma) \in B$ define $L(q, \gamma)$ as the following CFL: $L(q, \gamma) = \{x \in \Sigma^* \mid$ there exists a run of M on x s.t. $x: (q_0, Z_0) \overset{*}{\vdash}_M (q_1, \gamma_1) \overset{*}{\vdash}_M (q_2, Z\gamma_2) \vdash_M (q, \gamma\gamma_2)$, where $\gamma_1, \gamma_2 \in \Gamma^*$, $Z \in \Gamma$ and $q_1 \in F\}$ and let $L'(q, \gamma)$ be the ω-language in A1'-PDL$_\omega$, consisting of all ω-words for which there is a complete run of M, starting in configuration (q, γ). Then clearly $L = \bigcup_{(q,\gamma) \in B} L(q, \gamma) L'(q, \gamma)$. ∎

DEFINITION 3.1.7. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG. For $i = 1, 1', 2, 2'$, define $L_i(G) = \{\sigma \in \Sigma^\omega \mid$ there exists an infinite leftmost derivation d: $S \Rightarrow_G^\omega \sigma$ s.t. d_V is i-accepting w.r.t. $F\}$. $L_i(G)$ $(i = 1, 1', 2, 2')$ is the ω-language i-generated by G.

The following proposition is proved similarly to Theorem 4.1.3 in [3].

PROPOSITION 3.1.8. *For $i = 1, 1', 2, 2'$, Ai-PDL$_\omega$ equals the class of ω-languages i-generated by ω-CFG's.*

Note that for any ω-language L, if L is $1'$-generated by the ω-CFG with a single repetition set $G = (V_N, V_T, P, S, \{F\})$, then L is also generated by the unrestricted ω-CFG $G_1 = (V_N, V_T, P_1, S, 2^{P_1})$, where $P_1 = P(F)$ is the set of F-productions. On the other hand, generation by an unrestricted ω-CFG can be considered as a special case of $1'$-generation by an ω-CFG with a single repetition set. By Propositions 3.1.3, 3.1.8, every ω-language L in A$1'$-PDL$_\omega$ can also be $1'$-generated by an ω-CFG with a single repetition set. Thus we obtain

PROPOSITION 3.1.9. A$1'$-PDL$_\omega$ *coincides with the class of ω-languages generated by unrestricted ω-CFG's.*

PROPOSITION 3.1.10. *For $i = 1', 2'$, Ai-PDL$_\omega$ is closed under:* (a) *union,* (b) *finite substitution,* (c) *intersection with Ai-ω-regular languages, and* (d) GSM *mapping.*

Proof. (a) Obvious. (b) Let $L \in$ A$1'$-PDL$_\omega$, let $G = (V_N, V_T, P, S, F)$ be an ω-CFG that $1'$-generates L, and let h: $V_T \to 2^{\Sigma^*}$ be a finite substitution. Extend h to V_N by defining $h(A) = A$ $\forall A \in V_N$. Construct a new ω-CFG $G_1 = (V_N, \Sigma, P_1, S, F)$, where $P_1 = \{A \to \beta \mid \beta \in h(\alpha), A \to \alpha \in P\}$. Clearly. $L_{1'}(G) = h(L) \cap \Sigma^\omega$. The same construction will do for $L \in$ A$2'$-PDL$_\omega$. (c) Is proved by the standard direct product construction. (d) Follows from Proposition 1.11 and (b), (c) above.

3.2. *Real-Self-Embedding* ω-CFG's

The notion of real-self-embedding ω-CFG introduced below plays an important role in the proofs to follow.

DEFINITION 3.2.1. Let $G = (V_N, V_T, P, S)$ be a CFG. A variable X is called *self-embedding* iff $X \overset{*}{\Rightarrow}_G \alpha X \beta$ for some $\alpha, \beta \in V^+$. In case $\alpha \in V_T^+$, X is called *real-self-embedding*.

Every CFG which has a (real-)self-embedding variable is called *(real-)self-embedding.* Otherwise it is called *non-(real-) self-embedding.*

We say that variable X is reachable from variable Y if there exists a derivation d: $Y \overset{*}{\Rightarrow}_G xX\beta$, $x \in V_T^*$, $\beta \in V^*$. A variable that is reachable from S will be simply called *reachable*, and *nonreachable* otherwise.

DEFINITION 3.2.2. Let $L \subseteq V_T^*$ and X a symbol not in V_T. Define the following substitution $h(X, L)$: $V_T \cup \{X\} \to 2^{V_T^*}$, $h(X, L)(a) = a$ for $a \in V_T$ and $h(X, L)(X) = L$.

Substitution $h(X, L)$ will be frequently used in the following sections.

Notation 3.2.3. Let $G = (V_N, V_T, P, S)$ be a CFG. Let $X \in V_N$; we shall write $X \overset{*}{\Rightarrow} V_T^*[V_T^+]$ if there exists $x \in V_T^*[V_T^+]$ and a derivation $d: X \overset{*}{\Rightarrow} x$. If no such derivation exists we shall write $X \overset{*}{\not\Rightarrow} V_T^*[V_T^+]$.

$P_c(X)$ will denote $P - P(X)$ (the complement of $P(X)$ with respect to P).

The above definitions and notation will also be used w.r.t. ω-CFG's.

Remark 3.2.4. Note that if $G = (V_N, V_T, P, S)$ is non-real-self-embedding, then for each $X \in V_N$ such that $X \overset{*}{\Rightarrow}_G V_T^+$ and for each $S_1 \in V_N - \{X\}$, the grammar $G_1 = (V_N - \{X\}, V_T \cup \{X\}, P_c(X), S_1)$ is also non-real-self-embedding. For suppose $Y \overset{*}{\Rightarrow}_{G_1} xY\alpha$ for $Y \in V_N - \{X\}$, $\alpha \in V^*$, $x \in (V_T \cup \{X\})^+$; since by assumption $X \overset{*}{\Rightarrow}_G x_1 \in V_T^+$, substituting x_1 for X into x, we obtain $Y \overset{*}{\Rightarrow}_G x_2Y\alpha$, where $x_2 \in V_T^+$, contradicting the assumption that G is non-real-self-embedding.

The following technical lemma will be needed later.

LEMMA 3.2.5. *Let $G = (V_N, V_T, P, S)$ be a non-real-self-embedding, ϵ-free* CFG. *Then the following hold.*

(a) *$L(G)$ is a regular language.*

(b) *For each $Q \subseteq V_N$, the language $L_Q(G) = \{x \in V_T^* \mid$ there exists a derivation $d: S \overset{*}{\Rightarrow}_G x$ s.t. $Q \subseteq \mathrm{Var}(d)\}$ (Notation 0.3) is regular.*

(c) *For each $X \in V_N$, the language $L_X' = \{x \in V_T^* \mid$ there exists a derivation $d: S \overset{*}{\Rightarrow}_G xX\beta, \beta \in V^*\}$ is regular.*

Proof. (a) The proof is similar to that of the well-known theorem on self-embedding grammars, as it appears in [14, pp. 46]. Let L be generated by a non-real-self-embedding CFG $G = (V_N, V_T, P, S)$. Without loss of generality, assume that every X in V_N is reachable from S. We separate two cases:

Case 1. S is reachable from every variable in V_N. Every production in P containing at least one nonterminal on the right-hand side is of one of the four forms: (1) $X \to \alpha Y\beta$; (2) $X \to \alpha Y$; (3) $X \to Y_1\beta$; (4) $X \to Y$; where $X, Y, Y_1 \in V_N$ and $\alpha, \beta \in V^+$. If P contains a production of form (1), say $X \to \alpha_1 Y\beta_1$, then $\alpha_1 \not\Rightarrow^* V_T^*$, because otherwise $X \overset{*}{\Rightarrow} x_1Y\beta_1 \Rightarrow x_1xS\beta\beta_1 \overset{*}{\Rightarrow} x_1xx_2X\beta_2\beta\beta_1$; hence $\alpha_1Y\beta_1 = A\gamma$, where $A \in V_N$ and $A \not\Rightarrow^* V_T^+$. If P contains, for some X, productions of forms (2) and (3), then by a similar argument, if $X \to \alpha_1 Y$, then $\alpha_1 Y = A\gamma$ where $A \in V_N$ and $A \not\Rightarrow^* V_T^+$, so $X \to \alpha_1 Y$ is a production of form (3). If P contains, for some X, only productions of form (2), then there are $X \to xY$, where $x \in V_T^+$. We conclude that for each $X \in V_N$, the X-productions are either right linear or of the form $X \to A\alpha$, $A \in V_N$, $\alpha \in V^+$, where $A \not\Rightarrow^* V_T^*$. Let $G_1 = (V_N, V_T, P_1, S)$, where P_1 contains all of the right-linear productions of P; then $L(G_1) = L(G)$, hence $L(G)$ is regular in this case.

Case 2. There is a variable X_1 s.t. for no words x and w does $X_1 \overset{*}{\Rightarrow} xSw$. The proof that $L(G)$ is regular is in this case by induction on the number n of variables.

(b) If $Q = \{S\}$, then $L_Q(G) = L(G)$ is regular by (a). Suppose $Q = \{X\}$ and $X \neq S$. If $X \not\Rightarrow^* V_T^*$, then $L_Q(G) = \varnothing$ is regular; so assume $X \overset{*}{\Rightarrow} V_T^*$. Define $G_2 = (V_N - \{X\}, V_T', P_c(X), S)$ where $V_T' = V_T \cup \{X\}$ and $G_1 = (V_N, V_T, P, X)$. Let

$L_i = L(G_i)$ for $i = 1, 2$. By (a), L_i, $i = 1, 2$ is regular and hence also $L_{\{X\}}(G) = h(X, L_1)(L_2 \cap V_T'^* X V_T'^*)$ is regular. If $|Q| > 1$ then $L_Q(G) = \bigcap_{X \in Q} L_{\{X\}}(G)$ is regular by the above argument.

(c) Define $V_1 = \{A \in V_N \mid A \overset{*}{\Rightarrow} V_T^*\}$, and for every $A \in V_1$, define $L_A = L(G_A)$, where $G_A = (V_N, V_T, P, A)$. Then L_A is a regular language for each $A \in V_1$. Let $\bar{V}_1 = \{\bar{A} \mid A \in V_1\}$ and let $\bar{P}$ be obtained from P as follows. If $A \to \alpha \in P$, then for every decomposition $\alpha = \alpha_1 \beta_1 B$ s.t. $\alpha_1 \in (V_T \cup V_1)^*$, $B \in V_N$ and $\beta \in V^*$, $\bar{P}$ will include the productions $A \to h(\alpha_1)B$, and $A \to h(\alpha_1)$, where $\forall a \in V_T$, $h(a) = a$ and $\forall Y \in V_1$, $h(Y) = \bar{Y}$. Define the right-linear grammar $G_1 = (V_N, V_T \cup \bar{V}_1 \cup \{c\}, \bar{P} \cup \{X \to c\}, S)$, where c is a new symbol. Let $L' = L(G_1)/\{c\}$ and h_1 be the regular substitution $\forall a \in V_T$, $h(a) = a$ and $\forall \bar{A} \in \bar{V}_1$, $h(\bar{A}) = L_A$. It is easily verified that $L_X' = h_1(L')$ and hence L_X' is a regular language. ▌

Note that Lemma 3.2.5 above was stated only for ϵ-free CFG's; this weak version happens to be sufficient for our purposes and simplifies the proof to some extent. However, by modifying the above proof appropriately, it can be shown that the lemma holds for CFG's with ϵ-rules as well.

The next theorrem is a (not so straightforward) generalization of the well-known result from the classical theory [2, 14]. A weaker version of this theorem will be used in the proof of Proposition 3.3.1 below.

THEOREM 3.2.6. *For every non-real-self-embedding ϵ-free ω-CFG G, $L(G)$ is an ω-regular language.*

Proof. Let $G = (V_N, V_T, P, S, F)$ be a non-real-self-embedding ϵ-free ω-CFG; w.l.o.g. we may assume that every variable in V_N is reachable. For the purpose of this proof, a variable with no productions will be called *dummy*. As in [14] and Lemma 3.2.5 we distinguish two cases:

Case 1. S is reachable from every nondummy variable. Following the argument in Lemma 3.2.5(a), we conclude that for each $X \in V_N$, the X-productions are only in the forms (1) right linear, (2) $X \to A\alpha$, $A \in V_N$, $\alpha \in V^+$, where $A \not\Rightarrow^* V_T^*$, (3) $X \to xZ\gamma$, $x \in V_T^* \cup V_N$, $\gamma \in V^*$, and Z a dummy variable. Construct a new ω-CFG $G_1 = (V_N, V_T, P_1, S, F)$, where P_1 contains all right-linear productions in P, on top of the additional productions $X \to A \in P_1$ for each rule $X \to A\beta \in P$ s.t. $\beta \in V^*$ and $A \in V_N$. Clearly $L(G_1) = L(G)$ and since G_1 is right linear, the assertion follows.

Case 2. There is a nondummy variable X_1 s.t. for no $x \in V_T^*$ and $\alpha \in V^*$, does $X_1 \overset{*}{\Rightarrow} xS\alpha$. We shall prove that $L(G)$ is an ω-regular language by induction on the number n of nondummy variables in G. For $n = 1$, since $S \overset{*}{\Rightarrow} S$, S must be a dummy variable and $L(G) = \emptyset$. Assume that the assertion holds for $n = k$. Let $k + 1$ be the number of nondummy variables in G. Without loss of generality we may assume that F consists of only one set, denoted by F itself.

First assume $X_1 \overset{*}{\Rightarrow} V_T^*$.

Subcase a. $X_1 \notin F$. Define CFG $G_1 = (V_N, V_T, P_c(S), X_1)$ and two ω-CFG's:

$\bar{G}_1 = (V_N, V_T, P_c(S), X_1, F)$, $\bar{G}_2 = (V_N - \{X_1\}, V_T \cup \{X_1\}, P_c(X_1), S, F)$. Let $\bar{L}_1 = L(\bar{G}_1) \subseteq V_T^{\omega}$, $\bar{L}_2 = L(\bar{G}_2) \cap (V_T \cup \{X_1\})^* V_T^{\omega}$ and $L_1 = L(G_1) \subseteq V_T^*$. By the induction hypothesis, $\bar{L}_1$ and $\bar{L}_2$ are ω-regular languages, and L_1 is regular by Lemma 3.2.5(a). Hence also $L_4 = h(X_1, L_1)(\bar{L}_2)$ (see Definition 3.2.2), representing the derivations in G which are not "blocked" by X_1, is ω-regular. Let $L_2' = \{x \in V_T^* \mid$ there exists a derivation d: $S \overset{*}{\Rightarrow}_G xX_1\alpha$, $\alpha \in V^*\}$. Then L_2' is regular by Lemma 3.2.5(c) and hence $L_3 = L_2'\bar{L}_1$ is an ω-regular language. Since $L = L_3 \cup L_4$, L is ω-regular.

Subcase b. $X_1 \in F$. Define: a set of new symbols $E = \{Y_Q \mid Q \subseteq F\}$; a finite substitution h_1 on $V_T \cup \{X_1\}$: $\forall a \in V_T$, $h_1(a) = a$ and $h_1(X_1) = \{X_1\} \cup E$; and a CFG $G_3 = (V_N, V_T, P_3, X_1)$, where $P_3 = \bigcup_{A \in F} P(A)$. For each $Q \subseteq F$, let $L_Q = \{x \in V_T^* \mid$ there exists a derivation d: $X_1 \overset{*}{\Rightarrow}_{G_3} x$ s.t. $Q \subseteq \mathrm{Var}(d)\}$ (Notation 0.3).

By Lemma 3.2.5(b) L_Q is regular for each $Q \subseteq F$. For each $H \subseteq F - \{X_1\}$, define the ω-CFG $G_H = (V_N - \{X_1\}, V_T \cup \{X_1\}, P_c(X_1), S, H)$; also define R_H to be the following regular language $R_H = \{x \mid$ for some $l \geqslant 1$, $x = \prod_{i=1}^{l} x_i Y_{Q_i}$, where $x_i \in V_T^*$, $Y_{Q_i} \in E$ for $1 \leqslant i \leqslant l$, and $F - H \subseteq \bigcup_{i=1}^{l} Q_i\}$. Then $L_H^{(1)} = h_1(L(G_H)) \cap (V_T \cup \{X_1\})^* R_H^{\omega}$ is an ω-regular language by the induction hypothesis. Let h_2 be the following regular substitution on $V_T \cup \{X_1\} \cup E$: $\forall a \in V_T$, $h_2(a) = a$, $h_2(X_1) = L(G_1)$, where G_1 is as defined in Subcase *a*, and $\forall Y_Q \in E$, $h_2(Y_Q) = L_Q$. Then clearly $L_H = h_2(L_H^{(1)})$ is also an ω-regular language. Since $L = L_3 \cup (\bigcup_{H \subseteq F - \{X_1\}} L_H)$, where L_3 is as in Subcase a, L is ω-regular.

Now assume $X_1 \not\Rightarrow^* V_T^*$. In Subcase a $L = L_3 \cup (L(\bar{G}_2) \cap V_T^{\omega})$ and in Subcase b $L = L_3$, hence the proof is completed. ▌

3.3. *Proof of* A1′-PDL$_\omega \subsetneqq$ A2′-PDL$_\omega \subsetneqq$ CFL$_\omega$

We are now in a position to prove the two main results of this section, namely, A1′-PDL$_\omega \subsetneqq$ A2′-PDL$_\omega \subsetneqq$ CFL$_\omega$.

First, we exhibit a family of ω-CFL's in A2′-PDL$_\omega$, which cannot be 1′-accepted by any ω-PDA; this family consists of all ω-languages of the form Ld^{ω}, where L is a nonregular CFL not containing d. We next exhibit a family of a ω-CFL's which cannot be 2′-accepted by any ω-PDA; this second family is the class of all languages of the form $(Ld)^{\omega}$, where L is as above.

PROPOSITION 3.3.1. *For each nonregular language L over alphabet Σ_1 and symbol $d \notin \Sigma_1$, $Ld^{\omega} \notin$ A1′-PDL$_\omega$.*

Proof. Assume $L_1 = Ld^{\omega} \in$ A1′-PDL$_\omega$ for some nonregular L. Let $\Sigma = \Sigma_1 \cup d$. By Proposition 1.10, L_1 is not ω-regular. By Proposition 3.1.9, there exists an unrestricted ω-CFG $G = (V_N, \Sigma, P, S, 2^P)$ that generates L_1. We may assume that G is ϵ-free. By Theorem 3.2.6, G is real-self-embedding. We shall construct from G a new equivalent unrestricted ω-CFG which is non-real-self-embedding, leading to the contradictory conclusion that L_1 must be an ω-regular language.

Let $A \in V_N$ be a real-self-embedding variable in G, and let $u \in \Sigma^+$, where $A \overset{*}{\Rightarrow}_G uA\alpha$ for some $\alpha \in V^*$. A is reachable, hence $S \overset{*}{\Rightarrow} xA\beta$ for some $x \in \Sigma^*$ and thus $xu^{\omega} \in L_1$. It follows that every $u \in \Sigma^+$, s.t. $A \overset{*}{\Rightarrow} uA\alpha$, $\alpha \in V^*$, is in d^+.

Define P_1 to be the set of productions obtained by modifying P as follows: Let B be a new variable.

(1) Add to P the rules $A \to dB$, $B \to dB$;

(2) Every production $A \to \alpha A\beta \in P$, where $\alpha \in (V - A)^+$, $\beta \in V^*$, is replaced by $A \to \alpha B$.

(3) Define

$$K = \{X \in V_N \mid X \neq A \,\&\, A \overset{*}{\Rightarrow} xX\alpha,\ x \in \Sigma^+,\ \alpha \in V^*\},$$

$$K_1 = \{X \in V_N \mid X \neq A, X \notin K \,\&\, A \overset{*}{\Rightarrow} X\alpha,\ \alpha \in V^*\}.$$

The productions of the variables in K and K_1 will be modified as follows.

(a) Substitute B for A in the productions of K.

(b) Substitute B for A in the productions of K_1, excluding appearances of A as the first symbol of a right-hand side of a rule as $X \to A\alpha$, $X \in K_1$, $\alpha \in V^*$.

Let $G_1 = (V_N \cup \{B\}, \Sigma, P_1, S, 2^{P_1})$, where P_1 is the above modified version of P. To show that $L(G_1) = L(G)$, note that $L(G)$ is not changed by (1) and (2) above because if $\alpha \overset{*}{\Rightarrow} x \in V_T{}^+$, then $x \in d^+$, therefore only d^ω will be generated from this point on. Hence $L(G) = L(G_1)$, and moreover, G_1 has one less real-self-embedding variable (namely A) than G. The above procedure can therefore be repeated for all real-self-embedding variables in G, yielding a new unrestricted ω-CFG G' which is non-real-self-embedding and equivalent to G. This concludes the proof. ∎

Since clearly $Ld^\omega \in$ A2′-PDL$_\omega$ for each CFL L over Σ_1, we have

COROLLARY 3.3.2. A1′-PDL$_\omega \subsetneqq$ A2′-PDL$_\omega$.

PROPOSITION 3.3.3. *For each nonregular language L over alphabet Σ and $d \notin \Sigma$, $(Ld)^\omega \notin$ A2′-PDL$_\omega$.*

Proof. Suppose $(Ld)^\omega \in$ A2′-PDL$_\omega$. By Theorem 3.1.6, there exist for some integer $l \geq 1$, $2l$ sets $L_i^{(1)}$ and $L_i^{(2)}$, where $L_i^{(1)}$ is a CFL and $L_i^{(2)} \in$ A1′-PDL$_\omega$ for $1 \leq i \leq l$, s.t. $(Ld)^\omega = \bigcup_{i=1}^{l} L_i^{(1)} L_i^{(2)}$. Clearly for every $x \in L$, there is some $x_1 \in \Sigma^* d \cup \{\epsilon\}$ s.t. $x_1(xd)^\omega \in \bigcup_{i=1}^{l} L_i^{(2)}$.

Construct the following GSM $S = (\{q_1, q_2, q_3\}, \Sigma \cup \{d\}, \Sigma \cup \{d\}, \delta, q_1)$ where $\forall a \in \Sigma$, $\delta(q_1, a) = (q_1, \epsilon)$, $\delta(q_1, d) = (q_2, \epsilon)$, $\delta(q_2, a) = (q_2, a)$, $\delta(q_2, d) = (q_3, d)$ and $\forall a \in \Sigma \cup \{d\}$, $\delta(q_3, a) = (q_3, d)$. Then $S(\bigcup_{i=1}^{l} L_i^{(2)}) = Ld^\omega$. But $Ld^\omega \notin$ A1′-PDL$_\omega$ by Proposition 3.3.1, and by Proposition 3.1.10 also $\bigcup_{i=1}^{l} L_i^{(2)} \notin$ A1′-PDL$_\omega$, a contradiction. It follows that $(Ld)^\omega \notin$ A2′-PDL$_\omega$. ∎

COROLLARY 3.3.4. A2′-PDL$_\omega \subsetneqq$ CFL$_\omega$.

Remark 3.3.5. Using the same proof techniques as in Propositions 3.3.1 and 3.3.3 above, we can also prove that (a) $L_0 a^\omega \notin$ A1′-PDL$_\omega$, and (b) $L_0{}^\omega, L_1{}^\omega \notin$ A2′-PDL$_\omega$, where $L_0 = \{a^n b^n \mid n \geq 1\}$ and $L_1 = \{wcw^R \mid w \in \{a, b\}^*\}$. We conjecture that for every strict deterministic [10] nonregular language L, $L^\omega \notin$ A2′-PDL$_\omega$, but we have been able to establish this only for certain types of strict deterministic languages, e.g. those with a "natural" endmarker, like Ld in Proposition 3.3.3, or those which are bounded CFL's [6], like L_0 in (b) above.

As a by-product of the proof techniques developed in this chapter, we now obtain a new and elegant proof for the inclusion of CFL$_\omega$ in the ω-Kleene closure of the context-free languages. Unlike our original proof in [3], the new proof is straightforward and algebraic in nature, with no reference to ω-PDA's. However, unlike our original proof, this proof relies on the characterization of the nondeterministic ω-FSA languages as ω-KC(Reg) [13] (a result which was rather easy to prove), whereas our original proof was independent of it.

THEOREM 3.3.6. CFL$_\omega \subseteq \omega$-KC(CF).

Proof. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG. With no loss of generality, we may assume that F consists of only one set, denoted by F itself. Define new sets of symbols: $\bar{V} = \{\bar{A} \mid A \in V_N\}$ and $E_A = \{A_Q \mid Q \subseteq F\}$ for every $A \in V_N$. Also define a substitution h_1 on V by $h_1(a) = a$, $\forall a \in V_T$ and $h_1(A) = \bar{A} \cup E_A$, $\forall A \in V_N$. For each $A \in V_N$, $Q \subseteq F$, let $L(A, Q) = \{x \in V_T^* \mid$ there exists a derivation $d\colon A \overset{*}{\Rightarrow}_G x$ s.t. $\mathrm{Var}(d) = Q\}$ (Definition 0.3), and let $\bar{L}_A = L(\bar{G})$, where $\bar{G} = (V_N, V_T, P, A)$. Let $\bar{P}$ be obtained from P as follows. If $A \to \alpha \in P$, then for every decomposition $\alpha = \alpha_1 B \alpha_2$ s.t. $\alpha_1, \alpha_2 \in V^*$, $B \in V_N$, $\bar{P}$ will include the set $\{A \to \beta B \mid \beta \in h_1(\alpha_1)\}$. For each $H \subseteq F$, define the ω-RLG $G_H = (V_N, V_T \cup \bar{V} \cup \{E_A\}_{A \in V_N}, \bar{P}, S, H)$; also define the regular language $R_H = \{x \mid$ for some $l \geq 1$, $x = \prod_{i=1}^{l} x_i A_{Q_i}^{(i)}$, where $x_i \in V_T^*$, $A^{(i)} \in F$, $A_{Q_i}^{(i)} \in E_{A^{(i)}}$, for $1 \leq i \leq l$, $F - H \subseteq \bigcup_{i=1}^{l} Q_i\}$. Then $L_H^{(1)} = L(G_H) \cap (V_T \cup \bar{V})^* R_H^\omega$ is an ω-regular language.

Define a context-free substitution h_2 on $V_T \cup \bar{V} \cup \{E_A\}_{A \in V_N}$, as $\forall a \in V_T$, $h_2(a) = a$; $\forall \bar{A} \in \bar{V}$, $h_2(\bar{A}) = \bar{L}_A$ and $\forall A \in V_N$, $\forall Q \subseteq F$, $h_2(A_Q) = L(A, Q)$. Now, $L = L(G) = h_2(\bigcup_{H \subseteq F} L_H^{(1)})$, $\bigcup_{H \subseteq F} L_H^{(1)}$ is ω-regular, thus belongs to ω-KC(Reg) and hence to ω-KC(CF). Since the context-free languages are closed under context-free substitution, we obtain $L \in \omega$-KC(CF). ∎

4. NON-LEFTMOST GENERATION IN ω-CFG'S

This section deals with non-leftmost (nl) derivation in ω-CFG's. The class of ω-languages nl-generated by ω-CFG's, nl-CFL$_\omega$, is investigated. The main result is, somewhat surprisingly, that certain ω-CFL's cannot be nl-generated by ω-CFG's; however, any ω-language nl-generated by an ω-CFG is an ω-CFL. Hence leftmost generation is strictly more powerful.

A problem arising with respect to 3-generation of ω-sequences by nl-derivations is that parts of the required repetition sets may be a contribution of some "unreached" part of the sentential form. Therefore, for each variable occurring in the nl-derivation, one has to determine whether it belongs to the "reached" or to the "unreached" part of the sentential form, and in the latter case its possible contribution to the repetition sets must be taken into account. This is accomplished with the aid of the new notions of "self-providing" and "transient" sets of variables introduced in Subsection 4.1 below.

4.1. *Analysis of Non-Leftmost Derivations*

In this section we develop the basic tools for dealing with non-leftmost derivations. Some preliminary result on nl-CFL$_\omega$ are derived and the inclusion of nl-CFL$_\omega$ in CFL$_\omega$ is established.

PROPOSITION 4.1.1. nl-CFL$_\omega$ *equals the class of ω-languages generated by non-leftmost derivations by ω-CFG's with production repetition sets.*

Proof. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG with production repetition sets, where $V_N = \{A_i\}_{i=1}^{l}$, $P(A_i) = \{A_i \to \alpha_i^{(j)} \in P \mid 1 \leqslant j \leqslant k_i\}$ are the A_i-productions in P. Define a new set of variables $\bar{V}_N = \bigcup_{i=1}^{l} \{B_i^{(j)} \mid 1 \leqslant j \leqslant k_i\}$. Let $P_2 = \{B_i^{(j)} \to \beta \mid A_i \to \alpha_i^{(j)} \in P, \beta \in h(\alpha_i^{(j)})\}$, where h is defined by $h(a) = a\ \forall a \in V_T$ and $h(A_i) = \{B_i^{(j)}\}_{j=1}^{k_i}$ for $1 \leqslant i \leqslant l$. Assuming $A_1 = S$, define $P_1 = P_2 \cup \{B \to B_1^{(j)} \mid 1 \leqslant j \leqslant k_1\}$, where B is a new symbol. For every $H \subseteq \bar{V}_N$, let $P_H = \{A_i \to \alpha_i^{(j)} \mid B_i^{(j)} \in H\}$. Define the ω-CFG $G_1 = (\bar{V}_N \cup \{B\}, V_T, P_1, B, F_1)$, where $F_1 = \{H \subseteq \bar{V}_N \mid P_H \in F\}$. Clearly $L_{\text{nl}}(G_1) = L_{\text{nl}}(G)$.

The other direction follows directly from the definitions. ∎

PROPOSITION 4.1.2. *For any ω-CFG with ϵ-rules there can be constructed an ϵ-free ω-CFG G_1 s.t. $L_{\text{nl}}(G_1) = L_{\text{nl}}(G)$.*

Proof. Let $G = (V_N, V_T, P, S, F)$. For every $\alpha \in V^+$, define $NL(\alpha) = \{D \subseteq V_N \mid$ there exists a derivation d: $\alpha \overset{*}{\Rightarrow}_G \epsilon$ s.t. $\text{Var}(d) = D\}$ (Notation 0.3). Define $V_1 = V_N \times 2^{V_N}$, and for each $A \in V_N$, let $h(A) = A \times 2^{V_N}$ if $NL(A) = \varnothing$, $h(A) = A \times 2^{V_N} \cup \{\epsilon\}$ if $NL(A) \neq \varnothing$, and $h(a) = a$ for $a \in V_T$. Let $\alpha = \prod_{i=1}^{l} A_i$ and $\beta = \prod_{i=1}^{l} B_i \in h(\alpha)$, where $A_i \in V$ and $B_i \in V_1 \cup V_T \cup \{\epsilon\}$ for $i = 1,..., l$. Define $D(\beta) = \{\bigcup_{i=1}^{l} H_i \mid H_i = \varnothing$ if $B_i = A_i \in V_T$ or $B_i \in A_i \times 2^{V_N}$, and $H_i \in NL(A_i)$ if $B_i = \epsilon\}$. Also let $P_1 = \{(A, H) \to \beta \mid A \to \alpha \in P, \epsilon \neq \beta \in h(\alpha), H \in D(\beta)\} \cup \{S_1 \to (S, H) \mid H \subseteq V_N\}$, where S_1 is a new symbol. For every $D \in F$, define $\bar{D} = \{K \subseteq V_1 \mid \text{var}(K) = D\}$, where $\forall (A, H) \in V_1$, $\text{var}((A, H)) = \{A\} \cup H$ and $\forall K_1, K_2 \subseteq V_1$, $\text{var}(K_1 \cup K_2) = \text{var}(K_1) \cup \text{var}(K_2)$. Let $G_1 = (V_1 \cup \{S_1\}, V_T, P_1, S_1, \bar{F})$, where $\bar{F} = \{\bar{D} \mid D \in F\}$. By definition, G_1 is an ϵ-free ω-CFG and it can be easily verified that $L_{\text{nl}}(G_1) = L_{\text{nl}}(G)$. ∎

DEFINITION 4.1.3. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG, let d be the following nonleft derivation d: $\alpha_0 \Rightarrow_G \alpha_1 \Rightarrow_G \cdots \Rightarrow \alpha_l$ and let $\alpha_0 = \beta_0\gamma_0$ be a decomposition of α_0. For each $1 \leqslant i \leqslant l$, α_i can be decomposed into $\beta_i\gamma_i$, where $\beta_i(\gamma_i)$ is generated from $\beta_{i-1}(\gamma_{i-1})$ in step i of d. Let $\mathbf{d}_{\beta_0}$ denote the derivation $\beta_0 \overset{*}{\Rightarrow}_G \beta_1 \overset{*}{\Rightarrow}_G \cdots \overset{*}{\Rightarrow}_G \beta_l$ and let $\mathbf{d}_{\gamma_0}$ denote the derivation $\gamma_0 \overset{*}{\Rightarrow}_G \gamma_1 \overset{*}{\Rightarrow}_G \cdots \overset{*}{\Rightarrow}_G \gamma_l$, where for each $1 \leqslant i \leqslant l$ either $\beta_{i-1} \Rightarrow_G \beta_i$ and $\gamma_{i-1} = \gamma_i$ or $\gamma_{i-1} \Rightarrow_G \gamma_i$ and $\beta_{i-1} = \beta_i$.

DEFINITION 4.1.4. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG. Let d be a nonleft infinite derivation, $d\colon \alpha \Rightarrow_G \alpha_1 \Rightarrow_G \cdots \Rightarrow_G \alpha_i \Rightarrow \cdots$. For each $i, j \geqslant 1$ s.t. $i \leqslant j$, $d(i,j)$ will denote the derivation $\alpha_i \Rightarrow_G \cdots \Rightarrow_G \alpha_j$. Every sentential form α_i can be decomposed into $\alpha_i = \beta_i \gamma_i$ s.t.

(1) for any variable A in β_i s.t. $\beta_i = \gamma A \gamma'$ for some $\gamma, \gamma' \in V^*$, there is $j \geqslant i$ for which $d_\gamma(i,j)\colon \gamma \stackrel{*}{\Rightarrow}_G V_T{}^*$;

(2) for any variable A in γ_i s.t. $\gamma_i = \gamma A \gamma'$ for some $\gamma, \gamma' \in V^*$, for every $j \geqslant i$, $d_{\beta_i\gamma}(i,j)\colon \beta_i \gamma \stackrel{*}{\Rightarrow}_G V^* V_N V^*$.

$\beta_i(\gamma_i)$ will be referred to as the *reached* (*unreached*) *part* of α_i.

If in some ω-derivation, a string γ appears in the unreached part of the sentential form, then its possible impact on the derivation lies only in its set of variables (with multiplicities) and the potential contribution of this set toward obtaining the repetition set INV(d). Informally, for a given string γ, a set of variables H is *self-providing* if there exists an infinite nl derivation d' starting from γ s.t. the set of variables rewritten infinitely many times in d' is precisely H (hence we may say that H reproduces itself infinitely many times during the derivation d'). A set of variables H is *transient* for γ if there exists a finite derivation starting from γ, in which precisely the variables in H are rewritten.

DEFINITION 4.1.5. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG. For any $\gamma \in V^+$, *the class of self-providing sets* (SP(γ)) and *the class of transient sets* (TR(γ)) of γ are defined as

SP(γ) $= \{D \subseteq V_N \mid$ there exists an infinite nl derivation d starting in γ s.t. INV(d) $= D\}$.

(Note that d is not required to generate a string from $V_T{}^\omega$; here we are only concerned with the repetition set INV(d).)

TR(γ) $= \{D \subseteq V_N \mid$ there exists a derivation $d\colon \gamma \underset{G}{\stackrel{*}{\Rightarrow}} \gamma'$ for some $\gamma' \in V^*$ s.t. Var(d) $= D\}$.

Remark 4.1.6. It follows from the definition, that SP($\alpha\beta$) $= \{H_1 \cup H_2 \mid H_1 \in$ SP(α), $H_2 \in$ SP(β)$\}$ for every $\alpha, \beta \in V^+$. This property of SP will be often utilized in the constructions to follow.

THEOREM 4.1.7. *For any* ω-CFG $G = (V_N, V_T, P, S, F)$, $L_{\text{nl}}(G)$ *is an* ω-CFL.

Proof. With no loss of generality we may assume F consists of only one (variable) repetition set, denoted by F itself. We may also assume that $P = P_1 \cap P_2$, where the rules in P_1 are of the form $A \to \alpha$, $\alpha \in V_N{}^+$ and the rules of P_2 are of the form $A \to a$, $a \in \Sigma \cup \{\epsilon\}$. Define G_1 to be the CFG $G_1 = (V_N, V_T, P(F), S)$, where $P(F)$ denotes the set of F-productions in P, and define for any $\alpha \in V^*$, TR$_1$(α) $= \{D \subseteq F \mid$ there exists a derivation $d\colon \alpha \stackrel{*}{\Rightarrow}_{G_1} \gamma$ for some $\gamma \in V^*$ s.t. $D \subseteq$ Var(d)$\}$. (Note the slight difference between the definition of TR(α) (Definition 4.1.5) and that of TR$_1$(α) above.)

Construct the following U-ω-PDA $M = (K_1, V_T, \Gamma, [q_0, \varnothing], S, \{\bar{q} \times 2^F\})$, where $K_1 = \{q_0\} \times B \cup \{q_1\} \times 2^F \times 2^F \cup \{\bar{q}\} \times 2^F$, $\Gamma = V_N \cup V_T \cup \{Z\}$, $B = 2^{2^F}$, and Z is a new symbol. δ is defined as follows.

Stage 1. $\forall H \in B$, $\forall A \in V_N$, $([q_0, H], \epsilon) \in \delta([q_0, H], a, A)$ if $A \to a \in P_2$ and

$([q_0, H], \gamma) \in \delta([q_0, H], \epsilon, A)$ if $A \to \gamma \in P_1$. If $A \to \gamma \in P_1$, then for every $\gamma_1, \gamma_2 \neq \epsilon$ s.t. $\gamma_1\gamma_2 = \gamma$, $([q_0, H_1], \gamma_1 Z) \in \delta([q_0, H], \epsilon, A)$, where $H_1 = \{D_1 \cup D_2 \mid D_1 \in H, D_2 \in \mathrm{SP}(\gamma_2)\}$.

In Stage 1 the machine M nondeterministically simulates some finite derivation d: $S \overset{*}{\Rightarrow} \alpha$ in G, and in each step nondeterministically chooses a decomposition $\gamma_1\gamma_2$ (denoted by Z on the pushdown store) of the r.h.s. of the rule into the reached and unreached part. The addition of γ_2 to the unreached part of α is only represented by its contribution to the collection H_1 of self-providing sets of the unreached part, rememebered by the second component of the finite control.

Stage 2. For every $A \in V_N$, $H \subseteq 2^N$ and for every $K \subseteq F$ s.t. $K \in H$, $([q_1, F - K, \varnothing], A) \in \delta([q_0, H], \epsilon, A)$.

In Stage 2 the machine nondeterministically chooses one of the self-providing sets, K, belonging to the collection H of self-providing sets of the unreached part of α and contained in F. From now on M only has to check that the variables in $F - K$ also will be rewritten infinitely many times. This is done using the third component of the state above. Initially this third component is set to ϕ. In Stage 3 below, we accumulate in this component the variables in F rewritten in the simulated derivation, and also those variables in F which could be rewritten in the newly generated unreached part (namely those appearing in transient sets of γ_2).

Stage 3. $K, H \subseteq F$, $\forall A \in F$, (i) if $A \to a \in P_2$, then $([q_1, F - K, H \cup \{A\}], \epsilon) \in \delta([q_1, F - K, H], a, A)$; (ii) if $A \to \gamma \in P_1$, then (a) $([q_1, F - H, K \cup \{A\}], \gamma) \in \delta(q_1, F - K, H], \epsilon, A)$, and (b) for every $\gamma_1, \gamma_2 \neq \epsilon$ s.t. $\gamma_1\gamma_2 = \gamma$, $([q_1, F - K, H \cup H_1 \cup \{A\}], \gamma_1 Z) \in \delta([q_1, F - K, H], \epsilon, A)$ for each $H_1 \in \mathrm{TR}_1(\gamma_2)$.

Stage 4. $K, H \subseteq F, \forall A \in V_N$, $([\bar{q}, F - K], A) \in \delta([q_1, F - K, H], \epsilon, A)$ if $F - K \subseteq H$ and $([q_1, F - K, \varnothing], A) \in \delta([\bar{q}, F - K], \epsilon, A)$.

In Stage 4, whenever the third component of the state contains all of $F - K$, M, enters state $[q, F - K]$, and then reenters $[q_1, F - K, \varnothing]$, where Stage 3 starts all over again. Thus entering state $[\bar{q}, F - K]$ denotes the event of having rewritten each of the variables in $F - K$ at least one more time since the last time M has been in this same state.

Clearly $T_2(M) = L_{\mathrm{nl}}(G)$, hence $L_{\mathrm{nl}}(G)$ is an ω-CFL. ▮

4.2. *Real-Self-Embedding in* nl-*Derivations*

Paralleling the proof for A2′-$\mathrm{PDL}_\omega \subsetneqq \mathrm{CFL}_\omega$ in Section 3, the proof for the non-equivalence of nl-CFL_ω and CFL_ω also relies heavily on the notion of real-self-embedding ω-CFG's. In this section we derive an analog of Theorem 3.2.6 for nonleft derivations, which will be utilized in the next subsection for proving the main result.

Definition 4.2.1. In an ω-CFG G, variables of the following types will be called *dummy*.

0.—Variables with no productions;

1.—Variables of which the only productions are of the form $A \to B$, where B is of type 0 above;

2.—Variables of which the only productions are of the form $A \to B$, where B is of type 3 below;

3.—Variables of which the only productions are of the form $A \to A$.

Let G be an ω-CFG. We now add to G a set of dummy variables which are duplicates of the variables of G and which will be used in self-providing and transient sets in the derivations of G. The next definition and notation will be frequently used in the rest of this section.

DEFINITION 4.2.2. *Let* $G = (V_N, V_T, P, S, F)$ be an ω-CFG. For $i = 1, 2, 3$, define $V_N^{(i)} = \{A^{(i)} \mid A \in V_N\}$ and $V_N^{(d)} = \{Z\} \cup \{\bigcup_{i=1}^{3} V_N^{(i)}\}$, where Z is a new symbol. For $V_N^{(d)}$ define the following set of productions P_d

$$P_d = \{A^{(1)} \to Z \mid A \in V_N\} \cup \{A^{(2)} \to A^{(3)} \mid A \in V_N\} \cup \{A^{(3)} \to A^{(3)} \mid A \in V_N\}.$$

With the aid of Pd, a distinction will be made between the inclusion of a variable in a transient set, in which case we use its duplicate from $V_N^{(1)}$, and its appearance in a self-providing set, in which case its duplicate from $V_N^{(3)}$ will be used. The phase in which a variable is generated to be included in a self-providing set is designated by the use of its duplicate in $V_N^{(2)}$. The above distinction will be particularly useful in the proof of Theorem 4.2.7 below.

DEFINITION 4.2.3. Let G, $V_N^{(d)}$, and P_d be as above. Let p_r be the projection $p_r(A^{(i)}) = A$ $\forall A^{(i)} \in V_N^{(i)}$, $i = 1, 2, 3$, and $\forall A \in V_N \cup \{Z\}$, $p_r(A) = A$. p_r is extended to strings and sets of strings in the standard way.

We now define a *concatenating function* to be any arbitrarily chosen function $\beta\colon 2^V \to V^*$ s.t. for any set $D \subseteq V$, $\beta(D)$ is a string made up of the variables of D in some arbitrary order. For $i = 1, 2, 3$, define $\beta_i\colon 2^V \to V_N^{(i)*}$, where $\forall D \subseteq V$, $p_r(\beta_i(D)) = \beta(D)$.

DEFINITION 4.2.4. Let G, $V_N^{(d)}$, and P_d be as above and let $A \in V_N$. Define $P_m(A) = \{X \to \beta \mid \beta \in h(\alpha),\ X \to \alpha \in P - P(A)\}$, where h is the substitution $h(X) = X$ for $X \in V - \{A\}$ and $h(A) = \{\beta_1(D) \mid D \in \mathrm{TR}(A)\} \cup \{\beta_2(D) \mid D \in \mathrm{SP}(A)\}$. $P_m(A)$ is a modification of the set of all productions in P, excluding those of A, obtained by substituting all appearances of A by the concatenated version of its self-providing and transient sets.

Remark 4.2.5. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG; then there exists an ω-CFG G_1 s.t. $L_{\mathrm{nl}}(G_1) = L_{\mathrm{nl}}(G)$ and all nondummy variables in G_1 are reachable. To show this, suppose X is a nondummy variable in G, which is not reachable from S. Clearly $X \neq S$. Construct a new ω-CFG $G_1 = (\{V_N - X\} \cup V_N^{(d)},\ V_T, P_m(X) \cup P_d, S, \bar{F})$, where $\bar{F} = \{D \mid D \subseteq \{V_N - \{X\}\} \cup V_N^{(d)},\ p_r(D) \in F\}$ and $V_N^{(d)}$, P_d, $P_m(X)$, and p_r are as in Definitions 4.2.2–4.2.4. Clearly $L_{\mathrm{nl}}(G_1) = L_{\mathrm{nl}}(G)$. If there are several nondummy nonreachable variables in G, the above procedure can be carried out simultaneously on all of them.

The following is a modification of Lemma 3.2.5(c).

LEMMA 4.2.6. *Let* $G = (V_N, V_T, P, S, F)$ *be a non-real-self-embedding* ω-CFG,

$X \in V_N$ and $K \subseteq V_N$; then $L_K = \{x \in V_T^* \mid$ there exists a derivation d: $S \overset{*}{\Rightarrow}_G xX\alpha$ for some $\alpha \in V^*$ s.t. *$K \in \mathrm{SP}(\alpha)\}$ is a regular language.*

Proof. Modifying the proof of part (c) of Lemma 3.2.5, define $V_1, L_A, \bar{V}_1$ and homomorphism h as in that proof, and also define $Q = \{H \mid H \subseteq 2^{V_N}\}$, $Q_K = \{H \subseteq 2^{V_N} \mid K \in H\}$, and the alphabet $C_K = \{c_H \mid H \in Q_K\}$. Let $\bar{P}$ be obtained from P as follows. For each $A \to \alpha \in P$ and for every decomposition $\alpha = \alpha_1 B \beta$, $\alpha_1 \in (V_T \cup V_1)^*$, $B \in V_N$ and $\beta \in V^*$, $\bar{P}$ will include, for each $D \in Q$, $[A, D] \to h(\alpha_1)[B, H]$ where $H = \{H_1 \cup H_2 \mid H_1 \in D, H_2 \in \mathrm{SP}(\beta)\}$. In case $B = X$ and H above is in Q_K, $\bar{P}$ will also include $[A, D] \to h(\alpha_1)\, c_H$. Define the right-linear grammar $G_1 = (V_N \times Q, V_T \cup \bar{V}_1 \cup C_K, \bar{P}, [S, \varnothing])$ and let h_1 be the regular substitution $h_1(a) = a$ for $a \in V_T$ and $h_1(\bar{A}) = L_A$ for $A \in V_1$. Then $L_K = h_1(L(G_1)/C_K)$ is a regular language. ∎

Following Theorem 3.2.6 we have

THEOREM 4.2.7. *For any non-real-self-embedding ϵ-free ω-CFG G, $L_{\mathrm{nl}}(G)$ is an ω-regular language.*

Proof. Let $G = (V_N, V_T, P, S, F)$. By Remark 4.2.5, we may assume that every nondummy variable in V_N is reachable. The proof parallels that of Theorem 3.2.6. As in that proof, we separate two cases.

Case 1. S is reachable from each nondummy variable. Following the argument in Lemma 3.2.5(a), we conclude that for each $X \in V_N$, the X-productions are in the forms (1) right linear, (2) $X \to A\alpha$, $A \in V_N$, $\alpha \in V^+$, where $A \not\Rightarrow^* V_T^*$, (3) $X \to xZ\gamma_1$ or $X \to AZ\gamma_1$, $x \in V_T^*$, $A \in V_N$, $\gamma_1 \in V^*$, and Z is a dummy variable. With no loss of generality we may assume that F consists of only one set, denoted by F itself. Let $\bar{P}$ include all right-linear productions in P with the following additions. For each $X \to A\gamma \in P$, $\bar{P}$ includes the set $\{X \to \beta_3(D)A \mid D \in \mathrm{SP}(\gamma)\}$ and in case $X \in F$, $\bar{P}$ also includes $\{X \to \beta_1(D)A \mid D \in \mathrm{TR}(\gamma)\}$, where $\mathrm{SP}(\gamma)$ and $\mathrm{TR}(\gamma)$ are as in Definition 4.1.5. For every $H \subseteq F$, define $G_H = (V_N, V_T \cup V_N^{(1)} \cup V_N^{(3)}, \bar{P}, S, H)$. $L_{\mathrm{nl}}(G)$ is an ω-regular language since G_H is an ω-RLG. For each $D \subseteq F$, define $R_D = \{x \in (V_T \cup V_N^{(3)})^* \mid D^{(3)}$ is the set of variables from $V_N^{(3)}$ in x and $p_r(D^{(3)}) = D\}$, $\bar{R}_D = \{x \in (V_T \cup V_N^{(1)})^* \mid D^{(1)}$ is the set of variables in x and $p_r(D^{(1)}) = D\}$. Then $L_H = \bigcup_{F-H \subseteq D \cup D', D, D' \subseteq F} (L_{\mathrm{nl}}(G_H) \cap R_D \bar{R}_{D'}^{\omega})$ is an ω-regular language. Since $L_{\mathrm{nl}}(G) = \bigcup_{H \subseteq F} L_H$, $L_{\mathrm{nl}}(G)$ is an ω-regular language.

Case 2. There is a nondummy variable X_1 s.t. for no $x \in V_T^*$, $\alpha \in V^*$, does $X_1 \overset{*}{\Rightarrow} xS\alpha$. We shall prove that $L_{\mathrm{nl}}(G)$ is ω-regular by induction on the number n of nondummy variables in G.

For $n = 1$ clearly S must be a dummy variable and $L(G) = \varnothing$. Assume the assertion holds for $n = k$, and let the number of the nondummy variables in V_N be $k + 1$. Without loss of generality we may assume that F consists of only one set, denoted by F itself.

First assume $X_1 \overset{*}{\Rightarrow} V_T^*$.

Subcase (a). $X_1 \notin F$. Define a CFG $G_1 = (V_N, V_T, P - P(S), X_1)$; then $L_1 = L(G_1) \subseteq V_T^*$ is regular by Lemma 3.2.5(a). For every $K \subseteq V_N$, L_K, defined as in Lemma 4.2.6, is regular. For every $H \subseteq F$, define $G_H = ((V_N - S) \cup V_N^{(a)}, V_T, P_m(S) \cup P_d, X_1, \bar{H})$, where $\bar{H} = \{K_1 \subseteq (V_N - \{S\}) \cup V_N^{(d)} \mid H \subseteq p_r(K_1) \subseteq F\}$ and $V_N^{(d)}$, P_d, $P_m(S)$ and p_r are as in Definitions 4.2.2–4.2.4. G_H has all nondummy variables of G except for S, thus bu induction hypothesis $\bar{L}_H = L_{\text{nl}}(G_H)$ is ω-regular for every $H \subseteq F$. Define $L_3 = \bigcup_{K \cup H = F} L_K \bar{L}_H$. Let $G_2 = ((V_N - \{X_1\}) \cup V_N^{(d)}, V_T \cup \{X_1\}, P_0(X) \cup P_m(X_1) \cup P_d, S, \bar{F})$, where $\bar{F} = \{F_1 \mid p_r(F_1) = F\}$. Again by induction hypothesis $\bar{L}_2 = L_{\text{nl}}(G_2) \cap (V_T \cup \{X_1\})^* V_T^{\omega}$ is an ω-regular language. Since $L = L_3 \cup L_4$, where $L_4 = h(\bar{L}_2)$ and $h = h(X_1, L_1)$ (see Definition 3.2.2), L is an ω-regular language.

Subcase (b). $X_1 \in F$. Let G_1 be as in Subcase (a) above and E, h_1, G_3, L_Q, R_H, h_2 as in Subcase (b) in Theorem 3.2.6. For every $H \subseteq F - \{X_1\}$, define the ω-CFG's $G_4 = ((V_N - \{X_1\}) \cup V_N^{(d)}, V_T \cup \{X_1\}, P_m(X_1) \cup P_d, S, F_{X_1})$, where $F_{X_1} = \{H_1 \mid H_1 \subseteq (V_N - \{X_1\}) \cup V_N^{(1)} \cup V_N^{(3)}, p_r(H_1) = F\}$ and $G_H' = ((V_N - \{X_1\}) \cup V_N^{(d)}, V_T \cup \{X_1\}, P_m(X_1) \cup P_d, S, F_H)$, where $F_H = \{H_1 \subseteq (V_N - \{X_1\}) \cup V_N^{(d)} \mid p_r(H_1) = H\}$.

Let $L_4' = L_{\text{nl}}(G_4) \cap (V_T \cup \{X_1\})^* V_T^{\omega}$ and $L_H' = h_1(L_{\text{nl}}(G_H')) \cap (V_T \cup \{X_1\})^* R_H^{\omega}$. L_4' and L_H' are ω-regular languages by hypothesis. Therefore $L_4 = \{\bigcup_{H \subseteq F - \{X_1\}} h_2(L_H')\} \cup h_2(L_4')$ is ω-regular. Again we have $L = L_3 \cup L_4$, where L_3 is as in Subcase (a). Hence L is ω-regular.

Now assume $X_1 \not\Rightarrow^* V_T^*$. In Subcase (a) $L = L_3 \cup (L_{\text{nl}}(G_2) \cap V_T^{\omega})$ and in Subcase (b), $L = L_3$; thus in both subcases L is ω-regular. This concludes the proof. ∎

4.3. *The Nonequivalence of* nl-CFL$_\omega$ *and* CFL$_\omega$

We now define a new type of generation in ω-CFG's, denoted 4-nl, which turns out to be a useful tool both for characterizing nl-CFL$_\omega$ and for proving the main result of this chapter, namely that there exists an ω-CFL which cannot be nl-generated by any ω-CFG.

DEFINITION 4.3.1. For any ω-CFG $G = (V_N, V_T, P, S, F)$, define $L_{4\text{-nl}}(G) = \{\sigma \in V_T^{\omega} \mid$ there exists an infinite nonleft derivation $d: S \Rightarrow_G^{\omega} \sigma$ s.t. $\exists H \in F$ for which $H \subseteq \text{INV}(d)\}$; $L_{4\text{-nl}}(G)$ is the *ω-language* 4-nl *generated by G.*

PROPOSITION 4.3.2. nl-CFL$_\omega$ *equals the class* $\mathscr{L}_4$ *of ω-languages of the form* $L = \bigcup_{i=1}^{l} L_i L_i'$, *where* $l \geqslant 1$ *and for each* $1 \leqslant i \leqslant l$, L_i *is a* CFL *and* L_i' *is an ω-language* 4-nl *generated by some ω-*CFG.

Proof. Following the definitions one can prove that $\mathscr{L}_4 \subseteq$ nl-CFL$_\omega$. Let $L = L_{\text{nl}}(G)$, where $G = (V_N, V_T, P, S, F)$ is an ω-CFG. Without loss of generality we may assume that F consists of only one set of variables, denoted by F itself. Let $C = \{(X, K) \mid X \in F, K \subseteq F\}$, and for each $(X, K) \in C$, define $L_{(X,K)} = \{x \in V_T^* \mid$ there exists a derivation $d: S \overset{*}{\Rightarrow}_G xX\alpha$ E $K \in \text{SP}(\alpha)\}$; also define the ω-CFG $G_{(X,K)} = (V_N, V_T, P(F), X, \{F - K\})$ and let $L'_{(X,K)} = L_{4\text{-nl}}(G_{(X,K)})$. Clearly for each $(X, K) \in C$, $L_{(X,K)}$ is a CFL, and $L = \bigcup_{(X,K) \in C} L_{(X,K)} L'_{(X,K)}$. Hence nl-CFL$_\omega \subseteq \mathscr{L}_4$ and the assertion follows. ∎

Following the lines of the proof of Proposition 4.1.2, we can also prove

LEMMA 4.3.3. *For any ω-*CFG *G with ϵ-rules, there can be constructed an ϵ-free ω-*CFG G_1 s.t. $L_{4\text{-nl}}(G_1) = L_{4\text{-nl}}(G)$.

LEMMA 4.3.4. *The family of ω-languages* 4-nl *generated by* ω-CFG's *is closed under* GSM *mappings.*

Proof. Let $G = (V_N, V_T, P, S, F)$ be an ω-CFG over Σ. Following (b) in Proposition 3.1.10, for any finite substitution $h: \Sigma \to 2^{\Delta^*}$, there exists an ω-CFG G_1 s.t. $L_{4\text{-nl}}(G_1) = h(L_{4\text{-nl}}(G)) \cap \Delta^\omega$. By Proposition 1.11 it suffices to show that the family is closed under intersection with A1′-ω-regular languages. Thus let R be an A1′-ω-regular language and let $M = (K, \Sigma, \delta, p_0, H)$ be an ω-DSFA with a single designated set H that 1′-accepts R. If $R = \varnothing$ we are done; thus assume $p_0 \in H$. Define $G_1 = (V_1, \Sigma, P_1, S_1, F_1)$, where $V_1 = H \times V \times H \cup \{S_1\}$ and S_1 is a new symbol. P_1 is defined as follows. For every $q \in H$, $S_1 \to (p_0, S, q) \in P_1$; for every $q, p \in H$ and $a \in \Sigma$, $(p, a, q) \to a \in P_1$ if $\delta(p, a) = q$ and for each $A \in V_N$, $p \in H$ $(p, A, p) \to \epsilon$ if $A \to \epsilon \in P$; also let $(p, A, q) \to \prod_{i=0}^{k-1} (q_i, B_{i+1}, q_{i+1}) \in P_1$ for every production $A \to \prod_{i=1}^{k} B_i \in P$, $B_i \in V$ for $1 \leqslant i \leqslant k$, and for every $q_i \in K$, $0 \leqslant i \leqslant k$, s.t. $q_0 = p$, $q_k = q$. Let $F_1 = \{D \subseteq H \times V \times H \mid \exists D' \in F \text{ s.t. } D' \subseteq P_2(D) \subseteq D' \cup V_T\}$, where $P_2(D)$ is the projection of the second component of D. Clearly $L_{4\text{-nl}}(G_1) = L_{4\text{-nl}}(G) \cap R$. By the above and Proposition 1.11 the assertion follows. ▌

LEMMA 4.3.5. *For no* ω-CFG G *does* $L_{4\text{-nl}}(G) = L = \{a^n b^n \mid n \geqslant 1\} d^\omega$.

Proof. Let $\Sigma = \{a, b, d\}$. Suppose $G = (V_N, \Sigma, P, S, F)$ is an ω-CFG for which $L = L_{4\text{-nl}}(G)$. By Lemma 4.3.3 we may assume that G is ϵ-free. Let $\bar{F} = \{K \subseteq V_N \mid \exists H \in F \text{ s.t. } H \subseteq K\}$; then $L_{4\text{-nl}}(G) = L_{\text{nl}}(\bar{G})$, where $\bar{G} = (V_N, \Sigma, P, S, \bar{F})$. Since L is not ω-regular by Proposition 1.10, $\bar{G}$, and therefore also G, are real-self-embedding. We now define a slight variation of SP(α), the class of self-providing sets (Definition 4.1.5). For every $\alpha \in V^*$, let

$$\text{SP}_1(\alpha) = \{D \subseteq V_N \mid \text{there exists an infinite nonleft derivation } d\text{: } \alpha \underset{(G)}{\overset{\omega}{\Rightarrow}} \sigma \in V_T^\omega$$
$$\text{s.t. } \text{INV}(d) = D\}.$$

Note that here d is only allowed to generate an infinite string. Let A be a real-self-embedding variable. By the argument in Remark 4.2.5 we may assume that A is reachable from S in the derivation of some $\sigma \in L$. We claim that every word $u \in \Sigma^+$ s.t. $A \overset{*}{\Rightarrow}_G uA\beta$ for some $\beta \in V^*$, is in d^+. We have two cases.

Case 1. There is a derivation in G $S \overset{*}{\Rightarrow} xA\alpha \overset{*}{\Rightarrow} xx_1\alpha \Rightarrow^\omega xx_1\sigma \in L$. As the derivation is nonleft also $xu^\omega \in L$. Hence $u \in d^+$.

Case 2. In every derivation of G of the form $S \overset{*}{\Rightarrow} xA\alpha \Rightarrow^\omega x\sigma \in L$, $\alpha \in V^*$ is never reached; in this case $xu^*\sigma \subseteq L$. By the structure of L, $u \in d^+$.

We shall now construct from G a non-real-self-embedding ω-CFG G_1 s.t. $L_{4\text{-nl}}(G) = L_{4\text{-nl}}(G_1)$, contradicting the fact that L is not ω-regular. The rest of the proof will follow the lines of Proposition 3.3.1. Let P_1 be obtained by modifying P as follows. Let B be a new variable.

(1) Add to P the productions $B \to dB$, $A \to dB\beta_2(D)$ for each $D \in \text{SP}_1(A)$.

(2) Every production $A \to \alpha A\beta \in P$, where $\alpha \in (V - A)^+$, $\beta \in V^*$, will be replaced by all productions of the form $A \to \alpha h(A\beta)$, where h is the substitution $h(A) = \{B\} \cup \{\beta_2(D) \mid D \in \mathrm{SP}(A)\} \cup \{\beta_1(D_1) \mid D_1 \in \mathrm{TR}(A)\} \cup \{B\beta_2(H) \mid H \in \mathrm{SP}_1(A)\}$ and $h(X) = X$ for all $X \in V$, $X \neq A$.

(3) Define $K = \{X \in V_N \mid X \neq A \;\&\; A \overset{*}{\Rightarrow} xX\alpha,\; x \in \Sigma^+,\; \alpha \in V^*\}$ and $K_1 = \{X \in V_N - K \mid X \neq A \;\&\; A \overset{*}{\Rightarrow} X\alpha, \alpha \in V^*\}$.

The productions of the variables in K and K_1 will be modified by substituting $h(A)$ for A in (a) all productions of the variables in K; (b) all productions of the variables in K_1, excluding occurrences of A at the beginning of a right-hand side of a rule of the form $X \to A\alpha$, $X \in K_1$, $\alpha \in V^*$, where A remains unaltered.

(4) All the other rules in P remain unchanged in P_1.

In (2) and (3) above, substituting B for A takes care of the case $A \overset{*}{\Rightarrow}_G V_T^+$ in the derivation, substituting $B\beta_2(H)$, where $H \in \mathrm{SP}_1(A)$, for A takes care of the case $A \Rightarrow_G^\omega d^\omega$ and substituting A by $\beta_2(D)$ and $\beta_1(D_1)$, where $D \in \mathrm{SP}(A)$, $D_1 \in \mathrm{TR}(A)$, serves the case A is never reached in the course of the derivation. Let $G_1 = (V_N \cup \{B\} \cup V_N^{(d)}, \Sigma, P_1 \cup P_d, S_1, F_1)$, where P_1 is the above-modified version of P and $F_1 = \{D \subseteq V_N \cup V_N^{(d)} \mid p_r(D) \in F\} \cup \{D \cup \{B\} \mid D \subseteq V_N \cup V_N^{(d)}, A \notin D \text{ and } p_r(D \cup \{A\}) \in F\}$. It can be verified that $L_{4\text{-nl}}(G_1) = L_{4\text{-nl}}(G)$. Moreover, G_1 has one less real-self-embedding variable (namely A) than G. The above procedure can therefore be repeated for all real-self-embedding variables in G, until a new ω-CFG G' is obtained, which is non-real-self-embedding and s.t. $L_{4\text{-nl}}(G') = L_{4\text{-nl}}(G)$, leading to the above contradiction. This concludes the proof. ∎

We are finally ready to exhibit an ω-CFL which cannot be generated by any ω-CFG by nonleft derivations.

PROPOSITION 4.3.6. $\{a^n b^n \mid n \geqslant 1\}^\omega \notin \text{nl-CFL}_\omega$.

Proof. Let $L = \{a^n b^n \mid n \geqslant 1\}$ and suppose $L^\omega \in \text{nl-CFL}_\omega$. By Proposition 4.3.2 there exist for some $l \geqslant 1$, $2l$ sets $L_i^{(1)}$ and $L_i^{(2)}$ s.t. $L^\omega = \bigcup_{i=1}^{l} L_i^{(1)} L_i^{(2)}$, where for $1 \leqslant i \leqslant l$, $L_i^{(1)}$ is a CFL and $L_i^{(2)} = L_{4\text{-nl}}(G_i)$ for some ω-CFG $G_i = (V_N^{(i)}, \Sigma, P_i, S_i, H^{(i)})$ and $\Sigma = \{a, b\}$. We may assume $V_N^{(i)} \cap V_N^{(j)} = \varnothing$ for $i \neq j$. Let $\bar{G} = (\bar{V}_N \cup \{S\}, \Sigma, \bar{P}, S, H)$, where $\bar{V}_N = \bigcup_{i=1}^{l} V_N^{(i)}$, S a new symbol, $\bar{P} = \{\bigcup_{i=1}^{l} P_i\} \cup \{S \to S_i \mid 1 \leqslant i \leqslant l\}$, $H = \bigcup_{i=1}^{l} H^{(i)}$; then $L_{4\text{-nl}}(\bar{G}) = \bigcup_{i=1}^{l} L_i^{(2)}$. The proof proceeds similarly to that of Proposition 3.3.3. Since for every $y \in L$, $\exists x \in a^* b^+$ s.t. $xy^\omega \in \bigcup_{i=1}^{l} L_i^{(2)}$, one can construct a GSM which maps $\bigcup_{i=1}^{l} L_i^{(2)}$ onto $\{a^n b^n \mid n \geqslant 1\}\, d^\omega$.

By Lemma 4.3.4, there exists an ω-CFG G' s.t. $L_{4\text{-nl}}(G') = Ld^\omega$, which contradicts Lemma 4.3.5 above. We conclude that $L^\omega \notin \text{nl-CFL}_\omega$. ∎

From the above and Theorem 4.1.7 we have

THEOREM 4.3.7. $\text{nl-CFL}_\omega \subsetneqq \text{CFL}_\omega$.

Remark 4.3.8. In Section 2 we saw that nothing is added to CFL_ω by using ω-regular control sets in leftmost derivations of ω-CFG's. However, though $\text{nl-CFL}_\omega \subsetneqq \text{CFL}_\omega$, using regular control sets in nonleft derivations leads us outside the class CFL_ω, as the following example (motivated by [14]) will show.

Let $G = (\{X_0, X, Y, Z, W\}, \{a, b, c\}, P, X_0, 2^P)$ be an unrestricted ω-CFG, where P is

$$p_1\colon X_0 \to XYZW, \quad p_2\colon X \to aX, \quad p_3\colon Y \to bY,$$
$$p_4\colon Z \to cZ \quad p_5\colon X \to a, \quad p_6\colon Y \to b,$$
$$p_7\colon Z \to c, \quad p_8\colon W \to aW,$$

Let $C = p_1(p_2 p_3 p_4)^* p_5 p_6 p_7 p_8^\omega$. Clearly C is an ω-regular set but $L_C(G) = \{a^n b^n c^n \mid n \geq 1\} a^\omega$ is not an ω-CFL. ∎

Conclusion and Preview of Further Work

In this paper the theory of ω-languages and ω-machines, initiated in [3], was further developed, providing a deeper insight into the properties of ω-grammars, ω-CFL's. and ω-PDA's. Emphasis was placed upon the study and comparison of various generation modes in ω-grammars, and of the various recognition types in ω-PDA's. The families Ai-PDL_ω, $i = 1, 1', 2, 2', 3$ were investigated and type 3 (also type 2) acceptance shown to be the most powerful acceptance mode in ω-PDA's.

We were particularly concerned with the analysis of non-leftmost generation in ω-CFG's, providing the tools for establishing the proper inclusion of nl-CFL_ω within CFL_ω. The following hierarchy of families of ω-CFL's was obtained

$$\omega\text{-Regular} \subsetneqq A1'\text{-}PDL_\omega \subsetneqq A1\text{-}PDL_\omega = A2'\text{-}PDL_\omega \subseteqq \text{nl-}CFL_\omega \subsetneqq CFL_\omega .$$

A problem which remains open is whether the inclusion of $A2'$-PDL_ω in nl-CFL_ω is proper. Another open problem concerns the validity of the conjecture stated in Remark 3.3.5, namely, that for every strict deterministic nonregular language L, the ω-CFL L^ω cannot be $2'$-accepted by any ω-PDA.

A subsequent paper [4] is devoted to the deterministic variants of ω-PDA's. A rich hierarchy of deterministic and quasi-deterministic ω-CFL families is obtained, differing in structure from the above rather simple hierarchy of nondeterministic ω-CFL families. An extensive study of the families is made, algebraic characterizations are derived, and certain problems, generally undecidable, are shown to be decidable within some of the families.

Still another paper [5] presents the theory of ω-type Turing machines and type 0 ω-languages. The theory differs considerably from the classical theory of Turing machines and due to the nonterminating nature of the input tapes, a few rather peculiar results are obtained.

Acknowledgment

The authors wish to thank the anonymous referee for his helpful suggestions.

References

1. J. R. Buchi, "On a Decision Method in Restricted Second Order Arithmetic," International Congress Logic Methodology and Philosophy of Science, Stanford, Calif., 1960.
2. N. Chomsky, On certain formal properties of grammars, *Inform. Contr.* **2** (1959), 137–167.
3. R. S. Cohen and A. Y. Gold, Theory of ω-languages. I: Characterizations of ω-context-free languages, *J. Comput. System Sci.* **15** (1977), 169–184.
4. R. S. Cohen and A. Y. Gold, "ω-Computations on Deterministic Pushdown Machines," *J. Comput. System Sci.*, to appear.
5. R. S. Cohen and A. Y. Gold, "ω-Computations on Turing Machines," TR. No. 76, July 1976, Computer Science Dept., Technion-IIT, Haifa, Israel.
6. S. Ginsburg, "The Mathematical Theory of Context free Languages," McGraw–Hill, New York, 1966.
7. S. Ginsburg and E. Spanier, Control sets on grammars, *Math. Systems Theory* **2** (1968), 177–177.
8. E. J. Hopcroft and J. D. Ullman, "Formal Languages and their Relation to Automata," Addison–Wesley, Reading, Mass., 1969.
9. R. Hosseley, "Finite Tree Automata and ω-automata," TR. MAC TR-102, MIT, Cambridge, 1972.
10. M. A. Harrison and I. M. Havel, Strict deterministic grammars, *J. Comput. System Sci.* **7** (1973), 237–277.
11. L. H. Landweber, Decision problems for ω-automata, *Math. Systems Theory* **3** (1969), 376–384.
12. G. H. Matthews, A note on asymmetry in phrase structure grammars, *Inform. Contr.* **7** (1964), 360–365.
13. R. McNaughton, Testing and generating infinite sequences by a finite automaton, *Inform. Contr.* **9** (1966), 521–530.
14. A. Salomaa, "Formal Languages," Academic Press, New York, 1973.