RAIRO. INFORMATIQUE THÉORIQUE

H. A. MAURER
A. SALOMAA
D. WOOD

Synchronized EOL forms under uniform interpretation

RAIRO. Informatique théorique, tome 15, nº 4 (1981), p. 337-353
<http://www.numdam.org/item?id=ITA_1981__15_4_337_0>

R.A.I.R.O. Informatique théorique/Theoretical Informatics
(vol. 15, n° 4, 1981, p. 337 à 353)

SYNCHRONIZED EOL FORMS UNDER UNIFORM INTERPRETATION (*) (**)

by H. A. MAURER (1), A. SALOMAA (2) and D. WOOD (3)

Communicated by J. BERSTEL

Abstract. — *The aim of this paper is the further study of EOL forms under uniform interpretation. We are naturally led to the consideration of synchronized EOL forms, which become the central theme of this paper. We prove: a "strongest" possible normal form result, the spanning normal form theorem; the non-existence of uni-generators for "reasonable" language families, which include the finite, regular and context-free families; and a characterization of those synchronized forms which, under uniform interpretation, yield all regular languages. In closing we remark on three topics which are worthy of separate investigation.*

Résumé. — *Le but de cet article est de poursuivre l'étude des EOL-formes sous des interprétations uniformes. Nous sommes conduits naturellement à considérer les EOL-formes synchronisées qui deviennent le thème central de cet article. Nous établissons : un résultat sur la forme normale la plus « serrée » possible, le théorème sur la forme normale recouvrante, la non-existence d'uni-générateurs pour des familles « raisonnables » de langages qui comprennent les langages finis, rationnels et algébriques; et une caractérisation des formes synchronisées qui donnent, par interprétation uniforme, tous les langages rationnels. Pour terminer, nous effleurons trois domaines qui méritent une étude séparée.*

I. INTRODUCTION

In [9] the notions of an EOL form and its interpretations was first introduced and in [10] this was followed up by a preliminary investigation of a restricted interpretation, namely uniform interpretation of terminal symbols. Given the productions:

$$A \to a B a \qquad \text{and} \qquad a \to abc,$$

then:

$$B \to c A c \qquad \text{and} \qquad d \to def,$$

are possible uniform interpretations, while neither of the following are:

(*) Received December 1979, revised August 1980.

(**) The work of the third author was supported partially by Natural Sciences and Engineering Research Council of Canada Grant No. A-7700 and partially by the Austrian Ministry for Science and Research.

(1) Institut für Informationsverarbeitung, TU Graz, Steyrergasse 17, A-8010, Graz, Austria.

(2) Department of Mathematics, University of Turku, SF-20500 Turku, Finland.

(3) Unit for Computer Science, McMaster University, Hamilton, Ontario L8S 4K1, Canada.

$$A \to b\,B\,a \quad \text{and} \quad a \to bcd.$$

This uniformity enables us to keep track of terminals. As is now known, not only can we characterize the family of context-free languages using only terminal productions of the type $a \to a$, under uniform interpretation (*see* section 4), but also under usual interpretations we cannot obtain this family [1]. Further, uniform interpretations correspond to the uniform substitutions usually studied in logic and in the two-level van Wijngaarten grammars. Hence we feel their study is especially well motivated. Recently uniform interpretations have been investigated in [2] and in [3] with respect to some decidability results. In the present paper our main theme is synchronized EOL forms under uniform interpretation, although we also present results in the general case. In particular we prove various normal form results in section 3, having noted in passing some basic results in section 2. For example, given a synchronized EOL form F, then $\mathscr{L}_u(F)$ is closed under intersection with regular sets. Further for each synchronized EOL form F there exists a uniform form equivalent EOL form G which is span-short, propagating and synchronized. In section 4 we discuss the generative capacity of EOL forms. In particular we show that there are *no* uni-generators for any "natural" language family, prove a "reduction" result, namely $\mathscr{L}_u(F) \subseteq \mathscr{L}$ iff $\mathscr{L}(F) \subseteq \mathscr{L}$, $\mathscr{L}$ a language family satisfying weak conditions, and also characterize when a synchronized EOL form generates all the regular sets. Finally, we conclude in section 5 by introducing three topics which we feel are worthy of further investigation.

2. DEFINITIONS AND FIRST RESULTS

Following [8] and [9] we introduce the notions of an EOL form and its interpretations and also uniform interpretations.

An *EOL system* is a quadruple $G=(V, \Sigma, P, S)$ where V is an alphabet, $\Sigma \subseteq V$ the terminal alphabet, $V-\Sigma$ the nonterminal alphabet, S in $V-\Sigma$ is the start symbol and P is a finite set of pairs (X, α) with X in V and α in V^* such that for each X in V there is at least one such pair in P. The elements of P are called *productions* and are usually written $X \to \alpha$. We say G is *propagating* if for all $X \to \alpha$ in P, $\alpha \neq \varepsilon$, the *empty word*. The yield relation $\Rightarrow$, as well as $\Rightarrow^+$, $\Rightarrow^*$ and $\Rightarrow^k$ (a k-step derivation) are defined in the usual way. The language generated by an EOL system G is defined by:

$$L(G)=\{x \text{ in } \Sigma^* : S \Rightarrow^* x\}.$$

To avoid unnecessary complications, languages which differ by at most the empty word, are said to be equal.

Families of languages which differ by at most the empty set are considered equal if for each non-empty language in the one family there is an equal language in the other family, and *vice versa*.

Let $F=(V, \Sigma, P, S)$ be an EOL system. A in $V-\Sigma$ is said to be a *blocking* symbol if for all α such that $A \Rightarrow^+ \alpha$ in F, α is not in Σ^*. Let $V_B \subseteqq V-\Sigma$ denote the set of blocking symbols in F. F is *synchronized* if for all a in Σ, $a \Rightarrow^+ \alpha$ implies α is not in Σ^*. F is *short* if the right-hand sides of all rules are of length $\leqq 2$, and *binary* if each production is one of the types:

$$A \to \varepsilon; \qquad A \to a; \qquad A \to B; \qquad A \to BC; \qquad a \to A,$$

where the capital letters are nonterminals.

Before turning to the definition of an EOL form and its interpretations we need the following notion. Let Σ and Δ be alphabets, then a substitution $f : \Sigma^* \to 2^{\Delta^*}$ is a *finite letter substitution* (*fl*-substitution) if for all a in Σ, $f(a) \subseteqq \Delta$. Moreover f is a *disjoint finite letter substitution* (*dfl*-substitution) if f is an *fl*-substitution and for all a, b in Σ, $a \neq b$ implies $f(a) \cap f(b) = \emptyset$.

An *EOL form* F is an EOL system $F=(V, \Sigma, P, S)$. An EOL system $F'=(V', \Sigma', P', S')$ is *an interpretation of* F, *modulo* μ, $F' \lhd F(\mu)$ (or simply $F' \lhd F$), if μ is a *dfl*-substitution on V such that the following conditions (i)-(iv) hold:

(i) $\mu(V-\Sigma) \subseteqq V'-\Sigma'$;

(ii) $\mu(\Sigma) \subseteqq \Sigma'$;

(iii) $p' \subseteqq \mu(P)$, where $\mu(P) = \{ \alpha' \to \beta' : \alpha \to \beta$ is in P, α' is in $\mu(\alpha)$ and β' is in $\mu(\beta) \}$;

(iv) S' is in $\mu(S)$.

F' is a *uniform interpretation* of F, in symbols, $F' \lhd_u F$, if in (iii) $P' \subseteqq \mu_u(P)$, where $\mu_u(P)$ is the subset of all productions $\alpha'_0 \to \alpha'_1 \ldots \alpha'_t$ obtained as follows. Assume that $p : \alpha_0 \to \alpha_1 \ldots \alpha_t$ is in P, then $\alpha'_0 \to \alpha'_1 \ldots \alpha'_t$ contained in $\mu(P)$ is in $\mu_u(P)$ iff $\alpha_r = \alpha_s$ in Σ implies $\alpha'_r = \alpha'_s$. (Thus the substitution has to be uniform on terminals.)

The family of EOL forms generated by F is defined by $\mathscr{G}(F) = \{ F' : F' \lhd F \}$, and the family of languages generated by F is defined by $\mathscr{L}(F) = \{ L(F') : F' \lhd F \}$. Similarly we define $\mathscr{G}_u(F)$ and $\mathscr{L}_u(F)$.

Two EOL forms F_1 and F_2 are (*uniform*) *form equivalent* if $\mathscr{L}(F_1) = \mathscr{L}(F_2)$ ($\mathscr{L}_u(F_1) = \mathscr{L}_u(F_2)$).

An EOL form $F = (V, \Sigma, P, S)$ is a *two-symbol* form if $V = \{ S, a \}$ and $\Sigma = \{ a \}$. It is *simple* if it is also short.

In the following we denote by $\mathscr{L}$(FIN), $\mathscr{L}$(REG), $\mathscr{L}$(LIN), $\mathscr{L}(CF)$, $\mathscr{L}$(OL) and $\mathscr{L}$(EOL) the families of finite, regular, linear, context-free, OL and EOL languages respectively.

A *deterministic complete sequential machine* (*dcsm*) M, is a quintuple $(Q, \Sigma, \Delta, \delta, q_0)$, where Q is a finite nonempty set of states, Σ is an input alphabet, Δ an output alphabet, q_0 in Q an initial state and δ the transition function $\delta : Q \times \Sigma \to Q \times \Delta$. δ is extended to $Q \times \Sigma^*$ in the standard manner. A map $f : \Sigma^* \to \Delta^*$ is a *dcsm map* if there is a *dcsm* M such that for all x in Σ^*,

$$f(x) = M(x) = \{ y : \delta(q_0, x) = (p, y) \text{ for some } p \text{ in } Q \}.$$

We extend M to sets of words in the natural way. Let $\mathscr{L}$ be a family of languages then by $\mathscr{M}(\mathscr{L})$ we denote the closure of $\mathscr{L}$ under *dcsm* maps, that is:

$$\mathscr{M}(\mathscr{L}) = \{ M(L) : L \text{ is in } \mathscr{L} \text{ and } M \text{ is a } dcsm \}.$$

We close this section by formulating some basic results which are given without proof.

PROPOSITION 2.1. — *The pre-orders* $\lhd$ *and* $\lhd_u$ *are decidable and transitive. For two* EOL *forms* F_1 *and* F_2, $\mathscr{G}(F_1) \subseteq \mathscr{G}(F_2)$ *iff* $F_1 \lhd F_2$ *and* $\mathscr{G}_u(F_1) \subseteq \mathscr{G}_u(F_2)$ *iff* $F_1 \lhd_u F_2$. *It is decidable for arbitrary* EOL *forms* F_1 *and* F_2 *whether or not* $\mathscr{G}(F_1) = \mathscr{G}(F_2)$ *or* $\mathscr{G}_u(F_1) = \mathscr{G}_u(F_2)$. *For an* EOL *form* F, $\mathscr{G}_u(F) \subseteq \mathscr{G}(F)$ *and* $\mathscr{L}_u(F) \subseteq \mathscr{L}(F)$.

PROPOSITION 2.2. — *For* F *a synchronized* EOL *form*, $\mathscr{L}(F)$ *and* $\mathscr{L}_u(F)$ *are both closed under intersection with regular sets. If* F *also has a single terminal symbol then* $\mathscr{L}(F)$ *and* $\mathscr{L}_u(F)$ *are both closed under union.*

3. NORMAL FORMS

It was observed in [9] that, in general, none of the reduction results in [8] for EOL forms under the usual interpretation mechanism carry over to EOL forms under uniform interpretation. For example, let F_n, $n \geqq 3$ be the form whose productions are: $S \to a^n; a \to N; N \to N$. Then $\mathscr{L}_u(F_n)$ consists of finite unions of singleton languages of the type $\{ b^n \}$. In [9] it is proved that there is no short (separated, binary) form which is uniform form equivalent to F_n, for each $n \geqq 3$. Intuitively it is clear that in reducing $s \to a^n$ to a short form, the forced uniformity of interpretation of a^n will be lost, giving rise to words which contain at least two symbols under an appropriate interpretation.

This dirth of reduction results, and hence of normal forms, under uniform interpretation is a major obstacle to any serious investigation of EOL forms under uniform interpretation. However the situation is not quite as bleak as [9] would have us believe. In the sequel we consider synchronized EOL forms under uniform interpretation. In this case there are positive reduction results of similar flavor to those in [8] for the usual interpretations. In particular, we prove that for every synchronized EOL form there exists a uniform form equivalent synchronized and propagating EOL form. Finally we prove a "spanning" normal form result.

DEFINITION: A synchronized EOL system $G=(V, \Sigma, P, S)$ is:

(i) *n-short* if for all $A \to \alpha$ in P, A in $V-\Sigma$, α in $(V-\Sigma)^*$ α is in $\{\varepsilon\} \cup (V-\Sigma) \cup (V-\Sigma)^2$;

(ii) *disjoint* if for all $A \to \alpha$ in P, either A is in $V-(\Sigma \cup V_B)$ in which case α is in $\Sigma^* \cup V_B \cup (V-(V_B \cup \Sigma))^*$ or A is in $\Sigma \cup V_B$ in which case α is in V_B.

We now obtain our first positive result:

LEMMA 3.1: Disjointness lemma. *For every synchronized* EOL *form* $F=(V, \Sigma, P, S)$ *there exists a uniform form equivalent synchronized* EOL *form* $\overline{F}=(\overline{V}, \Sigma, \overline{P}, S)$ *which is disjoint.*

Proof: Construct $\overline{P}$ as follows:

(i) for all a in Σ, take $a \to N$ and $N \to N$ into $\overline{P}$, N a new blocking nonterminal;

(ii) for all A in $V-\Sigma$, for all $A \to \alpha$ in P, if α is in $\Sigma^* \cup V_B \cup (V-(\Sigma \cup V_B))^+$ take $A \to \alpha$ into $\overline{P}$, otherwise take $A \to N$ into $\overline{P}$.

Finally letting $\overline{V}=V \cup \{N\}$, it is clear that not only is $L(\overline{F})=L(F)$ but also $\mathscr{L}_u(\overline{F})=\mathscr{L}_u(F)$. This follows by observing, firstly, that all terminals must give rise to "blocking words", because F is synchronized and secondly, for productions $A \to \alpha$, A in $V-\Sigma$, where α either contains a mixture of terminals and nonterminals or contains a blocking symbol, then it is a "blocking" production. □

We now prove two simulation lemmas for EOL forms under uniform interpretation. These are analogous to those in [8] for the usual interpretation mechanism. Note however that there are EOL forms F_1 and F_2 such that F_2 "simulates" F_1 with $\mathscr{L}(F_1)=\mathscr{L}(F_2)$ by [8] but $\mathscr{L}_u(F_2) \nsubseteq \mathscr{L}_u(F_1)$. For example, consider F_1 defined by $S \to aa$; $a \to a$; and F_2 by $S \to AA$; $A \to a$; $a \to B$; $B \to a$. Clearly $S \Rightarrow^2 aa$, $a \Rightarrow^2 a$ in F_2 and, in fact, i^* is easy to see that $\mathscr{L}(F_1)=\mathscr{L}(F_2)$. However, although $\mathscr{L}_u(F_1) \subseteq \mathscr{L}_u(F_2)$ the reverse inclusion does not hold since $\{ab\}$ is in $\mathscr{L}_u(F_2)$ but $\{ab\}$ is not in $\mathscr{L}_u(F_1)$. We could therefore prove the first simulation lemma in full generality, while proving the second simulation lemma for the restricted case of disjoint synchronized EOL forms. However since we only use the simulation lemmas to prove uniform form equivalence of two forms, we restrict our attention for both lemmas to the restricted case.

LEMMA 3.2: The first simulation lemma. *Let* $F_1=(V_1, \Sigma_1, P_1, S)$ *and* $F_2=(V_2, \Sigma_2, P_2, S)$ *be two disjoint synchronized* EOL *forms*, $\Sigma_1 \subseteq \Sigma_2$ *and* $l \geq 1$ *an integer.*

Suppose:

(i) $A \to \alpha$ *in* P_1 *with* A *in* $V_1-\Sigma_1$, α *in* $(V_1-\Sigma_1)^+ \cap V_1^*(V_1-V_{1,B})V_1^*$ *implies* $A \Rightarrow^l \alpha$ *in* F_2, *and;*

(ii) $A \to \alpha$ *in* P_1 *with* A *in* $V_1-\Sigma_1$, α *in* $\Sigma_1^* \cup V_{1,B}^*$ *implies* $A \to \alpha$ *is in* P_2.

Then $\mathscr{L}_u(F_1) \subseteq \mathscr{L}_u(F_2)$.

Proof: Observe firstly that $A \to \alpha$ in P_1 with A in $V_1-\Sigma_1$ and α in $(V_1-\Sigma_1)^+ \cap V_1^*(V_1-V_{1,B})V_1^*$ implies that for all β, $A \Rightarrow^k \beta \Rightarrow^{l-k} \alpha$, $0 \leq k < l$, β is not in Σ^*. Otherwise α consists of blocking symbols with respect to F_2, which means that α cannot derive, in F_2, any word containing a terminal. However in F_1, α can derive words containing a terminal, giving a contradiction.

Secondly, each "simulating" terminal derivation in F_2 consists of a multiple of l simulating steps plus a one-step derivation which introduces terminals, that is: $S \Rightarrow^k \alpha \Rightarrow x$ in F_1, x in Σ_1^* implies $S \Rightarrow^{kl} \alpha \Rightarrow x$ in F_2 and in both cases x then blocks.

Let $F_1' \lhd_u F_1(\mu_1)$ with $F_1'=(V_1', \Sigma_1', P_1', S')$ be an arbitrary uniform interpretation of F_1, we need to show that $L(F_1')$ is in $\mathscr{L}_u(F_2)$. It is sufficient to construct an $F_2' \lhd_u F_2(\mu_2)$ such that $L(F_1')=L(F_2')$.

For all $A' \to \alpha'$ in P_1' with $A' \to \alpha'$ in $\mu_1(A \to \alpha)$ take into P_2'.

If $A \to \alpha$ fulfills condition (i) of the lemma then sufficient unique productions such that $A' \Rightarrow^l \alpha'$ is an "isolated" derivation in F_2'.

Otherwise, $A \to \alpha$ fulfills condition (ii) in which case take $A' \to \alpha'$.

Now extend μ_1 to be over V_2, giving μ_2. Clearly, such an $F_2' \lhd_u F_2(\mu_2)$ can be so constructed. Further it is immediate that $L(F_2')=L(F_1')$, by the previous remarks and by observing that the uniform interpretation μ_1 only affects the terminal introducing productions $A \to x$. However these are transferred unchanged to F_2'. Hence the result. □

DEFINITION: Let $F=(V, \Sigma, P, S)$ be an EOL form and $l \geq 1$ an integer, we denote by $V(l)$ the set of all symbols derivable in F in a multiple of l steps from S.

We now have the second simulation lemma.

LEMMA 3.3: The second simulation lemma:

$$Let\ F_1=(V_1, \Sigma_1, P_1, S) \qquad and \qquad F_2=(V_2, \Sigma_2, P_2, S),$$

be two disjoint synchronized EOL *forms*, $\Sigma_2 \subseteq \Sigma_1$ *and* $l \geq 1$ *be an integer.*

Suppose that for all A *in* $V_2(l)$:

(i) $A \Rightarrow^l \alpha$ *in* F_2 *with* α *in* $(V_2-\Sigma_2)^* \cap V_2^*(V_2-V_{2,B})\, V_2^*$ *implies* $A \rightarrow \alpha$ *is in* P_1, *and*

(ii) $A \rightarrow \alpha$ *in* P_2 *with* α *in* $\Sigma_2^* \cup V_{2,B}^*$ *implies* $A \rightarrow \alpha$ *is in* P_1.

Then $\mathscr{L}_u(F_2) \subseteq \mathscr{L}_u(F_1)$.

Proof: As in the first simulation lemma observe that $A \Rightarrow^k \beta \Rightarrow^{l-k} \alpha$ in F_2 with α fulfilling condition (i) implies that β is not terminal. Hence the result follows in a similar fashion. □

We now apply these simulation lemmas to yield a short normal form result when only nonterminals are considered.

DEFINITION: Let $F=(V, \Sigma, P, S)$ be a disjoint synchronized EOL form, then define $\max nr(F)=\max(\{|\alpha| : \alpha \text{ in } (V-\Sigma)^+,\ A \rightarrow \alpha \text{ is in } P \text{ for some } A \text{ in } V-\Sigma\})$.

THEOREM 3.4: *Let* $F=(V_1, \Sigma, P_1, S)$ *be a disjoint synchronized* EOL *form. There exists a disjoint synchronized* EOL *form* $F_2=(V_2, \Sigma, P_2, S)$ *such that* $\mathscr{L}_u(F_1)=\mathscr{L}_u(F_2)$ *and* F_2 *is in n-short normal form.*

Proof: Construct F an EOL form from F_1. For all productions $p : X \rightarrow \alpha$ in P_1:

(i) if X is in $\Sigma \cup V_{1,B}$ take $X \rightarrow \alpha$ into P;

(ii) if X is in $V_1-(\Sigma \cup V_{1,B})$:

(*a*) if α is in Σ^* take $X \rightarrow \alpha$ into P,

(*b*) if α is in $V_1-(\Sigma_1 \cup V_{1,B})$ or α is in $(V_1-\Sigma_1)^2$ take $X \rightarrow p$; $p \rightarrow \alpha$ into P where p is a new nonterminal,

(*c*) if α is in $(V_1-\Sigma_1)^3\, V_1^*$ take $X \rightarrow p_1 p_2$; $p_1 \rightarrow \alpha_1$; $p_2 \rightarrow \alpha_2$ into P where p_1, p_2 are new nonterminals and $\alpha_1 \alpha_2=\alpha$, $\alpha_1 \neq \varepsilon \neq \alpha_2$.

By inspection $\mathscr{L}_u(F_1)=\mathscr{L}_u(F)$ by way of the two simulation lemmas. Clearly $\max nr(F) < \max nr(F_1)$ if $\max nr(F_1) > 2$. If $\max nr(F) > 2$ repeat the construction, otherwise let $F_2=F$. This is a terminating process. Hence the result. □

We are now in a position to apply directly the theorem in [8] on the transformation of a non-propagating synchronized EOL form to a propagating synchronized one. By inspection of the proof in [8] of theorem 4.6, we observe that terminal introducing productions are not affected by the transformation. Hence we obtain:

THEOREM 3.5: *Let* $F_1=(V_1, \Sigma, P_1, S)$ *be a synchronized* EOL *form. There exists a disjoint n-short synchronized and propagating* EOL *form* F_2 *with* $\mathscr{L}_u(F_1)=\mathscr{L}_u(F_2)$.

Proof: By lemma 3.1 and theorem 3.4 above, and theorem 4.6 in [8]. □

We now turn to the promised spanning normal form result.

Let μ be a *dfl*-substitution from Σ to Σ', then define μ_u the uniform *dfl*-substitution of Σ^* to Σ'^* by $\mu_u(\varepsilon)=\varepsilon$ and for all $a_1 \ldots a_m$, a_i in Σ, $1 \leqq i \leqq m$, $m>0$, $\mu_u(a_1 \ldots a_m)=\{b_1 \ldots b_m : b_1 \ldots b_m$ is in $\mu(a_1 \ldots a_m)$ and for all i, j, $1 \leqq i<j \leqq m$, $b_i=b_j$ iff $a_i=a_j\}$.

We say x in Σ^+ is *prime* if for all y, z in Σ^*, such that $x=yz$, $\text{alph}(y) \cap \text{alph}(z)=\emptyset$ implies $y=\varepsilon$ or $z=\varepsilon$.

Hence, for example, *aa* and *ababb* are prime, while *aabb* is not prime.

If $x=a_1 \ldots a_m \neq \varepsilon$ is not prime then there exists a unique *prime factorisation*, $x_1, \ldots, x_r$, $1<r \leqq m$ such that $a_1 \ldots a_m=x_1 \ldots x_r$ and $\text{alph}(x_i) \cap \text{alph}(x_j)=\emptyset$, $1 \leqq i<j \leqq r$. Hence *aabb* has prime factorisation *aa* and *bb*. We are now in a position to define the *span of* $a_1 \ldots a_m$, $m>0$ The span of ε is 0, while for all $x=a_1 \ldots a_m$, $m>0$, the span of $x=\max(\{|x_i| : x=x_1 \ldots x_r$, where $x_1, \ldots, x_r$ is the prime factorisation of $x\})$. In other words the span of a non-empty word is the length of its largest prime factor.

Let F be an EOL system (V, Σ, P, S) then *span* (F) is defined as the maximal span of the right hand sides of all terminal introducing productions. When $\text{span}(F)=0$, then $L(F)=\{\varepsilon\}$ or $\emptyset$ and when $\text{span}(F)=1$ and F is synchronized $\mathscr{L}_u(F)=\mathscr{L}(F)$. Letting $\max tr(F)=\max(\{|x| : A \to x$ in P, x in $\Sigma^*\})$ our final result shows that we can obtain an n-short disjoint synchronized G, uniform form equivalent to F such that $\max tr(G)=\text{span}(F)=\text{span}(G)$.

THEOREM 3.6: The spanning normal form theorem. *Let* $F=(V, \Sigma, P, S)$ *be a disjoint synchronized* EOL *form. Then there exists a uniform form equivalent* EOL *form* G *such that* G *is n-short, disjoint, propagating, synchronized and* $\max tr(G)=\text{span}(F)$.

Proof: If $\max tr(F)=\text{span}(F)$ then we obtain G by theorem 3.5 above. Otherwise $\max tr(F)>\text{span}(F)$. In this case letting $m=\text{span}(F)$ construct a disjoint synchronized EOL form $H=(V_H, \Sigma, P_H, S)$ from F as follows:

(1) take all productions $A \to \alpha$ in P, A in $V-\Sigma$, α in $(V-\Sigma)^+$, into P_H;

(2) take all productions $a \to A$ in P, a in Σ, A in $V-\Sigma$, into P_H;

(3) for all productions $p : A \rightarrow x$, x in Σ^*, $|x| \leqq m$, take $A \rightarrow p$; $p \rightarrow x$ into P_H; where p is a new nonterminal;

(4) for all productions $p : A \rightarrow x, x$ in $\Sigma^*, |x| > m, x$ can be split into r prime factors, $x_1, \ldots, x_r$ such that $x = x_1 \ldots x_r$, x_i in Σ^+, $|x_i| \leqq m$ and span $(x) = \max(\{|x_i|\})$ Take productions $A \rightarrow p_1 \ldots p_r$; $p_i \rightarrow x_i$ into P_H, where the p_i are new nonterminals.

Clearly $\mathscr{L}_u(F) = \mathscr{L}_u(H)$. Finally apply the previous theorems to H to give the required G. □

4. LANGUAGE FAMILY GENERATION

Given a (synchronized) EOL form F we are interested in the language family $\mathscr{L}_u(F)$. A basic notion introduced in [8] is that of completeness.

If $\mathscr{L}_u(F) = \mathscr{L}(\text{EOL})$ we say that F is *uni-complete* and if $\mathscr{L}(F) = \mathscr{L}(\text{EOL})$ we say that F is *complete*. Clearly if a form F is uni-complete it is also complete, since $\mathscr{L}_u(F) \subseteqq \mathscr{L}(F)$, for all EOL forms F. However the converse does not necessarily hold, for example, consider the form $F : S \rightarrow a \mid S \mid aa \mid a\,S$; $a \rightarrow SS$; which is shown in [6] to be complete using the chain-free normal form of [5]. Now the language $\{ab\}$ can only be obtained by isolating the derivation $S \Rightarrow^* S \Rightarrow aa$ in F. Immediately under uniform interpretation $\{ab\}$ cannot be obtained. Hence F is not uni-complete.

We may of course generalize the notion of (uni-) completeness for any subfamily of $\mathscr{L}(\text{EOL})$. We say that, for $\mathscr{L} \subseteqq \mathscr{L}(\text{EOL})$, an EOL form F is *(uni-) $\mathscr{L}$-complete* if $\mathscr{L}(F) = \mathscr{L}$ ($\mathscr{L}_u(F) = \mathscr{L}$). Subfamilies of particular interest are $\mathscr{L}(\text{REG})$, $\mathscr{L}(\text{LIN})$ and $\mathscr{L}(CF)$.

This investigation leads naturally to the related question: Is $\mathscr{L}_u(F) \subseteqq \mathscr{L}$ or does $\mathscr{L}_u(F)_2 \supseteqq \mathscr{L}$, for some subfamily $\mathscr{L}$ of $\mathscr{L}(\text{EOL})$? For example, [2] has shown it to be decidable whether $\mathscr{L}_u(F) \subseteqq \mathscr{L}(\text{REG})$ for an arbitrary OL form F, while in [3] it is shown to be decidable whether $\mathscr{L}_u(F) \subseteqq \mathscr{L}(CF)$, for F a simple EOL form.

A technique of interest in its own right, namely the notion of a generator, has been introduced in [10]. We show that, apart from trivial exceptions, generators do not exist for EOL forms under uniform interpretations. Returning to the theme of uni-completeness, we have a preliminary result.

LEMMA 4.1: *Let $F = (\{S, a\}, \{a\}, P_F, S)$ be a two-symbol EOL form with productions* (i) $S \rightarrow S$; $a \rightarrow a \mid S$; (ii) $S \rightarrow \gamma$ *for at least one γ with $m = |\gamma| \geqq 2$ and γ contains at least one S, and* (iii) $S \rightarrow a^i$, $1 \leqq i \leqq m$. *Then F is complete.*

Proof: We derive a suitable normal form for EOL systems. We know that every EOL language can be generated by a synchronized EOL system having only productions of the types $A \to B \mid BC \mid a$ and $a \to D$, D blocking. Let $G=(V, \Sigma, P, S)$ be such an EOL system. Construct a new EOL system G' as follows:

(1) for each word x in $L(G)$ with $|x| \leqq m$ take $S \to x$ into P';

(2) for each production $X \to \alpha$ in P with $X\alpha$ in V^2 take $X \to \alpha$ into P', and

(3) for each production $A \to \alpha$ in P with α in $(V-\Sigma)^2$, take $A \to N_\alpha$ into P', where N_α is a new nonterminal.

Let $\gamma=\gamma_1 S \gamma_2$, where $\gamma_1\gamma_2$ is in $\{S, a\}^*$, and $|\gamma_1|=k$, $|\gamma_2|=l$, $k+l=m-1$.

Now consider the new symbols N_α, $2 \leqq |\alpha| < m$, α in $(V-\Sigma)^*$. For each derivation $\alpha \Rightarrow \beta$ in G with $2 \leqq |\alpha| \leqq m$ we have $2 \leqq |\beta| \leqq 2m-2$. In the case β is in Σ^+ we have $2 \leqq |\beta| \leqq m-1$.

For each derivation $\alpha \Rightarrow \beta$ in G, since G is synchronized we need only consider two cases:

(4) β in Σ^+, take $N_\alpha \to \beta$ in P;

(5) β in $(V-\Sigma)^+$,

(i) if $|\beta| \leqq m$ take $N_\alpha \to N_\beta$ in P',

(ii) if $|\beta|=m$ take $N_\alpha \to C_1 \ldots C_m$, where $\beta=B_1 \ldots B_m$ and $C_i=B_i$ if the ith symbol of γ is an S, and $C_i=T_{B_i}$ (a new terminal) otherwise,

(iii) if $|\beta|>m$ take $N_\alpha \to C_1 \ldots C_k \; N_{B_{k+1} \ldots B_{r-l}} C_{r-l+1} \ldots C_r$, where $\beta=B_1 \ldots B_r$, and $C_i=B_i$ if the ith symbol of γ is an S and $C_i=T_{B_i}$ a new terminal otherwise.

Finally for the new terminal symbols T_A, A in $V-\Sigma$, take:

(6) $T_A \to B$ if $A \to B$ is in P;

(7) $T_A \to a$ if $A \to a$ is in P, and finally;

(8) $T_A \to N_{BC}$ if $A \to BC$ is in P.

We leave the reader the straightforward but tedious proof that $L(G')=L(G)$. It is clear, by the technique of the construction that $G' \lhd F$. □

As already demonstrated there are forms which are complete but not uni-complete. That this is not a rare occurrence is seen from the following theorem, which is a consequence of the above lemma.

THEOREM 4.2: *There are an infinitude of complete forms which are not uni-complete.*

Proof: For each $m \geqq 2$, F_m defined by the productions $S \to a \mid \ldots \mid a^m \mid S \mid^{m-1} S$; $a \to S$ is complete by lemma 4.1. However since $\{ab\}$ can only be obtained as an interpretation of $S \Rightarrow^* S \Rightarrow aa$ in F_m, we cannot obtain $\{ab\}$ from any F_m under uniform interpretation. Hence F_m is not uni-complete.

One open question is to characterize, under the usual interpretation, those EOL forms which are complete. Although little progress has been made in the general case, for the case of simple forms (two-symbol and short) [6] have classified a large number of forms. The question of such a characterization also arises for uni-completeness.

Since the general situation is, no doubt, as difficult for uni-completeness as it is for completeness, we restrict our attention to simple EOL forms. How much of the classification of [6] holds also for uni-completeness ?

Three trivial observations are worthy of mention. Firstly, a form which is not complete is also not uni-complete. Secondly, if each production in a complete form contains at most one a, then it is also uni-complete, and thirdly, if repeated a' s only occur in productions of the type $a \rightarrow a S$ or $a \rightarrow S a$, which are used for blocking, then the form is also uni-complete. Examples of these three situations are:

(1) $S \rightarrow S \mid SS \mid a$; $a \rightarrow a$ is not complete and not uni-complete;

(2) $S \rightarrow a \mid S \mid a S$; $a \rightarrow S$ is complete, and uni-complete;

(3) $S \rightarrow a \mid S \mid SS$; $a \rightarrow a S$ is complete and also uni-complete.

However the situation with a form such as: $S \rightarrow a \mid SS$; $a \rightarrow a \mid S$ which has been shown to be complete in [6] by way of the chain-free normal form of [5] is completely open. We conjecture that this form is *not* uni-complete. Note that it contains all $\mathscr{L}(CF)$ hence any counter-example has to be non-context-free.

Recall that an EOL form is *stable with respect to terminals* if for each terminal a, the only production for a is $a \rightarrow a$. A grammar form G is a context-free grammar (that is, an EOL system with no productions for terminals, and sequential rather than parallel rewriting). Now uniform interpretations of grammar forms can be defined analogously to the present definition of uniform interpretations of EOL forms. If G is a grammar form we obtain $\mathscr{L}_u(G)$, while under standard interpretations of grammar forms we obtain $\mathscr{L}_g(G)$, *see* [4] and [11].

We now obtain:

THEOREM 4.3: *Assume F is an* EOL *form stable with respect to terminals and F_C is the context-free grammar form obtained from F by omitting all productions for terminals. Then $\mathscr{L}_u(F_C) = \mathscr{L}_u(F)$.*

Conversely, if F is a reduced grammar form and F_L is the EOL *form obtained by adding the productions $a \rightarrow a$ for each terminal, then $\mathscr{L}_u(F_L) = \mathscr{L}_u(F)$.*

Proof: $\mathscr{L}_u(F) \subseteq \mathscr{L}_u(F_C)$ simply by observing that each $F' \lhd_u F$ is also stable with respect to terminals. On the other hand $\mathscr{L}_u(F_C) \subseteq \mathscr{L}_u(F)$ since for each

$F'_C \lhd_u F_C$ we can obtain an equivalent $F' \lhd_u F$ by adding the productions $a \to a$ for each terminal in F'_C to F'_C.

The converse result follows analogously. □

Since we know that for each grammar form G with $L(G)$ infinite, $\mathcal{L}_g(G) = \mathcal{L}_u(H)$ for some grammar form H [11]. We have:

COROLLARY 4.4: *There are an infinitude of full principal semi-*AFLs *represented by* EOL *forms under uniform interpretation.*

Proof: Since $\mathcal{L}_g(G)$ is a full principal semi-AFL for each grammar form G with $L(G)$ infinite [4] and $\mathcal{L}_g(G) = \mathcal{L}_u(H)$ for some grammar form H in this case, then by theorem 4.3, $\mathcal{L}_u(H) = \mathcal{L}_u(H_L)$. □

It also follows from theorem 4.3 that for all EOL forms F, stable with respect to terminals, $\mathcal{L}_u(F) \subseteqq \mathcal{L}(CF)$, (*see* [9]).

Continuing our investigation of the families of languages of EOL forms under uniform interpretation we recall the notion of a generator [10]. We say a language L is a (*uni-*) *generator* for a family $\mathcal{L}$, ($\mathcal{L} \subseteqq \mathcal{L}$ (EOL)), if for all synchronized forms F with $L(F) = L$, $\mathcal{L}(F) \supseteqq \mathcal{L}$ ($\mathcal{L}_u(F) \supseteqq \mathcal{L}$). For example, in [10] it was shown that a^* is a generator for $\mathcal{L}$ (REG). When the generator L is in $\mathcal{L}$, we say L is a *proper generator*. In [10] it was shown that no proper generator for $\mathcal{L}$ (FIN) exists, while a^* is, of course, a generator for $\mathcal{L}$ (FIN), as well as a proper generator for $\mathcal{L}$ (REG).

For uni-generators we have:

THEOREM 4.5: *Let* $\mathcal{L} \subseteqq \mathcal{L}$ (EOL) *be a family of languages containing all finite languages.*

Then $\mathcal{L}$ *has no uni-generator.*

Proof: Assume L is a uni-generator for $\mathcal{L}$, and H is a synchronized EOL system with $L = L(H)$. Now L is infinite, otherwise $L(H)$ is finite and hence $\mathcal{L}_u(H) \nsupseteq \mathcal{L}$ (FIN), a contradiction.

Let $L \subseteqq \Sigma^*$, $\# \Sigma = n$ and define the languages L_k, $k \geqq 0$ by: $L_k = \{x : x \text{ is in } L, |x| = k\}$. Choose the first $L_k \neq \emptyset$ such that $k > n$. Construct:

(i) $\overline{H}$, such that $L(\overline{H}) = L(H) \cap (\Sigma^* - \Sigma^k)$. This can be constructed because of proposition 2.2. Clearly $L(\overline{H}) = L - L_k$;

(ii) $\overline{\overline{H}}$ by adding the productions $S \to x$ to $\overline{H}$, where S is the start symbol, for all x in L_k.

Clearly $L(\overline{\overline{H}}) = L$ but $\mathcal{L}_u(\overline{\overline{H}}) \supseteqq \mathcal{L}$ since the language $\{a_1 \dots a_k\}$, $a_i \neq a_j$, $1 \leqq i < j \leqq k$ is not in $\mathcal{L}_u(\overline{\overline{H}})$. □

COROLLARY 4.6: *There are no uni-generators and hence, no proper uni-generators for* $\mathscr{L}$ (FIN), $\mathscr{L}$ (REG), $\mathscr{L}$ (LIN), $\mathscr{L}$ (CF) *and* $\mathscr{L}$(EOL).

Notice that this is a stronger result than under the usual interpretations. It is still open whether or not there is a generator for $\mathscr{L}$ (LIN) and $\mathscr{L}$ (CF). In [10] it is shown there is no proper generator for these families.

Finally we turn to the questions of when does $\mathscr{L}_u(F) \supseteq \mathscr{L}$ or $\mathscr{L} \supseteq \mathscr{L}_u(F)$ for a synchronized form F and *a* language family $\mathscr{L} \subseteq \mathscr{L}$ (EOL). We first show that for reasonable families $\mathscr{L}$ uniform interpretations are no more restrictive than the usual interpretations when considering the question: is $\mathscr{L}_u(F) \subseteq \mathscr{L}$?

THEOREM 4.7: *Let* $\mathscr{L} \subseteq \mathscr{L}$ (EOL) *be closed under dcsm maps and* $F=(V, \Sigma, P, S)$ *be a disjoint synchronized* EOL *form.*

Then $\mathscr{L}_u(F) \subseteq \mathscr{L}$ *iff* $\mathscr{L}(F) \subseteq \mathscr{L}$.

Proof: if is trivial since $\mathscr{L}_u(F) \subseteq \mathscr{L}(F)$.

Only if: Let $F' \lhd F(\mu)$ be an arbitrary interpretation of F and $F'=(V', \Sigma', P', S')$. We demonstrate that there exists an $F'' \lhd_u F$ such that $M(L(F''))=L(F')$ for an appropriate *dcsm* M. Since $\mathscr{L}$ is closed under *dcsm* maps, $M(\mathscr{L}_u(F)) \subseteq \mathscr{L}$ and since $\mathscr{L}(F) \subseteq M(\mathscr{L}_u(F))$ we obtain the result. The requirement that the *dcsm* M is complete implies that $|M(x)| = |x|$ for all x in $L(F'')$.

Construct $F''=(V'', \Sigma'', P'', S')$ as follows:

(1) for each production $A' \to \alpha'$ in P', $A'\alpha'$ in $(V'-\Sigma')^+$ take $A' \to \alpha'$ into p'';

(2) for each production p: $A' \to x'$ in P', x in Σ^+ with $A' \to x'$ in $\mu(A \to x)$, letting $x=a_1 \ldots a_m$, $m \geq 1$, take into P'' the production $A' \to pc_2 \ldots c_m$, where $c_i=a_i$ if $a_i \neq a_1$ and $c_i=p$ if $a_i=a_1$. p is considered to be a new terminal symbol, $\Sigma''=\Sigma \cup \{p\colon p \text{ is a production in } F'\}$;

(3) for each a in Σ'', take the appropriate blocking productions fom F'.

Now observe that each word z'' in $L(F'')$ has the form: $z''=p_{i_1}x_{i_1} \ldots p_{i_m}x_{i_m}$, where each $p_{i_j}x_{i_j}$ corresponds to the right hand side of the production $A \to ax$ in F where p_{i_j} is in F' and p_{i_j} is in $\mu(A \to ax)$. Also $|p_{i_j}x_{i_j}| = |ax|$ and moreover $F'' \lhd_u F$, by the method of construction.

Construct a *dcsm* M, which processes each word z'' in $L(F'')$ from left to right. On meeting p_{i_1} it reads each symbol of x_{i_1} and outputs the corresponding symbol in the right hand side of the production p_{i_1} in F'. M continues in a similar manner for the remainder of z''. This is clearly a deterministic process and moreover the original word z' in $L(F')$ is recovered in this way. Hence $L(F')=M(L(F''))$, since each word z' in $L(F')$ has been encoded under F'' as a word $z''=p_{i_1} \ldots x_{i_m}$. Hence the theorem is proved. □

It should be noted that we cannot replace "*dcsm*" by finite letter substitution. This is seen by considering the form $F : S \to aa;\ a \to N;\ N \to N$. Now $\mathscr{L}_u(F)$ consists of finite unions of singleton languages $\{aa\}$, whereas $\mathscr{L}(F)$ contains $L = \{ab, aa\}$, for example. It is clear that L cannot be obtained by any finite letter substitution on $\{aa : a \text{ in } \Sigma\}$, for any Σ.

We have a number of interesting corollaries:

COROLLARY 4.8: *Let F be a synchronized* EOL *form. Then* $\mathscr{L}_u(F) \subseteqq \mathscr{L}$ (REG) ($\mathscr{L}$ (CF)) *implies* $\mathscr{L}(F) \subseteqq \mathscr{L}$ (REG) ($\mathscr{L}(CF)$).

COROLLARY 4.9: *Let G be a grammar form in Chomsky Normal Form with a single terminal symbol, and F a synchronized* EOL *form.*

Then $\mathscr{L}_u(F) \subseteqq \mathscr{L}_u(G)$ *iff* $\mathscr{L}(F) \subseteqq \mathscr{L}_u(G)$, *and* $\mathscr{L}_u(F) \subset \mathscr{L}_g(G)$ *iff* $\mathscr{L}(F) \subset \mathscr{L}_g(G)$.

Proof: Since G has a single terminal symbol and is in Chomsky Normal Form, $\mathscr{L}_u(G)$ is closed under finite letter substitution and intersection with regular sets. Hence $\mathscr{L}_u(G)$ is closed under *dcsm* maps. This is trivially true for $\mathscr{L}_g(G)$. □

We finally reformulate the theorem to give an interesting "gap" result.

COROLLARY 4.10: *Let F be a synchronized* EOL *form. There is no* $\mathscr{L}$ *closed under dcsm maps such that* $\mathscr{L}_u(F) \subsetneqq \mathscr{L} \subsetneqq \mathscr{L}(F)$.

This follows from the observation that $\mathscr{M}(\mathscr{L}_u(F)) \supseteqq \mathscr{L}(F)$. We only obtain equality when F has a single terminal symbol.

Let us consider the situation when F is not necessarily synchronized. Theorem 4.7 no longer holds for $\mathscr{L}_u(F) \subseteqq \mathscr{L}(CF)$, since F defined by $S \to S \mid SS \mid a;\ a \to a$ has $\mathscr{L}_u(F) \subseteqq \mathscr{L}(CF)$ but $\mathscr{L}(F) \nsubseteq \mathscr{L}(CF)$. This latter result holds since $\{a^i b^i c^i d^j : i, j \geqq 1\}$ is in $\mathscr{L}(F)$ and hence, by [1], $\mathscr{L}(F)$ contains non-context-free languages. Hence we can strengthen this observation to: $\mathscr{L}_u(F) \subseteqq \mathscr{L}$ (2-LIN) does not imply $\mathscr{L}(F) \subseteqq \mathscr{L}$ (2-LIN), where $\mathscr{L}$ (2-LIN) is the family of languages equal to $\mathscr{L}(\text{LIN}).\mathscr{L}(\text{LIN})$.

If $\mathscr{L}_u(F)$ is either sub-regular or sub-linear then is $\mathscr{L}(F)$ sub-regular or sub-linear, respectively ? The following counter-example for the sub-regular case is due to Hagauer [7].

Let $F : S \to a, a \to ba; a \to ab; b \to b$. For each $F' \lhd_u F$, $L(F')$ is the finite union of languages of the form $\Sigma_1^* \, a \, \Sigma_2^*$ and hence is regular. Consider $F'' \lhd F$, F'' : $S \to a;\ a \to b\bar{a}, \bar{a} \to ab;\ b \to b$. Clearly $L(F'') \cap b^* ab^* = \{b^n ab^n : n \geqq 1\}$ which is not regular.

A similar counter-example can be obtained for the sub-linear case.

However if $\mathscr{L}_u(F)$ is either sub-regular or sub-linear then is $\mathscr{L}(F)$ at least context-free ? We conjectured that the following EOL form F provided a counter-example to this question.

Let F be defined by the productions $S \to a$; $a \to a \mid aS$. It is clear that $\mathscr{L}(F)$ contains non-context-free languages. Consider F' : $S \to a \mid b$; $a \to aS \mid b$: $b \to bN$; $N \to b$. Now $F' \lhd F$ and $L(F') = \{ b^{F(i)} : F(i)$ is the ith Fibonacci number$\}$.

On the other hand by [3], $\mathscr{L}_u(F) \subseteq \mathscr{L}(CF)$. Hagauer [7] has recently shown that $\mathscr{L}_u(F)$ contains non-regular languages. Hence F is not a counter-example.

We now turn to the second question, specifically, for F a synchronized EOL form, when does $\mathscr{L}_u(F) \supseteq \mathscr{L}$ (REG). First of all notice that there is no result corresponding to theorem 4.7. Indeed consider F defined by $S \to a \mid aa \mid S \mid aS$; $a \to SS$, then F is complete by [6] but is not uni-complete and in particular the regular language $\{ ab \}$ cannot be obtained from F under uniform interpretation.

We provide a characterization of those synchronized forms which do in fact generate all regular languages. Let $F = (V, \Sigma, P, S)$ be a synchronized EOL form and x a word in $L(F)$. We say x can be *generated singly* if there is a derivation $S \Rightarrow^+ x$ in F in which the only terminal introducing productions are of the type $A \to y$, y in $\Sigma \cup \{ \varepsilon \}$. Similarly we say $X \subseteq L(F)$ is generated singly if each word x in X is generated singly.

We now have:

THEOREM 4.11: *Let* $F = (V, \Sigma, P, S)$ *be a synchronized* EOL *form,* $L(F) \subseteq \{ a_1, \ldots, a_n \}^* = \Sigma^*$.

Then $\mathscr{L}_u(F) \supseteq \mathscr{L}$ (REG) *iff* a_i^* *can be generated singly, for some* a_i *in* Σ.

Proof: if. Since $\mathscr{L}_u(F)$ is closed under intersection with regular sets, we can assume $L(F) = a_i^*$, and a_i^* is generated singly. Let $R \subseteq \Delta^*$ be an arbitrary regular set. Construct $F' \lhd_u F$ such that $L(F') = \Delta^*$, and then construct $F'' \lhd_u F'$ such that $L(F'') = L(F') \cap R = R$.

Only if: Suppose none of a_i^* can be generated singly, $1 \leqq i \leqq n$. Let $a_i^{r_i}$ be the shortest word in a_i^* which cannot be generated singly, $1 \leqq i \leqq n$. Let $t = \max(\{ r_i : 1 \leqq i \leqq n \})$ and Δ be an alphabet with nt elements. Since $\mathscr{L}$ (REG) $\subseteq \mathscr{L}_u(F)$, we have Δ^* is in $\mathscr{L}_u(F)$. Hence there exists $F' \lhd_u F(\mu)$ with $L(F') = \Delta^*$. Now $\mu(a_i) = \Delta_i$, where $\Delta = \cup \Delta_i$ and not all Δ_i are empty. Therefore there is an i, $1 \leqq i \leqq n$, with $\# \Delta_i \geqq t$. Let $c_1, c_2, \ldots, c_{r_i}$ be distinct symbols of Δ_i (this is alright since $r_i \leqq t$), and consider $x = c_1 c_2 \ldots c_{r_i}$ in $L(F')$. Since $\Delta_i = \mu(a_i)$ we must have x in $\mu(a_i^{r_i})$. However since $a_i^{r_i}$ is not generated singly, not all symbols can be made distinct under uniform interpretation. This provides a contradiction, hence the result. □

As a corollary we now obtain the following reduction results:

COROLLARY 4.12: *Let F be a synchronized* EOL *form. It is decidable whether* $\mathscr{L}_u(F) \supseteq \mathscr{L}$ (REG) *iff it is decidable whether* $L(F) = a^*$.

Surprisingly this latter question is still open.

Under the usual interpretation mechanism we have the much simpler characterization from [10].

PROPOSITION 4.13: *Let* $F = (V, \Sigma, P, S)$ *be a synchronized* EOL *form. Then* $\mathscr{L}(F) \supseteq \mathscr{L}$ (REG) *iff* $L(F) \supseteq a^*$, *for some a in* Σ.

5. CONCLUDING REMARKS

In carrying out the work reported we re-discovered the fact that little is known about specific forms under a specific interpretation mechanism. A case in point is the form $F : S \to a; a \to a \mid a S$ introduced in the previous section. Under the usual interpretations it is easy to construct interpretations yielding non-context-free languages, while $\mathscr{L}_u(F)$ appeared to be a sub-regular. That $\mathscr{L}_u(F)$ is not sub-regular has been demonstrated in [7] by using a cleverly constructed interpretation. It can be shown that we need only consider the following related question. Let $G : b \to a; a \to a \mid ab$ be an OL form with axiom b. Then it can be shown that $\mathscr{L}_{stab}(G)$ is sub-regular iff $\mathscr{L}_u(F)$ is sub-regular.

We say $G' \lhd_{stab} G$ if $G' \lhd G(\mu)$ and for all symbols a such that $a \to a$ is in G we take, at least, $a' \to a'$, for all a' in $\mu(a)$. Note that G' is a pure grammar. We may, of course, consider $\mathscr{L}_{stab}(H)$ for any EOL form H. However, any further investigation of stability preserving interpretations is left for another paper.

Second, we introduce a concept which is dual to that of generator. We say L is a (*uni-*) $\mathscr{L}$*-destroyer*, for some L in $\mathscr{L}$ (EOL) and $\mathscr{L} \subseteq \mathscr{L}$ (EOL), if for all synchronized forms F, $L(F) = L$ implies $\mathscr{L}(F) \neq \mathscr{L}$ ($\mathscr{L}_u(F) \neq \mathscr{L}$). Clearly we can speak about the family of all $\mathscr{L}$-destroyers, for a given $\mathscr{L}$, this we denote by $\mathscr{D}(\mathscr{L})$. For example, letting $\mathscr{L} = \mathscr{L}(CF)$ we know from [1] that $\{a^i b^i c^j d^j : i, j \geqq 1\}$ is a $\mathscr{L}(CF)$-destroyer. Indeed we conjecture that $\mathscr{D}(\mathscr{L}(CF))$ is equal to $\mathscr{L}$ (EOL)-$\mathscr{L}$ (LIN).

Third, consider the following modified definition of generator. Let $\mathscr{L}_1$, $\mathscr{L}_2$ be sub-EOL families, we say $\mathscr{L}_1$ is a (*uni-*) $\mathscr{L}_2$*-generator* if for all *synchronized* EOL forms F such that $\mathscr{L}(F) \supseteq \mathscr{L}_1$ ($\mathscr{L}_u(F) \supseteq \mathscr{L}_1$) we have $\mathscr{L}(F) \supseteq \mathscr{L}_2$ ($\mathscr{L}_u(F) \supseteq \mathscr{L}_2$). For example, $\mathscr{L}$ (OL) is a (uni-) $\mathscr{L}$ (EOL)-generator, since $\mathscr{L}_u(F)$ and $\mathscr{L}(F)$ are closed under intersection with a regular set and $\mathscr{L}(\text{EOL}) = \{L \cap R : L$ in $\mathscr{L}$ (OL) and R in $\mathscr{L}$ (REG)$\}$. We conjecture that we can remove the synchronization restriction in this case. Finally, if $\mathscr{L}(F) \supseteq \mathscr{L}(CF)$ and F is synchronized does this imply F is complete?

REFERENCES

1. J. ALBERT and H. A. MAURER, *The Class of Context-free Languages is Not an EOL Family*, Information Processing Letters, Vol. 6, No. 6, 1978, pp. 190-195.
2. J. ALBERT, H. A. MAURER and Th. OTTMANN, *On Sub-regular OL Forms*, Fundamenta Informatica (1981), to appear.
3. J. ALBERT, H. A. MAURER and G. ROZENBERG, *Simple EOL Forms Under Uniform Interpretation Generating CF Languages*, Fundamenta Informatica 3 (1980), pp. 141-156.
4. A. CREMERS and S. GINSBURG, *Context-Free Grammar Forms*, Journal of Computer and System Sciences, Vol. 11, 1975, pp. 86-116.
5. K. CULIK II and H. A. MAURER, *Propagating Chain-Free Normal Forms for EOL Systems*, Information and Control, Vol. 36, 1978, pp. 309-319.
6. K. CULIK II, H. A. MAURER and Th. OTTMANN, *Two-Symbol Complete EOL Forms*, Theoretical Computer Science, Vol. 6, 1978, pp. 69-92.
7. J. HAGAUER, *A Simple Variable-Free CF Grammar Generating a non Regular Language*, Bulletin of the EATCS, Vol. 6, October 1978, pp. 28-33.
8. H. A. MAURER, A. SALOMAA and D. WOOD, *EOL Forms*, Acta Informatica, Vol. 8, 1977, pp. 75-96.
9. H. A. MAURER, A. SALOMAA and D. WOOD, *Uniform Interpretations of L forms*, Information and Control, Vol. 36, 1978, pp. 157-173.
10. H. A. MAURER, A. SALOMAA and D. WOOD, *On Generators and Generative Capacity of EOL Forms*, Acta Informatica, Vol. 13, 1980, pp. 257-268.
11. V. K. VAISHNAVI and D. WOOD, *An Approach to a Unified Theory of Grammar and L Forms*, Information Sciences, Vol. 15, 1978, pp. 77-94.