Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

62

Automata, Languages and Programming

Fifth Colloquium, Udine, July 1978

Edited by G. Ausiello and C. Böhm

Springer-Verlag
Berlin Heidelberg New York

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

62

Automata, Languages and Programming

Fifth Colloquium, Udine, Italy, July 17–21, 1978

EATCS

Edited by G. Ausiello and C. Böhm

Springer-Verlag
Berlin Heidelberg New York 1978

Editors
Giorgio Ausiello
C.S.S.C.C.A.-C.N.R.
Istituto di Automatica
Via Eudossiana 18
Roma 00184/Italy

Corrado Böhm
Istituto Matematico
„Guido Castelnuovo"
Università di Roma
Piazzale delle Scienze
Roma 00100/Italy

AMS Subject Classifications (1970): 68-XX
CR Subject Classifications (1974): 4.1, 4.2, 5.2, 5.3

ISBN 3-540-08860-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-08860-1 Springer-Verlag New York Heidelberg Berlin

Printed in Germany

Printing and binding: Beltz Offsetdruck, Hemsbach/Bergstr.
2145/3140-543210

PREFACE

The Fifth International Colloquium on Automata, Languages and Programming (I.C.A.L.P.) was preceded by similar colloquia in Paris (1972), Saarbrücken (1974; see LNCS vol. 14), Edinburgh (1976), Turku (1977; see LNCS vol. 52). This series of conferences is sponsored by the European Association for Theoretical Computer Science, and is to be held each year (starting in 1976) in a different European country. The series of conferences will be published in the Lecture Notes in Computer Science.

In addition to the main topics treated - automata theory, formal languages and theory of programming - other areas such as computational complexity and λ-calculus are also represented in the present volume.

The papers contained in this volume and presented at Fifth I.C.A.L.P. in Udine (Italy) from July 17 to 21, 1978 were selected among over 90 submitted papers. The Program Committee, consisted of G. Ausiello, C. Böhm, H. Barendregt, R. Burstall, J.W. De Bakker, G. Degli Antoni, E. Engeler, J. Hartmanis, I.M. Havel, J. Loeckx, U. Montanari, M. Nivat, M. Paterson, A. Paz, J.F. Perrot. The editors feel very grateful to the other members of the Program Committee and to all referees that helped the Program Committee in evaluating the submitted papers. In particular we like to thank: L. Aiello, H. Alt, K. Apt, G. Barth, J. Bečvář, J. Berstel, A. Bertoni, D.P. Bovet, A. Celentano, G. Cioffi, R. Cohen, S. Crespi Reghizzi, J. Darlington, A. de Bruin, P.P.Degano, G. De Michelis, S. Even, G. Germano, J.A. Goguen, M.J.C. Gordon, P. Greussay, J. Gruska, A. Itai, R. Kemp, J. Král, I. Kramosil, G. Levi, M. Lucertini, A. Martelli, G. Mauri, D.B. McQueen, K. Mehlhorn, L.G.L.T. Merteens, P. Miglioli, R. Milner, C. Montangero, A. Nijholt, D. Park, G.D. Plotkin, B. Robinet, E. Rosenschein, E. Rosinger, B. Rovan, C.P. Schnorr, J. Schwarz, M.B.Smith, S. Termini, F. Turini, D. Turner, L.C. Valiant, P. van Emde Boas, J. van Lamsveerde, J. van Leeuwen, P.M.B. Vitanyi, W.W. Wadge, C. Wadsworth, C. Whitby-Strevens, P. Yoeli.

Finally we would like to express our gratitude to the Italian National Research Council for providing their financial support to the Fifth I.C.A.L.P. and the Centro di Studio dei Sistemi di Controllo e Calcolo Automatici (C.S.S.C.C.A.-C.N.R., Rome), the Centro Internazionale di Scienze Meccaniche (C.I.S.M., Udine), the Istituto di Automatica of the University of Rome that took the charge of the organization of the Conference.

G. Ausiello C. Böhm Rome, May 1978

CONTENTS

SIMPLE EOL FORMS UNDER UNIFORM INTERPRETATION GENERATING CF LANGUAGES

Jürgen Albert [1], Hermann Maurer [2], Grzegorz Rozenberg [3]

[1] University of Karlsruhe, W-Germany

[2] Technische Universitaet Graz, Austria

[3] University of Antwerp, Belgium

Abstract. In this paper we consider simple EOL forms (forms with a single terminal and single nonterminal) under uniform interpretations. We present a contribution to the analysis of generative power of simple EOL forms by establishing easily decidable necessary and sufficient conditions for simple EOL forms to generate (under uniform interpretations) CF languages only.

1. Introduction

The systematic study of grammatical similarity was begun in the pioneering paper [2] and extended to L-systems in [7]. The central concept introduced in [7] is the notion of an EOL form and its interpretations: each EOL system F - if understood as EOL form - generates, via the interpretation mechanism, a family of structurally related languages. Variations of the basic interpretation mechanism are possible, with the so-called uniform interpretation as one of the most promising candidates [8].

One of the central problems of EOL form theory is the systematic examination of language families generated by EOL forms and the study of generative capacity of EOL forms. Consequently, much effort has been concentrated on this type of problem and significant insights have been obtained. Fundamental results concerning generative capacity have already been established in [7] and ideas introduced there have been pursued in more detail in later papers. In particular, the notion of a complete form (a form generating all EOL languages) has been thoroughly investigated for simple forms (forms with a single terminal and single nonterminal) in [3]; the notion of goodness (a property involving subfamilies of families generated by EOL forms) has been further pursued in [10] and [11]; the class of CF languages cannot be generated by EOL forms under (ordinary) interpretation according to [1]; and [12] takes a new approach to the study of language classes generated by EOL forms by introducing the notion of generators. The generative capacity of non EOL L forms is the central topic of the papers [9], [13], [16], [4] and [6]. In contrast to the large number of papers quoted dealing with topics suggested in [7] , the notion of uniform interpretation, also proposed in [7] has received little attention sofar beyond [8]. We believe that a further study of uniform interpretations is essential and will increase the understanding of both L forms and L systems.

This paper is to be seen as a first but crucial step in this direction.

We present a complete classification of all simple EOL forms which under uniform interpretation yield CF languages only, yield at least one non-CF-language, respectively. This classification result, as expressed in our main Theorem 4 in Section 3, shows that some surprisingly complicated EOL forms generate under uniform interpretation nothing but CF languages, while other surprisingly trivial EOL forms (such as the forms with productions $S\to a$, $S\to SS$, $a\to a$, $a\to\varepsilon$ or just $S\to a$, $S\to S$, $a\to\varepsilon$, $a\to aS$) generate, even under uniform interpretation, non-CF-languages.

The rest of this paper is structured as follows. The next section 2 reviews definitions concerning L systems and L forms. Section 3 contains the results. Theorems 1,2,3 are of modest interest in themselves (and thus stated as theorems rather than lemmata). They lead to the central classification theorem, Theorem 4, whose proof constitutes the remainder of the paper. The table, presented in this proof, of the 38 possible types of minimal EOL forms yielding non-CF-languages should be of independent interest for further investigations of uniform interpretations of EOL forms.

2. Definitions

In this section some basic definitions concerning EOL forms and their interpretations are reviewed. An EOL system G is a quadruple $G=(V,\Sigma,P,S)$, where V is a finite set of symbols, the alphabet of G, $\Sigma\subset V$ is called the set of terminals, $V-\Sigma$ is the set of nonterminals, P is a finite set of pairs (α,x) with $\alpha\in V$, $x\in V^*$ and for each $\alpha\in V$ there is at least one such pair in P. The elements (α,x) of P are called productions or rules and are usually written as $\alpha\to x$. S is an element of $V-\Sigma$ and is called the startsymbol.

For words $x=\alpha_1\alpha_2\cdots\alpha_n$, $\alpha_i\in V$ and $y=y_1y_2\cdots y_n$, $y_i\in V^*$, we write $x\underset{G}{\Longrightarrow}y$ if $\alpha_i\to y_i$ is in P for every i.

We write $x \overset{o}{\underset{G}{\Rightarrow}} x$ for every $x \in V^*$ and $x \overset{m}{\underset{G}{\Rightarrow}} y$ for some $m \geq 1$ if there is a $z \in V^*$ such that $x \underset{G}{\Rightarrow} z \overset{m-1}{\underset{G}{\Rightarrow}} y$.

By $x \overset{*}{\underset{G}{\Rightarrow}} y$ we mean $x \overset{n}{\underset{G}{\Rightarrow}} y$ for some $n \geq 0$, and $x \overset{+}{\underset{G}{\Rightarrow}} y$ stands for $x \overset{n}{\underset{G}{\Rightarrow}} y$ for some $n \geq 1$.
The language generated by G is denoted by L(G) and defined by
$L(G) = \{x \in \Sigma^* \mid S \overset{*}{\underset{G}{\Rightarrow}} x\}$.
The family of all EOL languages is denoted by $\mathcal{L}_{EOL}$, i.e. $\mathcal{L}_{EOL} = \{L(G) \mid G$ is an EOL system$\}$. Similarly, we denote by $\mathcal{L}_{FIN}$, $\mathcal{L}_{REG}$, $\mathcal{L}_{LIN}$ and $\mathcal{L}_{CF}$ the classes of finite, regular, linear and context-free languages, respectively. We now review the notion of an EOL form F and its interpretations as introduced in [7] for the definition of structurally related EOL systems.

An EOL form F is an EOL system $F=(V,\Sigma,P,S)$. An EOL system $F'=(V',\Sigma',P',S')$ is called an interpretation of F (modulo μ), symbolically $F' \lhd F(\mu)$ if μ is a substitution defined on V and (i) - (v) hold:

(i) $\mu(A) \subseteq V'-\Sigma'$ for each $A \in V-\Sigma$,
(ii) $\mu(a) \subseteq \Sigma'$ for each $a \in \Sigma$,
(iii) $\mu(\alpha) \cap \mu(\beta) = \phi$ for all α, β in V, $\alpha \neq \beta$,
(iv) $P' \subseteq \mu(P)$ where
$\mu(P) = \{\beta \to y \mid \alpha \to x \in P, \beta \in \mu(\alpha), y \in \mu(x)\}$
(v) S' is in $\mu(S)$.

The family of EOL forms generated by F is defined by $\mathcal{G}(F) = \{F' \mid F' \lhd F\}$ and the family of languages generated by F is $\mathcal{L}(F) = \{L(F') \mid F' \lhd F\}$.

An important modification of this type of interpretation, first introduced in [8] , is called uniform interpretation and defined as follows. F' is a uniform interpretation of F, in symbols $F' \lhd_u F$, if $F' \lhd F$ and in (iv) even $P' \subseteq \mu_u(P)$ holds, where
$\mu_u(P) = \{\alpha_o' \to \alpha_1' \ldots \alpha_t' \in \mu(P) \mid \alpha_o \to \alpha_1 \ldots \alpha_t \in P, \alpha_i' \in \mu(\alpha_i)$, and $\alpha_r = \alpha_s \in \Sigma$ implies $\alpha_r' = \alpha_s'\}$

Thus, the interpretation has to be uniform on terminals. In analogy to the above definitions we introduce $\mathcal{G}_u(F) = \{F' \mid F' \lhd_u F\}$ and $\mathcal{L}_u(F) := \{L(F') \mid F' \lhd_u F\}$.

For a word x we use alph(x) to denote the set of all symbols occuring in x and for a language M we generalize this notion by
$\text{alph}(M) = \bigcup_{x \in M} \text{alph}(x)$.
Given some finite set Z, card (Z) is the number of elements contained

in Z.
And finally, an EOL form $F=(V,\Sigma,P,S)$ is called simple, if card (Σ) = card $(V-\Sigma) = 1$.
For all other notions not explicitely defined in this section we refer to [15].

3. Results

The first three theorems of this section list conditions under which a simple EOL form under uniform interpretation generates CF languages only. Based on these theorems the main result of this paper, a "classification theorem", is presented as Theorem 4. The classification theorem gives necessary and sufficient conditions under which simple EOL forms under uniform interpretation generate CF languages only.

Theorem 1

Let $F=(\{S,a\},\{a\},P,S)$ be an EOL form where P does not contain rules of the types $S\to a^i$ for any $i\geq 1$ and $a\to a^j$ for any $j\geq 2$. Then $\mathcal{L}_u(F)\subset \mathcal{L}_{CF}$.

Proof

Suppose F' is a uniform interpretation of F, $F'=(V,\Sigma,P',S')$ and $x\in L(F')$. Then for each $b\in alph(x)$ the rule $b\to b$ must be in P', since b cannot be derived by rules $A\to b^i$, $i\geq 1$ or $b\to b^j$, $j\geq 2$. Thus, $b\to b\in P'$ for each $b\in alph(L(F'))$.

Hence $L(F')\in \mathcal{L}_{CF}$ by a result in [14] and [5].

Remark

It can be shown readily that for the EOL form $F_2=(\{S,a\},\{a\},P_2,S)$ with $P_2=\{S\to\varepsilon,\ S\to aS,\ a\to a,\ a\to S\}$ we have $\mathcal{L}_u(F_2)=\mathcal{L}_{CF}$. Thus the inclusion in Theorem 1 is not necessarily a proper inclusion.

Theorem 2

Let $F=(\{S,a\},\{a\},P,S)$ be an EOL form, where
$P\subset\{S\to a^i\mid i\geq 0\}\cup\{S\to S\}\cup\{a\to a\}\cup\{a\to v\mid alph(v)=\{S,a\}\}$.
Then $\mathcal{L}_u(F)\subsetneq\mathcal{L}_{CF}$.

Proof

We consider an arbitrary uniform interpretation $F'=(V,\Sigma,P',S')$ of F and show how F' can be simulated by a context-free grammar G. The main idea leading to the construction of G is that big terminal derivations in F' consist of "marco-steps", which can be executed in a sequential manner. More specifically, if b is a terminal symbol of Σ and b generates a terminal word x in F' observe that $b \to b$ must be in P'.

Furthermore, if $b \underset{F'}{\Longrightarrow} b^{i_0} A_1 b^{i_1} A_2 \ldots A_t b^{i_t} \overset{m}{\underset{F'}{\Longrightarrow}} x$ where

$A_1, \ldots, A_t$ are in $V-\Sigma$, $t \geq 1$, $\sum_{\nu=0}^{t} i_\nu \geq 1$ and $m >$ card $(V-\Sigma)$, then there is a derivation

$b \underset{F'}{\Longrightarrow} b^{i_0} A_1 b^{i_1} A_2 \ldots A_t b^{i_t} \overset{\tau}{\underset{F'}{\Longrightarrow}} y \overset{\tau'}{\underset{F'}{\Longrightarrow}} x$ where $\tau \leq$ card $(V-\Sigma)$ and

$C \overset{+}{\underset{F'}{\Longrightarrow}} C$ for each symbol $C \in$ alph(y).

This can be shown by the pigeon-hole-principle. We now construct a context-free grammar $G=(V_1, \Sigma, P_1, S')$ as follows:

$V_1 := \{S'\} \cup \{\bar{b} \mid b \to b \in P', b \in \Sigma\}$.

P_1 contains exactly the following rules:

1. For each derivation $S' \underset{F'}{\Longrightarrow} A_1 \underset{F'}{\Longrightarrow} A_2 \underset{F'}{\Longrightarrow} \ldots \underset{F'}{\Longrightarrow} A_k \underset{F'}{\Longrightarrow} b^j$

   where $k \geq 0$, $A_1, \ldots, A_k$ are in $V-\Sigma$, $b \in \Sigma$ and $j \geq 0$ let $S' \to \bar{b}^j$ be in P_1 if $b \to b$ is in P' or $S' \to b^j \in P_1$ if $b \to b \notin P'$.

2. For each rule $b \to b$ in P', $b \in \Sigma$, let $\bar{b} \to b$ be in P_1.

3. Finally let us consider the terminal derivations of the type

   $b \underset{F'}{\Longrightarrow} b^{i_0} A_1 b^{i_1} A_2 \ldots A_t b^{i_t} \overset{\tau}{\underset{F'}{\Longrightarrow}} b_1 b_2 \ldots b_k$

   where $t \geq 1$, $A_1, \ldots, A_t \in V-\Sigma$, $\sum_{\nu=0}^{t} i_\nu \geq 1$,

   $\tau \leq$ card $(V-\Sigma)$ and $b_1, b_2, \ldots b_k \in \Sigma$.

3.1 If there is any $j \in \{1, \ldots, k\}$ such that

   $b_j \to b_j \notin P'$ and in $b^{i_0} A_1 b^{i_1} A_2 \ldots A_t b^{i_t} \overset{\tau}{\underset{F'}{\Longrightarrow}} b_1 b_2 \ldots b_k$

   there is no symbol C generating b_j, for which $C \overset{+}{\underset{F'}{\Longrightarrow}} C$ holds, then let $\bar{b} \to b_1 b_2 \ldots b_k$ be in P_1.

3.2 Otherwise let $\bar{b} \to \tilde{b}_1 \tilde{b}_2 \ldots \tilde{b}_k$ be in P_1, where

$$\tilde{b}_\nu = \begin{cases} \bar{b}_\nu & \text{if } b_\nu \to b_\nu \text{ is in P'} \\ b_\nu & \text{if } b_\nu \to b_\nu \notin P'\text{, but } b_\nu \text{ is generated by a C for which } C \overset{+}{\underset{F'}{\Longrightarrow}} C \text{ holds.} \end{cases}$$

Following the construction of P_1 one can show by inductive proofs for any terminal $b \in \Sigma$ and any terminal word $x \in \Sigma^*$:

$b \overset{+}{\underset{F'}{\Longrightarrow}} x$ if and only if $\bar{b} \overset{+}{\underset{G}{\vdash}} x$.

Together with the fact that each terminal derivation $S' \overset{+}{\underset{F'}{\Longrightarrow}} z$ begins with $S' \underset{F'}{\Longrightarrow} S_1 \underset{F'}{\Longrightarrow} S_2 \underset{F'}{\Longrightarrow} \ldots \underset{F'}{\Longrightarrow} S_k \underset{F'}{\Longrightarrow} d^j \underset{F'}{\Longrightarrow} \ldots$

where $k \geq 0$, $j \geq 0$ and $S' \to d^j$ or $S' \to \bar{d}^j$ is in P_1, we get $L(F') = L(G)$ and hence $\mathcal{L}_u(F) \subset \mathcal{L}_{CF}$.

It remains to prove that this inclusion is proper. If we take m to be the maximum of the numbers j such that $S \to a^j$ is in P then $\{a^{m+1}\}$ clearly is not in $\mathcal{L}_u(F)$.

Remark

If the EOL form F fulfills the conditions of Theorem 2 $\mathcal{L}_u(F)$ may well contain non-linear languages.

For example, if we take

$F=(\{S,a\},\{a\},\{S\to aa, S\to S, a\to a, a\to SaS\},S)$ let us consider the interpretation:

$F'=(\{S',A,B,a,b,c\},\{a,b,c\},P',S')$ where P' consists of:

$S'\to bb$, $b\to AbB$, $b\to b$, $A\to A$, $A\to aa$,

$B\to B$, $B\to cc$, $a\to AaA$, $c\to BcB$.

Then $L(F')=\{a^{2m}bc^{2m}a^{2n}bc^{2n} \mid m,n\geq 0\}$

which is known to be not linear.

Theorem 3

Let $F=(\{S,a\},\{a\},P,S)$ be an EOL form where

$P \subset \{S\to a^j \mid j\geq 0\} \cup \{a\to\varepsilon\} \cup \{a\to v \mid alph(v) = \{S,a\}\}$.

Then $\mathcal{L}_u(F) \subsetneqq \mathcal{L}_{CF}$.

Proof

The following lemma will lead us to the construction of a context-free grammar G simulating a uniform interpretation F' of F.

Lemma

Suppose F is as in Theorem 3 and $F'=(V,\Sigma,P',S')$ is a uniform interpretation of F.

If $b \overset{k}{\underset{F'}{\Longrightarrow}} x$ for some $b\in\Sigma$, $x\in\Sigma^*$, $k\geq 3$ then we have $b \overset{k+\nu}{\underset{F'}{\Longrightarrow}} x$ for each $\nu\geq 0$.

Proof of the Lemma:

We note first that whenever a terminal symbol b generates a terminal word in F' then the rule $b\to\varepsilon$ must be in P'. Now let us consider the derivation

$b \underset{F'}{\Longrightarrow} X_1 \underset{F'}{\Longrightarrow} X_2 \underset{F'}{\Longrightarrow} X_3 \overset{k-3}{\underset{F'}{\Longrightarrow}} x$

If $X_1=\varepsilon$ then $b \overset{k+\nu}{\underset{F'}{\Longrightarrow}} \varepsilon$ holds trivially. Otherwise X_1 is of the form

$X_1=b^{i_o}A_1b^{i_1}A_2\ldots A_tb^{i_t}$,

$t\geq 1$, $A_1,A_2,\ldots,A_t\in V-\Sigma$, $\sum_{\nu=o}^{t} i_\nu \geq 1$.

And $X_2=W_o\, a_1^{j_1}W_1a_2^{j_2}\ldots a_t^{j_t}W_t$ where $b^{i_\nu} \underset{F'}{\Longrightarrow} W_\nu$ and

$A_\nu \underset{F'}{\Longrightarrow} a_\nu^{\;j_\nu}$, $j_\nu\geq 0$, $a_\nu\in\Sigma$ for $\nu=0, 1,\ldots,t$.

If $j_\nu \geq 1$ the rule $a_\nu \to \varepsilon$ must be in P' since a_ν appears in a terminal derivation. So $A_\nu \overset{2}{\underset{F'}{\Longrightarrow}} \varepsilon$ for each $\nu \in \{1,2,..,t\}$.

We can write $X_1 = U_1 b U_2$ because of $b \in alph(X_1)$ and using the fact $U_i \underset{F'}{\Longrightarrow} U_i' \underset{F'}{\Longrightarrow} \varepsilon$ for i=1,2 we have a derivation $U_1'U_1bU_2U_2' \underset{F'}{\Longrightarrow} U_1'U_1bU_2U_2'$ and therefore

$b \underset{F'}{\Longrightarrow} U_1bU_2 \underset{F'}{\Longrightarrow} U_1'U_1bU_2U_2' \overset{\nu}{\underset{F'}{\Longrightarrow}} U_1'U_1bU_2U_2' \underset{F'}{\Longrightarrow} X_2 \overset{k-2}{\underset{F'}{\Longrightarrow}} x$ for each $\nu \geq 0$.

This establishes our lemma.

We now define a context-free grammar $G=(V_1,\Sigma,P_1,\bar{S}')$ which will simulate F':

$V_1 := \{\bar{Z} \mid Z \in V, \exists x \in \Sigma^* ; \; Z \overset{3}{\underset{F'}{\Longrightarrow}} x\} \cup \{\bar{S}'\}$;

P_1 contains the following rules:

1. For each rule $S' \to b^j$ in P', $b \in \Sigma$, $j \geq 0$ let $\bar{S}' \to b^j$ be in P_1.
2. For each rule $Z_o \to Z_1 Z_2 \ldots Z_k$ in P' where $k \geq 0$ and $Z_o, Z_1, \ldots, Z_k$ are in V_1 let $\bar{Z}_o \to \bar{Z}_1 \bar{Z}_2 \ldots \bar{Z}_k$ be in P_1.
3. For each $Z \in V$, $x \in \Sigma^*$ such that $Z \overset{3}{\underset{F'}{\Longrightarrow}} x$ let $\bar{Z} \to x$ be in P_1.

To prove $L(F') \subset L(G)$ we show that $S' \overset{k}{\underset{F'}{\Longrightarrow}} x$, $k \geq 1$, $x \in \Sigma^*$ implies $\bar{S}' \overset{+}{\underset{G}{\vdash}} x$.

For $k \in \{1,2\}$ the rules $\bar{S}' \to x$ are in P_1. If $k \geq 3$ we consider the derivationparts $S' \overset{k-3}{\underset{F'}{\Longrightarrow}} Y_1 Y_2 \ldots Y_j \overset{3}{\underset{F'}{\Longrightarrow}} x_1 x_2 \ldots x_j = x$ where $Y_\nu \in V$ and $Y_\nu \overset{3}{\underset{F'}{\Longrightarrow}} x_\nu$ for $\nu = 1(1)j$.

For each rule $Z_o \to Z_1 Z_2 \ldots Z_r$ applied in $S' \overset{k-3}{\underset{F'}{\Longrightarrow}} Y_1 Y_2 \ldots Y_j$ there is $\bar{Z}_o \to \bar{Z}_1 \bar{Z}_2 \ldots \bar{Z}_r$ in P_1 and all the rules $\bar{Y}_\nu \to x_\nu$ can be found in P_1. So we have $\bar{S}' \overset{*}{\underset{G}{\vdash}} \bar{Y}_1 \bar{Y}_2 \ldots \bar{Y}_j \overset{j}{\underset{G}{\vdash}} x_1 x_2 \ldots x_j = x$.

For the reverse inclusion $L(G) \subset L(F')$ we proceed in a similar way.

A terminal derivation in one step $\bar{S}' \underset{G}{\vdash} x$ is possible only if $S' \overset{t}{\underset{F'}{\Longrightarrow}} x$ for t=1 or t=2.

If $\bar{S}' \overset{k}{\underset{G}{\vdash}} x$ for some $k>1$ then there is also a derivation in G of the type $\bar{S}' \overset{k-j}{\underset{G}{\vdash}} \bar{Y}_1 \bar{Y}_2 \ldots \bar{Y}_j \overset{j}{\underset{G}{\vdash}} x_1 x_2 \ldots x_j = x$, where $\bar{Y}_\nu \to x_\nu \in P_1$ for $\nu = 1,2,\ldots,j$.

Then $Y_\nu \overset{3}{\underset{F'}{\Longrightarrow}} x_\nu$ and by the above lemma even $Y_\nu \overset{\mu}{\underset{F'}{\Longrightarrow}} x_\nu$ holds for all $\mu \geq 3$.

So we can just apply the according rules to those used in $\bar{S}' \overset{k-j}{\underset{G}{\vdash}} \bar{Y}_1 \bar{Y}_2 \ldots \bar{Y}_j$ in the parallel mode of F' and equalize the derivation-length by taking a suitable μ_ν in each $Y_\nu \overset{\mu_\nu}{\underset{F'}{\Longrightarrow}} x_\nu$.

Thus $S' \overset{*}{\underset{F'}{\Longrightarrow}} x$ and $L(F')=L(G)$.

By the same reasoning as in Theorem 2 one can show that the inclusion of $\mathcal{L}_u(F)$ in $\mathcal{L}_{CF}$ is proper.

We are now prepared to give in our main theorem a complete list of EOL forms $F=(\{S,a\},\{a\},P,S)$ such that $\mathcal{L}_u(F)$ is contained in $\mathcal{L}_{CF}$.

For the sake of clarity we will present these results in tabular form using a partitioning of the set of all possible productions $\Pi=\{X\to Y \mid X\in\{S,a\}, Y\in\{S,a\}^*\}$, into 13 disjoint subsets.

The following table defines rule-sets $\Pi_1,\Pi_2,\ldots,\Pi_{11}$ as follows. If in line Π_i a "T" occurs in columns with production sets $x_1,x_2,\ldots,x_t$ then $\Pi_i=x_1\cup x_2\cup\ldots\cup x_t$. Thus, for example,

$\Pi_1=\{S\to\varepsilon\}\cup\{S\to a\}\cup\{S\to a^i \mid i\geq 2\}\cup\{S\to S\}\cup\{S\to S^i \mid i\geq 2\}\cup\{S\to a^iSa^j \mid i+j\geq 1\}\cup$
$\cup\{S\to w \mid \#_a(w)\geq 1, \#_S(w)\geq 2\}\cup\{a\to\varepsilon\}$

Each Π_i is a superset of sets of productions of simple EOL forms generating (under uniform interpretation) only CF languages.

Note that the table is nearly symmetric in S and a, since the last entry for productions with a on lefthand side is the union of last two entries for productions with S on lefthand side.

	$\{S\to\varepsilon\}$	$\{S\to a\}$	$\{S\to a^i \mid i\geq 2\}$	$\{S\to S\}$	$\{S\to S^i \mid i\geq 2\}$	$\{S\to a^iSa^j \mid i+j\geq 1\}$	$\{S\to w \mid \#_a(w)\geq 1, \#_S(w)\geq 2\}$	$\{a\to\varepsilon\}$	$\{a\to a\}$	$\{a\to a^i \mid i\geq 2\}$	$\{a\to S\}$	$\{a\to S^i \mid i\geq 2\}$	$\{a\to v \mid \#_a(v)\geq 1, \#_S(v)\geq 1\}$
Π1	T	T	T	T	T	T	T	T					
Π2	T	T	T	T	T	T	T		T				
Π3	T	T	T	T		T	T						T
Π4	T	T	T	T		T		T	T				
Π5	T	T	T	T					T				T
Π6	T	T	T					T					T
Π7	T	T		T				T	T		T		
Π8	T			T	T	T	T	T	T		T	T	T
Π9	T			T	T			T	T	T	T	T	T
Π10		T				T	T				T		T
Π11				T	T	T	T	T	T	T	T	T	T

Theorem 4 (Classification Theorem)

Let F=({S,a},{a},P,S) be an EOL form. Then $\mathcal{L}_u(F) \subset \mathcal{L}_{CF}$ if and only if $P \subset \Pi_i$ for some $i \in \{1,2,\ldots,11\}$.

Proof

The "if-statement" in the theorem above can be shown easily. The cases $P \subset \Pi_1$ and $P \subset \Pi_2$ are given by results in [17] and [14], [5], respectively. If $P \subset \Pi_3$ then L(F) is finite and so $\mathcal{L}_u(F) \subset \mathcal{L}_{FIN}$. If $P \subset \Pi_4$ then obviously $\mathcal{L}_u(F) \subset \mathcal{L}_{LIN}$ holds.

The cases $P \subset \Pi_5$, $P \subset \Pi_6$, $P \subset \Pi_8$ have been treated in Theorems 2,3 and 1, respectively. The remaining cases $P \subset \Pi_7$, $P \subset \Pi_9$, $P \subset \Pi_{10}$, $P \subset \Pi_{11}$ are trivial, since the generated languages consist at most of one-symbol-words or the empty word.

The "only if-statement" can be reformulated equivalently to:
Let F=({S,a},{a},P,S) be an EOL form such that
(*) $P \cap (\Pi - \Pi_i) \neq \phi$ for each $i \in \{1,2,\ldots,11\}$, then $\mathcal{L}_u(F) - \mathcal{L}_{CF} \neq \phi$.

Thus, if F satisfies condition (*) above then we can find a uniform interpretation F' of F, such that L(F') is not context-free.

For the evaluation of the expression
(*) $(P \cap (\Pi - \Pi_1) \neq \phi) \wedge (P \cap (\Pi - \Pi_2) \neq \phi) \ldots \wedge (P \cap (\Pi - \Pi_{11}) \neq \phi)$ we proceed as follows.

The production-sets $(\Pi - \Pi_1)$, $(\Pi - \Pi_2)$, ..., $(\Pi - \Pi_{11})$ are written according to the partitioning of Π.

Thus, e.g. $(P \cap (\Pi - \Pi_1) \neq \phi)$ equals
$(P \cap \{a \to a\} \neq \phi) \vee (P \cap \{a \to a^i \mid i \geq 2\} \neq \phi) \vee (P \cap \{a \to S\} \neq \phi) \vee (P \cap \{a \to S^i \mid i \geq 2\} \neq \phi) \vee$
$\vee (P \cap \{a \to v \mid \#_a(v) \geq 1, \#_S(v) \geq 1\} \neq \phi)$.

The same decomposition is done with the sets $P \cap (\Pi - \Pi_2), \ldots, P \cap (\Pi - \Pi_{11})$.
The reordering of these elementary expressions in (*) "from left to right" gives rise to 38 alternatives K_j such that (*) equals
$K_1 \vee K_2 \vee \ldots \vee K_{38}$.

And in $K_1, \ldots, K_5$ we have $P \cap \{S \to \varepsilon\} \neq \phi$, in $K_6, \ldots, K_{26}$ $P \cap \{S \to a\} \neq \phi$ holds and in $K_{27}, \ldots, K_{38}$: $P \cap \{S \to a^i \mid i \geq 2\} \neq \phi$.