# ON GOOD ET0L FORMS

by

Sven Skyum

DAIMI PB-74

April 1977

Institute of Mathematics   University of Aarhus

DEPARTMENT OF COMPUTER SCIENCE

Ny Munkegade - 8000 Aarhus C - Denmark

Phone 06-12 83 55

# ON GOOD ET0L FORMS

by

Sven Skyum
Department of Computer Science
University of Aarhus

Mailing Address:

Department of Computer Science
University of Aarhus
Ny Munkegade
8000 Aarhus C
Denmark

## ABSTRACT

This paper continues the study of ET0L forms and good E0L forms done by Maurer, Salomaa and Wood. It is proven that binary very complete ET0L forms exist, good synchronized ET0L forms exist and that no propagating or synchronized ET0L form can be very complete.

# INTRODUCTION

Maurer, Salomaa, and Wood introduced in [1] and [2] the notion of E0L and ET0L forms and their interpretations. An ET0L form F defines a family of ET0L systems $\mathcal{G}(F)$ which are "structurally" similar and a family of languages $\mathcal{L}(F)$, namely those languages generated by systems in $\mathcal{G}(F)$. In [3] Maurer, Salomaa and Wood study good and complete E0L forms. Many of the theorems in [3] are trivially valid for ET0L forms as well as for E0L forms but they will not appear in this paper. In this paper we show that in contrast to what is the case for E0L forms, there exist good synchronized ET0L forms. Finally we show that, (as for E0L forms) there exist binary vomplete (short for very complete) ET0L forms and that no propagating or synchronized ET0L form is vomplete.

## DEFINITIONS AND BASIC RESULTS

We follow the definitions of ETOL forms and interpretations given in [1], [2], and [3]. EOL forms will not be defined explicitly. We can consider an EOL form as an ETOL form with one table.

### Definition

An <u>ETOL system</u> (or <u>n-ETOL system</u>) is an (n+3)-tuple $G = (V, \Sigma, P_1, \ldots, P_n, S)$, where $n \geq 1$, $V$ is an alphabet, $\Sigma \subseteq V$ is the terminal alphabet, $S \in V-\Sigma$ is the <u>start symbol</u>. For all $i$, $1 \leq i \leq n$, $P_i$ is a finite set of pairs $(A,\alpha)$ with $A \in V$ and $\alpha \in V^*$ such that for each $A \in V$ at least one such pair is in $P_i$. The elements $(A,\alpha)$ are called <u>rules</u> or <u>productions</u> and are usually written $A \underset{P_i}{\rightarrow} \alpha$ or just $A \rightarrow \alpha$. The sets $P_i$ are called <u>tables</u>.

<u>Definition</u> Let $G = (V, \Sigma, P_1, \ldots, P_n, S)$ be an n-ETOL system. For words $x = A_1A_2\ldots A_m$ and $y = \alpha_1\alpha_2\ldots\alpha_m$ with $A_i \rightarrow \alpha_i$ in $P_j$ for $1 \leq i \leq m$ and some $P_j$ we write $x \underset{P_j}{\Rightarrow} y$ or $x \Rightarrow y$. $\Rightarrow^+$ (and $\Rightarrow^*$) are the transitive (and reflexive) closure of $\Rightarrow$. The language generated by G is

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

Notice that in contrast to the usual definition of ETOL systems, the start symbol cannot be a terminal.

<u>Definition</u> Let $G = (V, \Sigma, P_1, \ldots, P_n, S)$ be an n-ETOL system. For a word $x$, $|x|$ denotes the length of $x$ and <u>Alph(x)</u> denotes the minimal alphabet such that $x \in \mathrm{Alph}(x)^*$. For all $1 \leq i \leq n$ let <u>maxr($P_i$)</u> $= \max\{|\alpha| \mid A \rightarrow \alpha \text{ in } P_i\}$, and let <u>maxr(G)</u> $= \max\{\mathrm{maxr}(P_i) \mid 1 \leq i \leq n\}$. A symbol $B \in V$ is <u>reachable</u>

(from S) if $S \Rightarrow^* \alpha B \beta$ for some words $\alpha, \beta \in V^*$. G is <u>reduced</u> if each $B \in V$ is reachable. G is <u>separated</u> if for all productions $A \to \alpha$ in $P_1, \ldots, P_n$ $\alpha \in (V-\Sigma)^*$ if $A \in \Sigma$ and $\alpha \in \Sigma \cup (V-\Sigma)^*$ otherwise. G is <u>propagating</u> if for all productions $A \to \alpha$ in $P_1, \ldots, P_n$ $\alpha \neq e$, the empty word. G is <u>synchronized</u> if, for all $a \in \Sigma$, $a \Rightarrow^+ \alpha$ implies $\alpha \notin \Sigma^*$. G is <u>short</u>, if for all $P_i$, $A \to \alpha \in P_i$ implies $|\alpha| \leq 2$. Finally G is <u>binary</u> if each rule in $P_1, \ldots, P_m$ is one of the forms $A \to e$, $A \to a$, $A \to B$, $A \to BC$, $a \to A$, where $a \in \Sigma$ and $A, B, C \in V-\Sigma$.

<u>Definition</u> An <u>ETOL form</u> (or <u>n-ETOL form</u>) is an ETOL system $F = (V, \Sigma, P_1, \ldots, P_n, S)$. An ETOL system $G = (V', \Sigma', P_1', \ldots, P_n', S')$ is an <u>interpretation of F</u> (<u>modulo $\mu$</u>), $G \lhd F(\mu)$, or simply $G \lhd F$, if $\mu$ is a substitution defined on V satisfying (i)-(v):

(i) $\mu(A) \subseteq V'-\Sigma'$ for $A \in V-\Sigma$,

(ii) $\mu(a) \subseteq \Sigma'$ for $a \in \Sigma$,

(iii) $\mu(\alpha) \cap \mu(\beta) = \emptyset$ for any symbols $\alpha \neq \beta$,

(iv) for all $1 \leq i \leq n$, $P_i' \subseteq \mu(P_i)$, where
$\mu(P_i) = \{(A', \alpha') \in V' \times V'^* \mid A' \in \mu(A), \alpha' \in \mu(\alpha)$ for some $A \in V, \alpha \in V^*$ such that $A \to \alpha \in P_i\}$,

(v) $S' \in \mu(S)$.

<u>The family of ETOL systems generated by F</u>, denoted $\mathcal{G}(F)$, is:
$\mathcal{G}(F) = \{G \mid G \lhd F\}$.

The family of languages generated by F, denoted $\mathcal{L}(F)$, is:
$\mathcal{L}(F) = \{L(G) \mid G \in \mathcal{G}(F)\}$.

Since an ETOL form is an ETOL system, and conversely, we will allow ourselves to use the term "form" in the rest of this paper.

Definition Two ETOL forms $F_1$ and $F_2$ are equivalent if $L(F_1) = L(F_2)$ and form equivalent if $\mathcal{L}(F_1) = \mathcal{L}(F_2)$.

The following lemmata are either contained in [2] or are a slight modification of some in [2].

Lemma 1 Let $F = (V, \Sigma, P_1, \ldots, P_m, S)$ be an ETOL form and $F' = (V', \Sigma', P'_1, \ldots, P'_m, S') \lhd F(\mu)$. Then for each derivation $x_0 \underset{P'_{i_1}}{\Rightarrow} x_1 \underset{P'_{i_2}}{\Rightarrow} \cdots \underset{P'_{i_k}}{\Rightarrow} x_k$ in $F'$, $\mu^{-1}(x_0) \underset{P_{i_1}}{\Rightarrow} \mu^{-1}(x_1) \underset{P_{i_2}}{\Rightarrow} \cdots \underset{P_{i_k}}{\Rightarrow} \mu^{-1}(x_k)$ is a derivation in $F$.

Lemma 2 For all ETOL forms $F$ a form equivalent reduced ETOL form $F'$ can be constructed.

Because of this lemma we will always assume the forms in this paper to be reduced.

Lemma 3 Let $F = (V, \Sigma, P_1, \ldots, P_m, S)$ and $\overline{F} = (V \cup \overline{V}, \Sigma, \overline{P}_1, \ldots, \overline{P}_m, S)$ with $V \cap \overline{V} = \emptyset$. If there are integers $k_1, k_2, \ldots, k_m$, such that (a) holds if and only if (b) holds for some $i$, then $\mathcal{L}(F) = \mathcal{L}(\overline{F})$.

(a) $A \in V$, $A \underset{\overline{P}_{i_1}}{\Rightarrow} x_1 \underset{\overline{P}_{i_2}}{\Rightarrow} x_2 \underset{\overline{P}_{i_3}}{\Rightarrow} \cdots \underset{\overline{P}_{i_t}}{\Rightarrow} x_t$, $x_t \in V^*$, and $x_j \notin V^*$ for $1 \leq j < t$.

(b) $t = k_i$, $i_1 = i_2 = \ldots = i_t = i$, $A \to x_t$ in $P_i$, $x_j \in \overline{V}^+$ for $1 \leq j < t$, $A \to x_1$ in $\overline{P}_i$, and $x_{j-1} \underset{\overline{P}_i}{\Rightarrow} x_j$ for $1 < j \leq t$.

Definition  Let F be an ET0L form and $\mathfrak{F}$ a family of languages. We call F $\mathfrak{F}$-complete or complete for $\mathfrak{F}$ if $\mathfrak{L}(F) = \mathfrak{F}$; if $\mathfrak{F}$ is the family of ET0L languages, then we simply call F complete instead of ET0L-complete. We call F good, if for each ET0L form $\overline{F}$ with $\mathfrak{L}(\overline{F}) \subseteq \mathfrak{L}(F)$ an ET0L form F' exists such that $F' \lhd F$ and $\mathfrak{L}(F') = \mathfrak{L}(\overline{F})$. F is called bad if it is not good. F is called vomplete (short for very complete) if it is complete and good.

## RESULTS

Most of the theorems on good EOL forms in [3] are easily shown to be valid for ETOL forms as well. Properties which are different for EOL and ETOL forms are related to synchronization. There are two "canonical" ways to synchronize an ETOL form. The first is to introduce a marked version of the terminals and make these new nonterminals and then change the production by marking the terminals and add $a' \to a$, $a \to N$, $N \to N$ to all tables for all terminals a. N is a new nonterminal. The second, which has no counterpart in the EOL case, is to add a new table consisting of the productions $a' \to a$, $a \to N$, $N \to N$, $A \to N$ for terminals a and nonterminals A. $a \to N$, $N \to N$ is added to the rest of tables. The following lemma and theorem show different properties for synchronized EOL and ETOL forms.

Lemma 4 The synchronized ETOL form

$F = (\{S,a,N\}, \{a\}, \{S \to SS;\ a \to N;\ N \to N\}, \{S \to a;\ a \to N;\ N \to N\}, S)$

generates no nonempty finite languages.

Proof Immediate.

□

All synchronized EOL forms generating nonempty languages generate finite nonempty languages. This is used to prove that no good synchronized EOL form exists (Th. 2.6 in [3]). The following theorem shows that good synchronized ETOL forms exist. Surprisingly enough the form shown to exist generates finite languages only!

<u>Theorem 5</u> The synchronized ETOL form

$F = (\{S, a, N\}, \{a\}, \{S \to S;\ a \to N; N \to N\}, \{S \to a;\ a \to N;\ N \to N\}, S)$ is good.

<u>Proof</u> $\mathcal{L}(F)$ consists of all nonempty finite languages consisting of single letter words. Let $F'$ be an arbitrary ETOL form such that $\mathcal{L}(F') \subseteq \mathcal{L}(F)$. Assume $L(F') = \Sigma = \{a_1, \ldots, a_n\}$ and let $\mathcal{D}_{F'}$ denote the family of languages $\mathcal{D}_{F'} = \{L(G) \mid G \lhd F'(\mu),\ \mu(a) = \{a\} \text{ for all } a \in \Sigma\}$ . Since the languages in $\mathcal{L}(F')$ consist of singletons $\mathcal{L}(F')$ can be characterized by:

$L \in \mathcal{L}(F')$

if and only if there exist $h \in \mathcal{D}_{F'}$ and finite substitution $\mu$ on $h$ such that

(1) $\mu(a) \neq \emptyset$ for all a in $h$,

(2) $\mu(a) \cap \mu(b) = \emptyset$ for all $a \neq b$ in $h$, and

(3) $L = \bigcup\limits_{a \in h} \mu(a)$.

Because of this characterization it suffices to show that there exists an interpretation $\overline{F}$ of F such that the corresponding $\mathcal{D}_{\overline{F}}$ equals $\mathcal{D}_{F'}$. Let $K = \bigcap\limits_{h \in \mathcal{D}_{F'}} h$. K denotes the set of symbols in $\Sigma$, which occur in all languages of $\mathcal{D}_{F'}$. Let R be the relation on $\Sigma$ defined as follows: $(a,b) \in R$ if and only if for all $h \in \mathcal{D}_{F'}$ and $a \in h$ imply $b \in h$. Define $\overline{R}(a)$, for $a \in \Sigma$, to be the smallest set Q such that $a \in Q$,and $(b,c) \in R$ and $b \in Q$ imply $c \in Q$. Let $\overline{R}(M)$, for $M \subseteq \Sigma$, denote $\bigcup\limits_{a \in M} \overline{R}(a)$.

$\mathcal{D}_{F'}$ can then be characterized by:

$h \in \mathcal{D}_{F'}$

if and only if

$K \subseteq h$ and $\overline{R}(h) = h$.

Without loss of generality we can assume that $K = \{a_1, \ldots, a_k\}$ for some $k \leq n$. Finally let $\overline{R}(a_i) = \{a_i^{(1)}, a_i^{(2)}, \ldots, a_i^{(k)}\}$ with $a_i = a_i^{(1)}$. Note that $\overline{R}(K) = K$.

Construct the ET0L form $\overline{F} = (V, \Sigma, P_1, P_2, S_1) \lhd F(\mu)$ as follows:

i) $V = \Sigma \cup \{S_1, S_2, \ldots, S_k\} \cup \{N\} \cup \bigcup_{k<i\leq n} \{S_i^{(1)}, S_i^{(2)}, \ldots, S_i^{(k_i)}\}$.

ii) $P_1$:

$S_i \to S_{i+1}$ for $1 \leq i < k$,

$S_k \to S_1 \mid S_{k+1}^{(1)} \mid S_{k+2}^{(1)} \mid \ldots \mid S_n^{(1)}$,

$S_i^{(j)} \to S_i^{(j+1)}$ for $k < i \leq n$ and $1 \leq j < k_i$,

$S_i^{(k_i)} \to S_1$ for $k < i \leq n$,

$a_i \to N$ for $1 \leq i \leq n$,

$N \to N$.

iii) $P_2$:

$S_i \to a_i$ for $1 \leq i \leq k$,

$S_i^{(j)} \to a_i^{(j)}$ for $k < i \leq n$, $1 \leq j \leq k_i$,

$a_i \to N$ for $1 \leq i \leq n$,

$N \to N$.

iv) for all $a \in \Sigma$, $\mu(a) = \{a\}$.

$\mu(S) = V - \Sigma$.

From the construction of $\overline{F}$ it follows that if we define $\mathcal{D}_{\overline{F}} = \{L(G) \mid G \lhd \overline{F}(\mu),$ $\mu(a) = \{a\}$ for $a \in \Sigma\}$ then $\mathcal{D}_{\overline{F}} = \mathcal{D}_{F'}$ and therefore $\mathcal{L}(\overline{F}) = \mathcal{L}(F')$. □

<u>Corollary 6</u> If F is an ET0L form generating a nonempty finite language consisting of single letter words only, then there exists an integer k such that $\mathcal{L}(F)$ equals the family of all finite languages of size at least k and consisting of single letter words only.

Although there exist good synchronized ETOL forms the following shows that no synchronized ETOL form is vomplete.

Lemma 7  Let $F = (\{S,a\},\{a\},\{S \to a;\ a \to aa\},S)$. Then no synchronized ETOL form $F'$ is form equivalent to $F$.

Proof  Assume that $F'$ is an arbitrary synchronized ETOL form and that $\overline{F} = (\overline{V},\{a\},\overline{P}_1,\ldots,\overline{P}_m,\overline{S}) \lhd F'$ with $L(\overline{F}) = L(F)$ and let $\overline{S} \underset{\overline{P}_{i_1}}{\Rightarrow} x_1 \underset{\overline{P}_{i_2}}{\Rightarrow} \cdots \underset{\overline{P}_{i_q}}{\Rightarrow} x_q = a^4$ be a derivation of $a^4$ in $\overline{F}$. Since $\overline{F}$ is synchronized we have $x_j \in (\overline{V} - \{a\})^+$ for $1 \leq i < q$. By renaming the symbols in $x_1,x_2,\ldots,x_q$ and adding the appropriate productions to the tables $\overline{P}_1,\ldots,\overline{P}_m$ we can obtain an $\overline{F}' \lhd \overline{F}$ such that $L(\overline{F}') = L(F) \cup \{abab\} = \{a^{2^n} \mid n \geq 0\} \cup \{abab\}$. This language does not belong to $\mathcal{L}(F)$. Consequently $\mathcal{L}(\overline{F}) \nsubseteq \mathcal{L}(F)$. □

Theorem 8  No synchronized ETOL form is vomplete.

Theorem 9  No synchronized and good ETOL form is complete for EOL.

Proof  F in Lemma 7 is an EOL form. □

Theorem 10 There exist complete ETOL forms which are bad.

Proof  In [2] it is shown that the synchronized ETOL form $F = (\{S,a,N\},\{a\},\{S \to a \mid S \mid SS;\ a \to N;\ N \to N\},\{S \to S;\ a \to N;\ N \to N\},S)$ is complete. It follows from Theorem 8 that F is bad. □

Lemma 11  Let $F=(\{S,a,b\},\{a,b\},\{S \to a;\ a \to abba;\ b \to e\},S)$. Then no propagating ETOL form $F'$ is form equivalent to F.

Proof Assume $\mathcal{L}(F') = \mathcal{L}(F)$ and $F'$ propagating. Let $\bar{F} = (V, \{a,b\}, P_1, \ldots, P_m, S) \lhd F'$ such that $L(\bar{F}) = L(F)$. Let $D : S \underset{P_{i_0}}{\Rightarrow} x_0 \underset{P_{i_1}}{\Rightarrow} \cdots \underset{P_{i_j}}{\Rightarrow} x_j = abba$ be a derivation of abba in $\bar{F}$ such that $x_i \neq abba$ for $i < j$. If $x_i \neq a$ for $0 \leq i < j$ then by renaming all symbols in $x_0, x_1, \ldots, x_j$ and adding appropriate productions to $P_1, \ldots, P_m$ for the new symbols, we can obtain $\bar{F}' \lhd \bar{F} \lhd F'$ such that $L(\bar{F}') = L(\bar{F}) \cup \{abcd\} = \{a^{2^n} \mid n \geq 0\} \cup \{abcd\}$. If $x_k = a$ for some $0 \leq k < j$ then we can assume that $x_i \neq a$ for $i \neq k$, $1 \leq i < j$. By renaming $x_0, x_1, \ldots, x_j$ and adding productions to $P_1, \ldots, P_m$ for new symbols we can then obtain an $\bar{F}' \lhd \bar{F} \lhd F'$ such that $L(\bar{F}') = L(\bar{F}) \cup \{d, abcd\} \cup L$, where $L \subseteq \{a,b,c,d\}^*$ consists of the words we might be forced to produce from the word abcd. Since $\bar{F}$ is propagating abbaabba cannot be derived from abba in $\bar{F}$. Therefore $L$ contains no words of length 8. If an interpretation of $F$ generates a language $L'$ such that $a \in L'$ (and d) is (are) the only word(s) of length one and abcd $\in L'$ then abcdw $\in L'$ (or wabcd $\in L'$) for some w of length four. Consequently $L(\bar{F}') \notin \mathcal{L}(F)$ and $\mathcal{L}(F') \nsubseteq \mathcal{L}(F)$. □

Theorem 12 No propagating ETOL form is vomplete.

Theorem 13 No propagating and good ETOL form is complete for EOL.

Proof F in Lemma 10 is an EOL form. □

To prove completeness in the EOL case, we have to show that for an arbitrary EOL form we can reduce the length of the right hand sides of the productions below a certain limit without changing the family of languages generated. For ETOL forms we have to be able to reduce the number of tables as well. The next theorem shows that this is indeed possible. Similar theorems are proven in [2] for synchronized ETOL forms.

Theorem 14 Given an ET0L form $F$, a form equivalent 2-ET0L form $F'$ can be constructed.

Proof Let $F = (\{A_1,\ldots,A_n\}, \{A_1,\ldots,A_t\}, P_1,\ldots,P_m, A_n)$. We construct a form equivalent 2-ET0L form $F' = (V, \{A_1,\ldots,A_t\}, P_1', P_2', A_n)$ as follows:

i) $V = \{A_1,\ldots,A_n\} \cup \{A_i[j] \mid 1 \le i \le n,\ 1 \le j \le m\} \cup \{N\}$

ii) $P_1'$ consists of the productions:

$$\left.\begin{array}{l} A_i \to A_i[1], \\ A_i[j] \to A_i[j+1],\quad 1 \le j < m \\ A_i[m] \to N, \end{array}\right\} \qquad 1 \le i \le n$$

$$N \to N.$$

iii) $P_2'$ consists of the productions:

$$A_i \to N,\quad 1 \le i \le n$$

$$N \to N,$$

$$A_i[j] \to \alpha, \text{ where } A_i \to \alpha \text{ is a production in } P_i,$$

$$1 \le i \le n,\quad 1 \le j \le m.$$

From the construction it follows easily that $A \to \alpha \in P_j$ if and only if $A \in \{A_1,\ldots,A_n\}$, $A \underset{P_1'}{\Rightarrow} \alpha_1 \underset{P_1'}{\Rightarrow} \alpha_2 \underset{P_1'}{\Rightarrow} \cdots \underset{P_1'}{\Rightarrow} \alpha_j \underset{P_2'}{\Rightarrow} \alpha$, where $\alpha_i \notin \{A_1,\ldots,A_n\}^*$, $1 \le i \le j$.

Therefore $L(F) = L(F')$.

Now let $\overline{F} \lhd F(\mu)$ be an arbitrary interpretation. We will prove that there exists an $\overline{F}' \lhd F'(\mu')$ such that $L(\overline{F}) = L(\overline{F}')$ and therefore $\mathcal{L}(F) \subseteq \mathcal{L}(F')$. Let $\overline{F} = (\{B_1, \ldots, B_p\}, \{B_1, \ldots, B_q\}, \overline{P}_1, \ldots, \overline{P}_m, B_p)$. We construct $\overline{F}' = (\overline{V}, \{B_1, \ldots, B_q\}, \overline{P}'_1, \overline{P}'_2, B_p) \lhd F'(\mu')$ such that $L(\overline{F}') = L(\overline{F})$ as follows:

i) $\overline{V} = \{B_1, \ldots, B_p\} \cup \{B_i[j] \mid 1 \leq i \leq p,\ 1 \leq j \leq m\} \cup \{N\}$

ii) $\overline{P}'_1$ consists of the productions:

$$\left.\begin{array}{l} B_i \to B_i[1], \\ B_i[j] \to B_i[j+1], \quad 1 \leq j < m \\ B_i[m] \to N, \end{array}\right\} \quad 1 \leq i \leq p$$

$$N \to N.$$

iii) $\overline{P}'_2$ consists of the productions:

$$B_i \to N$$

$$N \to N$$

$$B_i[j] \to \beta, \text{ where } B_i \to \beta \text{ is a production in } \overline{P}_j,$$
$$1 \leq i \leq p, \quad 1 \leq j \leq m.$$

iv) $$\left.\begin{array}{l} \mu'(A_k) = \mu(A_k) \\ \mu'(A_k[j]) = \{B_i[j] \mid B_i \in \mu(A_k)\} \end{array}\right\} \quad 1 \leq k \leq n$$

$$\mu'(N) = \{N\}.$$

As above it follows easily that $L(\overline{F}) = L(\overline{F}')$. That $\overline{F}' \lhd F'(\mu')$ is as clear.

Now let $G' \lhd F'(\eta')$ be an arbitrary interpretation. We will prove that there exists a $G \lhd F(\eta)$ such that $L(G') = L(G)$ and therefore $\mathcal{L}(F') \subseteq \mathcal{L}(F)$ which will complete the proof of the theorem.

Let $G' = (W', \Sigma, T'_1, T'_2, S)$. We construct $G = (W, \Sigma, T_1, T_2, \ldots, T_m, S)$ as follows:

$\iota$) $\quad W = \bigcup_{1 \leq i \leq n} \eta'(A_i)$

$\iota\iota$) $\quad \eta = \eta' \big| \{A_1, \ldots, A_n\}$

$\iota\iota\iota$) $\quad c \to \gamma$ is a production in $T_j$, $1 \leq j \leq m$ if and only if $c \in W$, $\gamma \in W^*$ and $c \underset{T'_1}{\Rightarrow} \gamma_1 \underset{T'_1}{\Rightarrow} \gamma_2 \underset{T'_1}{\Rightarrow} \cdots \underset{T'_1}{\Rightarrow} \gamma_j \underset{T'_2}{\Rightarrow} \gamma$.

Note that $W' = W \cup \bigcup_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \eta'(A_i[j])$ and $\Sigma \subseteq W$.

Since $G' \lhd F'(\eta')$ we get from Lemma 1 and $\iota\iota\iota$) above that

$(*)\ \eta'^{-1}(c) \underset{P'_1}{\Rightarrow} \eta'^{-1}(\gamma_1) \underset{P'_1}{\Rightarrow} \cdots \underset{P'_1}{\Rightarrow} \eta'^{-1}(\gamma_j) \underset{P'_2}{\Rightarrow} \eta'^{-1}(\gamma)$ if $c \to \gamma$ in $T_j$.

Then $\gamma_i \notin W^*$, $1 \leq i \leq j$. Therefore $L(G') = L(G)$. To prove that $G \lhd F(\eta)$ we have to check points i) through v) in the definition page 5. i), ii), iii), and v) follow from $\iota\iota$). To prove iv) assume that $c \to \gamma$ is in $T_j$, $A_i = \eta^{-1}(c)$ for some $1 \leq i \leq n$, and $\delta = \eta^{-1}(\gamma)$. From $(*)$ above we get

$$A_i \underset{P'_1}{\Rightarrow} A_i[1] \underset{P'_1}{\Rightarrow} A_i[2] \underset{P'_1}{\Rightarrow} \cdots \underset{P'_1}{\Rightarrow} A_i[j] \underset{P'_2}{\Rightarrow} \delta$$

which implies that $A_i \to \delta$ is a production in $P_j$ (in $F$). $\square$

<u>Theorem 15</u> Given an ETOL form $F$ a form equivalent short ETOL form $F'$ can be constructed.

<u>Proof</u> The proof is very similar to the one in [1] for EOL forms. Let $F = (V, \Sigma, P_1, \ldots, P_m, S)$. If $\max r(F) \leq 2$ then $F$ is already short. If

$\mathrm{maxr}(F) > 2$ then it suffices to show that we can construct a form equivalent ET0L form $\overline{F} = (\overline{V}, \Sigma, \overline{P}_1, \ldots, \overline{P}_m, S)$ such that for some i $\mathrm{maxr}(\overline{P}_i) + 1 = \mathrm{maxr}(P_i) = \mathrm{maxr}(F)$ and for $j \neq i$ $\mathrm{maxr}(\overline{P}_j) = \max\{\mathrm{maxr}(P_j), 1\}$.

Now let $\mathrm{maxr}(P_i) = \mathrm{maxr}(F) > 2$. We construct $\overline{F}$ as follows:

*i)* $\overline{V} = V \cup \{N\} \cup \{B^{(p)}, C^{(p)} \mid p \in P_i\}$

*ii)* $\overline{P}_i$ consists of the productions:

$$\left.\begin{array}{l} A \to B^{(p)} \\ B^{(p)} \to \alpha \\ C^{(p)} \to N \end{array}\right\} \text{ if } |\alpha| \leq 2 \text{ and } p: A \to \alpha \text{ is a production in } P_i$$

$$\left.\begin{array}{l} A \to B^{(p)}C^{(p)} \\ B^{(p)} \to A_1 \ldots A_{k-1} \\ C^{(p)} \to A_k \end{array}\right\} \text{ if } \alpha = A_1 \ldots A_k \text{ for some } k > 2 \text{ and } p: A \to \alpha \text{ is a production in } P_i$$

$$N \to N$$

*iii)* $\overline{P}_j = P_j \cup \{B^{(p)} \to N \mid p \in P_i\} \cup \{C^{(p)} \to N \mid p \in P_i\} \cup \{N \to N\}$ for $j \neq i$, $1 \leq j \leq m$.

By using Lemma 3 with $k_i = 2$ and $k_j = 1$ for $j \neq i$ we get that $\mathcal{L}(F) = \mathcal{L}(\overline{F})$. That $\mathrm{maxr}(\overline{P}_i) = \mathrm{maxr}(P_i) - 1$ and $\mathrm{maxr}(\overline{P}_j) = \max\{\mathrm{maxr}(P_j), 1\}$ for $j \neq i$ is clear.

□

Theorem 16 Given an ET0L form F a form equivalent short 2-ET0L form F' can be constructed.

Proof Immediate from the proofs of Theorems 14 and 15. □

Theorem 17 The binary 2-ET0L form $F = (\{a, S, \{a\}, \{a \to S;\ S \to S\},$
$\{a \to S; S \to e|a|S|SS\}, S\})$ is vomplete.

Proof Completeness follows from Theorem 5.5 in [2].
Given an arbitrary ET0L form $F'$ we can construct a form equivalent ET0L form $F'_1$, which is reduced and separated using Lemma 4.1 and 4.2 in [2]. Then using the constructions occurring in the proofs of Theorems 14 and 15 we obtain a form equivalent 2-ET0L form $F'_2$ which is reduced and binary. $F'_2$ must then be an interpretation of $F$, so $F$ is therefore a good ET0L form. □

ACKNOWLEDGEMENT

I would like to thank The Baron of Rauhala for pointing out this subject for me.

## REFERENCES

[1] Maurer, H.A., Salomaa, A., Wood, D.: E0L forms. Acta Informatica 8, (1977), 75-96.

[2] Maurer, H.A., Salomaa, A., Wood, D.: ET0L forms. Journal of Computer and Systems Sciences (1978). To appear.

[3] Maurer, H.A., Salomaa, A., Wood, D.: On good E0L forms. SIAM Journal of Computing (1978). To appear.