ON GOOD ET0L FORMS

by

Sven Skyum

DAIMI PB-74
April 1977
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.
Maurer, Salomaa, and Wood introduced in [1] and [2] the notion of EOL and ETOL forms and their interpretations. An ETOL form $F$ defines a family of ETOL systems $\mathcal{Q}(F)$ which are "structurally" similar and a family of languages $\mathcal{L}(F)$, namely those languages generated by systems in $\mathcal{Q}(F)$. In [3] Maurer, Salomaa and Wood study good and complete EOL forms. Many of the theorems in [3] are trivially valid for ETOL forms as well as for EOL forms but they will not appear in this paper. In this paper we show that in contrast to what is the case for EOL forms, there exist good synchronized ETOL forms. Finally we show that, (as for EOL forms) there exist binary incomplete (short for very complete) ETOL forms and that no propagating or synchronized ETOL form is complete.
DEFINITIONS AND BASIC RESULTS

We follow the definitions of ET0L forms and interpretations given in [1], [2], and [3]. E0L forms will not be defined explicitly. We can consider an E0L form as an ET0L form with one table.

Definition

An ET0L system (or n-ET0L system) 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, \( S \notin V - \Sigma \) is the start symbol. 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 rules or productions and are usually written \( A \rightarrow \alpha \) or just \( A \rightarrow \alpha \). The sets \( P_i \) are called tables.

Definition Let \( G = (V, \Sigma, P_1, \ldots, P_n, S) \) be an n-ET0L system. For words \( x = A_1 A_2 \ldots A_m \) 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 \( x \Rightarrow y \) 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 ET0L systems, the start symbol cannot be a terminal.

Definition Let \( G = (V, \Sigma, P_1, \ldots, P_n, S) \) be an n-ET0L system. For a word \( x \), \( |x| \) denotes the length of \( x \) and \( \text{Alph}(x) \) denotes the minimal alphabet such that \( x \in \text{Alph}(x)^* \). For all \( 1 \leq i \leq n \) let \( \text{maxr}(P_i) = \max\{|\alpha| \mid A \rightarrow \alpha \text{ in } P_i\} \), and let \( \text{maxr}(G) = \max\{\text{maxr}(P_i) \mid 1 \leq i \leq n\} \). A symbol \( B \in V \) is reachable
(from S) if $S \Rightarrow^* \alpha \beta \beta$ for some words $\alpha, \beta \in V^*$. $G$ is reduced if each $B \in V$ is reachable. $G$ is separated if for all productions $A \rightarrow \alpha$ in $P_1, \ldots, P_n$ $\alpha \in (V \Sigma)^*$ if $A \in \Sigma$ and $\alpha \in \Sigma \cup (V \Sigma)^*$ otherwise. $G$ is propagating if for all productions $A \rightarrow \alpha$ in $P_1, \ldots, P_n$ $\alpha \neq e$, the empty word. $G$ is synchronized if, for all $a \in \Sigma$, $a \Rightarrow^+ \alpha$ implies $\alpha \notin \Sigma^*$. $G$ is short, if for all $P_i$, $A \rightarrow \alpha \in P_i$ implies $|\alpha| \leq 2$. Finally $G$ is binary if each rule in $P_1, \ldots, P_m$ is one of the forms $A \rightarrow e$, $A \rightarrow a$, $A \rightarrow B$, $A \rightarrow BC$, $a \rightarrow A$, where $a \in \Sigma$ and $A, B, C \in V \Sigma$.

Definition. An ET0L form (or n-ET0L form) is an ET0L system $F = (V, \Sigma, P_1, \ldots, P_n, S)$. An ET0L system $G = (V', \Sigma', P'_1, \ldots, P'_n, S')$ is an interpretation of $F$ (modulo $\mu$), $G \triangleleft F(\mu)$, or simply $G \triangleleft F$, if $\mu$ is a substitution defined on $V$ satisfying (i)-(v):

1. $\mu(A) \subseteq V' \Sigma'$ for $A \in V \Sigma$,
2. $\mu(a) \subseteq \Sigma'$ for $a \in \Sigma$,
3. $\mu(\alpha) \cap \mu(\beta) = \emptyset$ for any symbols $\alpha \neq \beta$,
4. for all $1 \leq i \leq n$, $P'_i \subseteq \mu(P_i)$, where $\mu(P_i) = \{ (A', \alpha') \in V' \times V' \ | \ A' \in \mu(A), \alpha' \in \mu(\alpha) \text{ for some } A \in V, \alpha \in V^* \text{ such that } A \rightarrow \alpha \in P_i \}$,
5. $S' \in \mu(S)$.

The family of ET0L systems generated by $F$, denoted $\mathcal{G}(F)$, is:

$\mathcal{G}(F) = \{ G \ | \ G \triangleleft F \}$.

The family of languages generated by $F$, denoted $\mathcal{L}(F)$, is:

$\mathcal{L}(F) = \{ L(G) \ | \ G \in \mathcal{G}(F) \}$.

Since an ET0L form is an ET0L system, and conversely, we will allow ourselves to use the term "form" in the rest of this paper.
Definition. Two ET0L forms $F_1$ and $F_2$ are equivalent if $L(F_1) = L(F_2)$ and form equivalent if $L(F_1) = \mathcal{E}(F_2)$.

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 ET0L form and $F' = (V', \Sigma', P'_1, \ldots, P'_m, S') \triangleleft F(\mu)$. Then for each derivation $x_0 \Rightarrow x_1 \Rightarrow \ldots \Rightarrow x_k$ in $F'$, $\mu^{-1}(x_0) \Rightarrow \mu^{-1}(x_1) \Rightarrow \ldots \Rightarrow \mu^{-1}(x_k)$ is a derivation in $F$.

Lemma 2. For all ET0L forms $F$ a form equivalent reduced ET0L 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 $L(F) = \mathcal{E}(\overline{F})$.

(a) $A \in V$, $A \Rightarrow x_1 \Rightarrow x_2 \Rightarrow \ldots \Rightarrow x_t$, $x_i \in V^*$ and $x_j \notin V^*$ for $1 \leq j < t$.

(b) $t = k_i$, $i_1 = i_2 = \ldots = i_t = i$, $A \Rightarrow x_t$ in $P_i$, $x_j \in \overline{V}^+$ for $1 \leq j < t$, $A \Rightarrow x_1$ in $\overline{P}_i$, and $x_{j-1} \Rightarrow x_j$ for $1 < j \leq t$. 
Definition Let $F$ be an ET0L form and $\mathcal{F}$ a family of languages. We call $F$ \textit{$\mathcal{F}$-complete} or \textit{complete} for $\mathcal{F}$ if $\mathcal{L}(F) = \mathcal{F}$; if $\mathcal{F}$ is the family of ET0L languages, then we simply call $F$ \textit{complete} instead of ET0L-complete. We call $F$ \textit{good}, if for each ET0L form $F'$ with $\mathcal{L}(F') \subseteq \mathcal{L}(F)$ an ET0L form $F'$ exists such that $F' \triangleleft F$ and $\mathcal{L}(F') = \mathcal{L}(F)$. $F$ is called \textit{bad} if it is not good. $F$ is called \textit{vomplete} (short for very complete) if it is complete and good.
RESULTS

Most of the theorems on good E0L forms in [3] are easily shown to be valid for ET0L forms as well. Properties which are different for E0L and ET0L forms are related to synchronization. There are two "canonical" ways to synchronize an ET0L 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 \( a^1 \rightarrow a, a \rightarrow N, N \rightarrow N \) to all tables for all terminals a. N is a new nonterminal. The second, which has no counterpart in the E0L case, is to add a new table consisting of the productions \( a^1 \rightarrow a, a \rightarrow N, N \rightarrow N, A \rightarrow N \) for terminals a and nonterminals A. a \( \rightarrow N, N \rightarrow N \) is added to the rest of tables. The following lemma and theorem show different properties for synchronized E0L and ET0L forms.

Lemma 4  The synchronized ET0L form
\[
F = (\{S, a, N\}, \{a\}, \{S \rightarrow SS; a \rightarrow N; N \rightarrow N\}, \{S \rightarrow a; a \rightarrow N; N \rightarrow N\}, S)
\]
generates no nonempty finite languages.

Proof  Immediate.

All synchronized E0L forms generating nonempty languages generate finite nonempty languages. This is used to prove that no good synchronized E0L form exists (Th. 2.6 in [3]). The following theorem shows that good synchronized ET0L forms exist. Surprisingly enough the form shown to exist generates finite languages only!
Theorem 5  The synchronized ET0L form 
\[ F = (\{ S, a, N \}, \{ a \}, \{ S \rightarrow S; a \rightarrow N; N \rightarrow N \}, \{ S \rightarrow a; a \rightarrow N; N \rightarrow N \}, S) \] 
is good.

Proof  \( \mathcal{L}(F) \) consists of all nonempty finite languages consisting of single letter words. Let \( F' \) be an arbitrary ET0L form such that \( \mathcal{L}(F') \subseteq \mathcal{L}(F) \).

Assume \( \mathcal{L}(F') = \mathcal{L} = \{ a_1, \ldots, a_n \} \) and let \( \mathcal{L}_F \) denote the family of languages 
\[ \mathcal{L}_F = \{ L(G) \mid G \subseteq F'(\mu), \mu(a) = \{ a \} \text{ for all } a \in \Sigma \} \]. Since the languages in \( \mathcal{L}(F') \) consist of singletons \( \mathcal{L}(F') \) can be characterized by:

\[
L \in \mathcal{L}(F') \quad \text{if and only if there exist } \mu \in \mathcal{L}_F \text{ and finite substitution } \mu \text{ on } L \text{ such that}
\]

\[ (1) \quad \mu(a) \neq \emptyset \text{ for all } a \in \mu, \]
\[ (2) \quad \mu(a) \cap \mu(b) = \emptyset \text{ for all } a \neq b \text{ in } \mu, \] and
\[ (3) \quad L = \bigcup_{a \in \mu} \mu(a). \]

Because of this characterization it suffices to show that there exists an interpretation \( \overline{F} \) of \( F \) such that the corresponding \( \mathcal{L}_{\overline{F}} \) equals \( \mathcal{L}_F \).

Let \( K = \bigcap_{\mu \in \mathcal{L}_F} \mu \). \( K \) denotes the set of symbols in \( \Sigma \), which occur in all languages of \( \mathcal{L}_F \). Let \( R \) be the relation on \( \Sigma \) defined as follows: \( (a, b) \in R \) if and only if for all \( \mu \in \mathcal{L}_F \) and \( a \in \mu \) imply \( b \in \mu \). 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 \( \bigcup_{a \in M} \overline{R}(a) \).

\( \mathcal{L}_F \) can then be characterized by:

\[ \mu \in \mathcal{L}_F \quad \text{if and only if} \]
\[ K \subseteq \mu \text{ and } \overline{R}(\mu) = \mu. \]
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 \( \overline{F} = (V, \Sigma, P_1, P_2, S_1) \cup F(\mu) \) 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) \( P_1 : \)

\( S_i \rightarrow S_{i+1} \) for \( 1 \leq i < k \),

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

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

\( S_i^{(k_i)} \rightarrow S_1 \) for \( k \leq i \leq n \),

\( a_i \rightarrow N \) for \( 1 \leq i \leq n \),

\( N \rightarrow N \).

iii) \( P_2 : \)

\( S_1 \rightarrow a_i \) for \( 1 \leq i \leq k \),

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

\( a_i \rightarrow N \) for \( 1 \leq i \leq n \),

\( N \rightarrow N \).

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

\( \mu(V) = \Sigma \).

From the construction of \( \overline{F} \) it follows that if we define \( \mathcal{E}_{\overline{F}} = \{ L(G) \mid G \cup F(\mu) \} \),

\( \mu(a) = \{ a \} \) for \( a \in \Sigma \) then \( \mathcal{E}_{\overline{F}} = \mathcal{E}_{\overline{F}^1} \) and therefore \( \mathcal{L}(\overline{F}) = \mathcal{L}(\overline{F}^1) \). □

Corollary 6 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 ET0L forms the following shows that no synchronized ET0L form is complete.

**Lemma 7**  Let $F = (\{S, a\}, \{a\}, \{S \to a; a \to aa\}, S)$. Then no synchronized ET0L form $F'$ is form equivalent to $F$.

**Proof**  Assume that $F'$ is an arbitrary synchronized ET0L form and that

$F = (\overline{V}, \{a\}, \overline{F}_1, \ldots, \overline{F}_m, \overline{S}) \triangleleft F'$ with $L(F) = L(F)$ and let $\overline{S} \triangleright x_1 \triangleright \ldots \triangleright x_q = a^4$ be a derivation of $a^4$ in $\overline{F}$. Since $\overline{F}$ is synchronized we have $x_i \in (\overline{V} - \{a\})^+$ 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 $\overline{F}_1, \ldots, \overline{F}_m$ we can obtain an $\overline{F}' \triangleleft \overline{F}$ such that $L(\overline{F}') = L(F) \cup \{abab\} = \{a^{2n} | n \geq 0\} \cup \{abab\}$. This language does not belong to $L(F)$. Consequently $L(\overline{F}) \not\subseteq L(F)$.

**Theorem 8**  No synchronized ET0L form is complete.

**Theorem 9**  No synchronized and good ET0L form is complete for E0L.

**Proof**  $F$ in Lemma 7 is an E0L form.

**Theorem 10**  There exist complete ET0L forms which are bad.

**Proof**  In [2] it is shown that the synchronized ET0L 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 ET0L form $F'$ is form equivalent to $F$. 
Proof Assume₤(F') = ⦃F⦄ and F' propagating. Let ⦃F⦄ = (V, {a, b}, P_1, ..., P_m, S) ⊨ F' such that L(F) = L(F). Let D : S ⊨ x_0 ⊨ P_0 ⊨ P_1 ⊨ ... ⊨ P_j ⊨ x_j = abba be a derivation of abba in ⦃F⦄ such that x_i ≠ abba for i < j. If x_i ≠ a for 0 ≤ i < j then by renaming all symbols in x_0, x_1, ..., x_j and adding appropriate productions to P_1, ..., P_m for the new symbols, we can obtain ⦃F'⦄ ⊨ ⦃F⦄ ⊨ F' such that L(F') = L(F) ∪ {abcd} = {a^{2n} | n ≥ 0} ∪ {abcd}. If x_k = a for some 0 ≤ k < j then we can assume that x_i ≠ a for i ≠ k, 1 ≤ i < j. By renaming x_0, x_1, ..., x_j and adding productions to P_1, ..., P_m for new symbols we can then obtain an ⦃F'⦄ ⊨ ⦃F⦄ ⊨ F' such that L(F') = L(F) ∪ {abcd} ∪ L, where L ∈ {a, b, c, d} *. consists of the words we might be forced to produce from the word abcd. Since ⦃F⦄ is propagating abbaabba cannot be derived from abba in ⦃F⦄. Therefore L contains no words of length 8. If an interpretation of F generates a language L' such that a ∈ L' (and d) is (are) the only word(s) of length one and abcd ∈ L' then abcdw ∈ L' (or wabcd ∈ L') for some w of length four. Consequently L(F') ⊨ ⦃F⦄ and ⦃F'⦄ ⊨ ⦃F⦄.

**Theorem 12.** No propagating ETOl form is complete.

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

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

ii) $P_1'$ consists of the productions:

\[
\begin{align*}
A_i & \rightarrow A_i[1], \\
A_i[j] & \rightarrow A_i[j+1], \ 1 \leq j < m \\
A_i[m] & \rightarrow N,
\end{align*}
\]

$1 \leq i \leq n$

$N \rightarrow N.$

iii) $P_2'$ consists of the productions:

\[
\begin{align*}
A_i & \rightarrow N, \ 1 \leq i \leq n \\
N & \rightarrow N,
A_i[j] & \rightarrow \alpha, \text{ where } A_i \rightarrow \alpha \text{ is a production in } P_1,
\end{align*}
\]

$1 \leq i \leq n, 1 \leq j \leq m.$

From the construction it follows easily that $A \rightarrow \alpha \in P_j$ if and only if $A \in \{A_1, \ldots, A_n\}$, $A \overset{P_1}{\rightarrow} \alpha_1 \overset{P_1'}{\rightarrow} \alpha_2 \overset{P_1'}{\rightarrow} \cdots \overset{\alpha_j}{\rightarrow} \overset{P_1'}{\rightarrow} \alpha$, where $\alpha_i \notin \{A_1, \ldots, A_n\} \ast$, $1 \leq i \leq j$.

Therefore $L(F) = L(F').$
Now let $\mathcal{F} \triangleleft F(\mu)$ be an arbitrary interpretation. We will prove that there exists an $\mathcal{F}' \triangleleft F'(\mu')$ such that $L(\mathcal{F}) = L(\mathcal{F}')$ and therefore $\mathcal{L}(\mathcal{F}) \subseteq \mathcal{L}(\mathcal{F'})$.

Let $\mathcal{F} = \{ B_1, \ldots, B_p \}, \{ B_1, \ldots, B_q \}, \mathcal{F}_1, \ldots, \mathcal{F}_m, B_p \}$. We construct $\mathcal{F}' = (\{ B_1, \ldots, B_p \}, \mathcal{F}_1, \mathcal{F}_2, B_p) \triangleleft F'(\mu')$ such that $L(\mathcal{F}') = L(\mathcal{F})$ as follows:

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

ii) $\mathcal{F}_1$ consists of the productions:

\[
\begin{align*}
B_i & \to B_i[1], \\
B_i[j] & \to B_i[j+1], \ 1 \leq j < m \\
B_i[m] & \to N, \\
N & \to N.
\end{align*}
\]

iii) $\mathcal{F}_2$ consists of the productions:

\[
\begin{align*}
B_i & \to N \\
N & \to N \\
B_i[j] & \to \beta, \text{ where } B_i \to \beta \text{ is a production in } \mathcal{F}_j, \\
1 \leq i \leq p, 1 \leq j \leq m.
\end{align*}
\]

iv) $\mu'(A_k) = \mu(A_k)$

\[
\begin{align*}
\mu'(A_k[j]) = \{ B_i[j] \mid B_i \in \mu(A_k) \} \\
\mu'(N) = \{ N \}.
\end{align*}
\]

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

Now let $G' \triangleleft F'(\eta')$ be an arbitrary interpretation. We will prove that there exists a $G \triangleleft F(\eta)$ such that $L(G') = L(G)$ and therefore $\mathcal{L}(\mathcal{F'}) \subseteq \mathcal{L}(\mathcal{F})$ which will complete the proof of the theorem.
Let $G' = (W', \Sigma, T_1', T_2', \ldots, T_m', S)$. We construct $G = (W, \Sigma, T_1, T_2, \ldots, T_m, S)$ as follows:

\begin{align*}
1) & \quad W = \bigcup_{1 \leq i \leq n} \eta'(A_i) \\
2) & \quad \eta = \eta'\mid_{\{A_1, \ldots, A_n\}} \\
3) & \quad c \Rightarrow \gamma \text{ is a production in } T_j, \ 1 \leq j \leq m \text{ if and only if} \\
& \quad c \in W, \ \gamma \in W^* \text{ and } c \Rightarrow T_1' \Rightarrow T_2' \Rightarrow \cdots \Rightarrow T_j' \Rightarrow \gamma.
\end{align*}

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

Since $G' \triangleleft F'(\eta')$ we get from Lemma 1 and 3 above that

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

Then $\gamma_1 \in W^*, \ 1 \leq i \leq j$. Therefore $L(G') = L(G)$. To prove that $G \triangleleft F(\eta)$ we have to check points i) through v) in the definition page 5. i), ii), iii), and v) follow from 3). To prove iv) assume that $c \Rightarrow \gamma$ 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

\[A_i \Rightarrow A_i[1] \Rightarrow A_i[2] \Rightarrow \cdots \Rightarrow A_i[j] \Rightarrow \delta\]

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

\[\square\]

**Theorem 15** Given an ETOL form $F$ a form equivalent short ETOL form $F'$ can be constructed.

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

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

\[\nu = \nu \cup \{N\} \cup \{B(p), C(p) \mid p \in P_i\}\]

\(\nu\) \(\vec{P}_i\) consists of the productions:

\[
\begin{align*}
A & \to B(p) \\
B(p) & \to \alpha \\
C(p) & \to N \\
A & \to B(p)C(p) \\
B(p) & \to A_1 \ldots A_{k-1} \\
C(p) & \to A_k \\
N & \to N
\end{align*}
\]

if \(|\alpha| \leq 2\) and \(p: A \to \alpha\) is a production in \(P_i\)

if \(\alpha = A_1 \ldots A_k\) for some \(k > 2\) and \(p: A \to \alpha\) is a production in \(P_i\)

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 \(\xi(F) = \xi(\vec{F})\).

That maxr(\(\vec{P}_i\)) = maxr(\(P_i\)) - 1 and maxr(\(\vec{P}_j\)) = max\{maxr(\(P_j\)), 1\} for \(j \neq i\) is clear.

**Theorem 16** Given an ETA form \(F\) a form equivalent short ETA form \(F'\) can be constructed.

**Proof** Immediate from the proofs of Theorems 14 and 15.
**Theorem 17**  The binary 2-ET0L form $F = ([a, S, S], \{a \to S; S \to S\}$, 
$[a \to S; S \to e|a|S|SS|, S])$ is complete.

**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

    Acta Informatica 8, (1977), 75-96.

    To appear.