<table>
<thead>
<tr>
<th>Title</th>
<th>Relations between language classes in terms of insertion and locality (New Trends in Algorithms and Theory of Computation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Fujioka, Kaoru</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2012), 1799: 57-59</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2012-06</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/173003">http://hdl.handle.net/2433/173003</a></td>
</tr>
<tr>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>
Relations between language classes in terms of insertion and locality

Kaoru Fujioka *

1 Introduction

Insertion systems use only insertion operations of the form $uveta$ for a given string $uv$ by inserting the string $x$ between $u$ and $v$. From the definition of insertion operations, using only insertion operations, we generate only context-sensitive languages.

Using insertion systems together with some morphisms, characterizing recursively enumerable languages is obtained in [8], [6]. Furthermore, similarly to the Chomsky–Schützenberger representation theorem [1], each recursively enumerable language can be expressed using an insertion system and a Dyck language in [7], and each context-free language can be expressed using an insertion system and a star language in [5].

In [2] and [3], within the framework of the Chomsky–Schützenberger representation theorem, some characterizations and representation theorems of languages in the Chomsky hierarchy have been provided by insertion system $\gamma$, strictly locally testable language $R$, and morphism $h$ such as $h(L(\gamma) \cap R)$.

The purpose of this paper is to clarify the relation between the classes of languages $h(L(\gamma) \cap R)$ using insertion systems of weight $(i,0)$ for $i \geq 1$ and those using insertion systems of weight $(i,1)$ for $i \geq 1$.

2 Preliminaries

For a string $x \in V^*$ with an alphabet $V$, $|x|$ is the length of $x$. For $0 \leq k \leq |x|$, let $\text{Pre}_k(x)$ and $\text{Suf}_k(x)$ respectively denote the prefix and the suffix of $x$ with length $k$. For $0 \leq k \leq |x|$, let $\text{Int}_k(x)$ be the set of intermediate substrings of $x$ with length $k$.

For a positive integer $k$, a language $L$ over $T$ is strictly $k$-testable if a triplet $S_k = (A, B, C)$ exists with $A, B, C \subseteq T^k$ such that, for any $w$ with $|w| \geq k$, $w$ is in $L$ iff $\text{Pre}_k(w) \in A$, $\text{Suf}_k(w) \in B$, $\text{Int}_k(w) \in C$. A language $L$ is strictly locally testable if there exists an integer $k \geq 1$ such that $L$ is strictly $k$-testable.

Note that, for an alphabet $T$, a language $T^+$ is a strictly 1-testable language.

Let $\text{LOC}(k)$ be the class of strictly $k$-testable languages. There is the following result.

**Theorem 1** [4] $\text{LOC}(1) \subset \text{LOC}(2) \subset \cdots \subset \text{LOC}(k) \subset \cdots \subset \text{REG}$.

We define an insertion system $\gamma = (T, P, A)$, where $T$ is an alphabet, $P$ is a finite set of insertion rules of the form $(u, x, v)$ with $u, x, v \in T^*$, and $A$ is a finite set of strings over $T$ called axioms.

We write $\alpha \xrightarrow{i}{\gamma} \beta$ if $\alpha = \alpha_1uv\alpha_2$ and $\beta = \alpha_1ux\alpha_2$ for some insertion rule $\gamma : (u, x, v) \in P$ with $\alpha_1, \alpha_2 \in T^*$. We write $\alpha \Longrightarrow \beta$ if no confusion exists. The reflexive and transitive closure of $\Longrightarrow$ is defined as $\Longrightarrow^*$.
A language generated by $\gamma$ is defined as

\[ L(\gamma) = \{w \in T^* \mid s \Rightarrow^*_\gamma w, \text{ for some } s \in A \}. \]

An insertion system $\gamma = (T, P, A)$ is said to be of weight $(m, n)$ if

\[ m = \max\{ |x| \mid (u, x, v) \in P \}, \quad n = \max\{ |u| \mid (u, x, v) \in P \} \text{ or } (v, x, u) \in P \].

For $m, n \geq 0$, let $INS^{m,n}_{\sigma}$ be the class of all languages generated by insertion systems of weight $(m', n')$ with $m' \leq m$ and $n' \leq n$. We use $\ast$ instead of $m$ or $n$ if the parameter is not bounded.

**Theorem 2** [8]

1. $INS^{i}_{0} \subseteq INS^{i'}_{0}$ (0 ≤ $i \leq i'$, 0 ≤ $j \leq j'$).
2. $INS^{i}_{1} \subseteq CF$.

A mapping $h : V^* \to T^*$ is called morphism if $h(\lambda) = \lambda$ and $h(xy) = h(x)h(y)$ hold for any $x, y \in V^*$. For any $a$ in $T$, if $h(a) = a$ holds, then $h$ is an identity morphism.

The following results related to Chomsky-Schützenberger like characterization are obtained using insertion systems of weight $(i, 0)$ or $(i, 1)$ for $i \geq 1$ and strictly $k$-testable languages ($k \geq 1$).

**Theorem 3** [2]

1. $H(INS^{i}_{0} \cap LOC(1)) \subseteq CF$.
2. $H(INS^{i}_{0} \cap LOC(k)) = REG$ (k ≥ 2).
3. $H(INS^{i}_{0} \cap LOC(1))$ and $REG$ are incomparable (i ≥ 2).
4. $H(INS^{i}_{1} \cap LOC(1)) \subseteq CF$ (i ≥ 2).
5. $H(INS^{i}_{1} \cap LOC(k)) = CF$ (i, k ≥ 2).

**Theorem 4** [3]

1. $H(INS^{i}_{1} \cap LOC(k)) = CF$ (i ≥ 1, k ≥ 2).
2. $H(INS^{i}_{1} \cap LOC(1)) \subseteq CF$ (i ≥ 1).

In the present paper, we specifically examine the relation between language classes $H(INS^{i}_{0} \cap LOC(k_0))$ and $H(INS^{i}_{1} \cap LOC(k_1))$ for $i_0, k_0, i_1, k_1 \geq 1$.

**3 Main Results**

For context-free languages, from Theorem 3 and Theorem 4, we obtain

\[ CF = H(INS^{i}_{0} \cap LOC(k_0)) = H(INS^{i}_{1} \cap LOC(k_1)) \]

with $i_0, k_0, k_1 \geq 2, i_1 \geq 1$.

We next examine the language class $H(INS^{2}_{2} \cap LOC(1))$. From Theorem 3, $H(INS^{2}_{0} \cap LOC(1))$ and $REG$ are known to be incomparable.

**Theorem 5** $H(INS^{2}_{2} \cap LOC(1))$ and $H(INS^{1}_{1} \cap LOC(1))$ are incomparable.

**Proof** Consider an insertion system $\gamma_1 = (T, \{(\lambda, ab, \lambda), \{\lambda\})$ of weight $(2, 0)$ with $T = \{a, b\}$, a strictly 1-testable language $R = T^+$, and an identity morphism $h : T^* \to T^*$. The above definition indicates directly that $L(\gamma) = h(L(\gamma) \cap R)$.

We can show that $L(\gamma_1)$ is not in $H(INS^{1}_{1} \cap LOC(1))$ by contradiction. We omit the proof here.

Now we consider an insertion system $\gamma_2 = (T, \{(a, a, a), (b, b, b), \{a, b\})$ of weight $(1, 1)$ with $T = \{a, b\}$, a strictly 1-testable language $R = T^+$, and an identity morphism $h : T^* \to T^*$. From the definition, we have $L(\gamma_2) = h(L(\gamma_2) \cap R) = \{a^i \mid i \geq 1 \} \cup \{b^i \mid i \geq 1 \}$.

From [2], $L(\gamma_2)$ is not in $H(INS^{2}_{2} \cap LOC(1))$.

**Corollary 1** $H(INS^{2}_{2} \cap LOC(1))$ and $H(INS^{1}_{1} \cap LOC(1)) \cap H(INS^{2}_{2} \cap LOC(2))$ are incomparable.
Corollary 2 \( H(IN_{S2}^2 \cap LOC(1)) \subseteq H(IN_{S1}^1 \cap LOC(1)) \) (\( i \geq 2 \)).

For the class of languages \( H(IN_{S1}^0 \cap LOC(1)) \), from the size of parameters, we have the inclusions \( H(IN_{S1}^0 \cap LOC(1)) \subseteq H(IN_{S1}^1 \cap LOC(1)) \) and \( H(IN_{S1}^0 \cap LOC(1)) \subseteq H(IN_{S1}^0 \cap LOC(2)) \). Next we present the following proper inclusion.

Theorem 6 \( H(IN_{S1}^0 \cap LOC(1)) \subseteq H(IN_{S1}^1 \cap LOC(1)) \cap H(IN_{S1}^0 \cap LOC(2)) \).

Proof To show the proper inclusion, we consider an insertion system \( \gamma_2 = \langle T, \{(a,a,\lambda), (b,b,\lambda)\}, \{a,b\} \rangle \) of weight \((1,1)\) with \( T = \{a,b\} \), a strictly 1-testable language \( R = T^+ \), and an identity morphism \( h : T^* \rightarrow T^* \).

In a similar way to Theorem 5, we can show that \( L(\gamma_2) \) is not in \( H(IN_{S1}^0 \cap LOC(1)) \). □

Corollary 3 \( H(IN_{S1}^0 \cap LOC(1)) \subseteq H(IN_{S1}^1 \cap LOC(1)) \cap H(IN_{S1}^0 \cap LOC(2)) \cap H(IN_{S2}^0 \cap LOC(1)) \).

4 Concluding Remarks

In the present paper, we specifically examined the language classes \( H(IN_{S1}^0 \cap LOC(k_0)) \) and \( H(IN_{S1}^1 \cap LOC(k_1)) \) for \( i_0, i_1, k_0, k_1 \geq 1 \) and considered the relations of those language classes.

The following remain as open problems:

- \( H(IN_{S2}^2 \cap LOC(1)) \cap H(IN_{S1}^1 \cap LOC(1)) \cap H(IN_{S1}^0 \cap LOC(1)) \) holds?
- \( H(IN_{S2}^2 \cap LOC(1)) \cap H(IN_{S1}^1 \cap LOC(1)) \cap H(IN_{S1}^0 \cap LOC(2)) \cap H(IN_{S1}^0 \cap LOC(1)) \cap H(IN_{S1}^1 \cap LOC(2)) \) holds?
- \( CF = H(IN_{S1}^2 \cap LOC(k)) \) holds for some \( m, k \geq 1 \)?

References


