<table>
<thead>
<tr>
<th>Title</th>
<th>Maximal avoidable sets of words (Algebraic Theory of Codes and Related Topics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Kobayashi, Yuji</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (1989), 697: 24-28</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1989-06</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/101442">http://hdl.handle.net/2433/101442</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
Maximal avoidable sets of words

Yuji Kobayashi, Tokushima University

We use the following notations.

$\Sigma$: an alphabet (a finite set of letters),
$\Sigma^*: \text{ the set of words over } \Sigma$,
$\Sigma^\omega: \text{ the set of infinite words (sequences)},$
$\Sigma^#: = \Sigma^* \cup \Sigma^\omega$,
$\Sigma^1 := \Sigma^* - \{1\}$, where $1$ is the empty word.

For $x = a_1a_2\ldots, y = b_1b_2\ldots \in \Sigma^#$ define the distance of $x$ and $y$ by

$$d(x, y) = \frac{1}{\min \{n! | a_n \neq b_n\}}.$$  

As is well known ([14]), $(\Sigma^#, d)$ is a compact totally disconnected metric space.

Let $x, y \in \Sigma^*$ and $X \subset \Sigma^#$. We say $y$ avoids $x$, if $y$ does not contain $x$ as subword, and $y$ avoids $X$, if $y$ avoids every $x$ in $X$. $X$ is called avoidable, if there is an infinite word $y$ avoiding $X$, otherwise $X$ is called unavoidable. Avoidability of sets of words called patterns were studied in [11].

Example 1. Let $X = \{v^2 \mid v \in \Sigma^1\}$. Then $y$ avoids $X$ if and only if $y$ is square-free. It is a famous fact that $X$ is avoidable if $|\Sigma| \geq 3$.
An avoidable set $X$ is **maximal** if any set properly containing $X$ is unavoidable.

**Theorem 1.** For any avoidable set $X$, there is a maximal avoidable set containing $X$.

For a given $X \subseteq \Sigma^*$, the set

$$\text{Min}(X) = \{ x \in X \mid \text{any } x' \in X \text{ is not a proper subword of } X \}$$

is called the base of $X$. Easily we see $y$ avoids $X$ if and only if $y$ avoids $\text{Min}(X)$. $X$ is **finitely based** if $\text{Min}(X)$ is a finite set. The base of a maximal avoidable set is called a **critical set** of words.

**Corollary.** For any avoidable set $X$ which is factor-free, that is, any word in $X$ is not a subword of another word in $X$, there is a critical set containing $X$.

**Example 2.** Let $\Sigma = \{a, b\}$. Then, $(a^2, ab, ba)$ and $(a^2, b^2)$ are critical sets.

An infinite word $x$ is **recurrent**, if for any subword $v$ of $x$, there is an integer $k(v) > 0$ such that any subword of length $k$ of $x$ contains $v$ as subword. In this situation $v$ is said to be recurrent in $x$. If $x = v^\omega$ for some $v \in \Sigma^*$, $x$ is called **periodic**; the shortest such $v$ is the **period** of $x$. A periodic infinite word is recurrent, but the converse is not true.

The **shift transformation** $\tau$ is a mapping from $\Sigma^\omega$ to itself defined by

$$\tau(x) = a_2a_3\cdots \quad \text{for } x = a_1a_2\cdots.$$
Obviously, τ is a surjective continuous mapping.

A subshift $S$ is a non-empty closed subset of $\Sigma^\omega$ invariant under τ. $S$ is minimal if it does not contain a subshift properly. For a given set $X \subseteq \Sigma^*$ of words, $S(X)$ is the set of infinite words avoiding $X$. For a given subshift $S \subseteq \Sigma^\omega$, $X(S)$ is the set of words which do not appear as subwords of elements of $S$.

Theorem 2. For an avoidable set $X$, $S(X)$ is a subshift. If $X$ is maximal, then $S(X)$ is minimal. Conversely, if $S$ is a subshift, then $X(S)$ is an avoidable set. If $S$ is minimal, then $X(S)$ is maximal. This gives a 1-1 correspondence between maximal avoidable sets and minimal subshifts.

Lemma 1. An avoidable set $X$ is maximal if and only if any word out of $X$ is recurrent in any infinite word avoiding $X$.

Theorem 3 (Morse-Hedlund [4]). $S \subseteq \Sigma^\omega$ is a minimal subshift if and only if

$$S = \{ \tau^n(x) \mid n = 0, 1, 2, \ldots \}$$

for some recurrent infinite word $x$. Moreover,

(1) $S$ is perfect, if $x$ is non-periodic. In this case every element in $S$ is non-periodic.

(2) $S$ is finite, if $x$ is periodic. In this case every element in $S$ is periodic.

Corollary (c.f. [3, Theorem 4.2]). Let $X$ be an avoidable set such that for any $v \in \Sigma^1$, $v^n$ does not avoid $X$ for $x \gg 0$, then $S(X)$ contains a perfect subset.
Theorem 4. Let $X$ be a maximal avoidable set. Then, $S(X)$ is finite if and only if $X$ is finitely based.

For an avoidable set $X$, the radical $\text{rad}(X)$ of $X$ is the intersection of all the maximal avoidable sets containing $X$. $X$ is called reduced, if $X = \text{rad}(X)$.

Lemma 2. A word $v$ is in $\text{rad}(X)$, if and only if any recurrent infinite word avoiding $X$ avoids $v$.

Corollary. Any word out of $\text{rad}(X)$ is extensible to a recurrent infinite word avoiding $X$.

Theorem 5. If $X$ is a reduced avoidable set, then every isolated point of $S(X)$ is periodic.

Corollary. If $X$ is a reduced avoidable set such that for any $v \in \Sigma^+$, $v^n$ does not avoid $X$ for $n \gg 0$, then $S(X)$ is perfect.

A set $X$ of words is quasi-maximal, if $\text{rad}(X)$ is maximal.

Theorem 6. Let $X$ be an avoidable set. Then following statements are equivalent.

(1) $X$ is quasi-maximal.

(2) $S(X)$ contains a unique minimal subshift.

(3) For any $n > 0$, there is a word $v$ of length $n$ such that $X \cup \{v\}$ is unavoidable.

(4) For any word $w$ such that $X \cup \{w\}$ is unavoidable and for any $n > 0$, there is a word $v$ of length $n$ such that $X \cup \{vw\}$ is unavoidable.

Example 3. Let $X = \{a^2, bab\} \subseteq \{a, b\}^\infty$. Then, $b^\omega$ and $ab^\omega$ are only infinite words avoiding $X$, and $X \cup \{b^n\}$ is unavoidable for any $n > 0$. 
Thus \( X \) is quasi-maximal.

An unavoidable set \( X \) is said to be \( \text{minimal} \), if \( X - \{v\} \) is avoidable for any \( v \in X \). As is easily seen ([2]), a minimal unavoidable set is finite.

Conjecture I (Ehrenfeucht, see [2]). For any unavoidable set \( X \), there is a word \( x \in X \) and a letter \( a \in \Sigma \) such that \( (X - \{x\}) \cup \{xa\} \) is unavoidable.

Conjecture II. For any minimal unavoidable set \( X \), there is a word \( x \) in \( X \) such that \( X - \{x\} \) is a quasi-maximal avoidable set.

Theorem 7. Conjecture I and Conjecture II are equivalent.

References