数理解析研究所講究録
第 697 巻 1989 年 24-28

Maximal avoidable sets of words

Yuji Kobayashi, Tokushima University

小林 斐治（徳島大・総合科学）

We use the following notations.

Σ : an alphabet (a finite set of letters),

Σ^* : the set of words over Σ,

Σ^ω : the set of infinite words (sequences),

$\Sigma^\# := \Sigma^* \cup \Sigma^\omega$,

$\Sigma^+ := \Sigma^* - \{1\}$, where 1 is the empty word.

For $x = a_1 a_2 \cdots$, $y = b_1 b_2 \cdots \in \Sigma^\#$ define the distance of x and y by

$d(x, y) = \dfrac{1}{\min\{n \mid a_n \neq b_n\}}$.

As is well-known ([4]), $(\Sigma^\#, d)$ 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 [1].

Example 1. Let $X = \{v^2 \mid v \in \Sigma^+\}$. Then y avoids X if and only if y is square-free. It is a famous fact that X is avoidable if $|\Sigma| \geq 3$

([5]).

An avoidable set X is <u>maximal</u>, 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 $X \subset \Sigma^*$, the set

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

is called the <u>base</u> of X. Easily we see y avoids X if and only if y avoids Min(X). X is <u>finitely based</u> if Min(X) is a finite set. The base of a maximal avoidable set is called a <u>critical set</u> 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 <u>recurrent</u>, 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 <u>periodic</u>; the shortest such v is the <u>period</u> of x. A periodic infinite word is recurrent, but the converse is not true.

The <u>shift transformation</u> τ is a mapping from Σ^ω 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 Σ^{ω} invariant under τ. S is *minimal*, if it does not contain a subshift properly. For a given set $X \subset \Sigma^{*}$ of words, S(X) is the set of infinite words avoiding X. For a given subshift $S \subset \Sigma^{\omega}$, 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]). $S \subset \Sigma^{\omega}$ is a minimal subshift if and only if

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

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 $v \in \Sigma^{+}$, 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 <u>radical</u> rad(X) of X is the intersection of all the maximal avoidable sets containing X. X is called <u>reduced</u>, if X = rad(X).

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

Corollary. Any word out of 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 $n >> 0$, then S(X) is perfect.

A set X of words is <u>quasi-maximal</u>, if 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 $X \cup \{wv\}$ is unavoidable.

Example 3. Let $X = \{a^2, bab\} \subset \{a, b\}^*$. Then, b^ω and ab^ω 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 <u>minimal</u>, 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

1. D.R.Bean, A.Ehrenfeucht and G.F.McNulty, Avoidable patterns in strings of symbols, Pacific J. Math. 85 (1979), 261-294.
2. C.Choffrut and K.Culik II, On extendibility of unavoidable sets, Discrete Applied Math. 9 (1984), 125-137.
3. Y.Kobayashi, Repetition free words, Theoret. Comp. Sci. 44 (1986), 175-197.
4. M.Morse and G.A.Hedlund, Symbolic dynamics I, Amer. J. Math. 60 (1938), 815-866.
5. A.Thue, Uber unendliche Zeichenreihe, Norske Vid. Selsk. Skr. I, Math. Nat. Kl. Christiana VII (1906), 1-22.