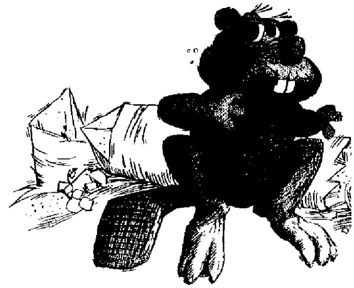
UNIVERSITY OF WATERLOO COMPUTER SCIENCE DEPARTMENT



On Extendibility of Unavoidable Sets

Christian Choffrut
Karel Culik II

CS-83-26

September, 1983

ON EXTENDIBILITY OF UNAVOIDABLE SETS *

Christian Choffrut

Laboratoire d'Informatique Théorique et Programmation
Université Paris VII
Paris, France

Karel Culik II

Department of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

ABSTRACT

A subset X of a free monoid A^* is said to be unavoidable if all but finitely many words in A^* contain some word of X as a subword. A. Ehrenfeucht has conjectured that every unavoidable set X is extendible in the sense that there exist $x \in X$ and $a \in A$ such that $(X-\{x\}) \cup \{xa\}$ is itself unavoidable. This problem remains open, we give some partial solutions and show how to efficiently test unavoidability, extendibility and other properties of X related to the problem.

1. INTRODUCTION

A subset X of a free monoid A^* is said to be *unavoidable*, if all but finitely many words in A^* contain some word of X as a subword. Ehrenfeucht has conjectured that provided A is finite, every unavoidable set X is *extendible* in the sense that there exist $x \in X$ and $a \in A$ such that $(X\backslash\{x\}) \cup \{xa\}$ is itself unavoidable.

The purpose of this paper is, after having introduced the main notions of unavoidability and extendibility in Section 2 and having shown that we can restrict ourselves to finite sets, to present the following results.

In Section 3 we consider the computational aspect of the problem. Indeed, we associate with every finite subset X a finite deterministic

---

* This research was supported by the Natural Sciences and Engineering Research Council of Canada, under the grant No. A-7403.

automaton and we show how to use it to deduce the properties of X related to the problem.

In Section 4 we give a partial solution to the conjecture in two special cases. As a consequence of the second case, we show that Ehrenfeucht's conjecture is equivalent to the statement where the word "extendible" need not necessarily mean extendible to the right as is implied by the above definition, but rather extendible either to the right or to the left.

Section 5 presents a reduction result which shows, via an encoding, that the conjecture need only be proved for highly restricted finite subsets of $\{a,b\}^*$.

2. PRELIMINARIES

Throughout this paper, A is a fixed finite alphabet containing at least two symbols. We denote by A^* the free monoid it generates and by 1 the unit or empty word. As usual we denote by $|u|$ the length of the word $u \in A^*$ and by A^+ the set of words of nonzero length: $A^+ = A^* \setminus \{1\} = \{u \in A^* \mid |u| > 0\}$.

The partial ordering on A^* "prefix of" is denoted by $\leq$: $u \leq v$ iff there exists $w \in A^*$ such that $v = uw$. We write $u < v$ if $u \leq v$ and $u \neq v$.

Assume we have $v = w_1 u w_2$ for some words u, v, w_1, w_2. Then u is a *subword* of v. If w_1 (resp. w_2) is the empty word, then u is a *prefix* (resp. a *suffix*) of v.

For any set S, we denote by $|S|$ its cardinality.

In the sequel, $X \subseteq A^+$ is a fixed set.

2.1. Unavoidability, Extendibility

We are interested in the set of all words in A^* which have no element of X as a subword. When this set is finite, X is unavoidable.

More formally, a word $w \in A^*$ *avoids* X if no subword of w belongs to X: $w \notin A^*XA^*$. Furthermore X is *avoidable* if there exist infinitely many words in A^* avoiding it. When X is not avoidable, it is *unavoidable* which amounts to saying that there exists an integer $n > 0$ such that:

$$(2.1) \qquad A^n A^* \subseteq A^* X A^* .$$

Assume X is unavoidable. An element $x \in X$ is *extendible by the letter* $a \in A$ (or simply *extendible*) if $Y = (X \setminus \{x\}) \cup \{xa\}$ is itself unavoidable. In this case, Y is *an extension* of X. Furthermore, X is *extendible* if it possesses some extendible element.

Example 2.1 With $A = \{a,b\}$, $X = \{aaa, ab, bbb\}$ is extendible since:

$$A^*\backslash A^*XA^* = (A^*\backslash A^*abA^*)\backslash A^*\{a^3,b^3\}A^* = (b^*a^*)\backslash A^*\{a^3,b^3\}A^* = \{a^ib^j \mid 0 \le i,j \le 2\}.$$

The word ab is not extendible. Indeed, $Y_a = \{aaa, aba, bbb\}$ is avoidable because of $(bba)^* \cap Y_a = \varnothing$ and so is $Y_b = \{aaa, abb, bbb\}$ because of $(ab)^* \cap Y_b = \varnothing$. However X is extendible in different ways. For example, $\{aaaa, ab, bbb\}$ is unavoidable.

We recall that the problem we are dealing with is the following:

Conjecture I *Every unavoidable set is extendible.*

We shall now show that the conjecture need only be proved for finite unavoidable sets.

Assume $X \subseteq A^+$ is unavoidable. Then it is *minimal* if no proper subset $Y \subsetneq X$ is unavoidable.

Example 2.2 With $A = \{a,b\}$, $X = A^2$ is unavoidable. However, it is not minimal since $Y = \{a^2, ab, b^2\}$ is unavoidable.

The following observation is straightforward:

(2.2) Let $Y \subseteq A^*XA^*$ be unavoidable. Then X is unavoidable.

As a consequence we have:

(2.3) The set $X \subseteq A^+$ is unavoidable iff the set $X^1 = X\backslash(A^*XA^+ \cup A^+XA^*)$ is unavoidable.

In particular we say that $X \subseteq A^+$ is *normal* if $X \cap (A^*XA^+ \cup A^+XA^*) = \varnothing$, that is if no word of X is a proper subword of an other word of X.

Thus we have:

(2.4) If $X \subseteq A^+$ is a minimal unavoidable set, it is normal.

Assume now that X is unavoidable and normal. Then by (2.1) all words of length n contain a subword in X. This shows that the length of the words of X is bounded by n.

In view of (2.4) we obtain:

(2.5) If $X \subseteq A^+$ is a minimal unavoidable set, it is finite.

With the help of this last observation we shall restrict ourselves, from now on, to finite unavoidable sets.

2.2. Preliminary Results on Unavoidable Sets

We first recall two known estimates on the number n appearing in (2.1) and on the cardinality of unavoidable sets.

Assume n is the minimum value for which (2.1) holds. Then the maximum length of a word avoiding X is equal to $n-1$. As we shall see in the next section, this number is bounded by $|A|^m$ where m is the maximum length of the words of X. In the case of unavoidable sets consisting only of words of the same length, we have the following result (cf. [Cr et al.]):

Theorem 2.1 *Let $m > 0$ be an integer and $X \subseteq A^m$. If there exists a word of length $|A|^{m-1}+m-1$ avoiding X, then X is avoidable.*

Furthermore, for all $m > 0$ there exists an unavoidable subset $X \subseteq A^m$ and a word of length $|A|^{m-1}+m-2$ which avoids it.

With respect to the cardinality of unavoidable subsets we have (cf. [Sch]):

Theorem 2.2 *If $X \subseteq A^m$ is unavoidable, then*

$$|X| \geq \frac{|A|^m}{m} .$$

We now relate unavoidable sets to some conditions involving sets of infinite words. These conditions are equivalent to condition (2.1) and prove useful in the sequel.

Let $A^\omega, {}^\omega A$ and ${}^\omega A^\omega$ be respectively the set of all right infinite, left infinite, and two-way infinite words. Thus, typically $a_0a_1a_2\cdots$, $\cdots a_{-2}a_{-1}a_0$ and $\cdots a_{-1}a_0a_1\cdots$ with $a_i \in A$ are elements of A^ω, ${}^\omega A$ and ${}^\omega A^\omega$ respectively. As usual the elements of ${}^\omega A^\omega$ are defined up to a translation: $\cdots a_{-1}a_0a_1\cdots = \cdots b_{-1}b_0b_1\cdots$ if there exists an integer $t \in \mathbb{Z}$ such that $a_i = b_{i+t}$ for all $i \in \mathbb{Z}$. Given $u \in A^+$, we let $u^\omega = uu\cdots \in A^\omega$, ${}^\omega u = \cdots uu \in {}^\omega A$ and ${}^\omega u^\omega = \cdots uuu\cdots \in {}^\omega A^\omega$.

Lemma 2.3 *$X \subseteq A^+$ is unavoidable iff it satisfies any one of the following conditions:*

(i) $A^\omega = A^*XA^\omega$

(ii) ${}^\omega A^\omega = {}^\omega AXA^\omega$

(iii) *For all* $u \in A^+$, $u^\omega \in A^*XA^\omega$.

(iv) *For all* $u \in A^+$, ${}^\omega u^\omega \in {}^\omega AXA^\omega$.

Proof: The implications (2.1) $\Rightarrow$ (*i*) and (*i*) $\Rightarrow$ (*iii*) are straightforward. We shall prove (*iii*) $\Rightarrow$ (2.1).

Assume by contradiction that X is avoidable. Since X is finite, the infinite set $A^*\backslash A^*XA^*$ is rational, and by the pumping lemma there exist two words $w \in A^*$ and $u \in A^+$ such that $wu^n \notin A^*XA^*$ for all $n \geq 0$. Thus, $u^\omega \notin A^*XA^\omega$ contradicting (iii).

The implications (2.1) $\Rightarrow$ (*ii*), (*ii*) $\Rightarrow$ (*iv*) and (*iv*) $\Rightarrow$ (2.1) can be proven similarly. □

Using condition (i) of the previous lemma we obtain:

(2.6) Let $x \in X$. Then X is unavoidable iff $(X\backslash\{x\}) \cup xA$ is itself unavoidable.

2.3. A Reformulation of the Conjecture

In order to give some equivalent statements of the conjecture, we introduce a partial ordering over all finite subsets of A^*.

Let $X,Y \subseteq A^+$ have the same cardinality $n \geq 0$. We write $X \leq Y$ whenever there exists an ordering $X = \{x_i\}_{1 \leq i \leq n}$ and $Y = \{y_i\}_{1 \leq i \leq n}$ such that x_i is a prefix of y_i for all $1 \leq i \leq n$. We write $X < Y$ whenever $X \leq Y$ and $X \neq Y$.

The following assertions are straightforward

(2.7) $X \leq Y$ implies $Y \subseteq A^*XA^*$.

(2.8) If Y is an extension of X, then $X \leq Y$.

Now we show that the minimality of unavoidable sets is preserved by extension.

Lemma 2.4 *Let X be a minimal unavoidable set. If Y is an unavoidable set such that $X \leq Y$, then Y is minimal.*

Proof: By hypothesis, for some integer $n \geq 0$ we have $X = \{x_i\}_{1 \leq i \leq n}$, $Y = \{y_i\}_{1 \leq i \leq n}$ and $x_i \leq y_i$ for $1 \leq i \leq n$.

Assume the contrary, that is there exists $y_i \in Y$ such that $Y\backslash\{y_i\}$ is unavoidable. Then because $X\backslash\{x_i\} \leq Y\backslash\{y_i\}$, (2.2) and (2.7) imply that $X\backslash\{x_i\}$ is unavoidable, a contradiction. □

We may now restate Ehrenfeucht's conjecture in terms of infinite extensions.

Lemma 2.5 *The following statements are equivalent:*

(i) *For every unavoidable set $X \subseteq A^+$ there exist $x \in X$ and $a \in A$ such that $Y = (X \backslash \{x\}) \cup \{xa\}$ is unavoidable.*

(ii) *For every unavoidable set $X \subseteq A^+$ there exist $x \in X$ and an infinite sequence $a_1, a_2, \ldots, a_i, \cdots$ where $a_i \in A$ such that $X_n = (X \backslash \{x\}) \cup \{xa_1a_2 \cdots a_n\}$ is unavoidable for all $n > 0$.*

(iii) *For every unavoidable set $X \subseteq A^+$, there exist $x \in X$ and an infinite sequence $x_1, x_2, \ldots, x_n, \cdots$ where $x_i \in A^+$, such that $X_n = (X \backslash \{x\}) \cup \{xx_1 \cdots x_n\}$ is unavoidable for all $n > 0$.*

Proof: Because of (2.2) and (2.8) statements (ii) and (iii) are equivalent. Clearly (ii) implies (i).

Now (i) implies that there exists an infinite sequence $X < Y_1 < \cdots < Y_i < \cdots$ of unavoidable sets. Since X is finite, there exist an infinite subsequence $i_1 < i_2 < \cdots < i_n < \cdots$ and an infinite sequence $x < x_1 < x_2 < \cdots < x_n < \cdots$ such that $x \in X$ and $x_n \in Y_{i_n}$ for all $n > 0$. Define $X_n = (X \backslash \{x\}) \cup \{x_n\} \subseteq Y_{i_n}$. Because of assertion (2.2), X_n is unavoidable, thus completing the proof. □

Whenever x satisfies condition (ii) or (iii), we say that it is *infinitely extendible*.

3. AN AUTOMATON RECOGNIZING $A^*\backslash A^*XA^*$

With every normal subset X, we associate a finite deterministic (in general non-minimal) automaton recognizing the set of words avoiding X. Next we provide an efficient algorithm to decide whether or not X is unavoidable, and moreover, when it is, whether or not it is extendible and minimal.

We denote by $P(X)$ the set of all prefixes of all words in X. Let s be the partial function undefined over A^*XA^+, which to every word $w \in A^*\backslash A^*XA^+$ assigns the longest word $u \in P(X)$ which is a suffix of w.

Let $\equiv$ be the equivalence relation defined on A^* by $w_1 \equiv w_2$ iff either w_1, w_2 both belong to A^*XA^* or $s(w_1) = s(w_2)$. Since X is normal, $\equiv$ is a right congruence, and we may consider a transition function $\lambda : P(X) \times A \to P(X)$ undefined on $X \times A$, and otherwise satisfying:

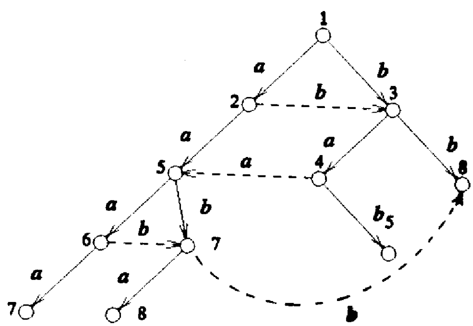
$$\lambda(w,a) = s(wa) .$$

Taking $P(X)$ as the set of states, $\{1\}$ as the initial state, $P(X)\backslash X$ as the set of final states and λ as the transition function, we obtain a finite determinstic automaton recognizing $L = A^*\backslash A^*XA^*$. Thus X is unavoidable iff L is finite. This again amounts to saying in terms of the state diagram of the automaton, that X is unavoidable iff there is no cycle in the state diagram.

Abusing terminology somewhat, we refer to the above automaton as the "automaton of X" and we shall denote it by $\mathbf{A}(X)$. If necessary, we shall write s_x and λ_x instead of λ and s, to remind ourselves to which set X these partial functions refer.

For the pictorial representation of $\mathbf{A}(X)$ we first draw the usual tree hanging from its root whose leaves are the elements of X and whose internal nodes are all proper prefixes $P(X)\backslash X$. In order to complete the automaton, i.e. to define the transitions $\lambda(w,a)$ where $w \in P(X)\backslash X$ and $wa \notin P(X)$, it helps to observe that if $c,d \in A$ and $cw \in P(X)\backslash X$ satisfy $cwd \notin P(X)$, then $\lambda(cw,d) = \lambda(w,d)$. The transitions on the prefixes may thus be easily computed by increasing lengths.

Example 3.1 $A = \{a,b\}$, $X = \{a^4, a^2ba, bab, b^2\}$.



$$\lambda(a,b) = \lambda(1,b) = b$$
$$\lambda(ba,a) = \lambda(a,a) = aa$$
$$\lambda(aaa,b) = \lambda(aa,b) = aab$$
$$\lambda(aab,b) = \lambda(ab,b) = b^2$$

3.1. Unavoidability

For the previous example, the state diagram immediately shows that there is no cycle. However for more complex examples, a visual inspection is insufficient. We need a more efficient procedure.

It is standard to associate with each letter $a \in A$ a square matrix indexed by $P(X)$, and having at position (p,p') a 1 if $\lambda(p,a) = p'$ and 0 otherwise.

The following result amounts to show how to simultaneously triangulate, if possible, the matrices associated with all $a \in A$, by a mere permutation of the columns.

Theorem 3.1 *A subset $X \subseteq A^+$ is unavoidable iff there exists a function $f : P(X) \to \mathbb{N}$ satisfying the two conditions:*

(i) $f(1) = 0$

(ii) $f(p) = 1+\max\{f(p') \mid \lambda(p',a) = p$ *for some* $a \in A\}$.

Proof: Condition (ii) implies $f(\lambda(p,u)) \geq f(p) + |u|$ whenever $\lambda(p,u)$ is defined, which shows that the automaton has no cycle, that is X is unavoidable.

Conversely, if X is unavoidable, for all $p \in P(X)$ let $f(p)$ be the length of the longest word $w \in A^*$ such that $\lambda(1,w) = p$. Then f satisfies conditions (i) and (ii). □

Example 3.1 (continued).
The corresponding values of the function f are shown in the diagram as node labels.

3.2. Extendibility

By definition, in order to check whether or not a given unavoidable set X is extendible, it suffices to verify that for some $x \in X$ and some $a \in A$, $(X\setminus\{x\}) \cup \{xa\} = Y$ is unavoidable. We know from the previous subsection, how to verify whether or not Y is unavoidable. It thus suffices to know how to construct its automaton, that is to show how the automaton $\mathbf{A}(X)$ is modified by an extension.

Proposition 3.2 *Let $X \subseteq A^+$ be a normal unavoidable set, $x \in X$ and $a \in A$. Consider $Y = (X\setminus\{x\}) \cup \{xa\}$ and observe that $P(Y) = P(X) \cup \{xa\}$.*

Then the transition function λ_Y of the automaton of Y satisfies for all $p \in P(X)\setminus\{x\}$ and $b \in A$:

$$\lambda_Y(p,b) = \lambda_X(p,b) \quad .$$

Proof: Assume the contrary, that is $\lambda_Y(p,b) \neq \lambda_X(p,b)$ for some $p \in P(X)\setminus\{x\}$ and $b \in A$. Then $\lambda_Y(p,b) = xa$, i.e. $p = ux$ for some $u \in A^+$ which violates the hypothesis that X is normal. □

3.3. Minimality

The following lemma shows that if X is a minimal unavoidable set, then every word of X which can be extended, can only be extended by one letter of the alphabet.

Lemma 3.3 *Let X be a minimal unavoidable set, $x \in X$ and $a, b \in A$. If $(X \backslash \{x\}) \cup \{xa\}$ and $(X \backslash \{x\}) \backslash \{xb\}$ are both unavoidable, then $a = b$.*

Proof: It suffices to prove that for every $x \in X$ there exist a letter $a \in A$ and an infinite word $s \in A^\omega$ having occurrences of x, no occurrence in $X \backslash \{x\}$ and such that all occurrences of x are followed by the same letter $a \in A$. In other words, all factorizations $s = s_1 y s_2$ with $s_1 \in A^*$, $y \in X$ and $s_2 \in A^\omega$ imply $y = x$ and $s_2 \in aA^\omega$.

Because of the minimality of X, there exists a word w having exactly two occurrences of x and no occurrence in $X \backslash \{x\}$: $w = w_1 x u w_2 = w_1 v x w_2$ for some $u, v \in A^+$. Equality $xu = vx$ implies $x = (zt)^r z$ $u = tz$ and $v = zt$ for some $z, t \in A^*$ and $r \geq 0$.

Consider the infinite word $s = (zt)^\omega$ and an occurrence of $y \in X$ in s : $s = s_1 y s_2$ with $s_1 \in A^*$ and $s_2 \in A^\omega$. Because X is normal, y must be a subword of $(zt)^{r+1} z = xu = vx$ which implies $y = x$. Now because the two occurrences of x in w are consecutive, we obtain $s_1 \in (zt)^*$ and therefore $s_2 \in (tz)^\omega$ which completes the proof. □

As a consequence, we obtain a characterization of minimal unavoidable sets, which via Theorem 3.1 and Proposition 3.2 provides an efficient procedure to test minimality.

Corollary 3.4 *Let X be an unavoidable set. Then it is minimal iff for each $x \in X$ there exists at most one letter $a \in A$ such that $X \backslash \{x\} \cup \{xa\}$ is unavoidable.*

4. PARTIAL SOLUTIONS

In this section we consider two different conditions under which unavoidable sets are extendible.

Among the words of an unavoidable set there is necessarily some power a^n of any letter $a \in A$. These words definitely play a special rôle and we are able to establish under which conditions they are extendible.

In the second case, we try to formalize the intuition that if a minimal unavoidable set possesses some word x which is very long compared with all other words in X, then this word must be extendible.

Theorem 4.1 *Let X be a minimal unavoidable set and $a^n \in X$ for some $n > 0$. Then a^n is extendible iff $X \cap A^*ba^{n-1}bA^* = \emptyset$.*

Proof: Assume first $X \cap A^*ba^{n-1}bA^* = \emptyset$. We shall show that any $s \in {}^\omega A^\omega$ contains an occurrence of a word in $Y = (X\setminus\{a^n\}) \cup \{a^{n+1}\}$. If it contains no occurrence of a^n then because X is unavoidable, s contains an occurrence of a word in $X\setminus\{a^n\} = Y\setminus\{a^{n+1}\}$. So we may assume from now on that s contains some occurrences of a^n and no occurrence of a^{n+1}.

Denote by $s' \in {}^\omega A^\omega$ the word obtained from s by substituting a^{n-1} for each occurrence of a^n in s:

$$s = \cdots w_{-1}x_{-1}w_0x_0w_1x_1 \cdots w_px_p \cdots$$

$$s' = \cdots w_{-1}x'_{-1}w_0x'_0w_1x'_1 \cdots w_px'_p \cdots$$

where for all $i \in \mathbb{Z}$ we have $w_i \in bA^* \cap A^*b\setminus A^*a^nA^*$

$$x_i = a^n \quad \text{and} \quad x'_i = a^{n-1} \quad .$$

Since X is unavoidable, s' has some occurrence x in $X\setminus\{a^n\}$. Because of the hypothesis, x is necessarily a subword of $x'_{i-1}w_ix'_i$ for some $i \in \mathbb{Z}$, i.e. a subword of $x_{i-1}w_ix_i$, thus proving one direction.

Conversely, assume by contradiction that X contains a word $x \in A^*ba^{n-1}bA^*$ and that a^n can be extended.

Since X is minimal there exists a two-way infinite word $s \in {}^\omega A^\omega$ which has some occurrences of x and no occurrence of any word from $X\setminus\{x\}$.

Denote by x' the word obtained from x by substituting ba^nb for the first occurrence of $ba^{n-1}b$ in x and by $s' \in {}^\omega A^\omega$ the two-way infinite word obtained from s by substituting x' for all occurrences of x in s. Formally we have:

$$s = \cdots w_{-1}x_{-1}w_0x_0 \cdots w_px_p \cdots$$

$$s' = \cdots w_{-1}x'_{-1}w_0x'_0 \cdots w_px'_p \cdots$$

where for all $i \in \mathbb{Z}$ we have

$$w_i \in bA^* \cap A^*b \,, \quad x'_i = a^n \,, \; x_i = a^{n-1} \,,$$

and there exist a suffix u of $\cdots w_{i-1}x_{i-1}w_i$ and a prefix v of $x_iw_{i+1}\cdots$ such that $vx_iu = x$ and $vx'_iu = x'$.

Since $Y = (X\setminus\{a^n\}) \cup \{a^{n+1}\}$ is unavoidable, s' contains some occurrence $y \in Y\setminus\{a^{n+1}\} = X\setminus\{a^n\}$. Because $Y \cap A^*a^nA^* = \{a^{n+1}\}$, y is necessarily a subword of $x_{i-1}w_ix_i$ for some $i \in \mathbb{Z}$, i.e. $y = x$. But this contradicts the fact that x'_{i-1} and x'_i are two consecutive occurrences of a^n. □

Theorem 4.2 *Let X be a minimal unavoidable set and assume there exist an integer $n > 0$ and a word $x \in X$ such that $|x| \geq 3.2^{n+1}$ and $|y| \leq n$ for* *all* $y \in X\backslash\{x\}$.

Then there exist $u = u_1u_2 \in A^+$ *and* $r > 0$ *satisfying the following conditions:*

(i) ${}^\omega u^\omega$ *is the only two-way infinite word avoiding* $X\backslash\{x\}$.

(ii) $x = (u_1u_2)^r u_1$.

(iii) *For all* $p \geq 0$, $X_p = (X\backslash\{x\}) \cup \{(u_1u_2)^{r+p}u_1\}$ *is unavoidable.*

Proof: Observe first that (i) trivially implies (iii). Further, if ${}^\omega u^\omega$ is the only two-way infinite word avoiding $X\backslash\{x\}$, then x is a subword of this word. This means that there exists a word $u'u''$ such that $u = u''u'$ and $x = (u'u'')^r u'$ for some integer $r \geq 0$. But ${}^\omega(u'u'')^\omega = {}^\omega u^\omega$ which shows that (i) implies (ii).

We now turn to prove assertion (i). Given any state q of $\mathrm{A}(X\backslash\{x\})$ we define:

$$F_q = \{w \in A^+ \mid \lambda(q,w) = q\}$$

and we denote by E_q the subset of words in F_q which define an elementary cycle in the state diagram of the automaton, i.e. the words which satisfy:

$$w = w_1w_2w_3 \ , \quad \lambda(q,w_1) = p \ , \quad \lambda(p,w_2) = p \ , \quad \lambda(p,w_3) = q \ ,$$
$$\text{and } w_1w_3 \neq 1 \text{ implies } w_2 = 1 \ .$$

Observe that all words in E_q are of length less than 2^{n+1} .

Claim 1. There exists, up to a conjugacy class, a unique primitive word $u \in A^+$, such that $E_q \subseteq u_q^+$, where u_q is a conjugate of u depending only on q.

Let q,p be two states in the automaton $\mathrm{A}(X\backslash\{x\})$, $v \in E_q$ and $w \in E_p$. Then ${}^\omega v^\omega$ and ${}^\omega w^\omega$ avoid $X\backslash\{x\}$. Thus these two words have x as a common subword. Since $|x| \geq |v|+|w|-1$, by ([LeSch] Cor. 1) v and w are powers of two conjugate primitive words. Thus there exist $u_1, u_2 \in A^*$ with $u_1 \neq 1$, and $i,j > 0$ such that u_1u_2 is primitive and:

$$v = (u_1u_2)^i \ , w = (u_2u_1)^j \ . \tag{4.1}$$

It now suffices to prove that $p = q$ implies $u_2 = 1$. Thus, assume that $p = q$, and therefore $v,w \in E_q$. Then the words ${}^\omega vw^\omega$ and ${}^\omega v^\omega$ have a common subword x of length $|x| \geq |v|+|vw|-1$ which by the same result quoted above implies that $vw = (u_1u_2)^i(u_2u_1)^j$ is a power of some conjugate of u_1u_2 . We obtain $u_1u_2 = u_2u_1$, i.e., $u_2 = 1$ which proves the claim.

Claim 2. $F_q \subseteq u_q^+$ holds for all $q \in P(X)$.

Using Claim 1, assume by contradiction that for some state q and some word of minimal length $w \in F_q$ we have $w \notin u_q^+$. Then there exists a factorization $w = w_1w_2w_3$, with $w_2 \neq 1$ and $w_1w_3 \neq 1$ and a state p such that the following holds:

$$\lambda(q, w_1) = p \ , \lambda(p, w_2) = p \text{ and } \lambda(p, w_3) = q \ .$$

By the minimality of $|w|$, and the fact that w_2 and w_3w_1 belong to F_p we have $w_2 = u_p^i$ and $w_3w_1 = u_p^j$ for some $i, j > 0$. Furthermore $w_1w_3 \in F_q$ implies $w_1w_3 = u_q^j$. Now equality $w_3u_q^j = u_p^jw_3$ implies $w_3u_q^k = u_p^kw_3$ for any $k \geq 0$. Thus:

$$w_3w_1w_2w_3 = u_p^ju_p^iw_3 = w_3u_q^{i+j} \ , \text{i.e.}, \ w \in u_q^+ \ , \text{a contradiction} \ .$$

In order to complete the proof, it suffices to observe that for any primitive word v such that "v" avoids $X\backslash\{x\}$, there exists a state q and some integer $i > 0$ such that $v^i \in F_q$, which by Claim 2 shows that v is a conjugate of u. □

As a consequence of this last result, we will show that Conjecture I is equivalent to its "two-way" version where instead of extending to the right, we may extend in either direction.

Conjecture II : *For every finite unavoidable set $X \subseteq A^+$ there exist $x \in X$ and $a \in A$ such that either $(X\backslash\{x\}) \cup \{xa\}$ or $(X\backslash\{x\}) \cup \{ax\}$ is unavoidable.*

Theorem 4.3 *Conjectures I and II are equivalent.*

Proof: Obviously Conjecture I implies Conjecture II. We are going to prove that the reverse also holds.

Let us say that an unavoidable set $Y \subseteq A^+$ is a two-way extension of the unavoidable set $X \subseteq A^+$ if there exist $x \in X$ and $a \in A$ such that either

$$Y = (X\backslash\{x\}) \cup \{xa\} \text{ or } Y = (X\backslash\{x\}) \cup \{ax\} \ .$$

If Conjecture II holds, then for any unavoidable set X there exists an infinite sequence $Y_0, Y_1, \cdots, Y_k, \cdots$ of subsets such that $Y_0 = X$, and Y_{k+1} is a two-way extension of Y_k for each $k \geq 0$. Since X is finite there exist $x \in X$, two sequences $s_0, s_1, \cdots, s_k, \cdots$ and $p_0, p_1, \cdots, p_k, \cdots$ and a subsequence $i_0, i_1, \cdots, i_k, \cdots$ such that the following conditions hold:

(i) $p_k \cdots p_0xs_0 \cdots s_k \in Y_{i_k}$

(ii) $|p_k \cdots p_0 s_0 \cdots s_k|$ is strictly increasing.

Assume first the the sequence $s_0, s_1, \cdots, s_k, \cdots$ contains infinitely many elements different from the empty word. Then, if necessary by considering a subsequence, we may assume that all elements are different from the empty word. By (2.2) all subsets $X_k = (X \backslash \{x\}) \cup \{xs_0 \cdots s_k\} \subseteq Y_{i_k}$ are unavoidable, showing thus that x is infinitely extendible in the usual way.

Assume next that the sequence $p_0, p_1, \cdots, p_k, \cdots$ contains infinitely many elements different from the empty word. Then, as in the previous case, we may assume that they are all different from the empty word. By (2.2) all subsets $X_k = (X \backslash \{x\}) \cup \{p_k \cdots p_0 x\} \subseteq Y_{i_k}$ are unavoidable. Thus, the word x is infinitely left extendible, and since Theorem 4.2 dually applies to left extendiblility there exist $u = u_1 u_2 \in A^+$ and $r > 0$ such that $x = u_1(u_2u_1)^r$ and such that all $X_p = (X \backslash \{x\}) \cup \{u_1(u_2u_1)^{r+p}\}$ are extendible. Now it suffices to observe that $u_1(u_2u_1)^{r+p} = u_1(u_2u_1)^r (u_2u_1)^p = x(u_2u_1)^p$ to show that $X < X_1 < \cdots < X_p < \cdots$ holds. □

5. A REDUCTION RESULT

Let $A = \{a_i\}_{1 \le i \le n}$ and $B = \{a, b\}$. Denote by $\psi: A^* \to B^*$ the morphism defined by $\psi(a_i) = a^i b$ for $1 \le i \le n$ and extend it in the usual way to A^{ω}, ${}^{\omega}A$, and ${}^{\omega}A^{\omega}$ (e.g. $\psi(a_0 a_1 \cdots) = \psi(a_0)\psi(a_1)\cdots$). Since the image $\psi(A)$ is a comma free code, ψ maps ${}^{\omega}A^{\omega}$ bijectively onto $I = {}^{\omega}B^{\omega} \backslash {}^{\omega}B\{a^{n+1}, b^2\}B^{\omega}$.

Now with the set $X \subseteq A^+$, we associate the set $Y = \theta(X) = b\psi(X) \cup \{a^{n+1}, b^2\}$. This "encoding", preserves the main properties of X as is now shown.

Lemma 5.1 *X is unavoidable (resp. unavoidable and minimal) iff Y is unavoidable (resp. unavoidable and minimal).*

Proof: We first verify that the following holds for all X :

$$\psi({}^{\omega}AXA^{\omega}) = I \cap {}^{\omega}Bb\psi(X)B^{\omega} \quad . \tag{5.1}$$

The inclusion $\subseteq$ is obvious. Thus it suffices to prove $I \cap {}^{\omega}Bb\psi(X)B^{\omega} \subseteq \psi({}^{\omega}AXA^{\omega})$. Indeed, if $w = zb\psi(x)t \in I$ for some $z \in {}^{\omega}B$, $x \in X$ and $t \in B^{\omega}$, then $z \in {}^{\omega}Ba \backslash {}^{\omega}B\{a^{n+1}, b^2\}B^*$ and $t \in aB^{\omega} \backslash B^*\{a^{n+1}, b^2\}B^{\omega}$, i.e. z and t may be factorized in elements of $\psi(A)$. Therefore there exist $w_1 \in {}^{\omega}A$ and $w_2 \in A^{\omega}$ such that $\psi(w_1) = zb$ and $\psi(w_2) = t$ which implies $w = \psi(w_1 b w_2)$.

Now observe that ${}^{\omega}B^{\omega}$ is partitioned into:

$${}^{\omega}B^{\omega} = I \cup {}^{\omega}B\{a^{n+1}, b^2\}B^{\omega} \quad .$$

Thus, the set $Y = b\psi(X) \cup \{a^{n+1}, b^2\}$ is unavoidable iff $I \subseteq {}^*Bb\psi(X)B^*$, i.e. because of (4.1), and the fact that ψ maps ${}^*A^*$ bijectively onto I, iff ${}^*A^* = {}^*AXA^*$. This proves the first part of the proposition.

The second part relies upon the fact that if $Y' \subseteq Y$ is unavoidable then $Y' \cap (a^+ \cup b^+) = Y \cap (a^+ \cup b^+) = \{a^{n+1}, b^2\}$. Thus there exists an unavoidable proper subset $X' \subseteq X$ iff there exists an unavoidable proper subset $Y' \subseteq Y$. □

Using the same notations we have

Lemma 5.2 *Assume X is unavoidable. Then an element $x \in X$ is infinitely extendible iff the element $b\psi(x) \in Y$ is infinitely extendible.*

Proof: If x is infinitely extendible, then there exists an infinite sequence $a_1, a_2, \cdots, a_k, \cdots$ where $a_k \in A$ for all $k > 0$, such that $X_k = (X \backslash \{x\}) \cup \{xa_1 \cdots a_k\}$ is unavoidable. By the previous lemma, this shows that $Y_k = (Y \backslash \{b\psi(x)\}) \cup \{b\psi\ (xa_1 \cdots a_k)\}$ is unavoidable, which proves one direction.

Now if $b\psi(x)$ is infinitely extendible, there exists an infinite sequence $u_1, u_2, \cdots, u_k, \cdots$ where $u_k = a^{i_k}b$ for some $1 \leq i_k \leq n$, such that $Y_k = (Y \backslash \{b\psi(x)\}) \cup \{b\psi(x)u_1 \cdots u_k\}$ is unavoidable. Because of the preceding lemma, $X_k = (X \backslash \{x\}) \cup \{x\psi^{-1}(u_1) \cdots \psi^{-1}(u_k)\}$ is unavoidable, which completes the proof. □

Example 5.1 Consider $A = \{a, b\}$ and $X = \{a^2, abab, b^2\}$. Then X is unavoidable (basically for the same reasons as in Example 2.1) and $abab$ is the only infinitely extendible element of X. Then by Theorem 4.1 and the two previous lemmas, the subsets $\theta^k(X) = \theta(\theta^{k-1}(X))$ are unavoidable for all $k > 0$ and have a unique infinitely extendible word $\psi^k(x)$. Furthermore the cardinality of $\theta^k(X)$ is equal to $2k+3$. In other words this shows that there are unavoidable sets of arbitrary cardinality having a unique infinitely extendible word.

The following surprising result shows that Conjecture I need only be proven in the case $B = \{a, b\}$ and $X \cap (a^+ \cup b^+) = \{a^3, b^2\}$.

Theorem 5.3 *Ehrenfeucht's conjecture holds iff it holds for all unavoidable sets X over a binary alphabet $B = \{a, b\}$ such that $a^3, b^2 \in X$.*

Proof: Consider an unavoidable set $X \subseteq A^+$ where $|A| = n \geq 2$. If $X = A$, then every element of X is infinitely extendible. Therefore assume $X \neq A$.

Let $A = \{a_i\}_{1 \leq i \leq n}$ be an enumeration of the alphabet and consider $Y = b\psi(X) \cup \{a^3, b^2\}$.

Because of Lemma 5.1, Y is unavoidable. By Theorem 4.1 a^{n+1} and b^2 are not extendible. Thus, by hypothesis there exists some $x \in X$ such that $b\psi(x)$ is infinitely extendible. By Lemma 5.2, $x \in X$ is itself infinitely extendible.

If $n = 2$, then we are done. Otherwise we repeat the same argument with $Y \subseteq B^+$. □

REFERENCES

[Cr et al] M. Crochemore, M. le Rest, and P. Wender, Sur les ensembles mevitables, *Information Proc. Letters*, to appear.

[Le Sch] A. Lentin and M.P. Schutzenberger, A Combinatorial Problem in the Theory of Free Monoids, *Proceedings of the Conference* held at the University of North Carolina at Chapel Hill, Box & Bowling (ed.), Chapell Hill (1967), 128-144.

[Sch] M.P. Schutzenberger, On the Synchronizing Properties of Certain Prefix Codes, *Information and Control 7*, (1964), 23-36.