Acta Cybernetica, Tom. 10, Fasc. 1-2, Szeged, 1991

On strict codes

Nguyen Huong Lam and Do Long Van *

Abstract

This paper continues an earlier paper of the authors. The maximality, the decomposability etc. for infinitary strict codes are considered. Every infinitary strict code is shown to be included in an infinitary strict-maximal one. The so-called Theorem of Defect for infinitary strict codes is proved. Some conditions for an infinitary strict code to be written by an indecomposable ones are stated.

1 Preliminaries

The concept of strict codes has been first mentioned in [3]. The classical definition of a code says that the (finite) identity relations are the only (finite) relations satisfied by the code. For a strict code the demand is stronger: any relation, finite or infinite, which is satisfied by the code is an identity one. So, strict codes form a subclass of codes. In [7] a particular case of strict codes, namely, that of finitary strict codes was considered. In our paper [6] we studied infinitary strict codes; we proposed some procedures to verify whether a given language is a strict code. Also, we characterized strict codes by ∞-submonoids generated by them. In the present article, which is a sequel to [6], we mostly adapt some well-known notions and properties of codes for strict one such as maximality, decomposability, Theorem of Defect, etc.

In what follows we mostly use standard terminology and notation (see, e.g. [5], [1]). Let A be an alphabet, finite or countable, A^* the free monoid generated by A whose elements are called finite words. We denote A^N the set of infinite words over A and $A^\infty = A^* \cup A^N$ whose elements we call simply words. We make A^∞ a monoid equipping it with the product defined as:

For any words α, β of A^∞,

$$\alpha\beta = \begin{cases} \alpha & \text{if } \alpha \in A^N, \beta \in A^\infty \\ \alpha\beta & \text{if } \alpha \in A^*, \beta \in A^\infty \end{cases}$$

where $\alpha\beta$ means the catenation of α and β (see [6]). Clearly, the empty word, denoted by ϵ, is the unit of A^∞.

We call a subset X of A^∞ (respectively, of A^*) an infinitary language (respectively, finitary language). For a finite subset X, Card$X$ denotes its cardinality; also, to simplify the notation we often identify a singleton set with its element. For a word $x \in A^*$, $|x|$ denotes the length of x and we say by convention that $|\epsilon| = 0$ and $|x| = \omega$ if $x \in A^N$.

*Institute of Mathematics, P.O.Box 631, Bo Ho, Hanoi, Vietnam

For any infinitary language X we denote:
$X_{\text{fin}} = X \cap A^*, X_{\text{inf}} = X \cap A^N$
$XY = \{\alpha\beta : \alpha \in X, \beta \in Y\}$, $Y \subseteq A^\infty$

(the product extended to languages).
$X^2 = XX$
$X^{n+1} = XX^n$, $n = 1, 2, \ldots$
$X^* = \bigcup_{n \geq 0} X^n$

(the smallest submonoid of A^∞ containing X).
$X^+ = X^* - \epsilon$
$X^\omega = \{x_1 x_2 \ldots : x_i \in X_{\text{fin}}, i = 1, 2, \ldots\}$

(the set of all infinite products of elements in X_{fin}).
$X^\infty = X^* \cup X^\omega$
$X^{+\infty} = X^\infty - \epsilon$.

For every $n \geq 1$ we introduce the set $X_{(n)}$ of n-tuples defined as:
$$X_{(n)} = \{(x_1, x_2, \ldots, x_n) : x_1, \ldots, x_{n-1} \in X_{\text{fin}}, x_n \in X\}$$
and the set $X_{(\omega)}$ of ω-tuples defined as:
$$X_{(\omega)} = \{(x_1, x_2, \ldots) : x_i \in X_{\text{fin}}, i = 1, 2, \ldots\}$$
and we put
$X_{(*)} = \bigcup_{n \geq 1} X_{(n)}$
$X_{(\infty)} = X_{(*)} \cup X_{(\omega)}$

We say that a word x of A^∞ admits a *$*$-factorization* (resp. an *ω-factorization*) $(x_1, x_2, \ldots)$ over X provided $x = x_1 x_2 \ldots$ with $(x_1, x_2, \ldots) \in X_{(*)}$ (resp. $(x_1, x_2, \ldots) \in X_{(\omega)}$); we say that X admits an *∞-factorization* over X if it admits either a $*$-factorization or an ω-factorization over X. A given subset X is said to be an *infinitary code* (resp. an *infinitary strict code*) if every word of A^∞ admits at most one $*$-factorization (resp. one ∞- factorization) [6].

Finally, for any two subsets X, Y of A^∞, we define:
$XY^{-1} = \{\alpha \in A^\infty : \exists \beta \in Y : (\alpha\beta \in X) \wedge ((|\alpha| = \omega) \Longrightarrow \beta = \epsilon)\}$.
$Y^{-1}X = \{\alpha \in A^\infty : \exists \beta \in Y : (\beta\alpha \in X) \wedge ((|\beta| = \omega) \Longrightarrow \alpha = \epsilon)\}$.

2 Maximality

In this section we consider maximality properties of strict codes and ∞-submonoids generated by them. First, we show that each strict code is included in a strict-maximal one. A strict code X is called *strict-maximal* if it is not contained properly in any other strict code. X is called *relatively strict-maximal* if for every finite word $w \in A^*$, $X \cup w$ is no more a strict code.

Theorem 2.1 *Every strict code is contained in a strict-maximal one over A.*

Proof: First, we prove that every strict code is included in a relatively strict-maximal one. To do this we enumerate all finite nonempty words in some order

$$A^+ = \{w_1, w_2, \ldots\}$$

and define an increasing sequence of strict codes $X_0 \subseteq X_1 \subseteq X_2 \subseteq \ldots$ as follows:

Put $X_0 = X$ and suppose for some $n \geq 0, X_n$ has been defined. Let $i(n)$ be the smallest integer such that $X_n \cup w_{i(n)}$ is a strict code (if no such $i(n)$ exists, we put $X_{n+1} = X_n$). Put $X_{n+1} = X_n \cup u_n$, where u_n is any word in $\{X_n \cup w_{i(n)}\}^* w_{i(n)}^\omega$. Since $X_n \cup w_{i(n)}$ is a strict code, so is X_{n+1}. Thus X_{n+1} is defined and by induction our sequence is built.

Consider the set $Y = \bigcup_{n \geq 1} X_n$. Since $Y = X \cup \{u_n : n \geq 0\}$ and all u_n's are infinite words it follows that every ∞-factorization over Y is also one over X_n for some n. This means that Y is a strict code because each X_n is a strict code. Also, it is easy to see that Y is relatively strict-maximal by construction of the sequence X_n. Thus every strict code is contained in a relatively strict-maximal one.

Next, we prove that the class of relatively strict-maximal codes is inductive, i.e. every chain (by inclusion) has an upper bound. Indeed, let

$$X_\alpha \subseteq X_\beta \subseteq X_\gamma \subseteq \dots$$

be a chain in this class, indexed by a set I. Since each member of this chain is relatively strict-maximal, we have

$$X_{\alpha \mathrm{fin}} = X_{\beta \mathrm{fin}} = X_{\alpha \mathrm{fin}} = \dots$$

Putting

$$Z = \bigcup_{\gamma \in I} X_\gamma$$

we have $Z_{\mathrm{fin}} = X_{\gamma \mathrm{fin}}$ for every $\gamma \in I$ that means Z is relatively strict-maximal and thus Z is an upper bound of the chain. So, by Zorn's Lemma, every relatively strict-maximal code is included in a maximal one, but it is easy to see that every maximal element of this class is a strict-maximal code. Theorem is proved.

It has been known that there exists an algorithm to decide whether a given finitary language is a maximal code (in the class of finitary codes) [1]. Below we state a similar result for finitary strict codes. Recall that a finitary language X is said to be *complete* if for any w of A^* : $A^* w A^* \cap X^* \neq \emptyset$. A word w of A^* is called *overlapping* if $w = ux = yu$ for some $u, x, y \in A^+$. It is easy to see that for every w of A^* (Card $A \geq 2$) there exist x and y of A^* such that xwy is not overlapping. Consider now in the class of finitary strict codes a maximal one by inclusion. We call such a code *finitary-strict-maximal,* or, following just defined terminology, X is finitary relatively strict-maximal code. We have then:

Proposition 2.2 *Every finitary-strict-maximal code X is complete.*

Proof: Let $w \in A^* - X$. As noted $xwy = w'$ is not overlapping for some x, y in A^*. If $w' \in X^*$, nothing is to be proved. Suppose $w' \notin X^*$, then $X \cup w'$ is not a strict code, it has to exist an infinite equality

$$x_1 x_2 \dots = y_1 y_2 \dots$$

over $X \cup w'$ with $x_1 \neq y_1$. Let $x_k = w', k \geq 1$ and m and n be respectively the largest and the smallest integers such that $x_k = w'$ is a subword of $y_m y_{m+1} \dots y_n$. Since w' is not overlapping, there is no w' among $y_m, \dots, y_n$ meaning that they are all in X. Consequently, w' is a subword of $y_m y_{m+1} \dots y_n \in X^*$ and so is w. This completes the proof.

As a consequence of Proposition 2.2., we have

Theorem 2.3 *There exists an algorithm to decide whether a given finitary finite subset X is a finitary-strict-maximal code.*

Proof: First, by Theorem 2.6. of [6], we can verify whether X is a strict code. Second, if X is finitary-strict-maximal code it must be complete. But it has been known that a finite (strict) code is complete if and only if it is maximal as a code (and therefore finitary-strict-maximal as a strict code). Thus it suffices to test the maximality of X as a code and this could be done, for example, by means of a Bernoulli distribution. Because X is finite the test is always effective. Theorem is proved.

We now return to the general framework of infinitary words and languages. We state some properties of strict-maximal codes analogous to the case of ordinary ones, but let us first define some notions.

A language $X \subseteq A^\infty$ is said to be *dense* if for every α of A^∞, $A^\infty \alpha A^\infty \cap X \neq \emptyset$, or, which amounts to the same, for every w of A^N : $A^* w \cap X_{\inf} \neq \emptyset$; X is said to be *complete* (resp. *weakly complete*) if X^* (resp. X^∞) is dense.

Theorem 2.4 *Every strict-maximal code is weakly complete.*

Proof.: If the alphabet A is a singleton $A = \{a\}$ then a language X is a strict code if and only if $X = \{a^n\}$ for some positive interger n or $X = \{a^\omega\}$. So every strict code is strict-maximal and weakly complete in this case.

Suppose now Card$A \geq 2$ and X is a strict-maximal code. We prove that $A^*\alpha \cap X_{\inf} \neq \emptyset$ for every α in A^N. If $\alpha \in X_{\inf}$, we are done, otherwise $X \cup \alpha$ is not a strict code, so that we have an equality:

$$x_1 x_2 \ldots = y_1 y_2 \ldots \tag{1}$$

with two possibilities:

(i) $(x_1, x_2, \ldots) \in (X \cup \alpha)_{(n)}$ and $(y_1, y_2, \ldots) \in (X \cup \alpha)_{(m)}$ for $m, n \geq 1$.

(ii) $(x_1, x_2, \ldots) \in X_{(\omega)}$ and $(y_1, y_2, \ldots) \in (X \cup \alpha)_{(m)}$.

If (ii) holds, we are done. Now suppose that we have (i). If $x_n \neq \alpha$, we are through again, otherwise, from (1) and from the fact that α must occur among y_i's, we have $y_m = \alpha$. Hence $\alpha = p^\omega$ for some primitive word $p \in A^*$. We choose a letter b different from the last letter of p. Consider the word $bp^\omega = b\alpha$, which we can suppose not to belong to X (in the contrary, we are done), therefore the set $X \cup bp^\omega$ is not a strict code. But now with bp^ω playing the role of α, we have the equality (1) for $X \cup bp^\omega$. The case (i) with $x_n = bp^\omega = b\alpha = y_m$ is already impossible, otherwise $bp^\omega = q^\omega$ for some another primitive word q. Certainly $q = bq', q' \in A*$, hence $p^\omega = (q'b)^\omega$. Since $q'b$ is also primitive, we get $p = q'b$ which contradicts the fact that the last letter of p is not b. So we have now either (i) with $x_n \neq b\alpha, y_m = b\alpha$ (or $x_n = b\alpha, y_m \neq b\alpha$) or (ii). Consequently, $A^*b\alpha \cup X \neq \emptyset$ and the theorem follows.

It is well-known that every recognizable complete finitary code is a maximal one, but we cannot state such an analog for strict codes as shown in the next example.

Example 2.5 *Consider the language $X = \{a^2, bA^\omega\}$ over the binary alphabet $A = \{a, b\}$. It is easy to see that X is a recognizable weakly complete strict code (even more so, complete one), but not a strict-maximal code, since $X \cup abA^\omega$ is still a strict code.*

Given two languages X, Y of A^∞, X is said to be *written* by Y, in notation $X \prec Y$, if $X \subseteq Y^\infty$ and no proper subset Z of Y has this properties i.e. $\forall Z \subset Y, X \not\subseteq Z^\infty$. It is easy to verify that in the class of strict codes $\prec$ is a partial order. A strict code X is said to be *indecomposable* over A if for any strict code $Y, X \prec Y$ implies $Y = X$ or $Y \subseteq A$.

Example 2.6 *(i) Consider the strict code* $X = \{a^2, b, (ba)^\omega\}$ *over the alphabet* $A = \{a, b\}$; X *is not indecomposable because* $X \prec \{a^2, b, (ab)^\omega\}$.

(ii) Consider now the strict code $X = \{b^3a, b^2, a, abab^\omega\}$. *We show that* X *is indecomposable. Indeed, if* X *is written by a strict code* $Y \not\subseteq A$ *then* Y_{fin} *equals either* $\{b^3a, b^2, a\}$ *or* $\{ba, b^2, a\}$, *and* $abab^\omega = y_1y_2$ *for some* $y_1 \in Y^*_{\text{fin}}, y_2 \in Y_{\text{fin}}$ *(the case* $abab^\omega \in Y^\omega_{\text{fin}}$ *is impossible). If* $Y_{\text{fin}} = \{b^3a, b^2, a\}$ *then* y_1 *must be* ϵ, *otherwise* $y_2 = bab^\omega$, *which is impossible because we would have then* $b^3a(b^2)^\omega = b^2y_2$. *Hence* $abab^\omega = y_1y_2 = y_2 \in Y$ *meaning that* $Y = X$.

If now $Y_{\text{fin}} = \{ba, b^2, a\}$, then $y_1 = \epsilon, y_1 = a$ or $y_1 = aba$, correspondingly $y_2 = abab^\omega, y_2 = bab^\omega$, or $y_2 = b^\omega$. In all cases, it is easy to see that Y is not a code, which is a contradiction.

We recall some notations introduced in [6]. A subset M of A^∞ is called *∞-submonoid* if $M^\infty = M$ and it is called freeable if $M^{-1}M \cap MM^{-1} = M$. Every subset X of an ∞-submonoid M such that $X^\infty = M$ is c alled a *generator set* of M. It was proved in [6] that every ∞-submonoid possesses a smallest generator set in the sense that it is contained in every generator set of M and freeable ∞-submonoids are always generated by strict codes which are the smallest generator sets of them. An ∞-submonoid of A^∞ is said to be *freeable-maximal* provided it is not contained in any freeable ∞-submonoid other than itself and A^∞.

We state now a result relating indecomposability off a strict code and freeable-maximality of the ∞-submonoid generated by it.

Theorem 2.7 *An ∞-submonoid* M *is freeable-maximal if and only if it is generated by an indecomposable strict-maximal code.*

Proof.: Let M be freeable-maximal and X its smallest generator set which is a strict code. If X is not strict-maximal, then there exists a word x of $A^\infty - X$ such that $X \cup x$ is a strict code. Hence $M = X^\infty \subset (X \cup x)^\infty$ which implies $(X \cup x)^\infty = A^\infty$. Since $X \cup x$ is a strict code, we have $X \cup x = A$, therefore x belongs to A. On the other hand $X \cup x^\infty$ is also a strict code and we have

$$M \subset (X \cup x^\omega)^\infty \subset A^\infty$$

i.e. M is not freeable-maximal: a contradiction. That X is not indecomposable also leads to a contradiction. Indeed, if $X \prec Y$ and $X \neq Y, Y \not\subseteq A$ then $M \subset Y^\infty \subset A^\infty$ which is a contradiction.

To prove the converse, we suppose that X is an indecomposable strict-maximal code and M' is a freeable ∞-submonoid such that $M \subseteq M' \subseteq A^\infty$. This yields $X \subseteq X^\infty = M \subseteq M' = X'^\infty$ with X' being a strict code generating M'. Thus $X \prec X''$ for some subset $X'' \subseteq X'$. We show that X'' is also a strict-maximal code. If it is not so, then $X'' \cup x$ is a strict code for some x in $A^\infty - X''$. Since X is strict-maximal, $X \cup x (x \notin X$, because $x \notin (X'')^\infty \supseteq X)$ is not a strict code, we have then two different ∞- factorisations $(x_1, x_2, \ldots)$ and $(y_1, y_2, \ldots)$ over $X \cup x$ of some word of A^∞. Since every x_i, y_i that differ from x are in $X \subseteq X''^\infty$ they admit then ∞-factorizations over X''. Now, we replace every entry $x_i \neq x$ and $y_i \neq x$ in $(x_1, x_2, \ldots)$ and $(y_1, y_2, \ldots)$ with their ∞-factorizations over X'' and as a result we obtain two ∞-factorizations over $X'' \cup x$ of the same word. Since $X'' \cup x$ is a strict code, they must be identical, from which it follows that either $x \in X''$ or X is not a strict code. This contradiction shows that X'' must be strict-maximal, therefore $X'' = X'$ and by indecomposability of X we have either $X' = X$ or $X' \subseteq A$. Hence $X' = X$ or $X' = A$, in other words, $M' = M$ or $M' = A^\infty$. The proof is completed.

Example 2.8 *Consider the subset $M = \{w \in A^\infty : |w| \geq p\}$, p is a prime number. It is clear that M is a freeable ∞-submonoid generated by the uniform (strict) code A^p which is strict-maximal and indecomposable. Therefore M is freeable-maximal by Theorem 2.7.*

3 Decomposition

In this section we study the relation $\prec$, namely, we are concerned with the question, does there exist for an arbitrary infinitary strict code X an indecomposable one by which X is written? Such a problem is not to be posed for finitary codes simply because the Zorn's Lemma guarantees this for them. In general, we do not know the answer to this question, but below we state some conditions under which a strict code can be written by an indecomposable one.

Theorem 3.1 *If X is a strict code with X_{fin} a finite finitary maximal code then X can be written by an indecomposable strict code.*

Proof.: First, we observe that for each strict code Y such that $X \prec Y$, we have $X_{\text{fin}} \subseteq Y_{\text{fin}}^*$ and it is not difficult to see that if X_{fin} is finite maximal code so is Y_{fin} and $X_{\text{fin}} \prec Y_{\text{fin}}$ (see [5], for example). Next, we denote $S(X) = \{Y \subseteq A^\infty : Y$ is a strict code : $X \prec Y\}$. Also, we denote $\|Y\| = \sum_{y \in Y_{\text{fin}}} |y|$, so that for each $Y \in S(X)$ we have $\|Y\| < \infty$. Let N be the smallest value of $\|Y\|$ as Y runs through $S(X)$: $N = \min\{\|Y\| : Y \in S(X)\}$. We can see that if $Y_1 \prec Y_2$ with Y_1, Y_2 in $S(X)$ then $\|Y_1\| \geq \|Y_2\|$. Let Y be a strict code of $S(X)$ with $\|Y\| = N$. We show that Y is written by an indecomposable strict code and since $\prec$ is an order relation, so does X.

Consider $S(Y)$. Certainly $S(Y) \subseteq S(X)$. We use Zorn's Lemma to show that $S(Y)$ contains a maximal element which, therefore, is an indecomposable strict code and by which X can be written. For each $Z \in S(Y) \subseteq S(X)$, we have $X \prec Y \prec Z$ and $\|Z\| \leq \|Y\|$. As noted above $Y_{\text{fin}} \prec Z_{\text{fin}}$, and thus both $Z_{\text{fin}}, Y_{\text{fin}}$ are maximal codes and since $\|Y\|$ is of minimum value, it follows that $\|Z\| = \|Y\|$ and $Z_{\text{fin}} = Y_{\text{fin}}$.

Consider first an arbitrary countable chain in $S(Y)$: $Y_1 \prec Y_2 \prec \ldots$. We have $Y_{1\text{fin}} = Y_{2\text{fin}} = \ldots$ For each $s \geq 1$ and $x \in Y_{s\text{inf}}$, x does not belong to $Y_{s+1\text{fin}}^\omega$, since $Y_{s+1\text{fin}} = Y_{s\text{fin}}$ and Y_s is a strict code. From $Y_s \prec Y_{s+1}$ it follows

$$x = x_{\text{fin}}^{(s+1)} x_{\text{inf}}^{(s+1)},$$

where $x_{\text{fin}}^{(s+1)} \in Y_{s+1\text{fin}}^*, x_{\text{inf}}^{(s+1)} \in Y_{s+1\text{inf}}$. By the same argument, we have

$$x_{\text{inf}}^{(s+1)} = x_{\text{fin}}^{(s+2)} x_{\text{inf}}^{(s+2)},$$

where $x_{\text{fin}}^{(s+2)} \in Y_{s+2\text{fin}}^*, x_{\text{inf}}^{(s+2)} \in Y_{s+2\text{inf}}$, and so on. Clearly, there exist only finitely many i such that $x_{\text{fin}}^{(s+1)} \neq \epsilon$, otherwise $x \in Y_{\text{fin}}^\omega$. Thus there must be an integer $n(x,s)$ such that for a.e. $m \geq n(x,s)$

$$x_{\text{inf}}^{(m)} = x_{\text{inf}}^{(m+1)} = x(s).$$

Therefore $x(s) \in Y_{m\text{inf}}$ for a.e. $m \geq n(x,s)$.

Now, consider the set P of A^∞ with $P_{\text{fin}} = Y_{1\text{fin}}, P_{\text{inf}} = \{x(s) : x \in Y_{s\text{inf}}, s = 1, 2, \ldots\}$. We verify that
(i) P is a strict code. In fact, if we have a relation, for example

$$x_1 \ldots x_m \alpha = y_1 \ldots y_n \beta$$

$(x_1, \ldots, x_m, \alpha) \in P_{(m+1)}, (y_1, \ldots, y_n, \beta) \in P_{(n+1)}$, then there exist $s, t \geq 1, x \in Y_{s\text{inf}}, y \in Y_{t\text{inf}}$ such that $\alpha = x(s), \beta = y(t)$. Let $l = \max\{n(x,s), n(y,t)\}$, then the words $x_1, \ldots x_m, \alpha, \beta, y_1, \ldots, y_n$ all are in Y_l and thus the above relation must be an identity. The other cases are treated similarly.
(ii) For all $i : Y_i \prec P$. As shown above, each $x \in Y_i$ is written in the form $x = x'x(i)$, where $x' \in Y^*_{1\text{fin}} = Y^*_{\text{fin}}$, so $x \in P^*_{\text{fin}} P_{\text{inf}}$. Thus $Y_i \subseteq P^\infty$ (moreover, $Y_i \subseteq P^*$). Further, for every $\alpha \in P_{\text{inf}}$, there is a positive integer s such that $x \in Y_{s\text{fin}}$ and $\alpha = x(s)$. If $n(x,s) \leq i$ then

$$\alpha = x(s) = x_{\text{inf}}^{(n(x,s))} = x_{\text{inf}}^{(i)} \in Y_{i\text{inf}}.$$

If $n(x,s) > i$, since $Y_i \prec Y_{n(x,s)}$ it follows $\alpha = x(s) \in Y_{n(x,s)}$ is present in the expression of some element of Y_i as a product of elements of $Y_{n(x,s)}$. Hence $Y_i \prec P$.

Thus we have proved that there exists an upper bound for any countable chain. Let now

$$Y_\alpha \prec Y_\beta \prec \ldots$$

be an uncountable chain (with cardinality at most continuum). We obviously can derive from this chain a countable subchain

$$Y_1 \prec Y_2 \prec \ldots$$

such that for each Y_γ from the uncoutable chain there exists Y_n from the countable subchain satisfying $Y_\gamma \prec Y_n$. For the latter one there exists a maximal element Y and it is easy to see that Y is also a maximal element for the uncountable chain. Now, in virtue of Zorn's Lemma Y is followed by a maximal element, i.e. Y is written by an indecomposable strict code. Proof is completed

In the next proposition we try to weaken the heavy demand of maximality of X_{fin}, but in compensation to this, the finiteness of X is required. We call X an *alphabetical code* if X is a subset of the alphabet A, otherwise we call it *nonalphabetical* one.

Theorem 3.2 *Each finite finitary nonalphabetical strict code is written by an indecomposable nonalphabetical strict code.*

Proof.: We note that for any finite finitary strict codes $X, Y : X \prec Y$ implies $\|X\| \geq \|Y\|$ as mentioned in the proof of the preceding theorem. The equality holds if and only if every word of Y occurs just once in the $*$-factorization of just one word of X, or equivalently, there exists a partition $Y = Y_1 \cup \ldots \cup Y_n (n \geq 1)$ such that for any $i : 1 \leq i \leq n$ there exists a word x (thus uniquely) such that x is a product of the words in Y_i (in some order). Hence if $\|X\| = \|Y\|$, we have $n = \text{Card } X \leq \text{Card } Y$ and in addition to this, if Card $X =$ Card Y then each Y_i is a singleton, which means $X = Y$.

We turn now to the proof. Let X be a finite finitary strict code. If X is indecomposable, we are done. Otherwise, we assume the contrary that X cannot be written by any indecomposable nonalphabetical strict code and as a consequence

of this, we have an infinite chain of finite finitary codes: $X = X_0 \prec X_1 \prec X_2 \prec \ldots$, where $X_i \neq X_{i+1}$ and X_i is nonalphabetical for all $i = 1, 2, \ldots$ Certainly, we have: $\|X_0\| \geq \|X_1\| \geq \|X_2\| \geq \ldots$. Since $\|X_0\| < \infty$ it follows that for some integer N : $\|X_N\| = \|X_{N+1}\| = \ldots$. On the other hand, as we noted above: Card $X_N \leq$ Card $X_{N+1} \leq \ldots$. But for every $i \geq 1$: $\|X_{N+1}\| \geq$ Card $X_{N+i} \min\{|x| : x \in X_{N+i}\} \geq$ Card X_{N+i}. Hence, there must be an integer M such that Card $X_{M+N} =$ Card X_{M+N+1}, therefore $X_{N+M} = X_{N+M+1}$ which is a contradiction with the assumpion that $X_i \neq X_{i+1}$ for all i. The proof is completed.

As a consequence of the preceding theorem, we shall have a decomposition theorem for finite finitary strict codes, but we recall some notations first. For more details one can consult [5] or [1]. Let X, Y be finitary codes over A and $X \prec Y$. Consider an alphabet B of the same cardinality as Y and a bijection $f : B \longrightarrow Y$. Because Y is a code, we can extend f to an isomorphism of B^* and Y^*, which we denote by the same $f, f : B^* \longrightarrow Y^*$. Let $Z = \{f^{-1}(x) : x \in X \subseteq Y^*\}$ and it is not difficult to see that Z is a code over B and it is a strict code if X and Y are strict codes. In this case we write $X = Y \bigotimes_B Z$. Conversely, if $Y \subseteq A^*, Z \subseteq B^*$ are codes (resp. strict codes) then the expression $X = Y \bigotimes_B Z$ stands for the following: there exists an isomorphism $f : B^* \longrightarrow Y^*$ such that $X = f(Z) \subseteq Y^*$. In this way X becomes a code (resp. strict code) and we have $X \prec Y$ if and only if B is the least alphabet such that $Z \subseteq B^*$. It is noteworthy that $\bigotimes$ is associative. Now we state our theorem.

Theorem 3.3 *Every finite finitary strict code X of A^* admits a finite decompostion:*

$$X = X_1 \bigotimes_B X_2 \bigotimes_C \ldots \bigotimes_D X_n,$$

where $X_1, X_2, \ldots X_n$ are indecomposable strict codes over the corresponding alphabets $A, B, \ldots, D$.

Proof.: The proof is proceeded by induction on $\|X\|$. If $\|X\| = 1$ then X is a letter, so it is indecomposable by definition. Suppose now that for every strict code X with $\|X\| < k$ the assertion is valid. Let $\|X\| = k$. If X is indecomposable, we are done; if not, by the preceding theorem, X is written by a nonalphabetical indecomposable strict code Y over A^* : $X \prec Y$. Thus, using the notations mentioned above, we have

$$X = Y \bigotimes_B Z.$$

Certainly, $|f^{-1}(x)| \leq |x|$ for each $x \in X$, therefore $\|Z\| \leq \|X\| = k$ and the equality holds if and only if $|f^{-1}(x)| = |x|$ for every x, i.e. when each word of Y is a letter meaning that $Y \subseteq A$ which is a contradiction. Thus we must have $\|Z\| < \|X\| = k$. By induction hypothesis Y admits a finite decomposition and so does X. Theorem is proved.

4 Theorem of Defect

In this concluding section we establish for strict codes a result, which is an analog of Theorem of Defect in the theory of finitary codes [2]. Note that Theorem of Defect was also proved for infinitary codes [4].

Theorem 4.1 *For any language X of A^∞, if X is not a strict code then X is written by a strict code of cardinality at most Card $X - 1$.*

Proof.: If Card $X = \infty$ nothing is to be done because X is always written by A or a subset of A and Card $A - 1 \leq \infty$. So we can assume Card $X < \infty$ and the proof is done by induction on Card X.

If Card $X = 1$ then X is a singleton strict code. Now suppose that for every X of cardinality not exceeding n the assertion is true. Let now X be a language of cardinality $n + 1$ and X not be a strict code. We have then two different ∞-factorizations $(u_1, u_2, \ldots)$ and $(v_1, v_2, \ldots)$ over X with $u_1 \neq v_1$ of a word $\alpha \in A^\infty$. Further, we can suppose that $|v_1| > |u_1|$. Hence $v_1 = u_1\beta$ for some $\beta \in A^{+\infty}$. Consider two cases:

(i) If $v_1 \in A^N$, then we have $v_1 = u_1u_2\ldots$, therefore $\beta = u_2u_3\ldots$. Consider the language $X_1 = X - v_1$. If v_1 occurs among $u_2, u_3, \ldots$, say, $v_1 = u_k$ with k the smallest possible, $k > 1$; then we have

$$v_1 = u_1u_2\ldots u_{k-1}v_1$$

hence $v_1 = (u_1\ldots u_{k-1})^\omega \in X_1^\infty$.

If $v_1 \neq u_i$ for $i = 2, 3, \ldots$ then v_1 is obviuosly in X_1^∞. So we have $X \subseteq X_1^\infty$ and Card $X_1 <$ Card X. If X_1 is not a strict code, then by the induction hypothesis, X_1 is written by a strict code of cardinality $<$ Card$X_1 <$ Card$X$.

(ii) If $v_1 \in A^*$, then $\beta \in A^+$ and we put $X_1 = X - v_1 \cup \beta$. Clearly, $X \subseteq X_1^\infty$ and $\|X_1\| < \|X\|$. There are two possibilities

(ii.1) Replacing all the occurrences of v_1 by $u_1\beta$ in the equation:

$$u_1u_2\ldots = v_1v_2\ldots, \tag{2}$$

it becomes an identity. This means that $\beta = u_2 \in X - v_1$ and therefore Card$X_1 \leq$ Card$X - 1 = n$. The assertion follows by induction hypothesis applied to X_1.

(ii.2) If (2) does not become an identity after the replacement (ii.1), then we repeat the argument with X_1 until (i) or (ii.1) occurs. The process cannot go into infinity avoiding (i) or (ii.1) since $\|X_i\|, i = 1, 2, \ldots$ decreases strictly each time the argument is repeated. Thus we should obtain a finite sequence $X = X_0, X_1, \ldots X_s$ with $X_i \subset X_{i+1}^\infty$, Card $X_{i+1} \leq$ Card X_i for $i = 1, 2, \ldots, s - 1$ and $X_s \subset C$, where C is some finite strict code with Card $C <$ Card X_s. So $X \subseteq C^\infty$ and Card $C \leq$ Card $X - 1$. The proof is complete.

Corollary 4.2 *Every two-element language is a strict code if and only if it is a code.*

Proof.: If $X = \{\alpha, \beta\}$ is not a code then X is not a strict code. Conversely, if X is not a strict code then by Theorem 4.1 $X \subseteq \{\gamma\}^\infty$ for some $\gamma \in A^\infty$. Since $\alpha \neq \beta, \gamma$ must belong to A^*. Hence X is not a code.

5 Acknowledgement

The authors express their sincere gratitude to the referee for his scrupulous work, especially, for his comments and suggestions, which helped them to improve the paper.

References

[1] Berstel J., Perrin D., "Theory of Codes," Academic Press, New York, 1985.

[2] Berstel J., Perrin D., Perrot J. F., Restivo A., *Sur la théorème du défaut*, Journal of Algebra **60** (1979), 169-180.

[3] Do Long Van, "Contribution to Combinatorics on Words," Thesis, Humboldt University, Berlin 1985.

[4] Do Long Van, *Languages écrits par un code infinitaire. Théorème du défaut*, Acta Cybernetica **7** (1986), 247-257.

[5] Lallement G., "Semigroups and Combinatorial Applications," John Wiley, New York, 1979.

[6] Nguyen Huong Lam, Do Long Van, *On a Class of Infinitary Codes*, Theoretical Informatics and Applications 24 (1990), 441-458.

[7] Staiger L., *On Infinitary Finite Length Codes*, Theoretical Informatics and Applications **20** (1986), 483-494.

(Received January 2, 1990)