# ON WEAKLY PREFIX SUBSEMIGROUPS OF A FREE SEMIGROUP

Renato M. Capocelli

Laboratorio di Cibernetica del CNR - 80072 Arco Felice, Naples, Italy.

ABSTRACT. *A remarkable family of free subsemigroups of a free semigroup, the family of weakly prefix subsemigroups, is considered. An algorithm for obtaining the minimal weakly prefix subsemigroup containing a given finite subset is proposed.*

## INTRODUCTION.

In this paper weakly prefix subsemigroups of a free semigroup and some of their properties are considered. A weakly prefix subsemigroup Z of a free semigroup $X^+$ is a subsemigroup of $X^+$ satisfying the condition: for all $a \in Z, x, y \in X^+$; $ax, xy, yx \in Z$ imply $x, y \in Z$. Weakly prefix subsemigroups and their bases, weakly prefix codes, have been the object of recent investigations (4), (5). They appear to be a remarkable family of free subsemigroups exibiting properties of relevant interest in information transmission. In particular, in the finite case their bases coincide with codes having finite decipherability delay.

The optimal factorization of a subsemigroup into a weakly prefix subsemigroup is also considered. It is shown that the intersection of the family of weakly prefix subsemigroups containing a given subset A of $X^+$ is itself a weakly prefix subsemigroup. It provides the minimal weakly prefix subsemigroup containing A. In the finite case, a procedure for constructing the basis F of such a subsemigroup is proposed. Obviously if A is a weakly prefix code F coincides with A, and viceversa. Therefore the procedure for obtaining F provides a criterion to test whether a finite code is decipherable with finite delay. When $A^+$ is not a weakly prefix subsemigroup of $X^+$, F provides



*D. G. Lainiotis and N. S. Tzannes (eds.), Advances in Communications, 123–129.*

somewhat optimal code, with the finite decipherability property, to deal with A-messages. We remark that the proposed procedure is very simple and based explicitly on certain fundamental properties of codewords and their sequences.

## WEAKLY PREFIX SUBSEMIGROUPS.

Let X be a finite nonempty set and let $X^+$ and $X^*$ be the *free semigroup* and the *free monoid* generated by X, respectively. We call *letters* the elements of X, *words* the elements of $X^+$ and denote by $l(w)$ the *length* of the word $w \in X^+$ and by $X^n$ the n-fold concatenation of X with itself. Given $p, r, w \in X^+$; if $pr = w$ then p is a *prefix* of w and r is a *suffix* of w. Let A be a subset of $X^+$. A is a *uniquely decipherable code,* or simply a *code,* iff $A^+$ is a free subsemigroup of $X^+$ having A as basis. A code A has *finite decipherability delay* iff there exists an integer p such that if $x \in X^+$, $l(x) \geq p$, and $xy \in A^+$, then x has a decomposition $x = x_1 x_2$ such that whenever $xz \in A^+$, $x_1 \in A^+$ and $x_2 z \in A^*$.

Given a subset A of $X^+$, a sequence of words $w_1, w_2, \ldots, w_n$ ($w_i \in A$) is called an *L(linked) sequence* of prefix $s_o$ and suffix $s_n$ (3) if

$$
\begin{array}{lll}
w_1 = s_o \; s_1 & & \\
w_2 = s_1 \; s_2 & \text{or} & s_1 = w_2 \, s_2 \\
w_3 = s_2 \; s_3 & \text{or} & s_2 = w_3 \, s_3 \\
\cdots\cdots\cdots & & \\
w_n = s_{n-1} s_n & \text{or} & s_{n-1} = w_n \, s_n, \quad (s_i \in X^+).
\end{array}
\qquad [1]
$$

Given an L-sequence, let us denote by $a_k$ the kth prefix of [ 1 ] (i.e., if $w_{k+1} = s_k s_{k+1}$, then $a_k = s_k$; if $s_k = w_{k+1} s_{k+1}$, then $a_k = w_{k+1}$) and let us consider the sequence

$$s_o \; a_1 \; a_2 \; \ldots a_k \; \ldots a_{n-1} \; s_n \qquad [2]$$

obtained by concatenating to $s_o$ the prefixes $a_k$ of [ 1] and $s_n$. It can be immediately seen that, in general, either

$$s_o a_1 \ldots a_{n-1} \in A^+ \quad \text{and} \quad a_1 a_2 \ldots a_{n-1} s_n \in A^+, \text{ or}$$

$$s_o a_1 \ldots a_{n-1} s_n \in A^+ \text{ and } \quad a_1 \ldots a_{n-1} \in A^+ \quad .$$

In particular, if $s_o$, $s_n \in A^+$ the above provide two different factorizations of [ 2 ].