REMARKS ON SEMAPHORE CODES *

István Babcsányi

Department of Algebra, Mathematical Institute,
Budapest University of Technology and Economics,
1111 Budapest, Egry József u. 1, HUNGARY
E-mail: babcs@math.bme.hu

Abstract

In [4] M. Satyanarayana raised the following two problems:
1. Find necessary and sufficient conditions when a right complete code L over an alphabet A satisfies one of the conditions $A^*L \subseteq LA^*$, $A^+L \subseteq LA^+$.
2. Prove or disprove that if L is a code over A satisfying the condition $A^+L \subseteq LA^+$ then L is a full uniform code over A.
In this paper we solve these two problems by proving the following more general results. A code $L \neq \emptyset$ over A is a semaphore prefix [suffix] code over A if and only if $A^*L \subseteq LA^*$ [$LA^* \subseteq A^*L$]. A code $L \neq \emptyset$ over A is a full uniform code over A if and only if $A^+L \subseteq LA^+$ [$LA^+ \subseteq A^+L$].

Let A^* and A^+ be the free monoid and the free semigroup over the alphabet A, respectively. In this paper, we assume that a language L is always a language over A and $\emptyset \subset L \subseteq A^+$.

If $L \cap LA^+ = \emptyset$ [$L \cap A^+L = \emptyset$] then the language L is a *prefix* [*suffix*] *code*. L is a *biprefix code* when L is both a prefix and a suffix code. A code [prefix, suffix, biprefix code] L is a *maximal* [*maximal prefix, maximal suffix, maximal biprefix code* if for each code [prefix code, suffix cod , biprefix code] L', $L \subseteq L'$ implies $L = L'$. For a language $L \subseteq A^+$, the language $A^*L - A^*LA^+$ [$LA^* - A^+LA^*$] is a maximal prefix [suffix] code named a

*Research supported by the Hungarian NFSR grant No 67639.
Mathematical Subject Classification: 94A45, 20M05.
Key words and phrases: Prefix code, semaphore code, right complete code, uniform code.

semaphore prefix [*suffix*] *code*. For every positive integer n, A^n is *the* (*n-length*) *full uniform code*. Clearly, every full uniform code is a semaphore prefix [suffix] code.

The language L is *complete* [*right complete, left complete*] if, for all $p \in A^*$, $L^* \cap A^*pA^* \neq \emptyset$ $[L^* \cap pA^* \neq \emptyset, L^* \cap A^*p \neq \emptyset]$. Every right [left] complete language is complete. The language L is *thin* if there exists $p \in A^+$ such that $L \cap A^*pA^* = \emptyset$.

For a language L, denote $\underline{L}$ and $\overline{L}$ the set of proper prefixes and proper suffixes of L, respectively.

For notation and notions not defined here, we refer to [1] and [4]. The following four lemmas will be used in our examination:

Lemma 1 *(Proposition 3.2 of [3]) For every prefix [suffix] code L, the following conditions are equivalent:*
(i) L is a maximal prefix [suffix] code.
(ii) $A^* = \underline{L} \cup LA^*$ $[A^* = \overline{L} \cup A^*L]$.
(iii) L is a right [left] complete code.
(iv) $A^* = L^*\underline{L}$ $[A^* = \overline{L}L^*]$.

Lemma 2 *(Theorem I5.7 of [1]) A thin code is a maximal code if and only if it is a complete code.*

Lemma 3 *(Theorem II5.2 of [1]) A prefix [suffix] code L is a semaphore code if and only if $A^*L \subseteq LA^*$ $[LA^* \subseteq A^*L]$.*

Lemma 4 *(Proposition II5.6 of [1]) A semaphore code is a thin code.*

Lemma 5 *A language L satisfies the condition $A^*L \subseteq LA^*$ $[LA^* \subseteq A^*L]$ if and only if $L - LA^+$ $[L - A^+L]$ is a semaphore prefix [suffix] code.*

Proof. Let L be a language satisfying condition $A^*L \subseteq LA^*$. If $X = L - LA^+$ then
$$A^*X \subseteq A^*L \subseteq LA^* = XA^*.$$
By Lemma 3, X is a semaphore prefix code.

Conversely, assume that $X = L - LA^+$ is a semaphore prefix code. Then, by Lemma 3, $A^*X \subseteq XA^*$. If $p \in A^*L$ then there are $u \in A^*$, $q \in X = L - LA^+$ and $v \in A^*$ such that $p = uqv$. Thus
$$p \in A^*XA^* \subseteq XA^*A^* = XA^*.$$
But $XA^* = LA^*$, and so $A^*L \subseteq LA^*$. For semaphore suffix codes, the proof is similar. ■

Theorem 1 *A code L is a semaphore prefix [suffix] code if and only if $A^*L \subseteq LA^*$ [$LA^* \subseteq A^*L$].*

Proof. Let L be a code satisfying condition $A^*L \subseteq LA^*$. By Lemma 5, $X = L - LA^+$ is a semaphore prefix code, and so it is a maximal prefix code. From Lemma 1 it follows that X is right complete and so a complete code. By Lemma 4, X is a thin code and so, by Lemma 2, it is a maximal code. But $X \subseteq L$, therefore $L = X$.

Conversely, if L is a semaphore prefix code then, by Lemma 3, $A^*L \subseteq LA^*$. For semaphore suffix codes, the proof is similar. ■

The following corollary gives an answer for the first problem of M. Satyanarayana.

Corollary 1 *For a right [left] complete code L, the following conditions are equivalent:*
(i) $A^*L \subseteq LA^*$ [$LA^* \subseteq A^*L$].
(ii) L is semaphore prefix [suffix] code.
(iii) $\underline{L}L \subseteq L\overline{L}$. [$L\overline{L} \subseteq \underline{L}L$].

Proof. By Theorem 1, (i) and (ii) are equivalent.

If the right complete code L is a semaphore prefix code then, by Theorem 1, $A^*L \subseteq LA^*$. Let $u \in \underline{L}$ and $p \in L$. There exist $q \in L$ and $v \in A^*$ such that $up = qv$. Since L is a prefix code, then $v \in \overline{L}$. Thus $\underline{L}L \subseteq L\overline{L}$. So (ii) implies (iii).

Conversely, assume that $\underline{L}L \subseteq L\overline{L}$ is satisfied for a right complete code L. If $p \in A^*L$ then there are $u \in A^*$ and $q \in L$ such that $p = uq$. Since L is right complete then $pA^* \cap L^* \neq \emptyset$. Thus $u \in LA^*$ or $u \in \underline{L}$. If $u \in LA^*$ then $p \in LA^*L \subseteq LA^*$. If $u \in \underline{L}$ then $p \in \underline{L}L \subseteq L\overline{L} \subseteq LA^*$. From these it follows that $A^*L \subseteq LA^*$, that is, (iii) implies (i). The proof is similarly for suffix codes. ■

Lemma 6 *For a language L, let n be the minimal length of words in L. Then $A^+L \subseteq LA^+$ [$A^+L \subseteq LA^+$] if and only if $A^n \subseteq L$.*

Proof. In the first part of the proof, we use the idea of the proof of Proposition 4.9 in [2]. Assume that the language L satisfies the condition $A^+L \subseteq LA^+$. Consider a word $p \in L$ of minimal length. Assume that $|p| = n$ in L. If $p = a_1a_1 \ldots a_n$ then, for every $a \in A$, we have $ap = qa_n$, where $q \in L$. As $n \leq |q|$, we get $|q| = n$ and so $q = aa_1a_2 \ldots a_{n-1}$. Then $Aa_1a_2 \ldots a_{n-1} \subseteq L$. If $p \in Aa_1a_2 \ldots a_{n-1}$ then, for every $a \in A$, we have $ap = qa_{n-1}$ ($q \in L$). Thus $A^2a_1a_2 \ldots a_{n-2} \subseteq L$. Continuing this procedure, we get (in step n) that $A^n \subseteq L$.

Conversely, if $A^n \subseteq L$ then

$$A^+L = \cup_{k=n+1}^{\infty} A^k = A^n A^+ = LA^+$$

The proof is similar if L satisfies the condition $LA^+ \subseteq A^+L$. ■

The following theorem gives a positive answer for the second problem of M. Satyanarayana. This is a generalization of Proposition 4.9 in [2] from prefix [suffix] codes to arbitrary codes.

Theorem 2 *A code L is a full uniform code if and only if $A^+L \subseteq LA^+$ [$LA^+ \subseteq A^+L$].*

Proof. It is evident that a full uniform codes $L = A^n$ satisfies the condition $A^+L \subseteq LA^+$. Let L be a code with the condition $A^+L \subseteq LA^+$. By Lemma 6, if n is the minimal length of words in L then $A^n \subseteq L$. But the full uniform code A^n is a maximal code. Therefore, $L = A^n$. The proof is similar in that case when L satisfies the condition $LA^+ \subseteq A^+L$. ■

Corollary 2 *For a code L, the following conditions are equivalent:*
(i) L is a semaphore biprefix code.
(ii) $A^+L = LA^+$.
(iii) L is a full uniform code.

Proof. Assume that L is a semaphore biprefix code. By Theorem 1, $A^*L = LA^*$. Since L is a prefix and suffix code, we have $L \cap LA^+ = \emptyset$ and $L \cap A^+L = \emptyset$. These mean that $A^+L = LA^+$, that is, (i) implies (ii). (ii) implies (iii) by Theorem 2. It is evident that (iii) implies (i). ■

Corollary 3 *A prefix [suffix] code L is a full uniform code if and only if $LA^* \subseteq A^*L$ [$A^*L \subseteq LA^*$].* ■

References

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

[2] Ito, M., H. Jürgensen, H.J. Shur and G. Thierrin, *Outfix and infix codes and related classes of languages*, Journal of Computer and System Sciences, 43(1991), 484-508.

[3] Lallement, G., *Semigroups and Combinatorial Applications*, John Wiley and Sons, New York, 1979.

[4] Satyanarayana, M., *Right complete codes*, Semigroup Forum, 64 (2002), 443-452.