Vietnam Journal of Mathematics **39**:4(2011) 425-442

Vietnam Journal of MATHEMATICS

On Codes Defined by Binary Relations Part II: Characterizations and Maximality

Do Long Van[1] **and Kieu Van Hung**[2]

[1] *Institute of Mathematics, 18 Hoang Quoc Viet, 10307 Hanoi, Vietnam*

[2] *The Office of Government, 01 Hoang Hoa Tham, Hanoi, Vietnam*

Received February 25, 2011

Abstract. Superinfix codes (p-superinfix codes, s-superinfix codes), sucypercodes and supercodes have been introduced and considered by the authors in earlier papers. In particular, it has been proved that the embedding problem for these classes of codes has positive solution in both the finite and regular cases. In this paper, characterizations of these codes, especially of the maximal ones, by means of Parikh vectors and their appropriate generalizations are given. Also a procedure to generate all the maximal supercodes on a two-letter alphabet is exhibited.

2000 Mathematics Subject Classification: 94A45, 68Q45.

Key words: Code, Parikh vector, embedding problem, independent set.

1. Introduction

Defining codes by binary relations was initiated by H. J. Shyr and G. Thierrin in the middle of 1970s [10] and then developed by several authors (see [5-8, 14, 15]). It appeared that this is a good method in introducing new classes of codes. The idea of this comes from the notion of independent sets in universal algebra [2].

One of the interesting problems in the theory of codes is that of embedding a code in a given class C of codes into a code maximal in the same class (not necessarily maximal as a code) which preserves some property (usually, the finiteness or the regularity) of the given code. This is called the *embedding problem* for the class C of codes.

Until now the answer for the embedding problem is known only for several cases using different combinatorial techniques. In [11] (see also [12, 13]) it was proposed a general embedding schema for the classes of codes, which can be defined by length-increasing transitive binary relations. This allows to solve positively, in a unified way, the embedding problem for many classes of codes, well-known as well as new (see [3, 4, 11-13]).

In this paper, we consider in detail several among the new classes of codes mentioned above, namely those of p-superinfix codes, s-superinfix codes, superinfix codes, sucypercodes and supercodes, which can all be defined by length-increasing transitive binary relations. Characterizations of these codes, especially of the maximal ones, by means of Parikh vectors and their appropriate generalizations are established. For the case of two-letter alphabets, a procedure to generate all the maximal supercodes and an algorithm to embed a supercode in a maximal one, are proposed.

We now recall some notions and notation, which will be used in the sequel. Let A throughout be a finite alphabet. We denote by A^* the free monoid generated by A whose elements are called *words* on A. The *empty word* is denoted by 1 and $A^+ = A^* - 1$. The number of all occurrences of letters in a word u is the *length* of u, denoted by $|u|$.

A word u is a *prefix* (*suffix*) of a word v if $v = ux$ ($v = xu$, respectively), for some $x \in A^*$. If $x \neq 1$ then u is a *proper prefix* (*proper suffix*, respectively) of v. An *infix* or *factor* of a word v is a word u such that $v = xuy$ for some $x, y \in A^*$; the infix is *proper* if $xy \neq 1$. We say that u is a *subword* of v if $u = u_1 \dots u_n, v = x_0u_1x_1 \dots u_nx_n$ for some $n \geq 1$ and $u_1, \dots, u_n, x_0, \dots, x_n \in A^*$. If $x_0 \dots x_n \neq 1$ then u is called a *proper subword* of v. If u is a subword (proper subword) of v we also say that v is a *superword* (*proper superword*) of u. A word u is a *permutation* of a word v if u and v are commutatively equivalent (see [1]), i.e. if $|u|_a = |v|_a$ for all $a \in A$, where $|u|_a$ denotes the number of occurrences of a in u. We say that u is a *cyclic permutation* of v if u and v are conjugate, i.e. if there exist x, y such that $u = xy$ and $v = yx$.

Any subset of A^* is a *language* over A. A language X is a *code* over A if for all integers $n, m \geq 1$ and for all $x_1, \dots, x_n, y_1, \dots, y_m \in X$, the equality

$$x_1x_2 \dots x_n = y_1y_2 \dots y_m$$

implies $n = m$ and $x_i = y_i$ for $i = 1, \dots, n$. A code X is *maximal* over A if it is not properly contained in another code over A. Let C be a class of codes over A and $X \in C$. The code X is *maximal in C* (not necessarily maximal as a code) if X is not properly contained in another code in C. For further details of the theory of codes we refer to [1, 6, 9].

Given a binary relation $\prec$ on A^*. A subset X in A^* is an *independent set* with respect to the relation $\prec$ if no two elements of X are in this relation. We say that a class C of codes is *defined by* $\prec$ if the codes in this class are exactly the independent sets with respect to $\prec$. Then we denote the class C by $C_\prec$. Very often, the relation $\prec$ characterizes some property α of words. In this case, we write $\prec_\alpha$ instead of $\prec$ and also C_α instead of $C_{\prec_\alpha}$. The relation $\prec$ is said to be *length-increasing* if for any $u, v \in A^*, u \prec v$ implies $|u| < |v|$. We denote by $\preceq$ the reflexive closure of $\prec$, i.e. for any $u, v \in A^*, u \preceq v$ if and only if $u = v$ or $u \prec v$.

It is easy to verify that the following binary relations on A^* are transitive (except for $\prec_b$) and length-increasing.

$u \prec_p v \Leftrightarrow v = ux$ with $x \neq 1$;
$u \prec_s v \Leftrightarrow v = xu$ with $x \neq 1$;
$u \prec_b v \Leftrightarrow (u \prec_p v) \vee (u \prec_s v)$;
$u \prec_{p.i} v \Leftrightarrow v = xuy$ with $y \neq 1$;
$u \prec_{s.i} v \Leftrightarrow v = xuy$ with $x \neq 1$;
$u \prec_i v \Leftrightarrow v = xuy$ with $xy \neq 1$;
$u \prec_{p.h} v \Leftrightarrow \exists n \geq 1 : u = u_1...u_n \wedge v = x_0u_1x_1...u_nx_n$ with $x_1...x_n \neq 1$;
$u \prec_{s.h} v \Leftrightarrow \exists n \geq 1 : u = u_1...u_n \wedge v = x_0u_1x_1...u_nx_n$ with $x_0...x_{n-1} \neq 1$;
$u \prec_h v \Leftrightarrow \exists n \geq 1 : u = u_1...u_n \wedge v = x_0u_1x_1...u_nx_n$ with $x_0...x_n \neq 1$;
$u \prec_{p.si} v \Leftrightarrow \exists w \in A^* : w \prec_p v \wedge u \preceq_h w$;
$u \prec_{s.si} v \Leftrightarrow \exists w \in A^* : w \prec_s v \wedge u \preceq_h w$;
$u \prec_{si} v \Leftrightarrow \exists w \in A^* : w \prec_i v \wedge u \preceq_h w$;
$u \prec_{p.scpi} v \Leftrightarrow (\exists v' : v' \prec_p v)(\exists v'' \in \sigma(v')) : u \preceq_h v''$;
$u \prec_{s.scpi} v \Leftrightarrow (\exists v' : v' \prec_s v)(\exists v'' \in \sigma(v')) : u \preceq_h v''$;
$u \prec_{scpi} v \Leftrightarrow (\exists v' : v' \prec_i v)(\exists v'' \in \sigma(v')) : u \preceq_h v''$;
$u \prec_{p.spi} v \Leftrightarrow (\exists v' : v' \prec_p v)(\exists v'' \in \pi(v')) : u \preceq_h v''$;
$u \prec_{s.spi} v \Leftrightarrow (\exists v' : v' \prec_s v)(\exists v'' \in \pi(v')) : u \preceq_h v''$;
$u \prec_{spi} v \Leftrightarrow (\exists v' : v' \prec_i v)(\exists v'' \in \pi(v')) : u \preceq_h v''$;
$u \prec_{scp} v \Leftrightarrow \exists v' \in \sigma(v) : u \prec_h v'$;
$u \prec_{sp} v \Leftrightarrow \exists v' \in \pi(v) : u \prec_h v'$;

where $\pi(v)$ and $\sigma(v)$ are the sets of all permutations and all cyclic permutations of v respectively. In the sequel, for any $X \subseteq A^*$, we put $\pi(X) = \bigcup_{u \in X} \pi(u)$ and $\sigma(X) = \bigcup_{u \in X} \sigma(u)$.

The relations mentioned above define corresponding classes of codes which are named respectively as the classes C_p of *prefix codes*, C_s of *suffix codes*, C_b of *bifix codes*, $C_{p.i}$ of *p-infix codes*, $C_{s.i}$ of *s-infix codes*, C_i of *infix codes*, $C_{p.h}$ of *p-hypercodes*, $C_{s.h}$ of *s-hypercodes*, C_h of *hypercodes*, $C_{p.si}$ of *p-subinfix codes*, $C_{s.si}$ of *s-subinfix codes*, C_{si} of *subinfix codes*, $C_{p.scpi}$ of *p-sucyperinfix codes*, $C_{s.spci}$ of *s-sucyperinfix codes*, C_{spci} of *sucyperinfix codes*, $C_{p.spi}$ of *p-superinfix codes*, $C_{s.spi}$ of *s-superinfix codes*, C_{spi} of *superinfix codes*, C_{scp} of *sucypercodes* and C_{sp} of *supercodes*.

To facilitate understanding we give now intuitive definitions of the classes of codes introduced above which are the main research subject of this paper. This explains also the way we named these kinds of codes.

A subset $X \subseteq A^+$ is a *superinfix* (*p-superinfix, s-superinfix*) *code*, $X \in C_{spi}$ ($X \in C_{p.spi}, X \in C_{s.spi}$, respectively), if no word in X is a **su**bword of a **per**mutation of a proper **infix** (i.e. factor) (**p**refix, **s**uffix, respectively) of another word in X. A subset X of A^+ is a *supercode* (*sucypercode*), $X \in C_{sp}$ ($X \in C_{scp}$, respectively), if no word in X is a proper **su**bword of a **per**mutation (**cy**clic **per**mutation, respectively) of another word in X. Thus supercodes and sucypercodes are hypercodes. Hence all the supercodes and sucypercodes over a finite alphabet are finite.

2. Characterizations

Let $A = \{a_1, a_2, \ldots, a_k\}$ and $K = \{1, 2, \ldots, k\}$. For every $u \in A^*$, we denote by $p(u)$ the Parikh vector of u, namely

$$p(u) = (|u|_{a_1}, |u|_{a_2}, \ldots, |u|_{a_k}),$$

where $|u|_{a_i}$ denotes the number of occurrences of a_i in u. Thus p is a mapping from A^* into the set V^k of all the k-vectors of non-negative integers. On V^k, one defines the relation $<$ as follows. Let $\xi = (x_1, x_2, \ldots, x_k)$ and $\eta = (y_1, y_2, \ldots, y_k)$ be two elements in V^k. We say that $\xi < \eta$ if $x_i \leq y_i$ for all $i \in K$ and there is at least one $i \in K$ such that $x_i < y_i$. Then, $\xi \leq \eta$ means that $\xi < \eta$ or $\xi = \eta$.

The following fact is useful in the sequel.

Lemma 2.1. *For any $u, v \in A^+$, the following conditions are equivalent:*

(i) *u is a subword* (*a proper subword, respectively*) *of a permutation of v;*

(ii) *v is a superword* (*a proper superword, respectively*) *of a permutation of u;*

(iii) $p(u) \leq p(v)$ ($p(u) < p(v)$, *respectively*).

Proof. (i) ⇔ (iii): Let u be a subword of a permutation of v. By definition, there exists $v' \in \pi(v)$ such that u is a subword of v'. Then we have $p(u) \leq p(v') = p(v)$. Conversely, let $p(u) \leq p(v)$. We shall prove by induction on $|v|$ that u is a subword of a permutation of v. If $|v| = 1$ then $u = v \in A$, the assertion is trivial. Let $|v| = n + 1$ and suppose that the assertion is true for all v' with $|v'| = n$. If $p(u) = p(v)$ then u is a permutation and therefore a subword of a permutation of v. Let now $p(u) < p(v)$. There exists then $a \in A$ such that $|u|_a < |v|_a$. Then there is $v' \in \pi(v)$ such that $v' = v''a$. We have $p(u) \leq p(v'')$. By the induction hypothesis, u is a subword of a permutation of v''. Hence u is a subword of a permutation of v' and therefore of v.

(ii) ⇔ (iii): The argument is similar. ■

For any subset $X \subseteq A^*$ we denote by $p(X)$ the set of all Parikh vectors of the words in X, $p(X) = \{p \in V^k \mid p = p(u) \text{ for some } u \in X\}$.

The following result gives a characterization of supercodes.

Theorem 2.2. *For any subset $X \subseteq A^+$ the following assertions are equivalent:*

(i) *X is a supercode;*

(ii) *$\pi(X)$ is a supercode;*

(iii) *$p(X)$ is an independent set with respect to the relation $<$ on V^k.*

Proof. (i) $\Leftrightarrow$ (iii): By definition, X is a supercode if and only if it is an independent set with respect to the relation $\prec_{sp}$. The latter is equivalent to the fact that $\forall u, v \in X$: $p(u) \not< p(v)$, which in turn is equivalent to the fact that $p(X)$ is an independent set with respect to the relation $<$ on V^k.

(iii) $\Rightarrow$ (ii): Let $p(X)$ be an independent set with respect to $<$. Since $p(X) = p(\pi(X))$, by the above, $\pi(X)$ is a supercode.

(ii) $\Rightarrow$ (i): Evident. ■

The following fact, proved in [13], allows us to establish a simple characterization of sucypercodes.

Lemma 2.3. *For any $u, v \in A^*$ we have $\exists v' \in \sigma(v) : u \preceq_h v'$ if and only if $\exists u' \in \sigma(u) : u' \preceq_h v$.*

Proposition 2.4. *For any subset X of A^+, X is a sucypercode if and only if so is $\sigma(X)$.*

Proof. The sufficiency is trivial. Let X be a sucypercode. If $\sigma(X)$ were not a sucypercode, there would exist $u, v \in \sigma(X)$ and $v' \in \sigma(v)$ such that $u \prec_h v'$. Then we have $u \in \sigma(x), v \in \sigma(y)$ for some $x, y \in X$. Thus $u \prec_h v'$ with $v' \in \sigma(y)$. By Lemma 2.3, there exists $u' \in \sigma(u)$ such that $u' \prec_h y$. Since $u' \in \sigma(x)$, again by Lemma 2.3, there exists $y' \in \sigma(y)$ such that $x \prec_h y'$. This means that $x \prec_{scp} y$, a contradiction. ■

Now, to every $u \in A^+$ we associate two elements of the Cartesian product $V^k \times K$, denoted by $p_L(u)$ and $p_F(u)$, and one element of $V^k \times K^2$, denoted by $p_{LF}(u)$, which are defined as follows

$$p_L(u) = (p(u), l); \quad p_F(u) = (p(u), f); \quad p_{LF}(u) = (p(u), l, f),$$

where l and f are the indices of the last and the first letter in u, respectively. Thus p_L and p_F are mappings from A^+ into $V^k \times K$, while p_{LF} is a mapping from A^+ into $V^k \times K^2$. These mappings are then extended to languages in a standard way: $p_L(X) = \{p_L(u) \mid u \in X\}$, $p_F(X) = \{p_F(u) \mid u \in X\}$ and $p_{LF}(X) = \{p_{LF}(u) \mid u \in X\}$.

Put $U = \{(\xi, i) \in V^k \times K \mid p_i(\xi) \neq 0\}$ and $W = \{(\xi, i, j) \in V^k \times K^2 \mid p_i(\xi), p_j(\xi) \neq 0\}$. To each of the sets U and W we associate a binary relation, denoted both by $\prec$, which are defined by

$$(\xi, i) \prec (\eta, j) \Leftrightarrow (\xi \leq \eta) \wedge (p_j(\xi) < p_j(\eta)),$$

$$(\xi, m, n) \prec (\eta, i, j) \Leftrightarrow (\xi \leq \eta) \wedge (p_i(\xi) < p_i(\eta) \vee p_j(\xi) < p_j(\eta)),$$

where $p_i(\xi), 1 \leq i \leq k$, denotes the i-th component of ξ. These relations on U and on W, as easily verified, are transitive. Notice that for all language $X \subseteq A^+$, $p_L(X)$ and $p_F(X)$ are subsets of U while $p_{LF}(X)$ is a subset of W.

The following fact is easily verified.

Lemma 2.5. *For any $u, v \in A^+$, we have*

(i) $u \prec_{p.spi} v$ *if and only if* $p_L(u) \prec p_L(v)$*;*
(ii) $u \prec_{s.spi} v$ *if and only if* $p_F(u) \prec p_F(v)$*;*
(iii) $u \prec_{spi} v$ *if and only if* $p_{LF}(u) \prec p_{LF}(v)$*.*

To every subset X of A^+ we associate the sets

$$E_X = \{x \in X \mid \exists y \in X : p(y) < p(x)\} \text{ and } O_X = X - E_X.$$

Clearly, if $E_X = \emptyset$ then X is a supercode.

Let u be a word in A^+; we define the following operations

$$\pi_L(u) = \pi(u')b, \text{ with } u = u'b, b \in A;$$
$$\pi_F(u) = a\pi(u'), \text{ with } u = au', a \in A;$$
$$\pi_{LF}(u) = \begin{cases} a\pi(u')b, & \text{if } |u| \geq 2 \text{ and } u = au'b, \text{ with } a, b \in A; \\ u, & \text{if } u \in A; \end{cases}$$

which are extended to languages in a normal way: $\pi_L(X) = \bigcup_{u \in X} \pi_L(u)$, $\pi_F(X) = \bigcup_{u \in X} \pi_F(u)$ and $\pi_{LF}(X) = \bigcup_{u \in X} \pi_{LF}(u)$.

Example 2.6. Consider the languages $X = \{a^2ba, aba^2, ab^3, ba^3, bab^2, b^2ab, a^2b^2a, a^2b^3, ababa, abab^2, ab^2a^2, ab^2ab, ba^2ba, ba^2b^2, baba^2, babab, b^2a^3, b^2a^2b\}$, $Y = \{a^3, b^2a, bab, ab^2, b^3, aba^2, b^2a^2, a^2ba, baba, a^3b, ba^2b\}$, $Z = \{a^3, a^2ba, aba^2, b^4, a^2b^2a, ababa, ab^2a^2, bab^3, b^2ab^2, b^3ab, a^2b^3a, abab^2a, ab^2aba, ab^3a^2, ba^2b^3, babab^2, bab^2ab, b^2a^2b^2, b^2abab, b^3a^2b\}$ over the alphabet $A = \{a, b\}$. Then we have $U = \{(\xi, j) \in V^2 \times \{1, 2\} \mid p_j(\xi) \neq 0\}$, $W = \{(\xi, i, j) \in V^2 \times \{1, 2\}^2 \mid p_i(\xi), p_j(\xi) \neq 0\}$. It is easy to check that the following holds true:

$$p_L(X) = \{((3,1),1), ((3,2),1), ((2,3),2), ((1,3),2)\};$$
$$p_F(Y) = \{((3,0),1), ((3,1),1), ((2,2),2), ((1,2),1), ((1,2),2), ((0,3),2)\};$$
$$p_{LF}(Z) = \{((3,0),1,1), ((3,1),1,1), ((3,2),1,1), ((3,3),1,1), ((2,4),2,2), ((1,4),2,2), ((0,4),2,2)\};$$
$$O_Y = \{a^3, b^2a, bab, ab^2, b^3\}, Y = \pi(O_Y) \cup \pi_F(E_Y);$$
$$O_Z = \{a^3, b^4\}, Z = \pi(O_Z) \cup \pi_{LF}(E_Z).$$

Lemma 2.7. *Let X be a subset of A^+. If $p_L(X)$ ($p_F(X)$) is an independent set with respect to the relation $\prec$ on U then so is $p_L(\pi(O_X) \cup \pi_L(E_X))$*

($p_F(\pi(O_X) \cup \pi_F(E_X))$, respectively). If $p_{LF}(X)$ is an independent set with respect to the relation $\prec$ on W then so is $p_{LF}(\pi(O_X) \cup \pi_{LF}(E_X))$.

Proof. We treat only the case of $p_L(X)$. The arguments in the other cases are similar. Let $p_L(X)$ be an independent set with respect to $\prec$ on U. If $p_L(\pi(O_X) \cup \pi_L(E_X))$ were not an independent set with respect to $\prec$ on U then there would exist $s, t \in p_L(\pi(O_X) \cup \pi_L(E_X))$ such that $s \prec t$. Since $s, t \in p_L(\pi(O_X) \cup \pi_L(E_X))$, we have $s = p_L(u), t = p_L(v)$ for some $u, v \in \pi(O_X) \cup \pi_L(E_X)$. Because $p_L(u) \prec p_L(v)$, we must have $v \in \pi_L(E_X)$. If $u \in \pi_L(E_X)$ then $p_L(u), p_L(v) \in p_L(\pi_L(E_X)) = p_L(E_X) \subseteq p_L(X)$, a contradiction. If $u \in \pi(O_X)$ then on one hand there exists $u' \in O_X$ such that $p(u') = p(u)$ with $p_L(u') \in p_L(O_X) \subseteq p_L(X)$, and on the other hand $p_L(v) \in p_L(E_X) \subseteq p_L(X)$. From $p_L(u) \prec p_L(v)$ it follows that $p_L(u') \prec p_L(v)$, which contradicts the hypothesis that $p_L(X)$ is an independent set with respect to $\prec$. ■

To end this section, we give characterizations of p-superinfix codes, s-superinfix codes and superinfix codes.

Theorem 2.8. *For any subset X of A^+, the following assertions are equivalent:*

(i) *X is a p-superinfix code (a s-superinfix code, a superinfix code, respectively);*

(ii) *$\pi(O_X) \cup \pi_L(E_X)$ is a p-superinfix code ($\pi(O_X) \cup \pi_F(E_X)$ is a s-superinfix code, $\pi(O_X) \cup \pi_{LF}(E_X)$ is a superinfix code, respectively);*

(iii) *$p_L(X)$ is an independent set with respect to the relation $\prec$ on U ($p_F(X)$ is an independent set with respect to the relation $\prec$ on U, $p_{LF}(X)$ is an independent set with respect to the relation $\prec$ on W, respectively).*

Proof. We treat only the case of p-superinfix codes. For the other cases the argument is similar.

(i) $\Leftrightarrow$ (iii): By definition, X is a p-superinfix code if and only if it is an independent set with respect to $\prec_{p.spi}$. By Lemma 2.5(i), the latter is equivalent to the fact that $p_L(u) \not\prec p_L(v)$ for all $u, v \in X$, or equivalently $p_L(X)$ is an independent set with respect to $\prec$ on U.

(iii) $\Rightarrow$ (ii): Let $p_L(X)$ be an independent set with respect to $\prec$ on U. According to Lemma 2.7, $p_L(\pi(O_X) \cup \pi_L(E_X))$ is also an independent set with respect to $\prec$ on U. Hence, by the above, $\pi(O_X) \cup \pi_L(E_X)$ is a p-superinfix code.

(ii) $\Rightarrow$ (i): It is evident because any subset of a p-superinfix code is also a p-superinfix code. ■

Example 2.9. Consider the languages X, Y and Z in Example 2.6. It is easy to check that $p_L(X)$, $p_F(Y)$ and $p_{LF}(Z)$ are all independent sets with respect to $\prec$. Hence, by Theorem 2.8, X (Y, Z) is a p-superinfix (s-superinfix, superinfix, respectively) code.

3. Maximality

First we formulate a characterization of the maximal supercodes by means of independent sets with respect to the relation $<$ on V^k.

Theorem 3.1. *For any subset X of A^+, X is a maximal supercode if and only if $p(X)$ is a maximal independent set with respect to $<$ on V^k and $\pi(X) = X$.*

Proof. Let X be a maximal supercode. If $\pi(X) \neq X$ then, by Theorem 2.2, $\pi(X)$ is a supercode containing strictly X, a contradiction with the maximality of X. Thus $\pi(X) = X$. Next, we prove that $p(X)$ is a maximal independent set with respect to $<$ on V^k. Indeed, by Theorem 2.2, $p(X)$ is an independent set with respect to $<$. If it is not maximal then $\exists p \notin p(X)$ such that $p(X) \cup \{p\}$ is still an independent set with respect to $<$. Choose u to be any word with $p(u) = p$ (such a word always exists). Then $p(X \cup \{u\}) = p(X) \cup \{p\}$. Again by Theorem 2.2 this implies that $X \cup \{u\}$ is still a supercode, a contradiction with the maximality of the supercode X.

Conversely, let $p(X)$ be a maximal independent set with respect to $<$ on V^k and $\pi(X) = X$. By Theorem 2.2, X is a supercode. Suppose X is not a maximal supercode. There exists a word u not in X and therefore not in $\pi(X)$ such that $X \cup \{u\}$ is still a supercode. Because $u \notin \pi(X)$, $p = p(u)$ is not in $p(X)$. Again by Theorem 2.2, $p(X \cup \{u\}) = p(X) \cup \{p\}$ is still an independent set with respect to $<$, a contradiction. ∎

Next, we characterize the maximal p-superinfix, s-superinfix and superinfix codes by means of independent sets with respect to the relation $\prec$ on U and on W.

Theorem 3.2. *For any subset X of A^+, we have*

(i) *X is a maximal p-superinfix (s-superinfix) code if and only if $p_L(X)$ ($p_F(X)$, respectively) is a maximal independent set with respect to the relation $\prec$ on U and $\pi(O_X) \cup \pi_L(E_X) = X$ ($\pi(O_X) \cup \pi_F(E_X) = X$, respectively);*

(ii) *X is a maximal superinfix code if and only if $p_{LF}(X)$ is a maximal independent set with respect to the relation $\prec$ on W and $\pi(O_X) \cup \pi_{LF}(E_X) = X$.*

Proof. (i): We prove only the case of p-superinfix codes. For the case of s-superinfix codes the argument is similar. Let X be a maximal p-superinfix code. If $\pi(O_X) \cup \pi_L(E_X) \neq X$ then, by Theorem 2.8, $\pi(O_X) \cup \pi_L(E_X)$ would be a p-superinfix code strictly containing X, a contradiction with the maximality of X. Hence $\pi(O_X) \cup \pi_L(E_X) = X$. We next show that $p_L(X)$ is a maximal independent set with respect to the relation $\prec$ on U. Indeed, by Theorem 2.8, $p_L(X)$ is an independent set with respect to $\prec$ on U. If $p_L(X)$ were not maximal then $\exists t \in U - p_L(X)$ such that $p_L(X) \cup \{t\}$ is still an independent set with respect to $\prec$. Let $t = (\xi, j)$. Since $p_j(\xi) \neq 0$, we can choose a word u such that $p(u) = \xi$ and the last letter of u has index j. Thus $p_L(u) = t$. Evidently $u \notin X$. We have

$p_L(X\cup\{u\}) = p_L(X)\cup\{t\}$. Again by Theorem 2.8, $X\cup\{u\}$ is still a p-superinfix code, a contradiction with the maximality of X.

Conversely, let $p_L(X)$ be a maximal independent set with respect to $\prec$ on U and $\pi(O_X)\cup\pi_L(E_X) = X$. By Theorem 2.8, X is a p-superinfix code. Suppose X is not maximal as a p-superinfix code. Then there exists $u \notin X$ such that $X\cup\{u\}$ is still a p-superinfix code. If $p_L(u) \in p_L(X)$ then $p_L(u) = p_L(x)$ for some $x \in X$. This implies $p(u) = p(x)$ and the last letters of u and x are the same. Therefore $u \in \pi_L(x) \subseteq \pi(O_X)\cup\pi_L(E_X) = X$, a contradiction. Thus $t = p_L(u) \notin p_L(X)$. Again by Theorem 2.8, $p_L(X\cup\{u\}) = p_L(X)\cup\{t\}$ is still an independent set with respect to $\prec$, a contradiction with the maximality of $p_L(X)$. Thus X must be maximal as a p-superinfix code.

(ii): Let X be a maximal superinfix code. If $\pi(O_X)\cup\pi_{LF}(E_X) \neq X$ then, by Theorem 2.8, $\pi(O_X)\cup\pi_{LF}(E_X)$ would be a superinfix code strictly containing X, a contradiction. So $\pi(O_X)\cup\pi_{LF}(E_X) = X$. Now we show that $p_{LF}(X)$ is a maximal independent set with respect to the relation $\prec$ on W. By Theorem 2.8, $p_{LF}(X)$ is an independent set with respect to $\prec$ on W. If $p_{LF}(X)$ were not maximal then $\exists t \in W - p_{LF}(X)$ such that $p_{LF}(X)\cup\{t\}$ is still an independent set. Let $t = (\xi, i, j)$. Since $p_i(\xi) \neq 0$ and $p_j(\xi) \neq 0$, we can choose a word u, whose the last and the first letters are a_i and a_j respectively, and such that $p(u) = \xi$. Thus $p_{LF}(u) = t$. Evidently $u \notin X$. We have $p_{LF}(X\cup\{u\}) = p_{LF}(X)\cup\{t\}$. Again by Theorem 2.8, $X\cup\{u\}$ is still a superinfix code, a contradiction with the maximality of X.

Conversely, let $p_{LF}(X)$ be a maximal independent set with respect to $\prec$ on W and $\pi(O_X)\cup\pi_{LF}(E_X) = X$. By Theorem 2.8, X is a superinfix code. Suppose X is not maximal as a superinfix code. Then there exists $u \notin X$ such that $X\cup\{u\}$ is still a superinfix code. If $p_{LF}(u) \in p_{LF}(X)$ then $p_{LF}(u) = p_{LF}(x)$ for some $x \in X$. This implies $p(u) = p(x)$, and that u and x have the same last and first letters. Therefore $u \in \pi_{LF}(x) \subseteq \pi(O_X)\cup\pi_{LF}(E_X) = X$, a contradiction. Thus $t = p_{LF}(u) \notin p_{LF}(X)$. Again by Theorem 2.8, $p_{LF}(X\cup\{u\}) = p_{LF}(X)\cup\{t\}$ is still an independent set with respect to $\prec$ on W, a contradiction with the maximality of $p_{LF}(X)$. Thus X must be maximal as a superinfix code. ■

Example 3.3. Consider the languages Y and Z in Example 2.6. On one hand, we have $Y = \pi(O_Y)\cup\pi_F(E_Y)$ and $Z = \pi(O_Z)\cup\pi_{LF}(E_Z)$. On the other hand, it is easy to see that $p_F(Y)$ ($p_{LF}(Z)$) is a maximal independent set with respect to $\prec$ on U (W, respectively). By Theorem 3.2, it follows that Y is a maximal s-superinfix code and Z is a maximal superinfix code over A.

Recall that a subset X of A^+ is an *infix* (a *p-infix*, a *s-infix*) *code* if no word in X is an infix of a proper infix (prefix, suffix, respectively) of another word in X. The subset X is called a *sucyperinfix* (*p-sucyperinfix, s-sucyperinfix*) *code* if no word in X is a **su**bword of a **cy**clic **per**mutation of a proper **infix** (**p**refix, **s**uffix, respectively) of another word in X.

The following result establishes relationship between maximal p-superinfix (s-superinfix, superinfix) codes and p-infix (s-infix, sucyperinfix, respectively) codes.

Theorem 3.4. *For any subset X of A^+, we have*

(i) *X is a maximal p-superinfix (s-superinfix) code if and only if X is a maximal p-infix (s-infix, respectively) code and $\pi(O_X) \cup \pi_L(E_X) = X$ ($\pi(O_X) \cup \pi_F(E_X) = X$, respectively);*

(ii) *X is a maximal superinfix code if and only if X is a maximal sucyperinfix code and $\pi(O_X) \cup \pi_{LF}(E_X) = X$.*

Proof. (i): We treat only the case of p-superinfix codes. Let X be a maximal p-superinfix code. By Theorem 3.2(i), $\pi(O_X) \cup \pi_L(E_X) = X$. If X is not a maximal p-infix code then there exists a word $y, 1 \neq y \notin X$, such that $Y = X \cup \{y\}$ is still a p-infix code. By Theorem 3.2(i), we have $\pi(O_X) \cup \pi_L(E_X) = X$ and $p_L(X)$ is a maximal independent set with respect to $\prec$ on U. If $p_L(y) \in p_L(X)$ then there is an $x \in X$ such that $p(y) = p(x)$ and the last letters of y and x are the same. Then $y \in \pi_L(x) \subseteq \pi(O_X) \cup \pi_L(E_X) = X$, a contradiction with $y \notin X$. Thus we must have $p_L(y) \notin p_L(X)$ and therefore $p_L(X) \cup \{p_L(y)\}$ is not an independent set with respect to $\prec$ on U, i.e. either $p_L(y) \prec p_L(x)$ or $p_L(x) \prec p_L(y)$, for some $x \in X$. Suppose $p_L(y) \prec p_L(x)$, and let a_j be the last letter of x. Since $p(y) \leq p(x)$ and $p_j(y) < p_j(x)$, there exists $x' \in \pi_L(x) \subseteq \pi(O_X) \cup \pi_L(E_X) = X$ such that x' is of the form $x' = zya_j$ with $z \in A^*$. This is impossible because Y is a p-infix code. Suppose now $p_L(x) \prec p_L(y)$. Without loss of generality we may assume $x \in O_X$. Let a_j be the last letter of y. We have $p(x) \leq p(y)$ and $p_j(x) < p_j(y)$. Therefore there exists $x'' \in \pi(x) \subseteq \pi(O_X) \subseteq X$ such that y has the form $y = zx''a_j$, a contradiction. Thus X must be maximal as a p-infix code that required to prove.

Conversely, let X be a maximal p-infix code with $\pi(O_X) \cup \pi_L(E_X) = X$. We first show that $p_L(X)$ is an independent set with respect to $\prec$ on U. Suppose the contrary that there exist $u, v \in X$ such that $p_L(u) \prec p_L(v)$, and let a_j be the last letter of v. By definition, $p(u) \leq p(v)$ and $p_j(u) < p_j(v)$. Therefore, there is $v' \in \pi_L(v) \subseteq X$ such that $v' = zua_j$, which contradicts the hypothesis that X is a p-infix code. Thus $p_L(X)$ must be an independent set with respect to $\prec$ on U and hence X is a p-superinfix code. The maximality of X as a p-superinfix code is then evident.

(ii): Let X be a maximal superinfix code. By Theorem 3.2(ii), $\pi(O_X) \cup \pi_{LF}(E_X) = X$. Assume that X is not a maximal sucyperinfix code. Then there exists a word $y, 1 \neq y \notin X$, such that $Y = X \cup \{y\}$ is a sucyperinfix code. By Theorem 3.2(ii), $\pi(O_X) \cup \pi_{LF}(E_X) = X$ and $p_{LF}(X)$ is a maximal independent set with respect to $\prec$ on W. If $p_{LF}(y) \in p_{LF}(X)$ then there exists $x \in X$ such that $p(y) = p(x)$, and the first and last letters of y and x are the same. Then $y \in \pi_{LF}(x) \subseteq \pi(O_X) \cup \pi_{LF}(E_X) = X$, a contradiction with $y \notin X$. Thus we must have $p_{LF}(y) \notin p_{LF}(X)$ and therefore $p_{LF}(X) \cup \{p_{LF}(y)\}$ is not an independent set with respect to $\prec$ on W, i.e. either $p_{LF}(y) \prec p_{LF}(x)$ or $p_{LF}(x) \prec p_{LF}(y)$, for some $x \in X$. Suppose $p_{LF}(y) \prec p_{LF}(x)$, and let a_i and a_j be the first letter and last letter of x respectively. Since $p(y) \leq p(x)$ and $p_i(y) < p_i(x)$ or $p_j(y) < p_j(x)$, either there exists $x' \in \pi_F(x)$ such that x' is of the form $x' = a_i yz$ or there is $x'' \in \pi_L(x)$, $x'' = zya_j$, with $z \in A^*$. Assume

$x' = a_i yz$, and let $yz = y_1 y_2$ with a_j being the last letter of y_1. Then the word $w = a_i y_2 y_1 \in \pi_{LF}(x) \subseteq \pi(O_X) \cup \pi_{LF}(E_X) = X$ and therefore $y \prec_{scpi} w$, which contradicts the fact that Y is a sucyperinfix code. Assume now $x'' = zya_j$, and let $zy = y'_1 y'_2$ with a_i being the first letter of y'_2. We have $w' = y'_2 y'_1 a_j \in \pi_{LF}(x) \subseteq X$ and hence $y \prec_{scpi} w'$, a contradiction. Next, suppose $p_{LF}(x) \prec p_{LF}(y)$. Without loss of generality we may assume $x \in O_X$. Let a_i and a_j be the first letter and last letter of y respectively. By definition, $p(x) \leq p(y)$ and $p_i(x) < p_i(y)$ or $p_j(x) < p_j(y)$. Therefore either there exists $u \in \pi(x) \subseteq \pi(O_X) \subseteq X$ or there is $v \in \pi(x) \subseteq X$ such that y has the form either $y = a_i uz$ or $y = z' v a_j$, a contradiction. Thus X must be maximal as a sucyperinfix code that required to prove.

Conversely, let X be a maximal sucyperinfix code with $\pi(O_X) \cup \pi_{LF}(E_X) = X$. We show that $p_{LF}(X)$ is an independent set with respect to $\prec$ on W. Assume the contrary that there exist $u, v \in X$ such that $p_{LF}(u) \prec p_{LF}(v)$, and let a_i and a_j be the first letter and last letter of v respectively. Then we have $p(u) \leq p(v)$ and $p_i(u) < p_i(v)$ or $p_j(u) < p_j(v)$. Therefore either there is $v' \in \pi_F(v)$ such that $v' = a_i uz$ or there exists $v'' \in \pi_L(v)$, $v'' = zua_j$, with $z \in A^*$. Suppose $v' = a_i uz$ and let $uz = u_1 u_2$ with a_j being the last letter of u_1. Then the word $w = a_i u_2 u_1 \in \pi_{LF}(v) \subseteq \pi(O_X) \cup \pi_{LF}(E_X) = X$ and therefore $u \prec_{scpi} w$, which contradicts the hypothesis that X is a sucyperinfix code. Suppose now $v'' = zua_j$ and let $zu = u'_1 u'_2$ with a_i being the first letter of u'_2. We have $w' = u'_2 u'_1 a_j \in \pi_{LF}(v) \subseteq X$ and hence $u \prec_{scpi} w'$, a contradiction. Thus $p_{LF}(X)$ must be an independent set with respect to $\prec$ on W and hence X is a superinfix code. The maximality of X as a superinfix code is then trivial. ■

A subset X in A^+ is a *subinfix* (*p-subinfix, s-subinfix*) *code* if no word in X is a **sub**word of a proper **infix** (**p**refix, **s**uffix, respectively) of another word in X. We have evidently $C_{spi} \subset C_{scpi} \subset C_{si} \subset C_i$ as well as similar inclusions for the corresponding p-classes and s-classes of codes.

As a direct consequence of Theorem 3.4 we obtain

Corollary 3.5. *For any subset X of A^+, X is a maximal p-superinfix (s-superinfix, respectively) code if and only if X is a maximal p-subinfix/p-sucyperifix (s-subinfix/s-sucyperinfix, respectively) code and $\pi(O_X) \cup \pi_L(E_X) = X$ ($\pi(O_X) \cup \pi_F(E_X) = X$, respectively).*

We have moreover

Corollary 3.6. *Every maximal p-superinfix (s-superinfix) code is a maximal code.*

Proof. Recall that a code X is *thin* if there is a word w, which cannot be a factor of any word in X. Any p-infix (s-infix) code X is thin because any word of the form axa with $x \in X, a \in A$ cannot be a factor of any word in X. Every maximal p-infix (s-infix) code is a maximal prefix (suffix, respectively) code [5]. Thus, by Theorem 3.4(i), every maximal p-superinfix (s-superinfix) code is a maximal prefix (suffix, respectively) code which is thin. As well-known, for a thin code X, it is a maximal prefix (suffix) code if and only if it is a maximal code (see [1]). Hence every maximal p-superinfix (s-superinfix) code is a maximal code. ∎

This corollary in combination with Theorems 2.1 and 2.2 in [3] gives us immediately

Corollary 3.7. *Every finite* (*regular*) *p-superinfix* (*s-superinfix*) *code is included in a finite* (*regular, respectively*) *p-superinfix* (*s-superinfix*) *code which is maximal as a code.*

Remark 3.8. While, as seen above, a maximal p-superinfix (s-superinfix) code is always a maximal prefix (suffix, respectively) code, a maximal superinfix code is not necessarily a maximal subinfix code. Indeed, consider the code $X = ab^*a$ over the alphabet $A = \{a, b\}$ which is easily verified to be a maximal superinfix code. But it is not a maximal subinfix code because $X \cup \{bab\}$ is still a subinfix code.

Now we consider some properties of maximal sucypercodes and their relationship with other kinds of codes, namely with supercodes and hypercodes. Recall that a subset X of A^+ is a *hypercode*, $X \in C_h$, if no word in X is a proper subword of another word in it. Note that $C_{sp} \subset C_{scp} \subset C_h$. Supercodes have been first considered in [12].

Theorem 3.9. *For any subset X of A^+ we have the following:*

(i) *X is a maximal supercode if and only if X is a maximal hypercode and $\pi(X) = X$;*

(ii) *X is a maximal sucypercode if and only if X is a maximal hypercode and $\sigma(X) = X$;*

(iii) *X is a maximal supercode if and only if X is a maximal sucypercode and $\pi(X) = \sigma(X)$.*

Proof. (i): Let X be a maximal supercode. We have $\pi(X) = X$ by Theorem 3.1. Suppose that X is not a maximal hypercode. Then there is a word u not in X such that $X \cup \{u\}$ is still a hypercode. By Theorem 3.1, $\pi(X) = X$. Thus $Y = \pi(X) \cup \{u\}$ is a hypercode. If Y is not a supercode then either $p(u) < p(v)$ or $p(v) < p(u)$ for some $v \in \pi(X)$. By Lemma 2.1, u must be a proper subword of a permutation of v or a proper superword of a permutation of v. This means that there exists $v' \in \pi(v)$ such that u is either a proper subword of v' or a proper superword of v'. But v' is in Y too, which contradicts the fact that Y is a hypercode. The set Y and therefore the set $X \cup \{u\}$ must be a supercode, a contradiction. Thus X is a maximal hypercode as was required to prove.

Conversely, let X be a maximal hypercode with $\pi(X) = X$. Being a hypercode, no word in X is a proper subword of another word in X. Moreover, since

$\pi(X) = X$, no word in X can be a proper subword of a permutation of another word in X, i.e. X is a supercode. The maximality of X as a supercode is then evident.

(ii): Let X be a maximal sucypercode. If $\sigma(X) \neq X$ then, by Proposition 2.4, $\sigma(X)$ is a sucypercode strictly containing X, a contradiction with the maximality of X. Next, if X is not a maximal hypercode then there exists $1 \neq y \notin X$ such that $X \cup \{y\}$ is still a hypercode. Hence $Y = \sigma(X) \cup \{y\}$ is a hypercode. We now prove that Y is still a sucypercode. Suppose the contrary that it is not the case. Then either $y \prec_{scp} x$ or $x \prec_{scp} y$ for some $x \in \sigma(X)$. If $y \prec_{scp} x$ then there is $x' \in \sigma(x) \subseteq Y$ such that $y \prec_h x'$, which contradicts the fact that Y is a hypercode. If $x \prec_{scp} y$ then there exists $y' \in \sigma(y)$ such that $x \prec_h y'$. By Lemma 2.3, there exists $x'' \in \sigma(x)$ such that $x'' \prec_h y$, again a contradiction. Thus $Y = X \cup \{y\}$ is a sucypercode, which contradicts the maximality of X as a sucypercode. This contradiction shows that X must be a maximal hypercode.

Conversely, let X be a maximal hypercode with $\sigma(X) = X$. As X is a hypercode, we have $u \not\prec_h v$ for all $u, v \in X$. Since $\sigma(X) = X$ this implies $u \not\prec_{scp} v$ for all $u, v \in X$. Thus X is a sucypercode and hence a maximal sucypercode.

(iii): By (i), X is a maximal supercode if and only if X is a maximal hypercode and $\pi(X) = X$. Since $X \subseteq \sigma(X) \subseteq \pi(X)$ and by (ii), the latter is equivalent to the fact that X is a maximal sucypercode and $\pi(X) = \sigma(X)$. ■

4. Supercodes over two-letter alphabets

Fix a two-letter alphabet $A = \{a, b\}$. On V^2 we introduce the relation $\prec_{2.v}$ defined by

$$u \prec_{2.v} w \Leftrightarrow p_1(u) > p_1(w) \wedge p_2(u) < p_2(w),$$

where $p_i(u)$ denotes the i-th component of u. For simplicity, in this section we write $\prec$ instead of $\prec_{2.v}$.

A finite sequence (may be empty) S: $u_1, u_2, \ldots, u_n$ of elements in V^2 is a *chain* if

$$u_1 \prec u_2 \prec \cdots \prec u_n.$$

The chain S is *full* if

$$\forall i, 1 \leq i \leq n-1, \nexists w : u_i \prec w \prec u_{i+1}.$$

If the full chain S satisfies, moreover, the condition

$$p_2(u_1) = p_1(u_n) = 0,$$

then it is said to be *complete*. A finite subset T of V^2 is *complete* if it can be arranged to become a complete chain. For $1 \leq i < j \leq n$ we denote by $[u_i, u_j]$ the subsequence $u_i, u_{i+1}, \ldots, u_j$ of the sequence S.

Theorem 4.1. *For any finite subset X of A^+, X is a maximal supercode if and only if $p(X)$ is complete and $X = \pi(X)$.*

Proof. Let X be a maximal supercode with $|p(X)| = n$. By Theorem 3.1, $p(X)$ is a maximal independent set with respect to $<$ on V^2 and $X = \pi(X)$. So, for any different u, v in $p(X)$, $p_1(u) \neq p_1(v), p_2(u) \neq p_2(v)$. Arrange $p(X)$ to become a sequence $u_1, u_2, \ldots, u_n$ such that $p_1(u_1) > p_1(u_2) > \cdots > p_1(u_n)$. We must have $p_2(u_1) < p_2(u_2) < \cdots < p_2(u_n)$. That is $u_1 \prec u_2 \prec \cdots \prec u_n$. If $p_2(u_1) \neq 0$ then, choosing u to be any 2-vector with $p_1(u) > p_1(u_1)$ and $p_2(u) = 0$, the set $p(X) \cup \{u\}$ is still an independent set with respect to $<$, a contradiction. Thus $p_2(u_1) = 0$. Similarly we have $p_1(u_n) = 0$. Now if there exists v such that $u_i \prec v \prec u_{i+1}$ for some $i, 1 \leq i \leq n-1$, then $p(X) \cup \{v\}$ is an independent set with respect to $<$, which contradicts again the maximality of $p(X)$. Thus the sequence $u_1, u_2, \ldots, u_n$ is a complete chain and, therefore, the set $p(X)$ is complete.

Conversely, since, as it is easily verified, every complete set is a maximal independent set with respect to $<$, and $X = \pi(X)$, again by Theorem 3.1, we have X is a maximal supercode. ■

Example 4.2. For any $n \geq 1$ the sequence

$$(n, 0), (n-1, 2), \ldots, (n-i, 2i), \ldots, (0, 2n)$$

is obviously a complete chain. Therefore the set $V_n = \{(n, 0), (n-1, 2), \ldots, (0, 2n)\}$ is complete. With $n = 3$ for example, $V_3 = \{(3, 0), (2, 2), (1, 4), (0, 6)\}$. By Theorem 4.1 it follows that the set $X = \pi(\{a^3, a^2b^2, ab^4, b^6\}) = \{a^3,\ a^2b^2,\ abab,\ ab^2a,\ ba^2b,\ baba,\ b^2a^2,\ ab^4,\ bab^3,\ b^2ab^2,\ b^3ab,\ b^4a,\ b^6\}$ is a maximal supercode.

By Theorem 4.1, in order to characterize the maximal supercodes over $A = \{a, b\}$ we may characterize the complete sets instead. For this we first consider some transformations on complete chains. Let S: $u_1, u_2, \ldots, u_n$ be a complete chain.

(T1) (*extension*). It consists in doing consecutively the following:

- Add on the left of S a 2-vector u with $p_1(u) > p_1(u_1)$;
- Delete from S all the u_is with $p_2(u_i) \leq p_2(u)$;
- If u_{i_0} is the first among the u_is remained, then insert between u and u_{i_0} any chain such that $[u, u_{i_0}]$ is a full chain;
- If there is no such a u_{i_0}, then add on the right of u any chain ending with a v, $p_1(v) = 0$, and such that $[u, v]$ is a full chain;
- Add on the left of u any chain beginning with a v, $p_2(v) = 0$, and such that $[v, u]$ is a full chain.

(T2) (*replacement*). The following steps will be done successively:

- Replacing some element u_i in S by an element u with $p_1(u) = p_1(u_i)$;

• If $p_2(u) < p_2(u_i)$, then delete all the u_js on the left of u with $p_2(u_j) \geq p_2(u)$;

• If u_{j_0} is the last among the u_j remained, then insert between u_{j_0} and u any sequence such that $[u_{j_0}, u]$ is a full chain;

• If there is no such a u_{j_0}, then add on the left of u any chain commencing with a v, $p_2(v) = 0$, and such that $[v, u]$ is a full chain;

• If $i < n$ then insert between u and u_{i+1} any chain such that $[u, u_{i+1}]$ is a full chain;

• If $p_2(u) > p_2(u_i)$, then delete all the u_js on the right of u with $p_2(u_j) \leq p_2(u)$;

• If u_{j_0} is the first among the u_js remained, then insert between u and u_{j_0} any chain such that $[u, u_{j_0}]$ is a full chain;

• If there is no such a u_{j_0}, then add on the right of u any chain ending with a v, $p_1(v) = 0$, and such that $[u, v]$ is a full chain;

• If $i > 1$ then insert between u_{i-1} and u any chain such that $[u_{i-1}, u]$ is a full chain;

• If $i = 1$ then add on the left of u any chain beginning with a v, $p_2(v) = 0$, and such that $[v, u]$ is a full chain.

(T3) (*insertion*). This consists of the following successive steps:

• For some i, insert in the middle of u_i and u_{i+1}, $1 \leq i \leq n-1$, an element u with $p_1(u_i) > p_1(u) > p_1(u_{i+1})$;

• If $p_2(u) \leq p_2(u_i)$, then delete all the u_js on the left of u with $p_2(u_j) \geq p_2(u)$;

• If u_{j_0} is the last among the u_js remained, then insert between u_{j_0} and u any chain such that $[u_{j_0}, u]$ is a full chain;

• If there is no such a u_{j_0}, then add on the left of u any chain commencing with a v, $p_2(v) = 0$, and such that $[v, u]$ is a full chain;

• Insert between u and u_{i+1} any chain such that $[u, u_{i+1}]$ is a full chain;

• If $p_2(u) \geq p_2(u_{i+1})$, then delete all the u_js on the right of u with $p_2(u_j) \leq p_2(u)$;

• If u_{j_0} is the first among the u_js remained, then insert between u and u_{j_0} any chain such that $[u, u_{j_0}]$ is a full chain;

• If there is no such u_{j_0}, then add on the right of u any sequence ending with a v, $p_1(v) = 0$, and such that $[u, v]$ is a full chain;

• Insert between u_i and u any chain such that $[u_i, u]$ becomes a full chain.

Theorem 4.3. *The following assertions hold true:*

(i) *The transformations* $(T1)$-$(T3)$ *preserve the completeness of a chain;*

(ii) *Any complete chain can be obtained from another one by a finite number of applications of the transformations* $(T1)$-$(T3)$*;*

(iii) *Every chain* S *can be embedded in a complete chain by a finite number of* *applications of the transformations* $(T1)$-$(T3)$.

Proof. (i): Easily seen by the definitions of (T1)-(T3).

(ii): Let $S : u_1, u_2, \ldots, u_n$ and $S' : v_1, v_2, \ldots, v_m$ be two complete chains. To obtain S' from S we can do as follows. According as $p_1(v_1)$ greater, equal or less than $p_1(u_1)$ we apply to S the transformations (T1), (T2) or (T3) in an appropriate way with $u = v_1$. In any case we obtain a complete chain $S^{(1)}$ commencing with v_1. Suppose $S^{(k)}$, $1 \le k \le m-1$, have been constructed, which is a complete chain commencing with $v_1, \ldots, v_k$. Let $S^{(k)} : v_1, \ldots, v_k, w_{k+1}, \ldots, w_r$. We construct $S^{(k+1)}$ as follows. If $p_1(v_{k+1}) > p_1(w_{k+1})$ then, since $p_1(v_{k+1}) < p_1(v_k)$, we may apply (T3) to insert v_{k+1} in the middle of v_k and w_{k+1}. Because S' is complete, in the chain obtained, v_{k+1} must be next to v_k. If $p_1(v_{k+1}) = p_1(w_{k+1})$, then we may apply (T2) to replace w_{k+1} by v_{k+1}. Again by the completeness of S', in the chain obtained, v_{k+1} must be next to v_k. Let now $p_1(v_{k+1}) < p_1(w_{k+1})$. There exists then an integer $t \ge 1$ such that $p_1(w_{k+t+1}) \le p_1(v_{k+1}) < p_1(w_{k+t})$. If $p_2(v_{k+1}) > p_2(w_{k+1})$ then it follows that $v_k < w_{k+1} < v_{k+1}$, a contradiction with the completeness of S'. So we have $p_2(v_{k+1}) \le p_2(w_{k+1})$. According as $p_1(w_{k+t+1}) = p_1(v_{k+1})$ or $p_1(w_{k+t+1}) < p_1(v_{k+1})$, we may apply (T2) or (T3) to replace w_{k+t+1} by v_{k+1}, or to insert v_{k+1} in the middle of w_{k+t} and w_{k+t+1}. Because $p_2(v_{k+1}) \le p_2(w_{k+1})$, w_{k+1} will be deleted and in the chain obtained, v_{k+1} must be next to v_k. Thus, in any case, the chain obtained is complete and commences with $v_1, v_2, \ldots, v_{k+1}$. We take this chain to be $S^{(k+1)}$. As $p_1(v_m) = 0$, $S^{(m)}$ must coincide with S'.

(iii): Given a chain $S : v_1, v_2, \ldots, v_n$. Choose S' to be any complete chain. Similarly as above, we may apply to S' appropriate transformations (T1)-(T3), to "enter" $v_1, v_2, \ldots, v_n$ consecutively. Notice that entering v_{i+1}, $i \ge 1$, does not delete any of $v_1, \ldots, v_i$ which have been entered previously. ■

Example 4.4. Consider the chain $S : (5,2), (3,4), (1,7)$. We try to embed S in a complete chain by using (T1)-(T3). For this, we choose an arbitrary complete chain S', say $S' : (2,0), (1,2), (0,4)$, and manipulate like this:

- Applying (T1) to S' with $u = (5,2)$ we obtain from step to step the following sequences, where underline indicates the 2-vectors added in every step.

$$\underline{(5,2)}, (2,0), (1,2), (0,4);$$
$$(5,2), (0,4);$$
$$(5,2), \underline{(2,3)}, (0,4);$$
$$\underline{(6,0)}, (5,2), (2,3), (0,4).$$

- Applying (T3) to the last chain with $u = (3,4)$ we obtain successively

$$(6,0), (5,2), \underline{(3,4)}, (2,3), (0,4);$$
$$(6,0), (5,2), (3,4);$$

$$(6,0),(5,2),(3,4),\underline{(1,5)},\underline{(0,6)};$$
$$(6,0),(5,2),\underline{(4,3)},(3,4),(1,5),(0,6).$$

- Applying (T2) to the last chain with $u = (1,7)$ we obtain

$$(6,0),(5,2),(4,3),(3,4),\underline{(1,7)},(0,6);$$
$$(6,0),(5,2),(4,3),(3,4),(1,7);$$
$$(6,0),(5,2),(4,3),(3,4),(1,7),\underline{(0,8)};$$
$$(6,0),(5,2),(4,3),(3,4),\underline{(2,6)},(1,7),(0,8).$$

The last chain is a complete chain containing S.

As a consequence of Theorem 4.3 we have

Theorem 4.5. *Let A be a two-letter alphabet. Then we have*

(i) *There exists a procedure to generate all the maximal supercodes over A starting from an arbitrary given maximal supercode;*

(ii) *There is an algorithm allowing to construct, for every supercode X over A, a maximal supercode Y containing X.*

Proof. (i): Let X be a given maximal supercode. Compute first $p(X)$, which is a complete set. Arrange $p(X)$ to become a complete chain S. By Theorem 4.3(ii), every possible complete chain, hence every complete set, can be obtained from S by a finite number of applications of the transformations (T1)-(T3). The inverse images of all such sets with respect to the morphism p give all the possible maximal supercodes.

(ii): $p(X)$ is an independent set with respect to $<$. So it can be arranged to become a chain S. By Theorem 4.3(iii), we can construct a complete chain S' containing S. Let T be the complete set corresponding to S'. Put $Y = p^{-1}(T)$. Evidently Y contains X and $p(Y) = T$. By Theorem 4.1, Y is a maximal supercode. ■

Example 4.6. Let $X = \{b^2a^2bab, a^3ba^2b, b^4ab^3\}$. Since $p(X) = \{(3,4),(5,2),(1,7)\}$ is an independent set with respect to $<$ on V^2, by Theorem 2.2, X is a supercode over $A = \{a,b\}$. The corresponding chain of $p(X)$ is S : $(5,2),(3,4),(1,7)$. As has been shown in Example 4.4, the sequence

$$S' : (6,0),(5,2),(4,3),(3,4),(2,6),(1,7),(0,8)$$

is a complete chain containing S. The corresponding complete set of S' is

$$T = \{(6,0),(5,2),(4,3),(3,4),(2,6),(1,7),(0,8)\}.$$

So $Y = p^{-1}(T)$ is a maximal supercode containing X. More explicitly, $Y = \pi(Z)$ with $Z = \{a^6, a^5b^2, a^4b^3, a^3b^4, a^2b^6, ab^7, b^8\}$.

Acknowlegements. The authors express their sincere thanks to the referees whose valuable remarks and comments helped them very much in improving the paper.

References

1. J. Berstel and D. Perrin, *Theory of Codes*, Academic Press, New York, 1985.
2. P. M. Cohn, *Universal Algebra*, Harper and Row, 1965 (2nd rev. ed. D. Reidel Publ. Co., Dordrecht, 1981).
3. K. V. Hung, P. T. Huy and D. L. Van, On some classes of codes defined by binary relations, *Acta Math. Vietnam.*, **29** (2004), 163–176.
4. K. V. Hung, P. T. Huy and D. L. Van, Codes concerning roots of words, *Vietnam J. Math.*, **32** (2004), 345–359.
5. M. Ito, H. Jürgensen, H. J. Shyr and G. Thierrin, Outfix and infix codes and related classes of languages, *J. Comput. Syst. Sci.*, **43** (1991), 484–508.
6. H. Jürgensen and S. Konstatinidis, Codes, In: G. Rozenberg, A. Salomaa (eds.), *Handbook of Formal Languages*, Springer, Berlin, 1997, 511–607.
7. H. Jürgensen and S. S. Yu, Solid codes, *J. Inform. Process. Cybernet., EIK*, **26** (1990), 563–574.
8. H. Jürgensen and S. S. Yu, Relations on free monoides, their independent sets, and codes, *Internat. J. Comput. Math.* **40** (1991), 17–46.
9. H. J. Shyr, *Free Monoids and Languages*, Hon Min Book Company, Taichung, 1991.
10. H. J. Shyr and G. Thierrin, Codes and binary relations, *Lecture Notes 586 "Sèminarie d'Algèbre, Paul Dubreil, Paris* (1975 − 1976)*"*, Springer, 180–188.
11. D. L. Van, Embedding problem for codes defined by binary relations, *Preprint 98/A22*, Inst. of Math., Hanoi, 1998.
12. D. L. Van, On a class of hypercodes, In: M. Ito, T. Imaoka (eds.), *Words, Languages and Combinatorics III.* World Scientific, 2003, 171–183.
13. D. L. Van and K. V. Hung, On codes defined by binary relations, Part I: Embedding problem, *Vietnam J. Math.*, **39**(2)(2011), 159–176.
14. S. S. Yu, *Codes and n-ary Relations*, Ph. D. Thesis, The Univ. of Western Ontario, 1990.
15. S. S. Yu, *Languages and Codes*, Tsang Hai Book Publishing Company, Taichung, Taiwan, 2005.