数理解析研究所講究録
第 1712 巻 2010 年 27-38

On constructions of extractable codes

Genjiro Tanaka

Dept. of Computer Science, Shizuoka Institute of Science and Technology,

Fukuroi-shi, 437-8555 Japan.

Abstract. This paper deals with the construction problem on extractable codes. The base of a free submonoid of a free monoid is called a code. The code C with the property that $z, xzy \in C^*$ implies $xy \in C^*$ is called an extractable code. For example this kind of code sometimes appears as a certain type of group code; at other times it appears as Petri net codes of type D. One of the useful methods for constructing extractable codes is a composition of codes. We examine under what conditions on codes Y and Z the composition $Y \circ_\pi Z$ is extractable when Y and Z are composable through some bijection π.

Key words: code, prefix code, extractable code, composition of codes, minimal set of generators, extractable submonoid, free monoid.

1. INTRODUCTION

An extractable submonoid is a free submonoid of a free monoid. It was first mentioned in [5] that the study of extractable submonoids of free monoids was a theme of interest. The extractable code is the base of extractable submonoid. The notion of the extractable code was formally introduced in [6] and [7].

Let A be an alphabet. We denote by A^+ and A^* the free semigroup and the free monoid generated by A, respectively. The empty word is denoted by 1. A word v is a factor of a word $u \in A^*$ if there exist $w, w' \in A^*$ such that $u = wvw'$. A word $v \in A^*$ is a *right factor* (resp. left factor) of a word $u \in A^*$ if there is a word $w \in A^*$ such that $u = wv$ (resp. $u = vw$). If v is a right factor of u, we write $v <_s u$. Similarly, we write $v <_p u$ if v is a left factor of u. The left factor v of a word u is said to be *proper* if $v \neq u$.

We denote by wA^{-1} and wA^- the set of all left factors of w and the set of all proper left factors of w, respectively. Let $X \subset A^*$, and set $XA^- = \cup_{w \in X} wA^-$. Namely, by XA^- we denote the set of all proper left factors of words in X. The subset XA^{-1} of A^* is defined by $XA^{-1} = XA^- \cup X$. We set $ps(X) = XA^- \cap A^- X$.

The length $|w|$ of w is the number of letters in w. $Alph(w)$ is the set of all letters occurring at least once in w.

This is an abstract and the details will be published elsewhere.

Two words x, y are said to be conjugate if there exists words u, v such that $x = uv$, $y = vu$. For $x \in A^*$ we set $Cl(x) = \{y \in A^* | \, y \text{ and } x \text{ are conjugate}\}$.

Let Z is a subset of A^*. For each $x \in A^*$, we define the set of all right contexts of x with respect to Z by

$$Cont_Z^{(r)}(x) = \{w \in A^* \,|\, xw \in Z\}.$$

The right principal congruence $P_Z^{(r)}$ of Z is defined by $(x, y) \in P_Z^{(r)}$ if and only if $Cont_Z^{(r)}(x) = Cont_Z^{(r)}(y)$. Let $u \in A^*$, by $[u]_Z$ we denote the $P_Z^{(r)}$–class of u by $[u]_Z$ or simply by $[u]$. That is,

$$[u]_Z = \{v \,|\, Cont_Z^{(r)}(v) = Cont_Z^{(r)}(u), v \in A^*\}.$$

We denote by $[w_\phi]$ the class of $P_Z^{(r)}$ consisting of all words $w \in A^*$ such that $wA^* \cap Z = \phi$. Namely, $[w_\phi]$ is the class of the nonleft factors of words in Z.

A nonempty subset C of A^+ is said to be a *code* if for $x_1, ..., x_p, y_1, ..., y_q \in C$, $p, q \geq 1$,

$$x_1 \cdots x_p = y_1 \cdots y_q \implies p = q,\, x_1 = y_1,\, \ldots,\, x_p{=}y_p.$$

A code $C \subset A^+$ is said to be *infix* if for all $x, y, z \in A^*$,

$$z,\, xzy \in C \implies x = y = 1.$$

A subset M of A^* is a *submonoid* of A^* if $M^2 \subseteq M$ and $1 \in M$. Every submonoid M of a free monoid has a unique minimal set of generators $C = (M - \{1\}) - (M - \{1\})^2$. C is called the *base* of M. A submonoid M is *right unitary* in A^* if for all $u, v \in A^*$,

$$u,\, uv \in M \implies v \in M.$$

M is called *left unitary* in A^* if it satisfies the dual condition. A submonoid M is *biunitary* if it is both left and right unitary. Let M be a submonoid of a free monoid A^*, and C its base. If $CA^+ \cap C = \emptyset$, (resp. $A^+C \cap C = \emptyset$), then C is called a *prefix* (resp. *suffix*) code over A. C is called a *bifix* code if it is a prefix and suffix code. It is obvious that an infix code is a bifix code. A submonoid M of A^* is right unitary (resp. biunitary) if and only if its minimal set of generators is a prefix code (resp. bifix code) (e.g., [1, p.46], [4, p.108]).

Let C be a nonempty subset of A^*. If $|x| = |y|$ for all $x, y \in C$, then C is a bifix code. We call such a code a *uniform code.* The uniform code $A^n = \{w \in A^* | \, |w| = n\}$, $n \geq 1$, is called a *full uniform code.*

A submonoid M of A^* is *extractable* in A^* if for all $x, y, z \in A^*$,

$$z,\, xzy \in M \implies xy \in M^*.$$

If a submonoid M is extractable, then $u, 1uv \in M$ implies $1v = v \in M$. Similarly $v, uv \in M$ implies $u \in M$. Hence M is biunitary. Therefore, its minimal set of generators C is a bifix code.

Definition 1. Let $C \subset A^*$ be a code. If C^* is extractable in A^*, then C is called an *extractable* code.

For the terms used but not explained in this paper, readers refer to [1] or [4].

Remark 1. Let $C \subset A^*$ be a code. The following conditions are equivalent:
(a) $z, uzv \in C^* \implies uv \in C^*$.
(b) $z \in C$, $uzv \in C^* \implies uv \in C^*$.

Remark 2. Let $C \subset A^*$ be an infix code. The following conditions are equivalent:
(a) $z, uzv \in C^* \implies uv \in C^*$.
(b) $z \in C$, $uzv \in C^2 \implies uv \in C$.

2. COMPOSITION OF CODES RELATED TO EXTRACTABLE CODES

We bigin with the constructions of extractable codes by using the concatenation of codes. Let $Z \subset A^*$ a code and $S \subset A$ a nonempty subset. We set

$$H = Z \cap (\cup_{a \in S} aA^*).$$

It is obvious that $uv \in H$, $uv' \in Z$ implies $uv' \in H$. First we present the following proposition.

Proposition 1. Let $Z \subset A^*$ be an infix extractable code and $S_i \subset A$, $1 \le i \le n$, be nonempty subsets. Let $H_i = Z \cap (\cup_{a \in S_i} aA^*)$, $1 \le i \le n$. Then $X = H_1 H_2 \cdots H_n$ is an extractable code. In particular, Z^n is an extractable code for any $n \ge 1$.

Example 1. (1). Let $Z = \{a^3, ab, ba\}$ and $H = Z \cap aA^* = \{a^3, ab\}$, then $X = ZH = \{a^6, a^4b, aba^3, (ab)^2, ba^4, ba^2b\}$ is extractable.
(2). Let $Z = \{a^3, ba\}$. Then both Z and $Z^2 = \{a^6, a^3ba, ba^4, (ba)^2\}$ are extractable.

Proposition 2. Let $Z \subset A^*$ be an infix code such that for fixed $m \ge 1$ and for all $u \in ZA^-$, $v \in A^*$ and $c_1, \cdots, c_m, d_1, \cdots, d_m \in Z$ the equality

$$ud_1d_2 \cdots d_m = c_1c_2 \cdots c_m v \qquad \text{implies} \qquad u = v.$$

Let K_r, $1 \le r \le mn$, $n \ge 1$, be nonempty subsets of Z. Then $X = K_1K_2 \cdots K_{mn}$ is extractable. In particular, Z^{mn} is an extractable code.

Example 2. (1) The code Z=$\{aba, bab\}$ is not extractable. For $m = 2$, Z satisfies the condition in Proposition 2. Hence Z^{2n} is extractable for any $n \ge 2$. Note, however, that Z^3 is not extractable, since

$$ab \cdot (aba)(bab)(aba) \cdot b(bab)(bab) = (aba)(bab)(aba)(bab)^3 \in Z^6, \quad ab^2(bab)^2 \notin Z^3.$$

Let $Z \subset A^*$ and $Y \subset B^*$ be two codes with $B = Alph(Y)$. Then the codes Y and Z are *composable* through π, if there is a bijection π from B onto Z. The set $X = \pi(Y)$ is denoted by

$$X = Y \circ_\pi Z \quad \text{or} \quad X = Y \circ Z,$$

when no confusion arises. If both Y and Z are prefix (suffix) codes, then $X = Y \circ Z$ is a prefix (suffix) code ([1, p.73, Prop.6.3]). Therefore, if both Y and Z are bifix codes, then X is a bifix code. We note that we can regard Z^n in Proposition 1 and Proposition 2 as the composition $X = B^n \circ_\pi Z$ of B^n and Z through some bijection $\pi : B \to Z$. The composition of codes depends essentially on the bijection π. For example, let $Y = \{aab, aba, baa\}$ and $Z = \{a, ba\}$, and let $\pi_1 : a \to a,\ b \to ba$, $\pi_2 : a \to ba,\ b \to a$. Then $Y \circ_{\pi_1} Z$ is extractable, but $Y \circ_{\pi_2} Z$ is not extractable. Even though both Y and Z are extractable, in general the composition of Y and Z are not necessarily extractable. In the study of extractable codes it is convenient to have a composition of codes Y and Z such that $Y \circ_\pi Z$ is extractable for any bijection $\pi : B \to Z$. Therefore, we examine under what conditions on Y and Z the composition $Y \circ_\pi Z$ can be extractable for an arbitrary bijection π.

Proposition 3. Let $Z \subset A^*$ and $Y \subset B^*$ be two composable codes. If $X = Y \circ_\pi Z$ is extractable, then Y is extractable.

Let Z be a bifix code. We define the *internal multiplicity* $\mu(Z)$ of Z as follows: $\mu(Z) = 0$ if Z is infix, $\mu(Z) = n$ if $Z \cap A^+Z^nA^+ \neq \emptyset$ and $Z \cap A^+Z^{n+m}A^+ = \emptyset$ for all $m \geq 1$, $\mu(Z) = \infty$ if for any $n \geq 1$ there exists $m \geq 1$ such that $Z \cap A^+Z^{n+m}A^+ \neq \emptyset$.

Let Y be a code. Then we set $m(Y) = min\{|y| \mid y \in Y\}$. That is, $m(Y)$ is the shortest length of elements in Y.

Proposition 4. Let $Z \subset A^*$ be a bifix code such that $ps(Z) = \{1\}$, and let $Y \subset A^*$ be an extractable code such that $m(Y) > \mu(Z)$. If Y and Z are composable, then $X = Y \circ Z$ is an extractable code.

A submonoid M of A^* is said to be *pure* if for all $x \in A^*$ and $n \geq 1$ the condition $x^n \in M$ implies $x \in M$. A submonoid N of A^* is *very pure* if for all $u,\ v \in A^*$ the condition $uv, vu \in N$ implies $u, v \in N$.

Corollary 5. Let Y and Z be composable bifix codes such that $m(Y) > \mu(Z)$ and $ps(Z) = \{1\}$. If Y^* is an extractable pure (resp. extractable very pure) monoid, then $X = Y \circ Z$ is an extractable pure (resp. extractable very pure) monoid.

Definition 1 ([8],[9]). Let $n \geq 1$ be an integer. A non-empty subset Z of A^* is called an *intercode of index n* if $Z^{n+1} \cap A^+Z^nA^+ = \emptyset$.

By the definition any nonempty subset of an intercode is also an intercode. An intercode of index n for some $n \geq 1$ is a bifix code. Let $Z \subset A^*$ be an intercode of index n, $n \geq 1$. Then for every m, $m \geq n$, Z is an intercode of index m ([9]).

Proposition 6. Let $Z \subset A^*$ be an intercode of index n and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Example 3 . Let $B = \{a_1, \ldots a_n\}$ be an alphabet and m an integer. For arbitrary $p_1, \ldots p_n \geq m$, the code $Y = \{a_1^{p_1}, \ldots, a_n^{p_n}\}$ is extractable. Let $Z = \{w_1, \ldots, w_n\}$ is an intercode of index m. $\pi : a_i \rightarrow w_i$, $i = 1, \ldots n$, is a bijection. Thus $X{=}Y \circ_\pi Z = \{w_1^{p_1}, \ldots, w_n^{p_n}\}$ is an extractable code.

A code $Z \subset A^*$ is *comma-free* if for all $z \in Z^+$, $u, v \in A^*$, $uzv \in Z^*$ implies $u, v \in Z^*$ ([1, p.336]). It is shown that a code Z is comma-free if and only if Z is an intercode of index 1 ([9]). It is obvious that a comma-free code is extractable.

Corollary 7. Let $Z \subset A^*$ be a comma-free code, and let Y be an extractable code. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Therefore, in particular, if both Y and Z are comma-free, then $Y \circ Z$ is extractable. In fact, it is known that $Y \circ Z$ is comma-free ([1, p.337]).

Definition 3. Let n be an integer. A code $Z \subset A^*$ is a Jn-code if for all c_i, $d_i \in Z$, $1 \leq i \leq n$, and $u \in ZA^-$, $v \in A^*$, the equality

$$ud_1 \cdots d_n = c_1 \cdots c_n v \quad \text{implies} \quad u = v = 1.$$

Remark 3. An infix Jn-code is an intercode of index n.

Definition 3. Let n be an integer. A code $Z \subset A^*$ is an Jn_a-code if for all c_i, $d_i \in Z$, $1 \leq i \leq n$, and $u \in ZA^-$, $v \in A^*$, the equality $ud_1 \cdots d_n = c_1 \cdots c_n v$ implies one of the following conditions:
(1) $u = v = 1$,
(2) $u, v \in A^+$, $d_1 = \cdots = d_n = c_1 = \cdots = c_n$.,
(3) $u, v \in A^+$, $v \in Z^*(ZA^-)$, $c_1 \neq c_2$, $d_1 = \cdots = d_n = c_2 = \cdots = c_n$.

Let Z be a code. If Z is not a prefix code, then there exist some $c, d \in Z$ and $w \in A^+$ such that $c = dw \in Z \cap ZA^+$. Since $1 \cdot (c \cdots c) = (c \cdots d)w$, $1 \notin A^+$, $w \in A^+$. Hence the code Z is not a Jn_a-code. If Z is not a suffix code, then Z is not a $J2_a$-code. Hence Jn_a-code is a bifix code.

Proposition 8. Let $Z \subset A^*$ be an infix Jn_a-code . Let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Let $C \subset A^*$ be a bifix code such that $A^+C^nA^+ \cap C^n = \emptyset$ for some $n \geq 1$. Then $\mu(C) \leq n-1$ and $A^+C^nA^+ \cap C^m = \emptyset$ for any $m \leq n$. However, in general, $A^+C^nA^+ \cap C^n = \emptyset$ does not imply $A^+C^nA^+ \cap C^{n+1} = \emptyset$.

Proposition 9. Let $Z \subset A^*$ be an $J2_a$-code such that $A^+Z^nA^+ \cap Z^n = \emptyset$ for some $n \geq 2$, and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Example 4. The code $Y = \{abc, bca, cab\}$ over $\{a, b, c\}$ is an extractable code such that $m(Y) = 3$. $Z = \{(ab)^2, ba^2b, a^2(ab)^2b^2\}$ ia an $J2_a$-code such taht $Z^2 \cap A^+Z^2A^+ = \emptyset$. Since $(ab)^2$, $a^2(ab)^2b^2 \in Z$ and $a^2b^2 \notin Z^*$, Z is not extractable. By Proposition 9

$$X = Y \circ_\pi Z = \{(ab)^2ba^2ba^2(ab)^2b^2,\ ba^2ba^2(ab)^2b^2(ab)^2,\ a^2(ab)^2b^2(ab)^2ba^2b\}$$

is an extractable code, where $\pi : a \to (ab)^2,\ b \to ba^2b,\ c \to a^2(ab)^2b^2$.

Definition 4. Let n be an integer. A code $Z \subset A^*$ is an $In-$code if for all $c_i, d_i \in Z$, $1 \leq i \leq n$, and $u \in ZA^-$, $v \in A^*$, the equality $ud_1d_2 \cdots d_n = c_1c_2 \cdots c_nv$ implies the one of the following
(1) $u = v = 1$,
(2) $u, v \in A^+$, $v \notin Z^*(ZA^-)$.

Note that any nonempty subset of an In-code is also an In-code. Let Z be an $In-$code. Suppose that $x, xy \in Z^*$, $y \in Z^+$. Then $1 \cdot (xx \cdots xy) = (xx \cdots x) \cdot y$ and $1 \notin A^+$. However, this contradicts our hypothesis that Z is an $In-$ code. Therefore Z must be prefix. Now, suppose that $x, yx \in Z$, $y \in A^+$. Then $y \cdot (xx \cdots x) = ((yx)x \cdots x) \cdot 1$ and $1 \notin A^+$. This is a contradiction. Hence Z is suffix. Thus Z is a bifix code. Therefore, an In-code Z is a bifix code.

Proposition 10. Let $Z \subset A^*$ be an $I2$-code such that $A^+Z^nA^+ \cap Z^n = \emptyset$ for some $n \geq 2$. Let $Y \subset B^*$ be an extractable code with $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Definition 5. Let n be an integer with $n \geq 2$. A code $Z \subset A^*$ is an In_a-code if for all $u \in ZA^-$, $v \in A^*$ and $c_i, d_i, \in Z$, $1 \leq i \leq n$, the equality $ud_1d_2 \cdots d_n = c_1c_2 \cdots c_nv$ implies one of the following conditions:
(1) $u = v = 1$,
(2) $u, v \in A^+$, $v \notin Z^*(ZA^-)$,
(3) $u, v \in A^+$, $d_1 = \cdots = d_n = c_1 = \cdots = c_n$.

Note that any nonempty subset of an In_a-code is also an In_a-code. Let Z be an In_a-code. If $d = cw \in Z \cap ZA^+$, $d, c \in Z$, $w \in A^+$, $1(c \cdots c)d = (c \cdots c)w$ and $1 \in ZA^-$, $w \neq 1$. This contradicts the fact that Z is an In_a-code. If $d = wc \in Z \cap A^+Z$, d, $c \in Z$, $w \in A^+$. This also yields a contradiction. Thus an $I2_a$-code is a bifix code.

Proposition 11. Let $Z \subset A^*$ be an infix In_a-code, and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Corollary 12. Let $A=\{a_1, a_2, \cdots, a_m\}$ and $Z = \{a_1^{p_1}, a_2^{p_2}, \cdots, a_m^{p_m}\}$, where p_i, $1 \leq i \leq m$, are arbitrary positive integers. Let Y be an extractable code such that $m(Y) \geq 2$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Example 5. $Y=\{a^3, ab, ba\}$ is an extractable code. $Z=\{a, b^2\}$ is an infix $I2_a$-code. Define the bijections $\pi_1 : a \to a$, $b \to b^2$, and $\pi_2 : a \to b^2$, $b \to a$. Then we obtain two extractable codes

$$X_1 = Y \circ_{\pi_1} Z = \{a^3, ab^2, b^2a\} \quad \text{and} \quad X_2 = Y \circ_{\pi_2} Z = \{ab^2, b^2a, b^6\}.$$

Proposition 13. Let $Z \subset A^*$ be an $I2_a$-code such that $A^+Z^nA^+ \cap Z^n = \emptyset$ for some $n \geq 2$, and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Definition 6. Let n be an integer with $n \geq 2$. A code $Z \subset A^*$ is an In_b-code if for all $u \in ZA^-$, $v \in A^*$ and c_i, $d_i \in Z, 1 \leq i \leq n$ the equality $ud_1d_2 \cdots d_n = c_1 \cdots c_nv$ implies one of the following conditions:
(1) $u = v = 1$,
(2) u, $v \in A^+$, $v \notin Z^*(ZA^-)$,
(3) u, $v \in A^+$, $v \in Z^*(ZA^-)$, $d_1 = d_2 = \cdots = d_n$.

It is easily shown that an In_b-code is a bifix code.

Proposition 14. Let $Z \subset A^*$ be an infix In_b-code, and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$ and $b^t \notin Y$ for all $b \in B$ and $t \geq 2$. If Y and Z are composable, then $X = Y \circ Z$ is extractable.

Lemma 15. Let $Y \subset B^*$ and $Z \subset A^*$ be composable codes, and $X = Y \circ_\pi Z$.
(1) If both Y and Z are pure, then X is pure.
(2) If both Y^* and Z^* are very pure, then X^* is very pure. ([1, p.328, Proposition1.9])

Corollary 16. Let $Z \subset A^*$ be an infix In_b-code, and let $Y \subset B^*$ be an extractable code such that $m(Y) \geq n$ and $b^t \notin Y$ for all $b \in B$ and $t \geq 2$.
(1) If Y and Z is composable, and if both Y and Z are pure, then $Y \circ Z$ is an extractable pure

code.
(2) If Y and Z is composable, and if both Y^* and Z^* are very pure, then $(Y \circ Z)^*$ is an extractable very pure submonoid of A^*.

Example 6. Let $A = \{a, b\}$. $Y = a^2(aba)^*b \subset A^*$ is an extractable pure code such that $m(Y) = 3$ and $c^p \notin Y$ for all $c \in A$ and $p \geq 1$. $\{ab, ba\}$ is a pure $I2_b$-code. Thus, for $\pi : a \rightarrow ab,\ b \rightarrow ba$,

$$X = Y \circ_\pi Z = (ab)^2(ab^2a^2b)^*ba$$

is an extractable pure code.

Proposition 17. Let A be an alphabet, and let K_i, $1 \leq i \leq n$, be nonempty subsets of A. Then $X = K_1K_2 \cdots K_n$ is an extractable code.

Proposition 18. Let Z be a code, and let H_i, $1 \leq i \leq m$, be nonempty subsets of Z. Furthermore, let $H = H_1H_2 \cdots H_m$.
(1) If Z is an intercode of index n, and if $m \geq n$, then H is an extractable code. In particular, the code Z^m is extractable.
(2) If Z is a $J2_a$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \geq n$, then H is an extractable code.
(3) If Z is an infix Jn_a-code such that $m \geq n$, then H is an extractable code.
(4) If Z is an $I2$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \geq n$, then H is an extractable code.
(5) If Z is an infix In_a-code uch that $m \geq n$, then H is an extractable code.
(6) If Z is an $I2_a$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \geq n$, then H is an extractable code.
(7) If Z is an infix In_b-code such that $m \geq n$, and if $\cap_{i=1}^n H_i = \emptyset$, then H is an extractable code.

Now, we examine the initial literal shuffles of codes related to extractable codes.

Definition 7 ([2]). Let $x, y \in A^*$. Then the *initial literal shuffle* $x \bullet y$ of x and y is defined as follows:
(1) If either $x = 1$ or $y = 1$, then $x \bullet y = xy$.
(2) Let $x = a_1a_2 \cdots a_m$ and let $y = b_1b_2 \cdots b_n$, $a_i, b_j \in A$. Then

$$x \bullet y = \begin{cases} a_1b_1a_2b_2 \cdots a_nb_na_{n+1}a_{n+2} \cdots a_m & \text{if } m \geq n, \\ a_1b_1a_2b_2 \cdots a_mb_mb_{m+1}b_{m+2} \cdots b_n & \text{if } m < n. \end{cases}$$

For two subsets C_1 and C_2 we set $C_1 \bullet C_2 = \{c_1 \bullet c_2 \mid c_1 \in C_1,\ c_2 \in C_2\}$.

For fundamental properties of initial literal shuffles of codes, refer to [3] and [6].

Proposition 19 ([6]). Let $C \subset A^n$. Then C is extractable if and only if $C \bullet C$ is extractable.

Definition 8. Let Z be a code and $x,\ y \in Z^*$. Then the word $x \bullet_Z y$ is defined as follows:
(1) If either $x = 1$ or $y = 1$, then $x \bullet_Z y = xy$.

(2) Let $x = a_1a_2\cdots a_m$, and let $y = b_1b_2\cdots b_n$, a_i, $b_j \in Z$ $(1 \le i \le m,\ 1 \le j \le n)$. Then

$$x \bullet_Z y = \begin{cases} a_1b_1a_2b_2\cdots a_nb_na_{n+1}a_{n+2}\cdots a_m & \text{if } m \ge n, \\ a_1b_1a_2b_2\cdots a_mb_mb_{m+1}b_{m+2}\cdots b_n & \text{if } m < n. \end{cases}$$

For two subsets $C_1 \subset Z^*$ and $C_2 \subset Z^*$ we set $C_1 \bullet_Z C_2 = \{c_1 \bullet_Z c_2 \mid c_1 \in C_1,\ c_2 \in C_2\}$.

Proposition 20. Let $Y \subset B^m$, $m \ge 2$, be an extractable uniform code, and let Z be a code. Assume that Y and Z are composable, and put $X = Y \circ Z$. Then
(1) If Z is an intercode of index n, and if $m \ge n$, then $X \bullet_Z X$ is extractable.
(2) If Z is a $J2_a$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \ge n$, then $X \bullet_Z X$ is extractable.
(3) If Z is an infix Jn_a-code such that $m \ge n$, then $X \bullet_Z X$ is extractable.
(4) If Z is an $I2$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \ge n$, then $X \bullet_Z X$ is extractable.
(5) If Z is an infix In_a-code such that $m \ge n$, then $X \bullet_Z X$ is extractable.
(6) If Z is an $I2_a$-code such that $Z^n \cap A^+Z^nA^+ = \emptyset$, and if $m \ge n$, then $X \bullet_Z X$ is extractable.
(7) If Z is an infix In_b-code such that $m \ge n$ and $b^m \notin Y$ for all $b \in B$, then $X \bullet_Z X$ is extractable.

3 SOME RELATED REMARKS

There are not a few examples in which for a bifix code Z and some suitable integer n the code Z^n becomes an extractable code. However there exists a code Z such that Z^n is not extractable for any $n \ge 1$.

Example 7. A reflective code Z is extractable if and only if the following condition holds: For any $[x], [y] \in A^*/P^{(r)}_{C^*}$

$$Cont^{(r)}_{C^*}(x) \cap Cont^{(r)}_{C^*}(y) \ne \phi \implies [x] = [y].$$

That fact has already been shown in [5, Proposition 8]. Let $Z = Cl((ab)^2a)$, and n be an arbitrary integer. Then $ab \in Cont^{(r)}_Z(aba) \cap Cont^{(r)}_Z(aab)$ and $ba \in Cont^{(r)}_Z((aba) - Cont^{(r)}_Z((aab)$. Therefore Z^n is not extractable for $n = 1$. For $n \ge 2$, we have

$$(ababa)^n \in Z^n,\ aab(ababa)^nba(ababa)^{n-1} = (aabab)(abaab)^{n-1}(ababa)^n \in (Z^n)^2.$$

However, $aabba(ababa)^{n-1} \notin (Z^n)^*$. Therefore Z^n is not an extractable code for any $n \ge 1$.

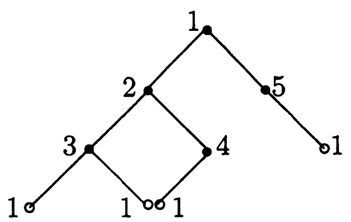
Let $C \subset A^*$ be a code, and let $u, v \in CA^-$. We write $[u] \downarrow [v]$ if $Cont^{(r)}_{C^*}(v)$ is not contained in $Cont^{(r)}_{C^*}(u)$. If $[v]_{C^*} = [w_\emptyset]_{C^*}$, then $Cont^{(r)}_{C^*}(v) = \emptyset$. In this case, $Cont^{(r)}_{C^*}(v)$ is contained in $Cont^{(r)}_{C^*}(u)$ for any $u \in A^*$. Therefore, if $[u] \downarrow [v]$ for some $u \in A^*$, the set $Cont^{(r)}_{C^*}(v)$ is not the emptyset.

Proposition 21. Let Z be an infix code. If there exist $z \in Z$ and $[u],[v] \in A^*/P^{(r)}_{Z^*}$ such that $[uz]=[v]$, $[u] \downarrow [v]$, and $[wzz]=[wz]$ for any $w \in A^*$, then Z^n is not extractable for any $n \geq 1$.

For a prefix code C we defined the automaton $\mathcal{A}(C^*) = (A^*/P^{(r)}_{C^*}, A, \delta, [1], \{[1]\})$, where δ is the transition function such that $\delta([w],x)=[wx]$ for $[w] \in A^*/P^{(r)}_{C^*}$ and $x \in A^*$. This automaton is [0]−transitive (for definition, see [4, p.213]). For each $x \in A^*$ the transformation $t(x)$ on the state set $A^*/P^{(r)}_{C^*}$ is defined by $t(x):[w] \rightarrow [wa]$, $[w] \in A^*/P^{(r)}_{C^*}$. The monoid $T(C^*) = \{t(x) \mid x \in A^*\}$ is called the *transition monoid* of the automaton $\mathcal{A}(C^*)$. $T(C^*)$ is isomorphic to the *syntactic monoid* of C^* (e.g., see [1] or [4]). Let $S_{[1]} = \{t(w) \mid w \in C^*\}$. Then $S_{[1]}$ is the *stablizer* of a state [1] in the automaton $\mathcal{A}(C^*)$. If $t(w) \in S_{[1]}$, and if $t(w)([u]) = [w_\emptyset]$ for all $[u] \in A^*/P^{(r)}_{C^*} - \{[1]\}$, then $t(w)$ is called the *zero-element* of $S_{[1]}$. By 0_1 we denote the zero element of $S_{[1]}$.

Let $T(Z^*)$ be the transition monoid of the automaton $\mathcal{A}(Z^*)$, and let $z \in Z$. The condition that $[wzz]=[wz]$ for all $w \in A^*$ means that the transformation $t(z)$ is an idempotent. Therefore, if there exists an idempotent $t(z)$, $z \in Z$, such that $t(z)([u]) = [v]$ for some $u, v \in ZA^-$ with $[u] \downarrow [v]$, then, by Proposition 19, Z^n is not extractable for all $n \geq 1$.

Example 8. Let $C = \{a^3,\, a^2b,\, aba,\, b^2\}$. Then $A^*/P^{(r)}_{C^*} = \{1, 2, 3, 4, 5, 0\}$, where $1=[1]$, $2=[a]$, $3=[a^2]$, $4=[ab]$, $5=[b]$, $0=[ba]$. The following figure is the tree of C.



$$t(a^3) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 0 \\ 1 & 2 & 3 & 3 & 0 & 0 \end{pmatrix}$$

Fig. 5.

The transformation $t(a^3)$ is an idempotent, and $4 \downarrow 3$ since $Cont^{(r)}_C(a^2) = \{a,\, b\}$ and $Cont^{(r)}_C(ab) = \{a\}$. Thus C^n is not extractable for any $n \geq 1$.

Remark 4. $T(C^*)$ is generated by the set $\{t(a) \mid a \in A\}$. For $z = a_1a_2\cdots a_n \in C$, $a_i \in A$, $1 \leq i \leq n$, we normally gain $t(z)$ by computing the products $t(a_1)t(a_2)\cdots t(a_n)$ of transformations. Without such computation, however, we can obtain $t(z)$ directly by using the tree of C. For instance, in Example 8, from the tree of C (Fig. 5) we have

$1 \xrightarrow{aba} 1$, $0 \xrightarrow{aba} 0$, $2 \xrightarrow{a} 3 \xrightarrow{b} 1 \xrightarrow{a} 2$, $3 \xrightarrow{a} 1 \xrightarrow{b} 5 \xrightarrow{a} 0$, $4 \xrightarrow{a} 1 \xrightarrow{b} 5 \xrightarrow{a} 0$, $5 \xrightarrow{a} 0 \xrightarrow{ba} 0$. Therefore,

$$t(aba) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Lastly we present a characterization of an intercode Z by using transformations in $T(Z^*)$.

Proposition 22. A code $Z \subset A^*$ is an intercode of index n if and only if $t(w) = 0_1$ holds for all $w \in Z^n$.

Corollary 23. Let $Z \subset A^*$ be a code, and let $T(Z^*)$ be the transition monoid of the automaton $\mathcal{A}(Z^*)$. If the subset $t(Z)$ of $T(Z^*)$ contains an idempotent which is neither the identity of $T(Z^*)$ nor the element 0_1 of $T(Z^*)$, then Z is not an intercode.

Corollary 24. Let $Z \subset A^*$ be an infix code. The following conditions are equivalent:
(1) Z is an In-code.
(2) Z is an intercode of index n.
(3) $t(w) = 0_1$ holds for all $w \in Z^n$.

As an elementary consequence of Proposition 22 we have the following assertion:

Example 9. The code Z is comma-free if and only if $t(Z) = \{0_1\}$.

Example 10. We show that $Z = \{a^2babc, acbabc^2, bab, cac\}$ is an intercode of index 4:

The tree of Z

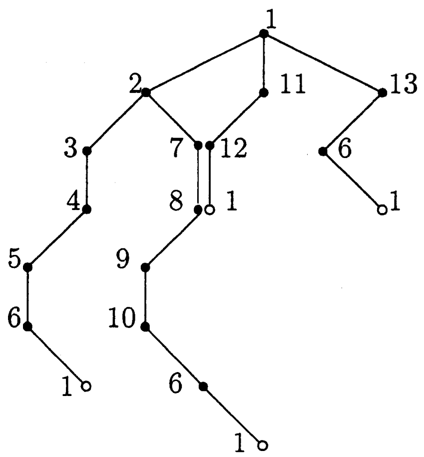


Fig. 6.

$t(a^2babc) = t(acbabc^2) = 0_1$.

$$x = t(bab) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 0 \\ 1 & 0 & 6 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$y = t(cac) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 0 \\ 1 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$xy = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 0 \\ 1 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$yx = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 0 \\ 1 & 0 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus $t(Z^2) = \{0_1, xy, yx\}$. Since

$$x \cdot yx = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 0 \\ 1 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad y \cdot t(Z^2) = \{0_1\},$$

we have $t(Z^3) = \{xyx, 0_1\}$. Since $t(w)(10) = 0$ for all $w \in Z$, we have $t(Z^4) = t(Z^3)t(Z) = \{0_1\}$. Thus Z is an intercode of index 4.

References

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

[2] Berard, B.,Literal shuffle, *Theoret. Comput. Sci.* 51 (1987), pp.281-299.

[3] Ito, M., and Tanaka, G. Dense property of initial literal shuffles, *Intern. J. Computer Math.*, Vol. 34 (1990), pp.161-170.

[4] Lallement, G. *Semigroup and Combinatorial Applications.* Wiley. 1979.

[5] Tanaka, G. Limited codes associated with Petri nets, *Acta Cybernetica*, 19 (2009), pp.217-230.

[6] Tanaka, G., and Kunimochi, Y. Initial literal shuffles of uniform codes, to appear.

[7] Tanaka, G., Kunimochi, Y., and Katsura, M. Remarks on extractable codes, In Kometa. J.(ed.) *Proc. Symposium on Algebras, Languages, Computations and their Applications, RIMS Kokyuroku*, No.1655, pp.106-110, (2009).

[8] Shyr, H, J., and Yu, S, S. Intercodes and Some Related Properties, *Soochow J. Math.*, Vol.20, No.3 (1990), pp.95-107.

[9] Yu, S. S. A Characterization of Intercodes, *Intern. J. Computer Math.*, Vol.36 (1990), pp.39-48.