B-CODES AND THEIR RELATIONSHIP WITH ALTERNATIVE CODES

NGO THI HIEN

Ha Noi University of Science and Technology; hien.ngothi@hust.edu.vn

Crossref
Similarity Check
Powered by iThenticate

Abstract. Codes of bounded words (b-codes, $\diamond$-codes), an extension of the notion of ordinary codes, have been first introduced and considered by P. T. Huy et al. in 2009. In this note, we consider some new subclasses of b-codes. In particular, characteristic properties of such codes are established. Also, relationships between b-codes, alternative codes and their subclasses are considered.

Keywords. Alternative code, b-code, norm b-code, two-side b-code, left b-code, strict b-code. Mathematics Subject Classification (2010): 94A45, 68Q45.

1. INTRODUCTION

The theory of length-variable codes has been initiated by M. P. Schützenberger in the 1950s and then developed by many others. A direction in this theory concentrates studying codes defined by control parameters or related with multi-value [1, 2, 4, 10, 13, 14, 16]. A code is a language such that every text encoded by words of the language can be decoded in a unique way. Codes are useful in many areas of application such as information processing, data compression, cryptography, information transmission and so on. For background of the theory of codes we refer to [3, 11, 15].

As mentioned above, the definition of codes is essentially based on unambiguity of the (catenation) product of words. Different modifications of such a product may lead to different extensions of the notion of codes. Such an approach has been proposed by P. T. Huy et al. which deals with the so-called even alternative codes and codes of bounded words (or $\diamond$-codes, for short) [5, 6, 7, 10, 17, 18, 19, 20]. They demonstrated the importance characterizations as well as algorithms to test for these classes. For simplicity, $\diamond$-codes, weak left $\diamond$-codes, weak right $\diamond$-codes and strict $\diamond$-codes in [17, 18] are called simply respectively b-codes, left b-codes, right b-codes and strict b-codes instead. Recently, in [8, 9, 19, 20], the authors proposed some classes of codes which concern alternative codes and b-codes. Several interesting characteristic properties and algorithms to test for such codes were established.

In this note, some new subclasses of b-codes are introduced and considered. In Section 2 necessary definitions are recalled, and several facts useful in the sequel are shown. Section 3 introduces four new classes of codes which are subclasses of b-codes. Also some characteristic properties and relationships of these classes are considered. The section ends with relations on b-codes, alternative codes and their subclasses. The final section discusses the problem investigations and presents a summary of the results obtained. Our work is motivated by the idea to define b-codes by equivalent factorizations on a b-language, and the way to solve the problem for b-codes as well as their well-known subclasses [6, 17, 18, 19].

2. PRELIMINARIES

Let A throughout be a finite alphabet, i.e. a non-empty finite set of symbols, which are called letters. Let A^* be the set of all finite words over A. The empty word is denoted by ε and A^+ stands for $A^* \setminus \{\varepsilon\}$. The number of all the occurrences of letters in a word u is the *length* of u, denoted by $|u|$. Any subset of A^* is a *language* over A. A language X is a *code* if for any $n, m \geq 1$ and any $x_1, \ldots, x_n, y_1, \ldots, y_m \in X$, the condition

$$x_1 x_2 \ldots x_n = y_1 y_2 \ldots y_m$$

implies $n = m$ and $x_i = y_i$ for $i = 1, \ldots, n$. Since $\varepsilon.\varepsilon = \varepsilon$, a code never contains the empty word ε.

The reversal of a word $w = a_1 a_2 \ldots a_n$, where $a_1, a_2, \ldots, a_n$ are letters, is the word $\widetilde{w} = a_n a_{n-1} \ldots a_1$. Similarly, for $X \subseteq A^*$, we denote $\widetilde{X} = \{\widetilde{x} \mid x \in X\}$.

Let X and Y be non-empty subsets of A^+. We say that $u_1 u_2 \ldots u_n$, $n \geq 2$, is an alternative factorization on (X, Y) if $u_i \in X$ implies $u_{i+1} \in Y$ and $u_i \in Y$ implies $u_{i+1} \in X$ for all $i = 1, 2, \ldots, n-1$. Two alternative factorizations $u_1 u_2 \ldots u_n$ and $v_1 v_2 \ldots v_m$ on (X, Y) are said to be left similar (right similar) if they begin (resp. end) with words in the same set X or Y, namely $u_1, v_1 \in X$ or $u_1, v_1 \in Y$ (resp. $u_n, v_m \in X$ or $u_n, v_m \in Y$); $u_1 u_2 \ldots u_n$ and $v_1 v_2 \ldots v_m$ are called dissimilar if they both begin and end with words in different sets, namely both $u_1 \in X, v_1 \in Y$ or $u_1 \in Y, v_1 \in X$ and $u_n \in X, v_m \in Y$ or $u_n \in Y, v_m \in X$. Two alternative factorizations on (X, Y) are called similar if they are both left similar and right similar. A pair (X, Y) is said to be norm form if no word in A^+ admits two different dissimilar alternative factorizations on (X, Y).

Definition 1. Let X and Y be non-empty subsets of A^+.

(i) (X, Y) is called an alternative code (a left alternative code, a right alternative code) over A if no word in A^+ admits two different similar (resp. left similar, right similar) alternative factorizations on (X, Y);

(ii) (X, Y) is called a two-sided alternative code over A if it is both a left alternative code and a right alternative code;

(iii) (X, Y) is called a norm alternative code (left norm alternative code, right norm alternative code) over A if it is an alternative code (resp. a left alternative code, a right alternative code) which has norm form;

(iv) (X, Y) is a strict alternative code over A if every word in A^+ admits at most one alternative factorization on (X, Y).

The classes of alternative codes, left alternative codes, right alternative codes, two-side alternative codes, left norm alternative codes, right norm alternative codes, norm alternative codes and strict alternative codes are denoted respectively by C_A, C_{lA}, C_{rA}, C_{tA}, C_{lnA}, C_{rnA}, C_{nA} and C_{sA}.

Next, let $B = \{0,1\}$, we define the sets of bounded words (b-words) on A as follows:

$$A_\Diamond = \{(l,a,r) \mid a \in A \cup \{\varepsilon\}, l, r \in B\}, \quad \varepsilon_\Diamond = \{(l,\varepsilon,r) \mid l, r \in B\};$$
$$A_\Diamond^* = \{(l,w,r) \mid w \in A^*, l, r \in B\} \cup \{\theta, e\}, \quad A_\Diamond^+ = A_\Diamond^* \setminus \{\theta, e\};$$
$$A_<^* = \{(0,w,1) \mid w \in A^*\} \cup \{\theta, e\}, \quad A_<^+ = A_<^* \setminus \{\theta, e\};$$
$$A_>^* = \{(1,w,0) \mid w \in A^*\} \cup \{\theta, e\}, \quad A_>^+ = A_>^* \setminus \{\theta, e\};$$
$$A_\wedge^* = \{(0,w,0) \mid w \in A^*\} \cup \{\theta, e\}, \quad A_\wedge^+ = A_\wedge^* \setminus \{\theta, e\};$$
$$A_\vee^* = \{(1,w,1) \mid w \in A^*\} \cup \{\theta, e\}, \quad A_\vee^+ = A_\vee^* \setminus \{\theta, e\}.$$

Each element $(l,w,r), w \in A^*$ is called a b-word (or a bounded word with borders l, r) extended from w. And e, θ are two new elements as the unit, the zero of the monoid $A_\Diamond^*$ of all b-words respectively, where $A_\Diamond^*$ is a monoid (it is easily seen this) by a product defined as follows:

$$z_1 z_2 = \begin{cases} (l_1, w_1 w_2, r_2), & \text{if } r_1 = l_2; \\ \theta, & \text{otherwise}; \end{cases}$$
$$z_1 \theta = \theta z_1 = \theta; \quad z_1 e = e z_1 = z_1.$$

where $z_1, z_2 \in A_\Diamond^*$, $z_1 = (l_1, w_1, r_1)$ and $z_2 = (l_2, w_2, r_2)$.

The reversal of a b-word $z = (l, a_1 a_2 \ldots a_n, r)$, where $a_1, a_2, \ldots, a_n$ are letters, $l, r \in B$, is the word $\widetilde{z} = (r, a_n a_{n-1} \ldots a_1, l)$. This notation is extended as usual to set by, $\widetilde{Z} = \{\widetilde{z} \mid z \in Z\}$.

A subset of $A_\Diamond^*$ is called an extended language (b-language) on A. For each $z = (l, w, r) \in A_\Diamond^*$, we denote by $p_L(z) = l$ the left border of z and $p_R(z) = r$ the right border of z. Whenever none of mistakes are made, we also use notation $|z|$ as the length of z,

$$|z| = \begin{cases} 0, & \text{if } z \in \varepsilon_\Diamond; \\ |w|, & \text{otherwise}. \end{cases}$$

By convention, $|\theta| = +\infty$ and $|e| = 0$.

We define a projection function $\text{Proj} : A_\Diamond^* \to A^* \cup \{0\}$, where $0 \notin A^*$ as the new zero of the monoid $A^* \cup \{0\}$, by

$$\text{Proj}(e) = \varepsilon, \text{Proj}(\theta) = 0 \text{ and } \text{Proj}(l, w, r) = w.$$

Two b-words s and t are said to be strict equivalent (left equivalent, right equivalent, equivalent), denoted by $s \approx_S t$ (resp. $s \approx_L t$, $s \approx_R t$, $s \approx t$), if $\text{Proj}(s) = \text{Proj}(t)$ (resp. $\text{Proj}(s) = \text{Proj}(t)$ and $p_L(s) = p_L(t)$, $\text{Proj}(s) = \text{Proj}(t)$ and $p_R(s) = p_R(t)$, $s \approx_L t$ and $s \approx_R t$). Two factorizations $s = s_1 s_2 \ldots s_n$ and $t = t_1 t_2 \ldots t_m$ on Z are called strict equivalent (left equivalent, right equivalent, equivalent) if $s \approx_S t$ (resp. $s \approx_L t$, $s \approx_R t$, $s \approx_R t$), where $s_i, t_j \in Z, i = 1, \ldots, n, j = 1, \ldots, m$.

Definition 2. A b-language Z is called a b-codes over A if any word in $A_\Diamond^+$ has at most one factorization in words in Z, and Z is a strict b-codes (left b-codes, right b-codes) over A if no word in $A_\Diamond^+$ admits two different strict (resp. left, right) equivalent factorizations on Z.

The classes of b-codes, left b-codes, right b-codes and strict b-codes are denoted respectively by C_B, C_{lB}, C_{rB} and C_{sB}. Over any alphabet consisting of at least two letters, the following holds true

$$C_{sB} \subset C_{lB} \cap C_{rB}, C_{lB} \cup C_{rB} \subset C_B.$$

For more details of alternative codes, b-codes and their subclasses we refer to [5, 6, 7, 8, 9, 10, 17, 18, 19, 20].

Now we formulate, in the form of lemmas, several facts which will be useful in the sequel.

Lemma 1 ([12]; see also [3], page 49). *Let $X = \{x_1, x_2\}$. Then, X is not a code if and only if x_1 and x_2 are powers of the same word, $x_1 = y^p, x_2 = y^q$ for some $y \in A^+, p, q \geq 0$.*

The following result which resumes relative positions of alternative codes and their subclasses.

Lemma 2 ([7]). *Over any alphabet consisting of at least two letters, the followings hold true*

(i) $C_{sA} \subset C_{lnA}, C_{sA} \subset C_{rnA}, C_{sA} \subset C_{tA}$,
   $C_{sA} = C_{nA} \cap C_{tA} = C_{lnA} \cap C_{rA} = C_{rnA} \cap C_{lA} = C_{lA} \cap C_{nA} \cap C_{rA}$
   $= C_{lnA} \cap C_{tA} = C_{tA} \cap C_{rnA} = C_{rnA} \cap C_{lnA} = C_{lnA} \cap C_{rnA} \cap C_{tA};$

(ii) $C_{lnA} \subset C_{lA}, C_{lnA} \subset C_{nA},\ C_{rnA} \subset C_{rA}, C_{rnA} \subset C_{nA},\ C_{tA} \subset C_{lA}$,
   $C_{tA} \subset C_{rA}, C_{lnA} = C_{lA} \cap C_{nA}, C_{rnA} = C_{rA} \cap C_{nA}, C_{tA} = C_{lA} \cap C_{rA};$

(iii) $C_{lA} \subset C_A, C_{nA} \subset C_A,\ C_{rA} \subset C_A.$

The following result claims relationships between b-codes and alternative codes.

Lemma 3 ([17]). *Let X, Y be non-empty subsets of A^+, and $Z = X_< \cup Y_>$ where $X_< \subseteq A_<^+$ and $Y_> \subseteq A_>^+$.*

(i) *Z is a b-code if and only if $X = Y$ and X is a code;*

(ii) *Z is a b-code (left b-code, right b-code, strict b-code) if and only if (X, Y) is an alternative code (resp. a left alternative code, a right alternative code, a strict alternative code).*

3.   RESULTS

We introduce in this section four new subclasses of b-codes and the b-codes with two elements. Several characterizations for b-codes and their subclasses are established.

Firstly, we introduce four new subclasses of b-codes. Let us begin with a new class of codes which is a subclass of left b-codes and right b-codes.

Definition 3. A b-language Z is called a two-side b-code over A if it is both a left b-code and a right b-code.

Let us take an example.

Example 1. Consider the set $Z = \{(0, aab, 1), (0, baa, 1), (1, aa, 0)\}$ over $A = \{a, b\}$. It is easy to check that Z is both a left b-code and a right b-code, and hence it is a two-side b-code. However, Z is not a strict b-code because the word $(0, aabaa, 1)$ has two distinct strict equivalent factorizations,

$$(0, aabaa, 1) \approx_S (0, aab, 1)(1, aa, 0) \approx_S (1, aa, 0)(0, baa, 1).$$

As an immediate consequence of Definition 3 and Lemma 3(ii), we obtain the following relationship between two-side b-codes and two-side alternative codes.

Theorem 1. *Let X, Y be non-empty subsets of A^+, and $Z = X_< \cup Y_>$ where $X_< \subseteq A^+_<$ and $Y_> \subseteq A^+_>$. Then, Z is a two-side b-code if and only if (X, Y) is a two-side alternative code.*

Let Z be a non-empty subset of $A^+_\Diamond$. Two b-words s and t are said to be opposite equivalent, denoted by $s \approx_O t$, if $\mathrm{Proj}(s) = \mathrm{Proj}(t)$, $p_L(s) \neq p_L(t)$ and $p_R(s) \neq p_R(t)$. Two factorizations $s = s_1 s_2 \ldots s_n$ and $t = t_1 t_2 \ldots t_m$ on Z are called opposite equivalent if $s \approx_O t$, where $s_i, t_j \in Z, i = 1, \ldots, n, j = 1, \ldots m$. A b-language Z is said to be norm form if no word in $A^+_\Diamond$ which admits two different opposite equivalent factorizations on Z.

Definition 4. A b-language Z is called a norm b-code (left norm b-code, right norm b-code) over A if it is a b-code (resp. left b-code, right b-code) which has norm form.

The classes of two-side b-codes, left norm b-codes, right norm b-codes and norm b-codes are denoted respectively by C_{tB}, C_{lnB}, C_{rnB} and C_{nB}.

Example 2. Consider the sets $Z_1 = \{(1, a, 0), (1, aba, 0), (0, b^n a, 1) \mid n \geq 1\}$ and $Z_2 = \{(0, a, 1), (1, b, 0), (1, ab, 0), (0, ba, 1)\}$ over $A = \{a, b\}$. Then, it is not difficult to check that no word in $A^+_\Diamond$ admits two different equivalent factorizations and two different opposite equivalent factorizations on Z_1. Thus, Z_1 is a norm b-code. However, Z_1 is not a left norm b-code because the word $(0, baaba, 0)$ has two different left equivalent factorizations on Z_1, $(0, baaba, 0) \approx_L (0, ba, 1)(1, a, 0)(0, ba, 1) \approx_L (0, ba, 1)(1, aba, 0)$.

Next, it is easy to verify that no word in A^+ admits two different left equivalent and two different opposite equivalent factorizations on Z_2. Hence, Z_2 is a left norm b-code. However, Z_2 is not a right norm b-code because the word $(0, baab, 0)$ has two different right equivalent factorizations on Z_2, $(0, baab, 0) \approx_R (0, ba, 1)(1, ab, 0) \approx_R (1, b, 0)(0, a, 1)(1, ab, 0)$.

In an entirely symmetric manner, we also obtain $\widetilde{Z}_1$ is a norm b-code but it is not a right norm b-code, and $\widetilde{Z}_2$ is a right norm b-code not being a left norm b-code.

Remark 1. For two different factorizations on Z, there are only separately possible cases as the following:

[00, 00] (1)   [11, 11] (2)   [01, 01] (3)   [10, 10] (4)   [00, 01] (5)

[11, 10] (6)   [00, 10] (7)   [01, 11] (8)   [01, 10] (9)   [00, 11] (10)

where $[01, 10]$ denotes that the first factorization has the form $(0, w, 1)$, and the second factorization has the form $(1, w, 0)$. By Definitions 2-4, the following holds true:

The class of codes	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
C_B	−	−	−	−	+	+	+	+	+	+
C_{lB}	−	−	−	−	−	−	+	+	+	+
C_{rB}	−	−	−	−	+	+	−	−	+	+
C_{tB}	−	−	−	−	−	−	−	−	+	+
C_{nB}	−	−	−	−	+	+	+	+	−	−
C_{lnB}	−	−	−	−	−	−	+	+	−	−
C_{rnB}	−	−	−	−	+	+	−	−	−	−
C_{sB}	−	−	−	−	−	−	−	−	−	−

where $+$ and $-$ denote that these classes accept and reject factorizations which have forms like $(1), (2), \ldots, (10)$ respectively.

As an immediate consequence of Definitions 2-4 and Remark 1, we can now formulate the following result which resumes relative positions of the classes of codes under consideration.

Theorem 2. *Over any alphabet consisting of at least two letters, the followings hold true*

(i) $C_{sB} \subset C_{lnB}, C_{sB} \subset C_{rnB}, C_{sB} \subset C_{tB}$,
$C_{sB} = C_{nB} \cap C_{tB} = C_{lnB} \cap C_{rB} = C_{rnB} \cap C_{lB} = C_{lB} \cap C_{nB} \cap C_{rB}$
$= C_{lnB} \cap C_{tB} = C_{tB} \cap C_{rnB} = C_{rnB} \cap C_{lnB} = C_{lnB} \cap C_{rnB} \cap C_{tB}$;

(ii) $C_{lnB} \subset C_{lB}, C_{lnB} \subset C_{nB}, C_{rnB} \subset C_{rB}, C_{rnB} \subset C_{nB}, C_{tB} \subset C_{lB}$,
$C_{tB} \subset C_{rB}, C_{lnB} = C_{lB} \cap C_{nB}, C_{rnB} = C_{rB} \cap C_{nB}, C_{tB} = C_{lB} \cap C_{rB}$;

(iii) $C_{lB} \subset C_B, C_{nB} \subset C_B, C_{rB} \subset C_B$.

Remark 2. By Theorem 2, $C_{sB} = C_{nB} \cap C_{tB}$. However, C_{nB} and C_{tB} are not comparable by inclusion. Indeed, the two-side b-code Z in Example 1 is not a norm b-code because the word $(0, aabaa, 0)$ has two different opposite equivalent factorizations on Z, $(0, aab, 1)(1, aa, 0) \approx_O (1, aa, 0)(0, baa, 1)$. On the other hand, the norm b-code Z_1 in Example 2 is easily verified to be not a two-side b-code.

By virtue of Lemma 2, Theorem 2 and Remark 2, the relative positions of the classes of codes under consideration can be illustrated in the Figure 1, where the arrow $\rightarrow$ stands for a strict inclusion, and $\leftrightarrow$ stands for a one-to-one correspondence (we only consider b-codes and their subclasses in $A_<^+ \cup A_>^+$). It is worthy to note that if we restrict ourselves to considering only one-letter alphabets then we have $C_{sA} = C_{lnA} = C_{rnA} = C_{nA} = \emptyset$, whereas the remaining classes represented in the Figure 1 coincide.

Theorem 3. *Let X, Y be non-empty subsets of A^+, and $Z = X_< \cup Y_>$ where $X_< \subseteq A_<^+$ and $Y_> \subseteq A_>^+$. Then, Z is a norm (left norm, right norm) b-code if and only if (X, Y) is a norm (resp. left norm, right norm) alternative code.*

Proof. We treat only the case of norm b-codes. For the other cases the arguments are similar. Let Z be a norm b-code. If (X, Y) is not a norm alternative code, then there exists $w \in A^+$ such that $w = u_1 u_2 \ldots u_n = v_1 v_2 \ldots v_m$, where both $u_1 \in X, v_1 \in Y$ or $u_1 \in Y, v_1 \in X$ and $u_n \in X, v_m \in Y$ or $u_n \in Y, v_m \in X$. We consider separately two possible cases.
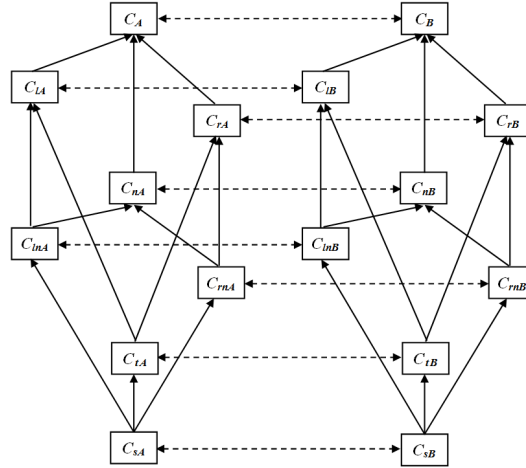
Figure 1. Relations on *b*-codes, alternative codes and their subclasses

Case 1: $u_1, v_m \in X, v_1, u_n \in Y$ or $u_1, v_m \in Y, v_1, u_n \in X$. Then, we have $(0, u, 0) \approx_O (1, v, 1)$ or $(1, u, 1) \approx_O (0, v, 0)$, with $u = u_1 u_2 \ldots u_n$, $v = v_1 v_2 \ldots v_m$.

Case 2: $u_1, u_n \in X, v_1, v_m \in Y$ or $u_1, u_n \in Y, v_1, v_m \in X$. Then, we have $(0, u, 1) \approx_O (1, v, 0)$ or $(1, u, 0) \approx_O (0, v, 1)$, with $u = u_1 u_2 \ldots u_n$, $v = v_1 v_2 \ldots v_m$.

Thus, in all the cases we have proved that whenever the word (l, w, r), with $l, r \in B$, admits two different opposite equivalent factorizations on Z, a contradiction. So, (X, Y) is a norm alternative code.

Conversely, suppose the contrary that Z is not a norm *b*-code. Then, if Z is not a *b*-code then, by Lemma 3(ii), (X, Y) is not a alternative code, which contradicts the hypothesis that (X, Y) is a norm alternative code. Otherwise, there exists a word z in $A_\Diamond^+$ admits two different opposite equivalent factorizations on Z,

$$s_1 s_2 \ldots s_n \approx_O t_1 t_2 \ldots t_m$$

where $s_i, t_j \in Z, 1 \le i \le n, 1 \le j \le m$, and $s_1 s_2 \ldots s_n$ and $t_1 t_2 \ldots t_m$ are two alternative factorizations on $(X_<, Y_>)$. We consider separately two possible cases.

Case 1: $p_L(s_1 \ldots s_n) = p_R(s_1 \ldots s_n) = 0$ or $p_L(s_1 \ldots s_n) = p_R(s_1 \ldots s_n) = 1$. Then, by definition of $\approx_O$, $p_L(t_1 \ldots t_m) = p_R(t_1 \ldots t_m) = 1$ or $p_L(t_1 \ldots t_m) = p_R(t_1 \ldots t_m) = 0$. Therefore, $s_1, t_m \in X_<, t_1, s_n \in Y_>$ or $s_1, t_m \in Y_>, t_1, s_n \in X_<$. Hence, $\mathrm{Proj}(z) = u_1 u_2 \ldots u_n = v_1 v_2 \ldots v_m$, where $u_i = \mathrm{Proj}(s_i), v_j = \mathrm{Proj}(t_j), 1 \le i \le n, 1 \le j \le m$, and $u_1, v_m \in X, v_1, u_n \in Y$ or $u_1, v_m \in Y, v_1, u_n \in X$.

Case 2: $p_L(s_1 \ldots s_n) = 0, p_R(s_1 \ldots s_n) = 1$ or $p_L(s_1 \ldots s_n) = 1, p_R(s_1 \ldots s_n) = 0$. Then, $p_L(t_1 \ldots t_m) = 1, p_R(t_1 \ldots t_m) = 0$ or $p_L(t_1 \ldots t_m) = 0, p_R(t_1 \ldots t_m) = 1$. Therefore, $s_1, s_n \in X_<, t_1, t_m \in Y_>$ or $s_1, s_n \in Y_>, t_1, t_m \in X_<$. Hence, $\mathrm{Proj}(z) = u_1 u_2 \ldots u_n = v_1 v_2 \ldots v_m$, where $u_i = \mathrm{Proj}(s_i), v_j = \mathrm{Proj}(t_j), 1 \le i \le n, 1 \le j \le m$, and $u_1, u_n \in X, v_1, v_m \in Y$ or $u_1, u_n \in Y, v_1, v_m \in X$.

Thus, in all the cases we have proved that whenever the word $\mathrm{Proj}(z)$ admits two different dissimilar alternative factorizations on (X, Y), a contradiction. So, Z is a norm *b*-code. ∎

In the rest of this section, we describe the *b*-codes with two elements. The case of sets with three words is much more complicated.

Theorem 4. *Let $Z = \{z_1, z_2\} \subseteq A_\diamondsuit^+$. Then, Z is not a b-code if and only if z_1 and z_2 have the form (1) or (2) in Remark 1, and they are powers of the same b-word, $z_1 = t^p, z_2 = t^q$ for some $t \in A_\diamondsuit^+, p, q \geq 0$.*

Proof. Suppose Z is not a b-code but the assertion is not true. Then, we consider separately two possible cases.

Case 1: z_1 and z_2 have not the form (1) and (2) in Remark 1. Then, z_1 and z_2 belong to one of the forms $(3), (4), \ldots, (10)$. It is not difficult to verify that no word in $A_\diamondsuit^+$ admits two different equivalent factorizations on Z. Thus, Z is a b-code, a contradiction.

Case 2: z_1 and z_2 have the form (1) or (2) in Remark 1 but they are not powers of the same b-word. Then, from Lemma 1 implies $\mathrm{Proj}(Z)$ is a code. Therefore, because $Z \subseteq A_\wedge^+$ or $Z \subseteq A_\vee^+$, it follows that Z is a b-code, a contradiction.

Conversely, since z_1 and z_2 are powers of the same b-word, by Lemma 1, $\mathrm{Proj}(Z)$ is not a code. Therefore, because $Z \subseteq A_\wedge^+$ or $Z \subseteq A_\vee^+$, by Lemma 3(i), Z is not a b-code. ∎

4.   CONCLUSIONS

The purpose of this note was to deal with the development of the classes of b-codes and alternative codes. Four new classes of codes, which are subclasses of b-codes, were introduced and considered. Several characteristics and relationships of these classes were proposed (Theorems 3-4). One of them provided us with a tool to describe the b-codes with two elements.

In future works, we hope we can find algorithms to test for subclasses of b-codes and many interesting problems for b-codes as well as their subclasses.

ACKNOWLEDGMENT

REFERENCES

[1] M. K. Ahmad and X. Augros, "Some results on codes for generalized factorizations," *Journal of Automaton, Languages and Combinatorics*, vol. 6, no. 3, pp. 239–251, 2001.

[2] M. Anselmo, "Automates et codes zigzag," *RAIRO Theoretical Informatics and Applications*, vol. 25, no. 1, pp. 49–66, 1991.

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

[4] C. Câmpeanu, K. Salomaa, and S. Vágvölgyi, "Shuffle Quotient and Decompositions," *Lecture Notes in Computer Science*, vol. 2295, pp. 186–196, 2002.

[5] N. D. Han and P. T. Huy, "On unambiguity of languages related to codes," in *Future Information Technology, Application, and Service*, J. P. James, C. M. L. Victor, L. W. Cho, and S. Taeshik, Ed. Springer Netherlands, pp. 32–38, 2012.

[6] N. D. Han, H. N. Vinh, D. Q. Thang, and P. T. Huy, "Quadratic algorithms for testing of codes and $\diamond$-codes," *Fundamenta Informaticae*, vol. 130, no. 2, pp. 163-177, 2014.

[7] N. T. Hien, "Characterizations for several classes of alternative codes," *Journal of Computer Science and Cybernetics*, vol. 32, pp. 273–283, 2016.

[8] N. T. Hien, "On strong alt-induced codes," *Applied Mathematical Sciences*, vol. 12, no. 7, pp. 327 – 336, 2018, https://doi.org/10.12988/ams.2018.8113.

[9] N. T. Hien and D. L. Van, "Codes induced by alternative codes," *Acta Mathematica Vietnammica*, pp. 1–15, 2018, https://doi.org/10.1007/s40306-018-0247-2.

[10] P. T. Huy and V. T. Nam, "Alternative codes and pre-context codes," in *Proceedings of the 7th National Conference: Selected problems about IT and Telecommunication*, 2004, pp. 188–197 (in Vietnamese).

[11] H. Jürgensen and S. Konstantinidis, "Codes," in *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa, Ed. Springer, Berlin, pp. 511–607, 1997.

[12] A. Lentin, "Equations dans les Monoïdes Libres," *Paris: Gauthier-Villars*, 1972.

[13] M. Madonia, S. Salemi, and T. Sportelli, "On z-submonoids and z-codes," *Theoretical Informatics and Applications*, vol. 25, pp. 305–322, 1991.

[14] M. Madonia, S. Salemi, and T. Sportelli, "Covering submonoids and covering codes," *Journal of Automaton, Languages and Combinatorics*, vol. 4, no. 4, pp 333–350, 1999.

[15] H. J. Shyr, *Free Monoids and Languages*. Hon Min Book Company, Taichung, 1991.

[16] D. L. Van, B. L. Saec, and I. Litovsky, "On coding morphisms for zigzag codes," *Theoretical Informatics and Applications*, vol. 26, no. 6, pp. 565–580, 1992.

[17] H. N. Vinh and P. T. Huy, "Codes of bounded words," in *Proceedings of the 3rd International Conference on Computer and Electrical Engineering*, vol. 2, pp. 89–95, 2010.

[18] H. N. Vinh, P. T. Huy, and D. L. Van, "Regular $\diamond$-languages and codes," in *Proceedings of the 4th Fundamental and Applied IT Research*, 2009, pp. 13–22 (in Vietnamese).

[19] H. N. Vinh, P. T. Huy, and D. L. Van, "Extension of codes and algorithms for alternative codes and codes of bounded words," *Journal of Computer Science and Cybernetics*, vol. 26, no. 4, pp. 301–311, 2010 (in Vietnamese).

[20] H. N. Vinh, V. T. Nam, and P. T. Huy, "Codes based on unambiguous products," *Lecture Notes in Computer Science*, vol. 6423, pp. 252–262, 2010.