MAPPED SHUFFLED FIBONACCI LANGUAGES

Hung-Kuei Hsiao
Libearl Arts Center, Da-Yeh University
112, Shan-Jeau Rd., Da-Tsuen, Chang-Hwa, Taiwan 515

Shyr-Shen Yu
Department of Applied Mathematics, National Chung-Hsing University
Taichung, Taiwan 402
(Submitted June 2001-Final Revision February 2002)

1. INTRODUCTION

The purpose of this paper is to study properties of mapped shuffled Fibonacci languages $F_{(a,b)}$ and $F_{(u,v)}$. Let $X = \{a, b\}$ be an alphabet and let X^* be the free monoid generated by X. Let 1 be the empty word and let $X^+ = X^* \setminus \{1\}$. The length of a word u is denoted by $\lg(u)$. Every subset of X^* is called a *language*. For two words $u, v \in X^+$, we consider the following type of Fibonacci sequence of words:

$$w_1 = u, w_2 = v, w_3 = uv, \ldots, w_n = w_{n-2}w_{n-1}, \ldots, n \geq 3.$$

Let $F_{u,v} = \{w_i | i \geq 1\}$. If $u = a$ and $v = b$, then $F_{u,v}$ is denoted by $F_{a,b}$. The sequence of Fibonacci words plays a very important role in the combinatorial theory of free monoids for the recursively defined structure and remarkable combinatorial properties of Fibonacci words can be shown. Some properties concerning the Fiboancci language $F_{u,v}$ have been investigated by De Luca in [2], by Fan and Shyr in [3] and by Knuth, Morris and Pratt in [6].

In [1], properties of Fibonacci words generated through the bicatenation operation, i.e., $F_i = F_{i-1}F_{i-2} \cup F_{i-2}F_{i-1} = \{uv, vu | u \in F_{i-1}, v \in F_{i-2}\}$ where $F_1 = \{a\}$ and $F_2 = \{b\}$, are investigated. Here we consider the shuffle operation. For $u, v \in X^*$, the *shuffle product* of u and v is the set $u \diamond v$ defined by:

$$u \diamond v = \{u_1v_1u_2v_2 \cdots u_nv_n | u_i, v_j \in X^*, 1 \leq i, j \leq n,\ u_1u_2 \cdots u_n = u,\ v_1v_2 \cdots v_n = v\}.$$

For $A, B \subseteq X^*$, the *shuffle product* of A and B is defined as: $A \diamond B = \bigcup_{u \in A, v \in B}(u \diamond v)$. We consider the following type of Fibonacci sequence of sets:

$$F_1 = \{a\},\ F_2 = \{b\},\ F_{n+2} = F_n \diamond F_{n+1} \text{ for } n \geq 1.$$

Let $F_{(a,b)} = \bigcup_{i \geq 1} F_i$. Remark that every word in the same F_i has the same length. For $u, v \in X^+$, let the homomorphism $h : X^* \to X^*$ be defined by $h(a) = u$ and $h(b) = v$. The *mapped shuffled Fibonacci language* $F_{(u,v)}$ is defined to be the language $h(F_{(a,b)}) = \{h(w) | w \in F_{(a,b)}\}$.

Section 2 concerns properties of the mapped shuffled Fibonacci language $F_{(u,v)}$ related to the theory of formal languages. We prove that $F_{(a,b)}$ is equal to the set of all combinations of words in the Fibonacci language $F_{a,b}$. In [3], Fan and Shyr show that $F_{a,b}$ is regular free. Then clearly $F_{a,b}$ is not a regular language. For more complicated cases, we show that $F_{(u,v)}$ is neither dense nor context-free for any $\{u, v\} \neq X$. In Section 2, we also show that $F_{(u,v)}$ is a context-sensitive language.

Section 3 is dedicated to investigate the relationships between Fibonacci words in $F_{(u,v)}$ and primitive words. In [3] and [5], the powers of a word which can be contained as a subword in a Fibonacci word are studied. Here we show that $F_{(a,b)}$ contains only primitive words. Some properties of words u and v such that $F_{(u,v)}$ contains primitive words are investigated in Section 3 too.

In Section 4, some conditions of u and v such that the homomorphism $h : X^* \to X^*$ defined by $h(a) = u$ and $h(b) = v$ is palindrome preserving or d-primitive preserving are studied. We also count the number of palindrome words in each F_i. Codes contained in $F_{(u,v)}$ are investigated in Section 5.

Items not defined here or in the subsequent sections can be found in [4] and [9].

2. THE MAPPED SHUFFLED FIBONACCI LANGUAGE $F_{(u,v)}$

In this paper we let the sequence of Fibonacci numbers m_i be defined by $m_1 = 1$, $m_2 = 1$ and $m_i = m_{i-1} + m_{i-2}$ for $i \geq 3$. We also let $m_0 = 0$. Let the Fibonacci language $F_{a,b}$ be ordered in the lexicographic order as $F_{a,b} = \{w_1, w_2, w_3, \ldots, w_n, \ldots\}$. For $u \in X^+$, $\mathcal{C}(u)$ denotes the set of all combinations of the word u.

Let $F_1 = \{a\}, F_2 = \{b\}$. Then

$$\begin{aligned} F_3 &= \{ab, ba\} = \mathcal{C}(ab) = \mathcal{C}(w_3), \\ F_4 &= \{bab, abb, bba\} = \mathcal{C}(abb) = \mathcal{C}(w_4), \\ F_5 &= \{abbab, babab, baabb, ababb, aabbb, abbba, babba, bbaba, bbaab, bbbaa\} \\ &= \mathcal{C}(aabbb) = \mathcal{C}(w_5). \end{aligned}$$

For $u \in X^*$ and $a \in X$, let $n_a(u)$ denote the number of a's in u. We shall show the above observations can be applied to all F_i. That is the following property:

Proposition 2.1: $F_1 = \{a\}$, $F_2 = \{b\}$ and $F_i = \mathcal{C}(a^{m_{i-2}}b^{m_{i-1}}) = \mathcal{C}(w_i)$ for $i \geq 3$.

Proof: From the previous observation, it is true for $i = 1,2,3,4,5$. Suppose that the hypothesis holds true for $i \leq n$ with an integer $n \geq 5$. Now consider sets F_{n+1} and $\mathcal{C}(a^{m_{n-1}}b^{m_n})$. From the facts that $F_{n-1} = \mathcal{C}(a^{m_{n-3}}b^{m_{n-2}})$ and $F_n = \mathcal{C}(a^{m_{n-2}}b^{m_{n-1}})$, it follows that $F_{n+1} = F_{n-1} \diamond F_n \subseteq \mathcal{C}(a^{m_{n-1}}b^{m_n})$. Next, let $w \in \mathcal{C}(a^{m_{n-1}}b^{m_n})$. Let $u \in \mathcal{C}(a^{m_{n-3}}b^{m_{n-2}}) = F_{n-1}$ be the word arranged in the same order as the first m_{n-3} a's and the first m_{n-2} b's of w. One can take $v \in X^+$ such that $w \in u \diamond v$. Then we get $n_a(v) = n_a(w) - m_{n-3} = m_{n-2}$ and $n_b(v) = n_b(w) - m_{n-2} = m_{n-1}$. Thus $v \in \mathcal{C}(a^{m_{n-2}}b^{m_{n-1}}) = F_n$. Therefore, $w \in u \diamond v \subseteq F_{n-1} \diamond F_n = F_{n+1}$. $\square$.

For $L \subseteq X^*$, let $\mathcal{C}(L) = \bigcup_{u \in L} \mathcal{C}(u)$. Proposition 2.1 derives that $F_{(a,b)} = \mathcal{C}(F_{a,b})$. A language L is said to be *dense* if $L \cap X^*uX^* \neq \emptyset$ for every $u \in X^*$.

Proposition 2.2: The language $F_{(a,b)}$ is dense.

Proof: It is clear that $n_a(w_i) = m_{i-2}$ and $n_b(w_i) = m_{i-1}$ for $i \geq 3$. For every $u \in X^*$, let $k = \lg(u), m = m_{k+2} - n_a(u)$ and $n = m_{k+3} - n_b(u)$. Then $a^m u b^n \in \mathcal{C}(w_{k+4}) \subseteq F_{(a,b)}$. Thus $F_{(a,b)}$ is dense. $\square$

For a given language $L \subseteq X^*$, the *principal congruence* P_L determined by L is defined as follows:

$$u \equiv v(P_L) \Longleftrightarrow (xuy \in L \Longleftrightarrow xvy \in L \ \forall x, y \in X^*).$$

It is well known that the language L is accepted by a finite automaton if and only if L has finite P_L congruence classes, that is P_L is a finite index. A language which is accepted by a finite automaton is called a *regular language* ([4]). We call a language L *disjunctive* if P_L is the equality. Clearly, a disjunctive language is not regular. It is known that every disjunctive language is dense (see [9]).

Corollary 2.3: The language $F_{(a,b)}$ is not disjunctive.

Proof: For any two distinct words $u, v \in X^*$ with $n_a(u) = n_a(v)$ and $n_b(u) = n_b(v)$, in view of Proposition 2.1, we have $xuy \in F_{(a,b)}$ if and only if $xvy \in F_{(a,b)}$ for $x, y \in X^*$. Hence the Fibonacci language $F_{(a,b)}$ is not disjunctive. □

Lemma 2.4: ([13]) Let $h : X^* \to X^*$ be a homomorphism. If $h(L)$ is dense for some $L \subseteq X^*$, then $h(X) = X$.

Corollary 2.5: For $u, v \in X^+$, if $\{u, v\} \neq X$, then $F_{(u,v)}$ is not dense.

Proof: If $\{u, v\} \neq X$, then by Lemma 2.4, $h(F_{(a,b)}) = F_{(u,v)}$ is not dense . □

Corollary 2.3 shows that $F_{(a,b)}$ is not disjunctive. Moreover, Corollary 2.5 shows that $F_{(u,v)}$ is not dense for $\{u, v\} \neq X$. In the following we shall show that $F_{(u,v)}$ is neither regular nor context-free for any $u, v \in X^+$. A language L is said to be *regular free* (*context-free free*) if every infinite subset of L is not a regular (context-free) language. Of course, if a language is context-free free, then it is also regular free. It is known that if L is an infinite context-free language, then there exist $x_1, x_2, x_3, x_4, x_5 \in X^*$ with $\lg(x_2x_4) \geq 1$ such that $\{x_1x_2^nx_3x_4^nx_5 | n \geq 0\} \subseteq L$ (see [4]). The language of the form $\{x_1x_2^nx_3x_4^nx_5 | n \geq 0\}$ is called a *context-free component.*

Proposition 2.6: For any $u, v \in X^+$, $F_{(u,v)}$ is context-free free.

Proof: Suppose on the contrary that $F_{(u,v)}$ is not context-free free. Then there is an infinite context-free subset of $F_{(u,v)}$. That is, there exist $x_1, x_2, x_3, x_4, x_5 \in X^*$ with $\lg(x_2x_4) \geq 1$ such that $\{x_1x_2^nx_3x_4^nx_5 | n \geq 0\} \subseteq F_{(u,v)}$. Remark that $F_1 = \{u\}, F_2 = \{v\}$, $F_i = F_{i-2} \diamond F_{i-1}$ for $i \geq 3$, $F_{(u,v)} = \bigcup_{i \geq 1} F_i$, and $m_i < m_{i+1}$ for every $i \geq 2$. There is $k \geq 3$ such that $x_1x_2^jx_3x_4^jx_5 \in F_k$ for some $j \geq 1$ and $m_{k-1} > \lg(x_2x_4)$. This implies that $m_{k+1} = m_{k-1} + m_k > \lg(x_1x_2^{j+1}x_3x_4^{j+1}x_5)$. Thus $x_1x_2^{j+1}x_3x_4^{j+1}x_5 \notin F_{(u,v)}$, which leads to a contradiction. Therefore, $F_{(u,v)}$ is context-free free. □

Moreover, we shall show that $F_{(u,v)}$ is a context-sensitive language. For definitions and properties of context-sensitive languages and linear bounded automata, one is referred to [4].

Proposition 2.7: For $u, v \in X^+$, $F_{(u,v)}$ is a context-sensitive language.

Proof: Here we consider the language $L = F_{(a,b)} \setminus \{a, b\}$. It is known that if L is context-sensitive, so is $F_{(a,b)}$. By Proposition 2.1, $F_{i+2} = \mathcal{C}(a^{m_i}b^{m_{i+1}})$ for $i \geq 1$. We construct a 5 track linear bounded automation such that the first track stores the input word w, the second track stores the number m_{i-1}, the third and fourth tracks store the number m_i and the fifth track stores the number m_{i+1}. This automation is initialized by $i = 1$, i.e., track 2 stores m_0, track 3 stores m_1, and so on. For any input word w in track 1, we check the number m_i stored in track 4 with a's in w. If $n_a < m_i$, then $w \notin L$. If $n_a > m_i$, then we put m_i from track 4 into track 5, put m_{i-1} from track 2 into track 4, replace the number in track 2 by m_i in track 3, replace the number in track 3 by the number in track 4, and compare the number in track 4 with a's in w again. If the number m_i in track 4 equals $n_a(w)$, then we compare the number m_{i+1} in track 5 with b's in w. If $m_{i+1} = n_b(w)$, then $w \in L$. Otherwise, $w \notin L$. This automation is a linear bounded automation which accepts L. Therefore, L is context-sensitive. As context-sensitive languages are closed under 1-free substitution, $F_{(u,v)}$ is also a context-sensitive language. □

Here, we consider one property of Fibonacci numbers. Then we shall study the difference between the shuffled Fibonacci language $F_{(a,b)}$ and the inserted Fibonacci language $I_{(a,b)}$.

Proposition 2.8: Let $i \geq 10$. Then
(1) $\lfloor m_i/(m_{i-2}+1)\rfloor = 2 = \lfloor m_{i-1}/(m_{i-3}-1)\rfloor$ and
(2) $0 < m_{i-2} - 2(m_{i-4}-1) \leq m_{i-4}-1$.

Proof: By definition, $m_5 = 5$, $m_i = m_{i-3} + 2m_{i-2}$ and $m_i = m_{i-1} + m_{i-2} \geq m_{i-1} + 5$ for $i \geq 7$. Let $i \geq 10$. Then $m_{i-2} + 1 > m_{i-3} - 2 > 0$ and $m_{i-3} - 1 > m_{i-4} + 2 > 0$. This together with the equalities $m_i/(m_{i-2}+1) = 2+(m_{i-3}-2)/(m_{i-2}+1)$ and $m_{i-1}/(m_{i-3}-1) = 2 + (m_{i-4}+2)/(m_{i-3}-1)$ imply that $\lfloor m_i/(m_{i-2}+1)\rfloor = 2 = \lfloor m_{i-1}/(m_{i-3}-1)\rfloor$. Moreover, $0 < m_{i-2} - 2(m_{i-4}-1) = m_{i-5} + 2 \leq m_{i-4} - 1$. □

For $A, B \subseteq X^*$, the *insertion* of B into A is defined as:

$$B \xrightarrow{i} A = \{uvw | u, w \in X^*, uw \in A, v \in B\}.$$

Let $I_1 = \{a\}, I_2 = \{b\}$ and $I_i = I_{i-2} \xrightarrow{i} I_{i-1}$ for $i \geq 3$. The *inserted Fibonacci language* $I_{(a,b)}$ is defined by $I_{(a,b)} = \cup_{i\geq 1} I_i$. Clearly, $I_i \subseteq \mathcal{C}(a^{m_{i-2}}b^{m_{i-1}}) = \mathcal{C}(w_i) = F_i$ for $i \geq 3$. By observation, $I_i = F_i$ for $i = 1,2,3,4,5,6,7,8,9$.

Proposition 2.9: $I_i \subset F_i$ for every $i \geq 10$.

Proof: It is clear that $I_i \subseteq F_i$ for $i \geq 1$. Let $w = a^7b^{14}a^7b^{14}a^7b^6$. Then $w \in F_{10}$ but $w \notin (I_8 \xrightarrow{i} I_9) = I_{10}$. Indeed, one can take $r = m_{i-2} - 2(m_{i-4}-1)$ and $s = m_{i-1} - 2(m_{i-3}+1)$ for $i \geq 10$. This is conjunction with Proposition 2.8 yields $0 < r \leq m_{i-4}-1$ and $0 < s < m_{i-3}+1$. Let $w = (a^{m_{i-4}-1}b^{m_{i-3}+1})^2 a^r b^s$. Then $w \in \mathcal{C}(a^{m_{i-2}}b^{m_{i-1}}) = F_i$ and $w \notin \mathcal{C}(a^{m_{i-4}}b^{m_{i-3}}) \xrightarrow{i} \mathcal{C}(a^{m_{i-3}}b^{m_{i-2}}) = F_{i-2} \xrightarrow{i} F_{i-1}$. Since $I_i = I_{i-2} \xrightarrow{i} I_{i-1} \subseteq F_{i-2} \xrightarrow{i} F_{i-1}$, we have $w \notin I_i$, which completes the proof. □

3. $F_{(u,v)}$ AND PRIMITIVE WORDS

A word $p \in X^+$ which is not a power of any other word is called a *primitive word.* Let Q be the set of all primitive words over X ([9]). It is known that every word in X^+ can be uniquely expressed as a power of a primitive word ([8]). In [3], Fan and Shyr have proved that the Fibonacci language $F_{a,b}$ is a subset of Q. Here we show that $F_{(a,b)} \subseteq Q$. We also want to find words u, v such that $F_{(u,v)} \subseteq Q$.

Proposition 3.1: $F_{(a,b)} \subseteq Q$.

Proof: We consider $w \in F_i$ for some $i \geq 3$ whenever $a, b \in Q$. By Proposition 2.1, $w \in \mathcal{C}(a^{m_{i-2}}b^{m_{i-1}})$. Since m_{i-2} and m_{i-1} are relatively prime, $w \in Q$. Therefore, $F_{(a,b)} \subseteq Q$. □

For $u \in X^+$, if $u = p^n$ and p is a primitive word, then $\sqrt{u} = p$ is called the *primitive root* of u. For a language $L \subseteq X^+$, let $\sqrt{L} = \{\sqrt{u} | u \in L\}$. A language $L \subseteq X^+$ is called *pure* if for any $u \in L^+$, $\sqrt{u} \in L^+$.

A non-empty language L is a *code* if for $x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m \in L, x_1x_2\cdots x_n = y_1y_2\cdots y_m$ implies that $m = n$ and $x_i = y_i$ for $i = 1, 2, \ldots, n$. Let $\{u, v\}$ be a code and let $h : X^* \to X^*$ be defined by $h(a) = u$ and $h(b) = v$. Then h being injective is derived directly from the definition of codes.

Proposition 3.2: ([10]) Let $h : X^* \to X^*$ be an injective homomorphism. If $h(X)$ is a pure code, then h preserves the primitive words.

Proposition 3.3: For two distinct words $u, v \in X^+$, if $\{u, v\}$ is a pure code, then $F_{(u,v)} \subseteq Q$.

Proof: By Proposition 3.1, $F_{(a,b)} \subseteq Q$. Let $\{u, v\}$ be a pure code. We define the injective homomorphism $h : X^* \to X^*$ as $h(a) = u$ and $h(b) = v$. Also, for a language $L \subseteq X^+$, let $h(L) = \{h(u) | u \in L\}$. Clearly, $F_{(u,v)} = h\left(F_{(a,b)}\right)$. From Proposition 3.2, one has that $F_{(u,v)} \subseteq Q$. □

The definition of pure codes makes checking whether $\{u, v\}$ is a pure code not easy. We are going to find some other properties of u and v related to the primitivity of $F_{(u,v)}$. A word u is a *conjugate* of a word w if there exist $x, y \in X^*$ such that $u = xy$ and $w = yx$. The following lemmas concerning basic properties of decompositions and catenations of words will be needed in the sequel.

Lemma 3.4: ([8]) For $x, y \in X^+, xy = yx$ implies that $\sqrt{x} = \sqrt{y}$.

Remark: In fact that for $x, y \in X^+, xy = yx$ if and only if $\sqrt{x} = \sqrt{y}$.

Lemma 3.5: ([11]) Let $xy = p^i, x, y \in X^+, p \in Q, i \geq 1$. Then $yx = q^i$ for some $q \in Q$.

Lemma 3.6: ([12]) Let $xq^m = g^k$ for some $m, k \geq 1, x \in X^+, q \in Q$ and $g \in Q$, with $x \notin q^+$. Then $q \neq g$ and $\lg(g) > \lg(q^{m-1})$.

If $u = xy$ for $x, y \in X^*$, then x is called a *prefix* of u and it is denoted by $x \leq_p u$; the word y is called a *suffix* of u and denoted by $y \leq_s u$.

Proposition 3.7: Let $u, v \in X^+$ with $\lg(u) = \lg(v)$ and $uv \in Q$, and let $h : X^* \to X^*$ be a homomorphism defined by $h(a) = u$ and $h(b) = v$ where $X = \{a, b\}$. Then h preserves primitive words except a and b. That is, $h(Q)\backslash Q \subseteq \{u, v\}$.

Proof: Let $u, v \in X^+$, $\lg(u) = \lg(v)$ and $uv \in Q$. By Lemma 3.5, $vu \in Q$. As $uv \in Q, u \neq v$. Define $h : X^* \to X^*$ by $h(a) = u$ and $h(b) = v$. Since $\{u, v\}$ is a uniform code, h is an injective homomorphism. We want to show that $h(w) \in Q$ whenever $w \in Q\backslash\{a, b, ab, ba\}$. Suppose on the contrary that there exists $w \in Q\backslash\{a, b, ab, ba\}$ such that $h(w) \notin Q$. As $w \in Q\backslash\{a, b, ab, ba\}, \lg(w) \geq 3$. Let w' be a conjugate of w. From Lemma 3.5, one has that $w \in Q$ if and only if $w' \in Q$. As $\lg(w) \geq 3$ and $w \in Q, n_a(w) \neq 0$ and $n_b(w) \neq 0$. If no conjugate of w contains any one of the following subwords b^2a or a^2b, then $w = (ab)^i$ or $w = (ba)^i$ for some $i \geq 2$. This implies that $w \notin Q$, a contradiction. Thus there is a conjugate of w that contains a subword b^2a or a^2b. In the other word, there exists a conjugate w' of w such that $a \leq_p w'$ and $b^2 \leq_s w'$, or $b \leq_p w'$ and $a^2 \leq_s w'$. Without loss of any generality, we let $a \leq_p w'$ and $b^2 \leq_s w'$. Clearly, $u \leq_p h(w')$ and $v^2 \leq_s h(w')$. Note that $h(w')$ is a conjugate of $h(w)$. This in conjunction with $h(w) \notin Q$ and Lemma 3.5 yields $h(w') \notin Q$. That is, there exist $p \in Q$ and $j \geq 1$ such that $h(w') = p^{j+1}$. Since $\lg(u) = \lg(v)$ and $v^2 \leq_s h(w')$, by Lemma 3.6, we get $\lg(p) > \lg(u)$. Hence there exists $y \in X^+$ such that $p = uy$.

If $y \in \{u, v\}^+$, then $h(w') = (uy)^{j+1}$ and $uy \in \{u, v\}^+$. This implies that $w' = h^{-1}(h(w')) = h^{-1}((uy)^{j+1}) = (h^{-1}(uy))^{j+1} \notin Q$, a contradiction. Hence, $y \notin \{u, v\}^+$. Since $(uy)^{j+1} \in \{u, v\}^+$, we have $y(uy)^j \in \{u, v\}^+$. Hence there exist $y_1 \in \{u, v\}^*$ and $y_2 \in X^+$ such that $y = y_1y_2$ and $\lg(y_2) < \lg(u)$. The fact $(uy_1y_2)^{j+1} = p^{j+1} = h(w') \in \{u, v\}^+$ implies that $w_1 = y_2(uy_1y_2)^j \in \{u, v\}^+$. Not that $\lg(w_1) = k \lg(v)$ for some positive integer k and that $\lg(w_1) > \lg(v)$. Hence $\lg(w_1) \geq 2\lg(v)$. This in conjunction with $v^2 \leq_s uy_1w_1 = h(w')$ yields $v^2 \leq_s w_1$. We consider the following cases:

(1) $u \leq_s uy_1$. As $v^2 \leq_s w_1$ and $\lg(v) > \lg(y_2)$, there exists $y_3 \in X^+$ such that $v = y_3y_2$. It follows that $w_1 = y_4(y_3y_2)^2$ for some $y_4 \in \{u,v\}^*$. Since $u \leq_s uy_1$, we obtain $u \leq_s y_2(uy_1y_2)^{j-1}uy_1 = y_4(y_3y_2)y_3$. This together with $\lg(u) = \lg(v)$ yields $u = y_2y_3$. Now we consider the following four subcases:

(1-a) $u^2 = y_2y_3y_2y_3 \leq_p w_1 = y_2(uy_1y_2)^j$. Then $y_3y_2 \leq_p uy_1y_2$. As $\lg(u) = \lg(y_2y_3)$, $u = y_3y_2 = v$. This implies that $uv \notin Q$, a contradiction.

(1-b) $uv = y_2y_3y_3y_2 \leq_p w_1 = y_2uy_1y_2(uy_1y_2)^{j-1}$. Then $y_2 \leq_p y_3y_3y_2 \leq_p (uy_1y_2)^j$. There exist $y_4 \leq_p y_3$ and $r \geq 0$ such that $y_2 = y_3^ry_4$. Thus $y_3^{r+1}y_4 \leq_p y_3^ry_4y_3y_1y_2(uy_1y_2)^{j-1}$, i.e., $y_3y_3y_4 \leq_p y_4y_3y_1y_3$. It follows that $y_3 = y_4y_5 = y_5y_4$ for some $y_5 \in X^*$. By Lemma 3.4, we have $\sqrt{y_4} = \sqrt{y_5} = \sqrt{y_3}$. This is conjunction with $y_2 = y_3^ry_4$ and $\lg(u) = \lg(v)$ yields $u = y_2y_3 = y_3y_2 = v$ and $uv \notin Q$; a contradiction.

(1-c) $vu = y_3y_2y_2y_3 \leq_p w_1 = y_2(y_2y_3y_1y_2)^j$. This implies that $y_3y_2y_2 = y_2y_2y_3$. By Lemma 3.4, $\sqrt{y_2} = \sqrt{y_2^2} = \sqrt{y_3}$. Thus $y_2y_3 = y_3y_2$ and $u = v$. Hence, $uv \notin Q$, a contradiction.

(1-d) $v^2 \leq_p w_1$. As $v \leq_p w_1 = y_2(uy_1y_2)^j$ and $v = y_3y_2$, there exists $y_4 \in X^+$ such that $v = y_2y_4$ with $\lg(y_4) = \lg(y_3)$. Since $v^2 = y_2y_4y_2y_4 \leq_p w_1 = y_2(uy_1y_2)^j$ and $\lg(u) = \lg(y_2y_3) = \lg(y_4y_2)$, $u = y_4y_2$. Consider the case that $\lg(y_4) \leq \lg(y_2)$. There exists $y_5 \in X^*$ such that $y_2 = y_4y_5$. Then $v = y_4y_5y_4$ and $u = y_4y_4y_5$. As $v^2 = vy_4y_5y_4 \leq_s w_1 = (y_2uy_1)^jy_2 = (y_2uy_1)^jy_4y_5, y_5y_4 = y_4y_5$ and $u = v$. Hence $uv \notin Q$, a contradiction. Now, let $\lg(y_4) > \lg(y_2)$. There exists $y_5 \in X^+$ such that $y_4 = y_2y_5$. Then $u = y_2y_5y_2$ and $v = y_2y_2y_5$. As $v^2 = vy_2y_2y_5 \leq_s w_1$ and $uy_2 = y_2y_5y_2y_2 \leq_s w_1, y_2y_2y_5 = y_5y_2y_2$. By Lemma 3.4, $\sqrt{y_2} = \sqrt{y_2^2} = \sqrt{y_5}$. This implies that $\sqrt{u} = \sqrt{v}$ and $uv \notin Q$, a contradiction.

(2) $v \leq_s uy_1$. As $v^2 \leq_s w_1$, there exists $y_3 \in X^+$ such that $v = y_3y_2 = y_2y_3$. By Lemma 3.4, $\sqrt{y_2} = \sqrt{y_3}$. That is, there exist $q \in Q$ and $r_1, r_2 \geq 1$ such that $y_2 = q^{r_1}, y_3 = q^{r_2}$ and $v = q^{r_1+r_2}$. We consider the following four subcases:

(2-a) $u^2 \leq_p w_1 = y_2(uy_1y_2)^j$. There exists $y_4 \in X^+$ such that $u = y_2y_4 = y_4y_2$. Thus $\sqrt{y_4} = \sqrt{y_2}$. This in conjunction with $\sqrt{y_2} = \sqrt{y_3}$ yields $u = q^{r_1+r_2} = v$ and $uv \notin Q$, a contradiction.

(2-b) $uv \leq_p w_1 = y_2(uy_1y_2)^j$. There exist $y_4, y_5, y_6 \in X^+$ such that $u = y_2y_4 = y_4y_5, v = y_5y_6, \lg(y_5) = \lg(y_2) = \lg(q^{r_1})$ and $\lg(y_4) = \lg(y_3)$. Thus $y_5 = q^{r_1} = y_2$. As $u = y_2y_4 = y_4y_2$, by Lemma 3.4, $\sqrt{y_4} = \sqrt{y_2}$. Thus $uv = y_2y_4 \notin Q$, a contradiction.

(2-c) $v^2 = y_2y_3y_2y_3 \leq_p w_1 = y_2(uy_1y_2)^j$. The condition $\lg(u) = \lg(v)$ implies that $u = y_3y_2 = v$ and $uv \notin Q$, a contradiction.

(2-d) $vu \leq_p w_1$. As $v = y_2y_3 \leq_p w_1 = y_2(uy_1y_2)^j$, there exist $y_4, y_5 \in X^+$ such that $u = y_3y_4 = y_4y_5$ with $\lg(y_4) = \lg(y_2)$. This implies that $y_4 = (y_3)^{r_3}y_6$ for some $r_3 \geq 0$ and $y_6 \leq_p y_3$. Since $\lg(y_4) = \lg(y_2) = \lg(q^{r_1})$ and $y_3 = q^{r_2}, y_4 = q^{r_1}$.
Thus $u = q^{r_1+r_2} = v$ and $uv \notin Q$, a contradiction. □

From the proof of Proposition 3.1, we have the following result immediately.

Corollary 3.8: Let $A = \{a,b\}$ and B a finite nonempty alphabet. If $h : A^* \to B^*$ is a homomorphism of A^* into B^* defined by $h(a) = u$ and $h(b) = v$ for some primitive words $u, v \in B^+$ such that $\lg(u) = \lg(v)$ and uv is a primitive word, then h preserves primitive words.

Corollary 3.9: $F_{(u,v)} \backslash Q \subseteq \{u,v\}$ for any two words $u, v \in X^+$ with $\lg(u) = \lg(v)$ and $uv \in Q$.

Proof: Let $u, v \in X^+$ with $\lg(u) = \lg(v)$ and $uv \in Q$. By Proposition 3.1, $F_{(a,b)} \subseteq Q$. From Proposition 3.7, one has that $F_{(u,v)} \backslash Q \subseteq \{u,v\}$. □

For $u, v \in X^+$, we conjecture that $\{uv, uv^2\} \subseteq Q$ if and only if $F_{(u,v)} \backslash Q \subseteq \{u, v\}$. This is left for our further research. The partially primitive-preserving homomorphisms is also an interesting research topic for our further work.

4. PALINDROME WORDS AND d-PRIMITIVE WORDS IN $F_{(u,v)}$

If $x = a_1 a_2 \cdots a_n$, where $a_i \in X$, then we define the *reverse* (or *mirror image*) of the word x to be $\hat{x} = a_n \cdots a_2 a_1$. A word x is called *palindromic* if $x = \hat{x}$ ([7]).

Proposition 4.1: Let n_i be the number of palindrome words in $F_i = \mathcal{C}(a^{m_{i-2}} b^{m_{i-1}})$. Then $n_1 = 1, n_2 = 1$, and for $i \geq 3$,

$$n_i = \begin{cases} 0, & \text{if } m_i \text{ is an even number,} \\ \frac{(k_1+k_2)!}{k_1! k_2!}, & \text{if } m_i \text{ is an odd number,} \end{cases}$$

where $k_1 = \lfloor \frac{m_{i-2}}{2} \rfloor$ and $k_2 = \lfloor \frac{m_{i-1}}{2} \rfloor$.

Proof: If w is a palindrome word with $\lg(w) \geq 2$, then there exist $u \in X^+$ and $v \in X \cup \{1\}$ such that $w = \hat{u} v u$. By the definition of reverses, we have $n_a(u) = n_a(\hat{u})$ and $n_b(u) = n_b(\hat{u})$. Thus at most one of $n_a(w)$ and $n_b(w)$ can be odd whenever w is a palindrome word. From definitions: $m_1 = 1, m_2 = 1$ and $m_i = m_{i-1} + m_{i-2}$ for $i \geq 3$, it follows that m_i is an even number if and only if m_{i-1} and m_{i-2} are odd numbers. Consider $i \geq 3$. Then $m_i \geq 2$. If $w \in F_i$ and m_i is an even number, then $\lg(w) = m_i$ and $w \in \mathcal{C}(a^{m_{i-2}} b^{m_{i-1}})$ where both m_{i-1} and m_{i-2} are odd numbers. This implies that there exists no palindrome word in F_i if m_i is an even number. Now we consider the case that m_i is an odd number. Let $w = \hat{u} v u \in F_i$ for some $u \in X^+$ and $v \in X$. Then $u \in \mathcal{C}(a^{k_1} b^{k_2})$, where $k_1 = \lfloor \frac{m_{i-2}}{2} \rfloor$, $k_2 = \lfloor \frac{m_{i-1}}{2} \rfloor$. This implies that $n_i = \frac{(k_1+k_2)!}{k_1! k_2!}$. $\square$

Lemma 4.2: ([7]) Let $u, v \in X^+$ be two distinct words and let $h : X^* \to X^*$ be defined by $h(a) = u$ and $h(b) = v$. Then u and v are palindrome words if and only if h is a palindrome preserving homomorphism.

It is known that $\{u, v\} \subseteq X^+$ is a code if and only if $\sqrt{u} \neq \sqrt{v}$ (see [9]). For two words $u, v \in X^+, \{u, v\}$ being a code implies that h is an injective homomorphism where $h(a) = u$ and $h(b) = v$.

Proposition 4.3: Let $u, v \in X^+$ be two palindrome words. Then $\sqrt{u} \neq \sqrt{v}$ if and only if L and $h(L)$ contain the same number of palindrome words for every $L \subseteq X^+$.

Proof: Let $u, v \in X^+$ be two palindrome words with $\sqrt{u} \neq \sqrt{v}$. For $w \in X^*$, by Lemma 4.2, $h(w)$ is a palindrome word whenever w is a palindrome word. Now, let $w = a_1 a_2 \cdots a_n$ be such that $h(w)$ is a palindrome word, where $a_i \in X, 1 \leq i \leq n$, i.e., $h(w) = \widehat{h(w)}$. Note that $\widehat{h(w)} = \widehat{h(a_n)} \widehat{h(a_{n-1})} \cdots \widehat{h(a_1)} = h(a_n) h(a_{n-1}) \cdots h(a_1) = h(\hat{w})$. This in conjunction with the fact that h is injective whenever $\sqrt{u} \neq \sqrt{v}$ yields $w = \hat{w}$, i.e., w is a palindrome word. Therefore, L and $h(L)$ contain the same number of palindrome words for every $L \subseteq X^+$. Conversely, we assume that for every $L \subseteq X^+, L$ and $h(L)$ contain the same number of palindrome words. Let $L_1 = \{a, b\}$ and $L_2 = \{ab, ba\}$. Then a, b being palindrome words, by Lemma 4.2, implies that both $h(a) = u$ and $h(b) = v$ are also palindrome words. Since ab and ba are not palindrome words, $uv \neq \widehat{uv} = \hat{v} \hat{u} = vu$. By the remark of Lemma 3.4, we obtain $\sqrt{u} \neq \sqrt{v}$. $\square$

Proposition 4.3 derives that for two palindrome words u and v, if $h(a) = u, h(b) = v$ and $\sqrt{u} \neq \sqrt{v}$, then F_i and $h(F_i)$ contain the same number of palindrome words for every $i \geq 1$. A word $d \in X^*$ is said to be a *proper d-factor* of a word $z \in X^+$ if $d \neq z$ and $z = dx = yd$ for some words x, y. The family of words which have i distinct proper d-factors is denoted by $D(i)$. A word $x \in X^+$ is *d-primitive* if $x = dy_1 = y_2d$, where $d \in X^+$ and $y_1, y_2 \in X^*$, implies that $x = d$ and $y_1 = y_2 = 1$. The set $D(1)$ is exactly the family of all d-primitive words. For the properties of $D(i)$, one is referred to [13]. For $u, v \in X^+$, let $d_{u,v}$ denote the maximal word in X^* being such that $u = xd_{u,v}$ and $v = d_{u,v}y$ for some $x, y \in X^*$.

Lemma 4.4: ([7]) Let $u, v \in X^+$ be two distinct d-primitive words such that $d_{u,v} = d_{v,u} = 1$ and let $h(a) = u$ and $h(b) = v$. Then h is d-primitive preserving.

Proposition 4.5: Let $u, v \in D(1)$ with $d_{u,v} = d_{v,u} = 1$ and let $h(a) = u$ and $h(b) = v$. Then the following two statements hold true:
(1) $w \in D(1)$ if and only if $h(w) \in D(1)$;
(2) L and $h(L)$ contain the same number of d-primitive words for any $L \subseteq X^+$.

Proof: By Lemma 4.4, h is d-primitive preserving. If $w \in D(1)$, then $h(w) \in D(1)$. Now assume that $w \in X^+ \backslash D(1)$. That is, there exist $d, x, y \in X^+$ such that $w = xd = dy$. Then $h(x)h(d) = h(xd) = h(w) = h(dy) = h(d)h(y)$. This implies that $h(d)$ is a non-empty d-factor of $h(w)$ and $h(w) \notin D(1)$. Thus statement (1) holds true. For any $L \subseteq X^+$, as h is injective and by (1), L and $h(L)$ contain the same number of d-primitive words. $\square$

Proposition 4.5 derives that for $u, v \in D(1)$ with $d_{u,v} = d_{v,u} = 1$, F_i and $h(F_i)$ contain the same number of d-primitive words where $h(a) = u$ and $h(b) = v$.

5. $F_{(u,v)}$ AND CODES

Proposition 2.1 derives that $F_{(a,b)} \supseteq \{a^{m_i}b^{m_{i+1}} | i \geq 2\}$ which is a bifix code. Let $F_{a,b}$ be ordered in the lexicographic order as $\{w_1, w_2, \ldots, w_n, \ldots\}$. In [3], Fan and Shyr show that languages $\{w_{2n} | n \geq 1\}$ and $\{w_{2n-1} | n \geq 1\}$ are codes. In [14], we show that for $k \geq 2$, $\{w_{nk} | n \geq 1\}$ is a code. Here we are going to find some other codes contained in $F_{(u,v)}$.

Example: For a given integer $k \geq 2$, let $L_n = \mathcal{C}(a^{m_{n-2}}b^{m_{n-1}-m_{n-k}})b^{m_{n-k}}$ for $n > k$. Then $L = \cup_{i \geq 2} L_{ik}$ is a suffix code contained in $F_{(a,b)}$.

Lemma 5.1 ([10]) Let $h : X^* \to X^*$ be a homomorphism. Then the following statements are equivalent:
(1) h is code preserving;
(2) h is injective;
(3) $|h(X)| = |X|$ and $h(X)$ is a code.

Corollary 5.2 For $u, v \in X^+$, let $h(a) = u$ and $h(b) = v$. Let $L \subseteq F_{(a,b)}$ be a code. Then $\{u, v\}$ being a code implies that $h(L)$ is a code.

According to Corollary 5.2, we then consider codes in $F_{(a,b)}$ instead of codes in $F_{(u,v)}$. We quote the following lemma from [14], which is needed in the sequel.

Lemma 5.3: ([14])
(1) For every $i \geq 1, w_i \not\leq_p w_{i+1}$;
(2) $w_i \leq_p w_j$ implies that $j - i$ is an even number;
(3) for $k \geq 5$ and $1 \leq i \leq k - 4, w_i \leq_p w_k$ implies that $w_iw_{i+1}w_{i+1}w_iw_{i+1} \leq_p w_k$;
(4) for each $k \geq 2, w_iw_i \not\leq_p w_k$ for every $i < k$.

Proposition 5.4: Let $L_i = w_{i-1}X^{m_{i-2}}$ for $i \geq 3$. For $k \geq 3$, let $L \subseteq \cup_{n\geq 1} L_{nk}$ be such that $|L \cap L_{ik}| = 1$ for each $i \geq 1$. Then L is a code.

Proof: Suppose there exists $k \geq 3$ such that there is $L \subseteq \bigcup_{n\geq 1} L_{nk}$ with $|L \cap L_{ik}| = 1$ for each $i \geq 1$ and that L is not a code. Then there exist $u_1, u_2, \ldots u_n, v_1, v_2, \ldots, v_m \in L$ for some finite integers $m, n \geq 1$ such that $u_1 \neq v_1$ and $u_1u_2\cdots u_n = v_1v_2\cdots v_m$. Since $u_1 \neq v_1$, without loss of generality, let $u_1 <_p v_1$. There exist $i_1 < j_1$ such that $u_1 \in L_{ki_1}$ and $v_1 \in L_{kj_1}$. This implies that $w_{ki_1-1} \leq_p w_{kj_1-1}$. By the definition of L and $i_1 \geq 1$, $kj_1 - 1 \geq ki_1 + k - 1 \geq 2k - 1 \geq 5$. Moreover, $kj_1 - ki_1 \geq 3$ which follows immediately from the inequalities $k \geq 3$ and $j_1 > i_1$. Then apply (2) of Lemma 5.3 to get $kj_1 - ki_1 \geq 4$, i.e., $(ki_1 - 1) \leq (kj_1 - 1) - 4$. This is the case considered in the following:

(*) By (3) of Lemma 5.3, $w_{ki_1-1}w_{ki_1}w_{ki_1}w_{ki_1-1}w_{ki_1} \leq_p w_{kj_1-1} <_p v_1$. This in conjunction with $u_1 \leq_p v_1, u_1 \in w_{ki_1-1}X^{m_{ki_1-2}}$ and $w_{ki_1} = w_{ki_1-2}w_{ki_1-1}$ yields $u_1 = w_{ki_1-1}w_{ki_1-2}$. Thus

$$u_1w_{ki_1+1}w_{ki_1+1} = u_1w_{ki_1-1}w_{ki_1}w_{ki_1-1}w_{ki_1} \leq_p v_1.$$

Let $u_2 \in L_{ki_2}$ for some $i_2 \geq 1$. If $i_2 > i_1$, then $\lg(u_2) = m_{ki_2-2} + m_{ki_2-1} \geq m_{ki_1+1} + m_{ki_1+2} \geq \lg(w_{ki_1+1}w_{ki_1+1})$. This together with $u_1u_2\cdots u_n = v_1v_2\cdots v_m$ and $u_1w_{ki_1+1}w_{ki_1+1} \leq_p v_1$ yields $w_{ki_1+1} \leq_p w_{ki_2-1} \leq_p u_2$. By (1) of Lemma 5.3, $ki_2 - 1 > ki_1 + 2$. This implies that $w_{ki_1+1}w_{ki_1+1} \leq_p w_{ki_2-1}$, in contradiction with (4) of Lemma 5.3. Thus $i_2 \leq i_1$. We consider the following two subcases:

(*1) $i_2 = i_1$. Then $u_2 = u_1 = w_{ki_1-1}w_{ki_1-2}$ and $u_1u_2w_{ki_1-1}w_{ki_1-1}w_{ki_1} \leq_p v_1$. Let $u_3 \in L_{ki_3}$ for some $i_3 \geq 1$. Then again by (4) of Lemma 5.3, $m_{ki_1+1} > 2m_{ki_1-1} = \lg(w_{ki_1-1}w_{ki_1-1}) > \lg(w_{ki_3-1}) = m_{ki_3-1}$. Thus $i_1 \geq i_3$ and $m_{ki_3} \leq 2m_{ki_1-1}$. It follows that $u_3 \leq_p w_{ki_1-1}w_{ki_1-1}w_{ki_1}$. If $i_3 = i_1, u_3 \in w_{ki_1-1}X^{m_{ki_1-2}}$ implies that $u_3 = w_{ki_1-1}w_{ki_1-3}w_{ki_1-4} \neq u_1$. This contradicts the fact that $|L \cap L_{ki_1}| = 1$. Thus one has the following case:

(*1') $i_3 < i_1$. Then we have $ki_1 - 1 \geq ki_3 + k - 1 \geq 5$. Since $ki_1 - ki_3 \geq 3$, by (2) of Lemma 5.3, $ki_1 - ki_3 \geq 4$. Note that $u_1u_2\cdots u_n = v_1v_2\cdots v_m, u_1u_2w_{ki_1-1}w_{ki_1-1}w_{ki_1} \leq_p v$ and $u_3 \in w_{ki_3-1}X^{m_{ki_3-2}}$. Hence $w_{ki_3-1} \leq pw_{ki_1-1} \leq_p v_1$. By (3) of Lemma 5.3, we obtain $w_{ki_3-1}w_{ki_3}w_{ki_3}w_{ki_3-1}w_{ki_3} \leq_p w_{ki_1-1} \leq_p v_1$. This is the same case as the case (*).

(*2) $i_2 < i_1$. This case is analogous to the case (*1') which is the same as the case (*). This implies that $u_1u_2\cdots u_n <_p v_1$, i.e. $u_1u_2\cdots u_n \neq v_1v_2\cdots v_m$, a contradiction, which completes the proof. □

Clearly, $L \subseteq F_{(a,b)} \cap \bigcup_{n\geq 1} L_{nk}$ with $|L \cap L_{nk}| = 1$ is also a code for any $k \geq 3$. Remark that the code L given in Proposition 5.4 can be neither a prefix code nor a suffix code. Furthermore, we conjecture that if we choose a word from each $F_{2n}, n \geq 2$, to form a set L, the L is a code. This is left for our further research.

REFERENCES

[1] W. Chuan. "Generating Fibonacci Words." *The Fibonacci Quarterly* **33.2** (1995): 104-112.

[2] A. De Luca. "A Combinatorial Property of Fibonacci Words." *Inform. Proc. Letters* **12** (1981): 193-195.

[3] Cheng-Ming Fan and H.J. Shyr. "Some Properties of Fibonacci Languages." *Tamkang Journal of Mathematics* **27.2** (1996): 165-182.

[4] John E. Hopcroft and Jeffrey D. Ullman. *Formal Languages and Their Relation to Automata.* Addison-Wesley, Massachusetts, 1969.

[5] J. Karhumäki. "On Cube-Free ω-Words Generated by Binary Morphism." *Discr. Appl. Math.* **5** (1985): 279-297.

[6] D. E. Knuth, J.H. Morris and V.R. Pratt. "Fast Pattern Matching in Strings." *SIAM J. Comput.* **6** (1977): 323-350.

[7] Yong-Yie Lin. *Properties of Words and the Related Homomorphisms.* Master Thesis, Institute of Applied Mathematics, Chung-Yuan Christian University, Chung-Li, Taiwan, 1995.

[8] R.C. Lyndon and M.P. Schützenberger. "The Equation $a^M = b^N c^P$ in a Free Group." *Michigan Math. J.* **9** (1962): 289-298.

[9] H.J. Shyr. *Free Monoids and Languages.* Second Edition, Hon Min Book Company, Taichung, Taiwan, 1991.

[10] H.J. Shyr and G. Thierrin. "Codes, Languages and MOL Schemes." *R.A.I.R.O. Informatique Theorique/Theoretical Computer Science* **1.4** (1977): 293-301.

[11] H.J. Shyr and G. Thierrin. "Disjunctive Languages and Codes." *Fundamentals of Computation Theory.* Proceedings of the 1977 Inter. FCT-Conference, Poznan, Poland, Lecture Notes in Computer Science. **56** Springer-Verlag (1977): 171-176.

[12] H.J. Shyr and S.S. Yu. "Non-Primitive Words in the Language p^+q^+." *Soochow Journal of Mathematics* **20.4** (1994): 535-546.

[13] H.L. Wu. "On the Properties of Primitive Words." Master Thesis, Institute of Applied Mathematics, Chung-Yuan Christian University, Chung-Li, Taiwan, 1992.

[14] S.S. Yu and Yu-Kuang Zhao. "Properties of Fibonacci Languages." *Discrete Mathematics* **22.4** (2000): 215-223.

AMS Classification Numbers: 68R15

✠ ✠ ✠