# EXTRACTION PROPERTY OF THE GOLDEN SEQUENCE

**Wai-fong Chuan***
Department of Mathematics, Chung-yuan Christian University,
Chung-li, Taiwan 320, Republic of China
(*Submitted June 1993*)

Let $a$ and $b$ be two distinct letters and let $\tau = (\sqrt{5}-1)/2$. Let $x$ be the infinite string whose $n^{\text{th}}$ term is "$a$" if $[(n+1)\tau]-[n\tau]=0$ and is "$b$" if $[(n+1)\tau]-[n\tau]=1$. Let $s_m$ be the left factor of $x$ of length $m$ and let $x_m$ be the corresponding right factor of $x$. Note that $x = x_0$ is the golden sequence. It is known that

$$x = c_1c_2c_3c_4\ldots \tag{1}$$

where $c_0 = a$, $c_1 = b$, and $c_{n+1} = c_{n-1}c_n$ $(n \geq 1)$. In the notation of [1]-[3], $x = F^1(b, ab)$, $c_n = w_{n+1}^1$ and $s_{F_n} = w_n^0$ $(n \leq 1)$, where $F_n$ denotes the $n^{\text{th}}$ Fibonacci number.

Hofstadter [6] formulated the concept of aligning two strings. By way of illustration, we present the procedure by which $x_m$ is aligned with $x = x_0$.

Starting from the $(m+1)^{\text{st}}$ term in $x$, an attempt is made to match each term in $x$ with a term in $x_m$. After a term in $x$ is matched with a term in $x_m$, one looks for the earliest match to the next term in $x$. Those terms in $x_m$ that are skipped over form the extracted string $y_{m,0}$. For example, when $m = 4$,

$$\begin{array}{lccccccccccccccccccc}
x_4: & a & b & a & b & b & a & b & b & a & b & a & b & b & a & b & a & b & b & \cdots \\
 & & | & | & | & | & | & | & & | & | & & | & & | & | & & | & & \\
x: & & b & a & b & b & a & b & & a & b & & b & & a & b & & b & & \ldots \\
y_{4,0}: & a & & & & & & & b & & & a & & b & & & a & & b & \ldots
\end{array}$$

It was Hendel and Monteferrante [4] who first reformulated Hofstadter's alignment concept in terms of a formal relation on strings. If $x_m$ aligns with $x_n$ with extraction $y_{m,n}$, then we notationally indicate this by

$$x_m \supset x_n;\ y_{m,n}. \tag{2}$$

[4] also introduced the idea of representing $x_m$ as a product of $c_\alpha$ with specific properties by using a canonical representation $x_m = c_{\alpha(1)}c_{\alpha(2)}\cdots$ where $\alpha(k)$ is an increasing function on the positive integers that can be derived from the Zeckendorf representation of $m$ as a sum of Fibonacci numbers. Using this, they were able to completely determine $y_{m,0}$ for all positive integers $m$.

The goal of this paper is to determine the remaining cases of $y_{m,n}$. In Section 2, $y_{0,m}$ is found to be precisely the reverse $R(s_m)$ of the left factor $s_m$ of $x$ of length $m$.

Here the *reverse operation* $R$ is defined by

$$R(a_1a_2\ldots a_k) = a_k\ldots a_2a_1,$$

where $a_1, a_2, \ldots, a_k$ are letters. The importance of the reversal operation in studying $x$ was first observed by Higgins [5]. In Section 4, it is shown that $y_{m,n}$ and $y_{m-1,n-1}$ differ by at most the first letter. From this, $y_{m,n}$ can easily be determined by $y_{m-n,0}$ (if $m > n$) or $y_{0,n-m}$ (if $n > m$).

---

* This research was supported in part by the National Science Council, R.O.C. Grant NSC 83-0208-M-033-013.

## 1. BASIC LEMMAS AND DEFINITIONS ON EXTRACTION

The following definitions come from [4, Definitions 1 and 2]. Suppose that $U = u_1 \dots u_n$, $V = v_1 \dots v_m$, and $E = e_1 \dots e_p$ with $u_i, v_j, e_k \in \{a, b\}, n, m > 0, p \geq 0$, and $n = m + p$. We say that $U$ *aligns* (with) $V$ with *extraction* $E$ if there exist integers $j(0), j(1), j(2), \dots, j(p)$ such that

$$U = (v_1 \dots v_{j(1)})e_1(v_{j(1)+1} \dots v_{j(2)})e_2 \dots e_p(v_{j(p)+1} \dots v_m),$$

with $v_i \dots v_k$ empty if $k < i$ and

(i) $0 = j(0) \leq j(1) \leq j(2) \leq \dots \leq j(p) < m,$

(ii) $e_i \neq v_{j(i)+1}$, for $1 \leq i \leq p$.

This relationship is called an *alignment* and is denoted by $U \supset V; E$. The strings $U$, $V$, and $E$ are called the *original*, *aligned*, and *extracted strings*, respectively. If $U = V$, we write $U \supset V; 1$, where 1 denotes the empty string.

Suppose that $U$, $V$, and $E$ are (possibly infinite) strings. Suppose that $U(n)$, $V(n)$, and $E(n)$, $n \geq 1$, are sequences of finite strings such that $U(n) \supset V(n)$; $E(n)$, $\lim U(n) = U$, $\lim V(n) = V$, and $\lim E(n) = E$. Then we say that *U aligns V* with *extraction E*. This alignment is also denoted by $U \supset V; E$.

***Lemma 1.1 [4, Lemmas 1 and 3]:***

***(a)*** (Uniqueness of extracted string) For given strings $U$ and $V$, there is at most one string $E$ such that $U \supset V; E$.

***(b)*** (Concatenation) If $U_i, V_i$, and $E_i$, $1 \leq i \leq m$, are strings of finite lengths and if $U_i \supset V_i; E_i$, $1 \leq i \leq m$, then

$$U_1U_2 \dots U_m \supset V_1V_2 \dots V_m; E_1E_2 \dots E_m.$$

***Lemma 1.2:***

***(i)*** $c_n \supset c_{n-1}; \; c_{n-2}, \; n \geq 2.$

***(ii)*** $c_n \supset c_n; \; 1, \; n \geq 1.$

***(iii)*** $c_n = c_{n-2}c_{n-1}, \; n \geq 2.$

***(iv)*** $c_nc_{n+1} \dots c_p \supset c_{n+1} \dots c_p; \; c_n, \; 1 \leq n < p.$

***(v)*** $c_nc_{n+2} \supset c_{n+2}; \; c_n, \; n \geq 1.$

***(vi)*** $c_nc_n \supset c_{n+1}; \; c_{n-2}, \; n \geq 2.$

***(vii)*** $c_nc_{n+3} \supset c_{n+1}c_{n+2}; \; c_n, \; n \geq 0.$

***Proof:*** Part (i) has been proved in [4] by induction. Parts (ii) and (iii) are trivial. According to (i) and (ii), we have

$$c_nc_{n+1} \supset c_{n+1}; \; c_n$$
$$c_{n+i} \supset c_{n+i}; \; 1, \; 2 \leq i \leq p-n.$$

Part (iv) now follows by concatenation [Lemma 1.1(b)]. The proofs of (v)-(vii) are similar to (iv).

***Lemma 1.3:*** Let $t \geq 1$. Let $\gamma(0) = 0$ and let $\gamma(1), \ldots, \gamma(t)$ be positive integers such that $\gamma(i) + 2 \leq \gamma(i+1)$, $1 \leq i \leq t-1$. Let

$$U = c_1 c_2 \ldots c_{\gamma(t)} c_{\gamma(t)+1}$$

$$V = \begin{cases} c_1 c_2 \ldots c_{\gamma(1)-1} c_{\gamma(1)+1}, & \text{if } t = 1, \\ (c_1 c_2 \ldots c_{\gamma(1)-1})(c_{\gamma(1)+1} \ldots c_{\gamma(2)-1}) \cdots (c_{\gamma(t-1)+1} \ldots c_{\gamma(t)-1}) c_{\gamma(t)+1}, & \text{otherwise} \end{cases}$$

$$E = c_{\gamma(1)} c_{\gamma(2)} \ldots c_{\gamma(t)},$$

where the factor $c_1 c_2 \ldots c_{\gamma(1)-1}$ does not appear if $\gamma(1) = 1$. Then $U \supset V;\ E$.

***Proof:*** By Lemma 1.2, we have

$$\begin{aligned} &c_1 c_2 \ldots c_{\gamma(1)-1} \supset c_1 c_2 \ldots c_{\gamma(1)-1};\ 1, && \text{if } \gamma(1) > 1, \\ &c_{\gamma(i)} c_{\gamma(i)+1} \ldots c_{\gamma(i+1)-1} \supset c_{\gamma(i)+1} \ldots c_{\gamma(i+1)-1};\ c_{\gamma(i)}, && 1 \leq i \leq t-1, \\ &c_{\gamma(t)} c_{\gamma(t)+1} \supset c_{\gamma(t)+1};\ c_{\gamma(t)}. \end{aligned}$$

The result now follows by concatenation.

***Lemma 1.4 [4, Lemma 5]:*** Let $m \geq 1$ have Zeckendorf representation

$$m = F_{k(1)} + F_{k(2)} + \cdots + F_{k(t)} \tag{3}$$

with $k(1) \geq 2$, $k(i) + 2 \leq k(i+1)$, $i = 1, \ldots, t-1$. Let $\gamma(i) = k(i) - 1$, $1 \leq i \leq t$, and let $V$ be as in Lemma 1.3. Then

$$x_m = V c_{\gamma(t)+2} c_{\gamma(t)+3} \cdots . \tag{4}$$

The ordered collection of indices $1, 2, \ldots, \gamma(1)-1, \gamma(1)+1, \ldots, \gamma(2)-1, \ldots, \gamma(t-1)+1, \ldots, \gamma(t)-1, \gamma(t)+1, \gamma(t)+2, \ldots$ is called the *canonical representation of* $x_m$. Actually [4, Definition 3] uses the term "canonical representation" to refer to the function of the positive integers enumerating this ordered collection. However, in the sequel, if there is no ambiguity, we will simply, by abuse of language, call (4) the canonical representation of $x_m$.

***Corollary 1.5:*** Let $x_m = c_{\alpha(1)} c_{\alpha(2)} \ldots$ be a canonical representation. Then

***(i)*** $(\alpha(1), \alpha(2)) \in \{(1,2), (1,3), (2,3), (2,4)\}$.

***(ii)*** $\alpha(k+1) \in \{\alpha(k)+1, \alpha(k)+2\}$, for all $k \geq 1$.

***(iii)*** There exists a positive integer $r$ such that $\alpha(k+1) = \alpha(k)+1$ for all $k \geq r$.

## 2. THE ALIGNMENTS $x \supset x_m;\ y_{0,m}$ AND $x_m \supset x;\ y_{m,0}$

We now express the extraction $y_{0,m}$ in terms of the $c_i$.

***Lemma 2.1:*** For $m \geq 1$, let $m$ have Zeckendorf representation (3). Let $\gamma(i) = k(i) - 1$, $1 \leq i \leq t$. Then

$$y_{0,m} = c_{\gamma(1)} \cdots c_{\gamma(t)},$$

where $y_{0,m}$ is defined by (2).

***Proof:*** The result follows from (1), (4), Lemma 1.3, and Lemma 1.2(ii) by concatenation.

Next, we look at the left factors of the golden sequence. Let

$$w_1 = a,\ x_2 = b,\ w_{n+1} = w_n w_{n-1},\ n \geq 2.$$

In the notation of [1]-[3], $w_n = w_n^0,\ n \geq 1$.

***Lemma 2.2:*** Let $n \geq 4$. Then $w_n w_n$ is a left factor of $x$.

***Proof:*** First, observe that

$$\begin{aligned} w_{n+2} &= w_{n+1} w_n \\ &= (w_n w_{n-1})(w_{n-1} w_{n-2}) \\ &= w_n w_{n-1} w_{n-2} w_{n-3} w_{n-2} \\ &= w_n w_n w_{n-3} w_{n-2}. \end{aligned}$$

By Lemma 1.4 of [3], $w_{n+2}$ is a left factor of $x$, for all $n \geq 4$. The result immediately follows.

***Lemma 2.3:*** Let $m \geq 1$ have Zeckendorf representation (3). Then

$$s_m = w_{k(t)} \cdots x_{k(2)} w_{k(1)}.$$

***Proof:*** The result clearly holds for $m = 1, 2, 3$. Suppose $m \geq 4$ and that the result is true for all positive integers less than $m$.

First, suppose $t = 1$ so that, by (3), $m = F_n$ for some $n$. By Lemma 2.2, $w_n$ is a left factor of $x$. By definition, $s_m$ is also a left factor of $x$. Since both these left factors of $x$ have the same length $F_n$, they are both equal.

Next, suppose that t > 1. Then, by (3),

$$F_{k(t)} < m < F_{k(t)+1} \leq 2F_{k(t)}.$$

Note that $s_{F_{k(t)}} = w_{k(t)}$ since they are both left factors of $x$ of the same length. let

$$s_m = s_{F_{k(t)}} s = w_{k(t)} s,$$

where $s$ has length $m - F_{k(t)}$. By Lemma 2.2, $w_{k(t)} w_{k(t)}$ is a left factor of $x$. Since $s_m = w_{k(t)} s$ is also a left factor of $x$, it follows that $s$ is a left factor of $w_{k(t)}$. Therefore, $s = s_{m-F_{k(t)}}$; hence,

$$s_m = w_{k(t)} s_{m-F_{k(t)}}.$$

By the induction hypothesis, the Zeckendorf representation

$$m - F_{k(t)} = F_{k(t-1)} + \cdots + F_{k(2)} + F_{k(1)}$$

gives the factorization

$$s_{m-F_{k(t)}} = w_{k(t-1)} \cdots w_{k(2)} w_{k(1)}.$$

Consequently, $s_m$ has the desired factorization.

***Theorem 2.4:*** For $m \geq 1$,

$$y_{0,m} = R(s_m).$$

***Proof:*** We have

$$\begin{aligned}
R(s_m) &= R(w_{k(1)})R(w_{k(2)})\cdots R(w_{k(t)}), && \text{by Lemma 2.3,}\\
&= c_{k(1)-1}c_{k(2)-1}\cdots c_{k(t)-1}, && \text{by the result, } R(w_n)=c_{n-1}, \text{ of Theorem 3 in [1],}\\
&= c_{\gamma(1)}\cdots c_{\gamma(t)}, && \text{using the notations of Lemma 1.4,}\\
&= y_{0,m}, && \text{by Lemma 2.1.}
\end{aligned}$$

***Theorem 2.5 (Modified Hofstadter's conjecture [4]):*** Let $m \ge 2$ have Zeckendorf representation (3). Then

$$\begin{aligned}
x_m &\supset x;\ ax_{m-1}, && \text{if } k(1)=2 \text{ and } k(2) \text{ is even;}\\
x_m &\supset x;\ x_{m-2}, && \text{otherwise.}
\end{aligned}$$

In other words, $y_{m,0} = ax_{m-1}$ in the first case (this is also true when $m = 1$) and $y_{m,0} = x_{m-2}$ in the second case.

## 3. SOME LEMMAS

The goal of this section is to prove that, under appropriate conditions, if $s \supset t;\, u$, then $c_p s \supset t;\, c_p u$. The precise statement and conditions are set forth in Lemma 3.5. The major tool in proving Lemma 3.5 will be Lemma 3.1, which considers three cases.

Throughout this section, we let $p \ge 2$ and we let

$$s = c_{\alpha(1)}c_{\alpha(2)}\cdots$$
$$t = c_{\beta(1)}c_{\beta(2)}\cdots$$

with

$$\alpha(1) = p+2 \text{ or } p+3, \qquad \beta(1) = p+1 \text{ or } p+2.$$

We suppose that $r$ is a positive integer such that, for $k < r$, we have

$$\alpha(k+1) \in \{\alpha(k)+1,\ \alpha(k)+2\}, \qquad \beta(k+1) \in \{\beta(k)+1,\ \beta(k)+2\},$$

while, for $k \ge r$, we have

$$\alpha(k+1) = \alpha(k)+1, \qquad \beta(k+1) = \beta(k)+1.$$

***Lemma 3.1:*** There is some $k$ such that either cases (i) and (ii) listed below hold, or else case (iii) below holds for all $k$.

**Case (i).** There exists a string $u_k$ such that

$$c_{\alpha(1)}\cdots c_{\alpha(k)} \supset c_{\beta(1)}\cdots c_{\beta(k)};\ u_k, \tag{5}$$

$$c_p c_{\alpha(1)}\cdots c_{\alpha(k)} \supset c_{\beta(1)}\cdots c_{\beta(k)};\ c_p u_k. \tag{6}$$

**Case (ii).** There exists a string $u_k$ such that

$$c_{\alpha(1)}\cdots c_{\alpha(k-1)}c_{\alpha(k)-2} \supset c_{\beta(1)}\cdots c_{\beta(k)};\ u_k, \tag{7}$$

$$c_p c_{\alpha(1)}\cdots c_{\alpha(k-1)}c_{\alpha(k)-2} \supset c_{\beta(1)}\cdots c_{\beta(k)};\ c_p u_k. \tag{8}$$

**Case (iii).**

$$\beta(k) = \alpha(k)-1, \tag{9}$$

and there exist strings $u_k$ and $v_k$ such that

$$v_k c_{\alpha(k)-1} = c_p u_k, \tag{10}$$

$$c_{\alpha(1)} \dots c_{\alpha(k)} \supset c_{\beta(1)} \dots c_{\beta(k)};\ u_k, \tag{11}$$

$$c_p c_{\alpha(1)} \dots c_{\alpha(k-1)} c_{\alpha(k)-2} \supset c_{\beta(1)} \dots c_{\beta(k)};\ v_k. \tag{12}$$

The factor $c_{\alpha(1)} \dots c_{\alpha(k-1)}$ in (7), (8), and (12) does not appear if $k = 1$.

***Proof:*** Lemma 3.1 follows immediately from the statements of Lemmas 3.2 and 3.3 which are proved below.

***Lemma 3.2:*** If $k = 1$, then one of the three cases listed in Lemma 3.1 holds.

***Proof:*** There are four cases to consider, according to the values of $\alpha(1)$ and $\beta(1)$.

**Case (a).** $\alpha(1) = p+2$ and $\beta(1) = p+1$

We show that case (iii) holds with $u_1 = c_p$ and $v_1 = c_{p-2}$. Clearly (9) holds. By Lemma 1.2(iii), (10) is satisfied. Alignment (11) follows from Lemma 1.2(i), while alignment (12) follows from Lemma 1.2(vi).

**Case (b).** $\alpha(1) = p+2$ and $\beta(1) = p+2$

We show that (i) holds with $u_1 = 1$. Then (5) follows from Lemma 1.2(ii) and (6) follows from Lemma 1.2(v).

**Case (c).** $\alpha(1) = p+3$ and $\beta(1) = p+1$

We show that (ii) holds with $u_1 = 1$. Then (7) follows from Lemma 1.2(ii) and (8) follows from Lemma 1.2(iv).

**Case (d).** $\alpha(1) = p+3$ and $\beta(1) = p+2$

We show that (iii) holds with $u_1 = c_{p+1}$ and $v_1 =$ . Clearly (9) holds. Lemma 1.2(iii) implies equation (10), alignment (11) follows from Lemma 1.2(i), and (12) follows from Lemma 1.2(iii).

***Lemma 3.3:*** Suppose, for some integer $k \geq 1$, case (iii) of Lemma 3.1 holds. Then, for $k+1$, one of the three cases of Lemma 3.1 holds.

***Proof:*** First, note that, by (9), $\beta(k+1) \in \{\alpha(k), \alpha(k)+1\}$. There are now four cases to consider, according to the values of $\alpha(k+1)$ and $\beta(k+1)$.

**Case (a).** $\alpha(k+1) = \alpha(k)+1$ and $\beta(k+1) = \alpha(k)$

Let

$$u_{k+1} = u_k c_{\alpha(k)-1} \text{ and } v_{k+1} = v_k c_{\alpha(k)-3}. \tag{13}$$

We show that (iii) holds with $k+1$ replacing $k$. Clearly $\beta(k+1) = \alpha(k+1) - 1$. By Lemma 1.2(iii) and (10), we have

$$\begin{aligned} v_{k+1} c_{\alpha(k+1)-1} &= v_k c_{\alpha(k)-3} c_{\alpha(k)} = v_k c_{\alpha(k)-3} c_{\alpha(k)-2} c_{\alpha(k)-1} \\ &= v_k c_{\alpha(k)-1} c_{\alpha(k)-1} = c_p u_k c_{\alpha(k)-1} = c_p u_{k+1}. \end{aligned}$$

This demonstrates that (10) holds with $k$ replaced by $k+1$.

To prove that (11) holds with $k+1$ replacing $k$, we concatenate the following two alignments: (11) as is, with $k$ and $c_{\alpha(k+1)} \supset c_{\beta(k+1)}$; $c_{\alpha(k)-1}$, the last alignment following from Lemma 1.2(i).

To prove that (12) holds with $k+1$ replacing $k$, we concatenate the following two alignments: (12) as is, with $k$ and $c_{\alpha(k)-1}c_{\alpha(k+1)-2} \supset c_{\beta(k+1)}$; $c_{\alpha(k)-3}$, the last alignment following from Lemma 1.2(vi) with $n = \alpha(k)$. Alignment (12) with $k+1$ replacing $k$ then holds since, by Lemma 1.2(iii), $c_{\alpha(k)-2}c_{\alpha(k)-1} = c_{\alpha(k)}$.

**Case (b).** $\alpha(k+1) = \alpha(k)+1$ and $\beta(k+1) = \alpha(k)+1$

Let $u_{k+1} = u_k$. We prove that (i) holds with $k+1$ replacing $k$.

To prove that (5) holds with $k+1$ replacing $k$, we concatenate the following two alignments: (11) and $c_{\alpha(k+1)} \supset c_{\beta(k+1)}$; 1, this last alignment holding by Lemma 1.2(ii).

To prove (6) with $k+1$ replacing $k$, we concatenate the following two alignments: (12) and $c_{\alpha(k)-1}c_{\alpha(k+1)} \supset c_{\beta(k+1)}$); $c_{\alpha(k)-1}$, the last alignment following from Lemma 1.2(v). Alignment (6) with $k+1$ replacing $k$ then follows from (10) and Lemma 1.2(iii) with $n = \alpha(k)$.

**Case (c).** $\alpha(k+1) = \alpha(k)+2$ and $\beta(k+1) = \alpha(k)$

Let $u_{k+1} = u_k$. We show that (ii) holds with $k+1$ replacing $k$.

To prove (7) with $k+1$ replacing $k$, we concatenate the following two alignments: (11) and $c_{\alpha(k+1)-2} \supset c_{\beta(k+1)}$; 1, the last alignment following from Lemma 1.2(ii).

To prove (8) with $k+1$ replacing $k$, we concatenate the following two alignments: (12) and $c_{\alpha(k)-1}c_{\alpha(k+1)-2} \supset c_{\beta(k+1)}$; $c_{\alpha(k)-1}$, the last alignment following from Lemma 1.2(iii) and (i). Alignment (8) with $k+1$ replacing $k$ then follows from (10) and Lemma 1.2(iii) with $n = \alpha(k)$.

**Case (d).** $\alpha(k+1) = \alpha(k)+2$ and $\beta(k+1) = \alpha(k)+1$

Let $u_{k+1} = u_k c_{\alpha(k)}$ and let $v_{k+1} = v_k$. We show that (iii) holds with $k+1$ replacing $k$. Clearly (10) with $k+1$ replacing $k$ follows from (10) as is and Lemma 1.2(iii).

To prove (11) with $k+1$ replacing $k$, we concatenate the following two alignments: (11) as is and $c_{\alpha(k+1)} \supset c_{\beta(k+1)}$; $c_{\alpha(k)}$, the last alignment following from Lemma 1.2(i).

To prove (12) with $k+1$ replacing $k$, we concatenate the following two alignments: (12) as is and $c_{\alpha(k)-1}c_{\alpha(k+1)-2} \supset c_{\beta(k+1)}$; 1, the last alignment following from Lemma 1.2(iii) and (i).

As already noted, Lemmas 3.2 and 3.3 provide an inductive proof to Lemma 3.1.

***Lemma 3.4:***

***(i)*** If cases (i) and (iii) of Lemma 3.1 do not hold for any $k$, then eventually (for all $k \geq r$) we are in case (a) of Lemma 3.3.

***(ii)*** In such a case, $v_k$ (resp. $u_k$) is a proper left factor of $v_{k+1}$ (resp. $u_{k+1}$).

***Proof:*** By the hypothesis of this lemma, Lemma 3.2, and Lemma 3.3, case (iii) of Lemma 3.1 must hold for all $k$. By the hypothesis at the beginning of the section, $\alpha(k+1) = \alpha(k)+1$ for all $k \geq r$. Hence, of the four cases of Lemma 3.3, case (d) cannot hold and, clearly, cases (b) and (c) also do not hold. This proves assertion (i).

Assertion (ii) follows from equation (13).

We are now in a position to state the main lemma.

***Lemma 3.5:*** Assume that the notations and assumptions stated at the beginning of this section hold. If $s \supset t;\, u$, then $c_p s \supset t;\, c_p u$.

***Proof:*** The proof of Lemma 3.5 follows directly from the proof of Lemmas 3.6 and 3.7 below.

***Lemma 3.6:*** If, for some $k$, case (i) or (ii) of Lemma 3.1 holds, then Lemma 3.5 is true.

***Proof:*** Let

$$s' = c_{\alpha(k+1)} c_{\alpha(k+2)} \cdots,$$
$$t' = c_{\beta(k+1)} c_{\beta(k+2)} \cdots.$$

Then

$$s = c_{\alpha(1)} \ldots c_{\alpha(k)} s',$$
$$t = c_{\beta(1)} \ldots c_{\beta(k)} t'.$$

If (i) holds, then define $u'$ so that $s' \supset t';\, u'$. Note that $u'$ exists because $s'$ and $t'$ each have an infinite number of "$a$"s and "$b$"s. By concatenating this alignment with (5) and (6), respectively, we obtain

$$s \supset t;\; u_k u'$$
$$c_p s \supset t;\; c_p u_k u'.$$

Hence, $u_k u' = u$ by uniqueness of extracted strings, $c_p u_k u' = c_p u$ and we are done.

If (ii) holds, let $c_{\alpha(k)-1} s' \supset t';\, u'$. Then

$$s \supset t;\; u_k u'$$
$$c_p s \supset t;\; c_p u_k u'$$

with $u_k u' = u$, $c_p u_k u' = c_p u$ and again we are done.

***Lemma 3.7:*** If cases (i) and (ii) of Lemma 3.1 do not hold for any $k$, then Lemma 3.5 is true.

***Proof:*** By Lemma 3.4(ii), both $v = \lim v_k$ and $u_0 = \lim u_k$ are infinite strings. Taking the limits of (11) and (12) as $k$ goes to infinity, it is clear that

$$s \supset t;\; u_0$$
$$c_p s \supset t;\; v.$$

By uniqueness of extracted strings, we have $u = u_0$. By Lemma 3.4 and (13), we have

$$v_{k+2} = v_{k+1} c_{\alpha(k+1)-3} = v_k c_{\alpha(k)-3} c_{\alpha(k)-2} = v_k c_{\alpha(k)-1} = c_p u_k \quad (k \geq r).$$

Consequently, $v = \lim v_{k+2} - \lim c_p u_k = c_p \lim u_k = c_p u$.

***Remark:*** Lemma 3.5 also holds when $p = 1$. The proof for this case is straightforward and is left for the reader.

## 4. THE ALIGNMENTS $x_m \supset x_n$; $y_{m,n}$

***Theorem 4.1:*** Either the two extracted strings $y_{m,n}$ and $y_{m+1,n+1}$ are equal or else they differ by the first letter only. Here, $y_{m,n}$ is defined by (2).

***Proof:*** Let $x_m = c_{\alpha(1)}c_{\alpha(2)} \dots$ and $x_n = c_{\beta(1)}c_{\beta(2)} \dots$ be the canonical representations of $x_m$ and $x_n$, respectively. By Corollary 1.5(i), we have three cases to consider, according to the values of $\alpha(1)$ and $\beta(1)$.

**Case (i).** $\alpha(1) = \beta(1)$

Clearly $y_{m,n} = y_{m+1,n+1}$ in this case.

**Case (ii).** $\alpha(1) = 2$ and $\beta(1) = 1$

By Corollary 1.5(i), there are three subcases to consider:

**(a)** If $x_m = c_2s$, $x_n = c_1c_2t$ and $s \supset c_2t; u$, then $y_{m,n} = au$ and $y_{m+1,n+1} = bu$.

**(b)** If $x_m = c_2c_3s$, $x_n = c_1c_3t$ and $s \supset t; u$, then $y_{m,n} = au$ and $y_{m+1,n+1} = bu$.

**(c)** If $x_m = c_2c_4s$, $x_n = c_1c_3t$ and $s \supset t; u$, then $y_{m,n} = ac_2u$ and $y_{m+1,n+1} = c_3u = bc_2u$ by Lemma 3.5.

**Case (iii).** $\alpha(1) = 1$ and $\beta(1) = 2$

**(a)** If $x_m = c_1c_2s$, $x_n = c_2t$ and $s \supset t; u$, then $y_{m,n} = bu$ and $y_{m+1,n+1} = au$.

**(b)** If $x_m = c_1c_3s$, $x_n = c_2t$ and $s \supset t; u$, then $y_{m,n} = bbu$ and $y_{m+1,n+1} = c_2u = abu$ by Lemma 3.5.

This theorem, together with Theorems 2.4 and 2.5 (the modified Hofstadter's conjecture) imply the followiing result.

***Corollary 4.2:*** Let $m$ and $n$ be two nonnegative integers.

*(a)* If $m > n$, then $y_{m,n}$ is an infinite string; for $m \geq n+2$ (resp. $m = n+1$) the strings $y_{m,n}$ and $x_{m-n-2}$ (resp. $ax$) differ by at most the first letter.

*(b)* If $n > m$, then $y_{m,n}$ is a finite string with length $n-m$; the strings $y_{m,n}$ and $R(s_{n-m})$ differ by at most the first letter.

The above corollary motivates determining the first letters of the strings $y_{m,n}$ $(m \neq n)$, $x_{m-n-2}$ $(m \geq n+2)$, and $R(s_{n-m})$ $(n > m)$, where $m$ and $n$ are nonnegative integers.

***Lemma 4.3:***

*(a)* Let $m \geq n+2$. Let $m-n-2 = \sum_{j=1}^{\infty} \varepsilon_j F_{j+1}$ be the Zeckendorf representation of $m-n-2$. Then the first letter of $x_{m-n-2}$ is an "$a$" or "$b$" depending on whether $\varepsilon_1$ equals 1 or 0, respectively.

*(b)* Let $n > m$. Let $n-m-1 = \sum_{j=1}^{\infty} \varepsilon_j F_{j+1}$ be the Zeckendorf representation of $n-m-1$. Then the first letter of $R(s_{n-m})$ is an "$a$" or "$b$" depending on whether $\varepsilon_1$ equals 1 or 0, respectively.

*(c)* Let $m \neq n$. Let $m = \sum_{j=1}^{\infty} \varepsilon_j F_{j+1}$ and $n = \sum_{j=1}^{\infty} \delta_j F_{j+1}$ be the Zeckendorf representations of $m$ and $n$, respectively. Let $k$ be the smallest positive integer such that $\varepsilon_k \neq \delta_k$. Then the first letter of $y_{m,n}$ is an "$a$" iff either $\varepsilon_k = 0$ with $k$ even or $\varepsilon_k = 1$ with $k$ odd.

***Proof:*** (a) and (b) follow from [8, p. 85]. A similar proof holds for (c) after noting that, by Lemma 1.4, the following statements are true:

If $\varepsilon_k = 0$ (resp. 1) and $\delta_k = 1$ (resp. 0), then $x_m = uc_k s$ (resp. $uc_{k+1}s$) and $x_n = uc_{k+1}t$ (resp. $uc_k t$) for some strings *u*, *s*, and *t*.

## ACKNOWLEDGMEWNT

The author wishes to express her gratitude to the referee for his/her valuable comments regarding this paper

## REFERENCES

1. W. Chuan. "Fibonacci Words." *The Fibonacci Quarterly* **30.1** (1992):68-76.
2. W. Chuan. "Symmetric Fibonacci Words." *The Fibonacci Quarterly* **31.3** (1993):251-55.
3. W. Chuan. "Embedding Fibonacci Words into Fibonacci Word Patterns." In *Applications of Fibonacci Numbers* **5:**113-22. Ed. G. E. Bergum, A. N. Philippou, & A. F. Horadam. Dordrecht: Kluwer, 1993.
4. R. J. Hendel & S. A. Monteferrante. "Hofstadter's Extraction Conjecture." *The Fibonacci Quarterly* **32.2** (1994):98-107.
5. P. M. Higgins. "The Naming of Popes and a Fibonacci Sequence in Two Noncommuting Indeterminates." *The Fibonacci Quarterly* **25.1** (1987):57-61.
6. D. R. Hofstadter. *Eta-Lore*. First presented at the Stanford Math Club, Stanford, California, 1963, p. 13.
7. V. E. Hoggatt, Jr. *Fibonacci and Lucas Numbers*. Boston: Houghton Mifflin, 1969; rpt. Santa Clara, CA: The Fibonacci Association, 1980.
8. D. E. Knuth. *The Art of Computer Programming*, Vol. I. New York: Addison-Wesley, 1973.

AMS Classification Numbers: 68R15, 20M05

❖❖❖