

# Collapse: A Fibonacci and Sturmian Game

Dennis D.A. Epple  
Jason Siefken\*

## Abstract

We explore the properties of *Collapse*, a number game closely related to Fibonacci words. In doing so, we fully classify the set of periods (minimal or not) of finite Fibonacci words via careful examination of the Exceptional (sometimes called singular) finite Fibonacci words.

*Collapse* is not a game in the Game Theory sense, but rather in the recreational sense, like the 15-puzzle (the game where you slide numbered tiles in an attempt to arrange them in order). It was created in an attempt to better understand Sturmian words (to be explained later). *Collapse* is played by manipulating finite sequences of integers, called *words*, using three rules. Before we introduce the rules, we need some notation.

For an alphabet  $\mathfrak{A}$ , a *word*  $w$  is one of the following:

a finite list of symbols  $w = w_1w_2w_3 \cdots w_n$  (*finite word*),

an infinite list of symbols  $w = w_1w_2w_3 \cdots$  (*infinite word*),

or a bi-infinite list of symbols  $w = \cdots w_{-1}w_0w_1w_2w_3 \cdots$  (*bi-infinite word*).

The  $w_i$  in each of these cases are called the *letters* of  $w$ . The number of letters of a finite word  $w$  is called the *length* of  $w$  and denoted by  $|w|$ .

A *subword* of the word  $w = w_1w_2w_3 \cdots$  is a finite word  $u = w_kw_{k+1}w_{k+2} \cdots w_{k+n}$  composed of a contiguous segment of  $w$  and denoted by  $u \subseteq w$ .

For words  $u = u_1u_2 \cdots u_m$  and  $w = w_1w_2 \cdots w_n$ , the *concatenation* of  $u$  and  $w$  is the word

$$uw = u_1u_2 \cdots u_mw_1w_2 \cdots w_n.$$

Similarly,  $w^k$  is the concatenation of  $w$  with itself  $k$  times.

---

\*I would like to acknowledge the University of Victoria for their support and Professor Robert Burton for introducing me to the structures in the Fibonacci word.

**Definition 1.** If  $w$  is a finite word, we denote by  $w^*$  the bi-infinite periodic word formed by repeating  $w$ , e.g.

$$(123)^* = \cdots 123123 \cdots .$$

When writing  $w^*$  we will always assume  $w$  to be as short as possible, e.g. instead of  $(1212)^*$  we would write  $(12)^*$ .

The goal of *Collapse* is to take a finite word on the alphabet  $\mathfrak{A} = \mathbb{N}$ , called a *starting word* (we explain which words will be allowed in a moment), and reduce it, using only the rules listed below, to a word of the form  $(r)^*$  where  $r$  is a single letter.

**Rules.**

- (i) You may replace any letter  $n$  with the letter  $n + 1$  followed by the letter  $n + 2$ . For example,  $434 \rightarrow 4454$  is a valid application.
- (ii) You may replace any pair of consecutive letters of the form  $n + 1$  followed by  $n + 2$  with the single letter  $n$ . For example,  $4454 \rightarrow 434$  is a valid application.
- (iii) You may replace a word  $S$  (or  $S^*$ ) by a periodic word  $T^*$  with  $S \subset T^*$  (respectively  $S^* \subseteq T^*$ ) as long as  $|T| \leq |S|$ . For example,  $3453 \rightarrow (345)^*$  or  $(345)^* \rightarrow (534)^*$  are valid applications.

To apply rule (i) or (ii) to a bi-infinite periodic word  $X^*$ , one should apply the rule to every occurrence of  $X$ . For example  $(343)^* \rightarrow (4543)^*$ .

A few samples of how to apply the above rules:

$$\begin{aligned} 5667 &\xrightarrow{(ii)} 45 \xrightarrow{(ii)} 3 \xrightarrow{(iii)} (3)^* \\ 656 &\xrightarrow{(iii)} (65)^* \xrightarrow{(iii)} (56)^* \xrightarrow{(ii)} (4)^* \\ 233 &\xrightarrow{(i)} 3433 \xrightarrow{(iii)} (343)^* \xrightarrow{(ii)} (23)^* \xrightarrow{(ii)} (1)^* \end{aligned}$$

**Exercise.** Reduce 64556.

You may have noticed that rule (iii) has some interesting consequences. For example, we can eliminate any prefix that also appears as a suffix, because if  $x$  and  $y$  are finite words then the word  $xyx$  is contained in the periodic word  $(xy)^*$ . Further, we may cyclically permute the letters since  $xy$  is contained in  $(yx)^*$ . It is worth noting that using rule (iii) in different ways can change the eventual solution, for example:

$$\begin{aligned} 343 &\rightarrow (343)^* \rightarrow (23)^* \rightarrow (1)^* \\ 343 &\rightarrow (34)^* \rightarrow (2)^* \end{aligned}$$

Let us now go about considering the starting words. The most naive way to obtain a word that can be reduced using the rules of *Collapse* is to start with any integer and expand it to a word using rule (i) repeatedly. For example, the rows of Figure 1 show the words obtained from 0 by repeatedly applying rule (i) to each letter.

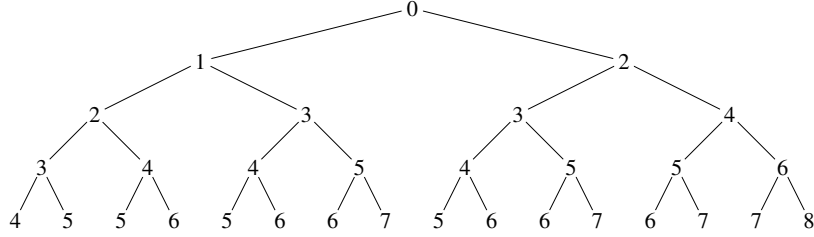


Figure 1: Applications of rule (i) to obtain starting words.

Of course, if these were the only starting words, it would make *Collapse* a fairly boring game since you could always win by using only rule (ii). However, a small modification will give this game a rich flavour.

**Definition 2.** A *starting word* of *Collapse* is any subword of the word  $X$  where  $X$  is obtained by repeatedly applying rule (i) to the word 0.

For example, if  $R_n$  is the  $n$ th row of the tree in Figure 1, any subword of  $R_n$  is a valid starting word.

**Exercise.** Reduce the following words: 33, 455, 6756, 667677.

Now that you are familiar with *Collapse*, we present the following questions.

**Question A.** Can all starting words be reduced?

**Question B.** Which starting words can be reduced in more than one way?

## 1 A Sturmian Excursion

It turns out that we can answer questions about *Collapse* by understanding certain properties of Sturmian words (or Sturmian sequences). The precise relationship between Sturmian words and *Collapse* is given in Section 3.

**Definition 3.** For any real number  $\alpha \in [0, 1]$ , the *characteristic Sturmian word*  $\mathcal{S}_\alpha = a_1 a_2 \dots$  is defined by

$$a_i = \lfloor (i+1)\alpha \rfloor - \lfloor i\alpha \rfloor.$$

An alternative definition for  $\mathcal{S}_\alpha$  is the following.

**Definition 4.** For  $\alpha \in [0, 1]$ , let  $[0; 1 + a_1, a_2, \dots]$  be the continued fraction expansion of  $\alpha$  with the convention that if  $\alpha \in \mathbb{Q}$ , then  $a_n = \infty$  is the last letter of the expansion. The (finite) standard Sturmian words of  $\mathcal{S}_\alpha$  are the finite words  $\mathcal{S}_i$  defined by

$$\begin{aligned}\mathcal{S}_0 &= 1, \\ \mathcal{S}_1 &= 0, \\ &\vdots \\ \mathcal{S}_n &= (\mathcal{S}_{n-1})^{a_{n-1}} \mathcal{S}_{n-2},\end{aligned}$$

and  $\mathcal{S}_\alpha = \lim_{i \rightarrow \infty} \mathcal{S}_i$ .

It should be noted that for  $\alpha \in [0, 1]$ , there is a larger set, the set of *all* Sturmian words associated with  $\alpha$ . These words arise as limits of  $\mathcal{S}_\alpha$ , but are not needed for our investigation. We refer the curious reader to Lothaire [3] and Fogg [2]. Now, considering the number of subwords of  $\mathcal{S}_\alpha$  of a particular length, we get the following interesting property.

**Proposition 5** (Fogg [2]). *Let  $\alpha \in [0, 1]$ . Then the number of distinct subwords of  $\mathcal{S}_\alpha$  of length  $n$  is at most  $n + 1$ , with equality if  $\alpha \notin \mathbb{Q}$ .*

We can say even more if  $n$  is the length of a standard Sturmian word of  $\mathcal{S}_\alpha$ .

**Proposition 6** (de Luca [4]). *All rotates (cyclic permutations) of a standard Sturmian word are distinct.*

*Proof.* Let  $p_i$  denote the number of 1's in  $\mathcal{S}_i$  and  $q_i$  the length of  $\mathcal{S}_i$ . The recursion for  $\mathcal{S}_i$  implies the following recursions for  $p_i$  and  $q_i$ :

$$\begin{aligned}p_0 &= 1, & q_0 &= 0, \\ p_1 &= 0, & q_1 &= 1, \\ p_n &= a_{n-1}p_{n-1} + p_{n-2}, & q_n &= a_{n-1}q_{n-1} + q_{n-2}.\end{aligned}$$

It is easy to verify by induction that  $|p_i q_{i-1} - p_{i-1} q_i| = 1$  and so  $\gcd(p_i, q_i) = 1$  for all  $i$ . Now suppose that  $\mathcal{S}_i$  can be written as a nontrivial rotate of itself. Then  $\mathcal{S}_i = w^k$  for some word  $w$  and  $k > 1$ . But then  $k$  divides both  $p_i$  and  $q_i$ , a contradiction.  $\square$

**Proposition 7.** *Let  $\alpha \in [0, 1] \setminus \mathbb{Q}$ . Then every rotate of the standard Sturmian word  $\mathcal{S}_i$  appears in  $\mathcal{S}_\alpha$  infinitely often.*

*Proof.* If  $i \leq 1$ , then the statement holds trivially. For  $i > 1$ , we will show that every rotate of  $\mathcal{S}_i$  appears in  $\mathcal{S}_{i+3}$  and thus in  $\mathcal{S}_\alpha$ . Using the recursion for Sturmian words, we have

$$\mathcal{S}_{i+3} = (\mathcal{S}_{i+2})^{a_{i+2}} \mathcal{S}_{i+1} = (\mathcal{S}_{i+2})^{a_{i+2}-1} (\mathcal{S}_{i+1})^{a_{i+1}} \mathcal{S}_i (\mathcal{S}_i)^{a_i} \mathcal{S}_{i-1}.$$

Since  $a_i \geq 1$  for  $i > 1$ , the word  $\mathcal{S}_i \mathcal{S}_i$ , which contains every rotate of  $\mathcal{S}_i$ , is a subword of  $\mathcal{S}_{i+3}$ . Lastly it is clear from the Sturmian recursion that  $\mathcal{S}_{i+3}$  appears infinitely often in  $\mathcal{S}_\alpha$ .  $\square$

This proof can be modified, using slightly more arithmetic, to show that all rotates of  $\mathcal{S}_i$  do in fact appear in  $\mathcal{S}_{i+2}$ . We remark that the previous proposition also holds for rational  $\alpha \in [0, 1]$  with the obvious exception that  $\mathcal{S}_1 = 0$  (respectively  $\mathcal{S}_1 = 1$ ) does not appear in  $\mathcal{S}_\alpha$  if  $\alpha = 1$  (respectively  $\alpha = 0$ ). As a corollary of the previous propositions, we observe that if  $n = |\mathcal{S}_i|$  is the length of a standard Sturmian word for  $\alpha \in [0, 1] \setminus \mathbb{Q}$ , then of the  $n + 1$  subwords of length  $n$  appearing in  $\mathcal{S}_\alpha$ , exactly one is not a rotate of  $\mathcal{S}_i$ . These words will be crucial in the investigation of *Collapse*.

## 2 A Fibonacci Analogy

Let us examine the characteristic Sturmian word that will be our main concern in this paper. The *infinite Fibonacci word* is the word  $\mathcal{F} = \mathcal{S}_\varphi$ , where  $\varphi = \left(\frac{1-\sqrt{5}}{2}\right)^2$  is the square of the inverse of the Golden Ratio. In this case, the standard Sturmian words generating  $\mathcal{F}$  are the Fibonacci words  $\mathcal{F}_n$  given by

$$\begin{aligned}\mathcal{F}_0 &= 1, \\ \mathcal{F}_1 &= 0, \\ &\vdots \\ \mathcal{F}_n &= \mathcal{F}_{n-1}\mathcal{F}_{n-2}.\end{aligned}$$

Notice that  $f_n = |\mathcal{F}_n|$  satisfies the recursion for the Fibonacci numbers with  $f_0 = f_1 = 1$ . The first few Fibonacci words are:

$$\begin{aligned}\mathcal{F}_0 &= 1, \\ \mathcal{F}_1 &= 0, \\ \mathcal{F}_2 &= 01, \\ \mathcal{F}_3 &= 010, \\ \mathcal{F}_4 &= 01001, \\ \mathcal{F}_5 &= 01001010.\end{aligned}$$

There is a strong connection between *Collapse* and Fibonacci words. Consider the word given by the  $n$ th row of the tree in Figure 1 (the row starting with the letter  $n$ ) and use rule (i) repeatedly on all letters less than  $2n - 1$  until all letters of the word are either  $2n - 1$  or  $2n$ . Now replace each  $2n - 1$  by 0 and each  $2n$  by 1. The result is the Fibonacci word  $\mathcal{F}_{2n}$ !

**Example.** The word 2334 becomes 34334 and upon replacement,  $01001 = \mathcal{F}_2$ .

Let us now consider this relationship with a little more rigour. For an alphabet  $\mathfrak{A}$ , let

$$\mathcal{W}_{\mathfrak{A}} = \bigcup_{n \in \mathbb{N}} \mathfrak{A}^n$$

be the set of all finite words on  $\mathfrak{A}$ .

Define

$$\begin{aligned} \Phi_k : \mathcal{W}_{\{k-1, k\}} &\rightarrow \mathcal{W}_{\{0, 1\}} && \text{by} \\ (k-1) &\mapsto 0 \\ k &\mapsto 1 \end{aligned}$$

and

$$\begin{aligned} T_k : \mathcal{W}_{\{k-1, k\}} &\rightarrow \mathcal{W}_{\{k, k+1\}} && \text{by} \\ k-1 &\mapsto k, (k+1) \\ k &\mapsto k. \end{aligned}$$

Notice that both  $\Phi_k$  and  $T_k$  have as a domain words consisting of  $k-1$  and  $k$  and that  $T_k$  is in bijection with its image. Further, applying  $T_k$  is analogous to applying rule (i) to every letter  $k-1$  in a word (consisting of  $(k-1)$ 's and  $k$ 's). Since  $\Phi_k$  is invertible, we have an induced map

$$T = \Phi_{k+1} \circ T_k \circ \Phi_k^{-1}$$

given by  $0 \mapsto 01$  and  $1 \mapsto 0$ , and the following diagram commutes.

$$\begin{array}{ccccccc} \cdots & \xrightarrow{T_{k-1}} & \mathcal{W}_{\{k-1, k\}} & \xrightarrow{T_k} & \mathcal{W}_{\{k, k+1\}} & \xrightarrow{T_{k+1}} & \cdots \\ & & \downarrow \Phi_k & & \downarrow \Phi_{k+1} & & \\ \cdots & \xrightarrow{T} & \mathcal{W}_{\{0, 1\}} & \xrightarrow{T} & \mathcal{W}_{\{0, 1\}} & \xrightarrow{T} & \cdots \end{array}$$

Figure 2: Commutation of  $T_k$  and  $\Phi_k$

**Proposition 8.** *If  $\mathcal{F}_n$  is the  $n$ th Fibonacci word,  $T(\mathcal{F}_n) = \mathcal{F}_{n+1}$ .*

*Proof.* The proof by induction is straightforward. Observe that  $T(\mathcal{F}_0) = T(1) = 0 = \mathcal{F}_1$ . Now, suppose the claim holds for  $i < n$ . Notice if  $w = ab$  is a word, then  $T(w) = T(a)T(b)$ , and so

$$T(\mathcal{F}_n) = T(\mathcal{F}_{n-1}\mathcal{F}_{n-2}) = T(\mathcal{F}_{n-1})T(\mathcal{F}_{n-2}) = \mathcal{F}_n\mathcal{F}_{n-1} = \mathcal{F}_{n+1}. \quad \square$$

We now have another way of writing  $\mathcal{F}_n$ , namely  $\mathcal{F}_n = T^n(1)$ . On the other hand, consider the word  $L_n$ , arising by applying rule (i) to 0 until it consists only of the letters  $n-1$  and  $n$ . Observe that

$$L_n = T_{n-1} \circ \cdots \circ T_1 \circ T_0(0) = \Phi_n^{-1} \circ T^n(1).$$

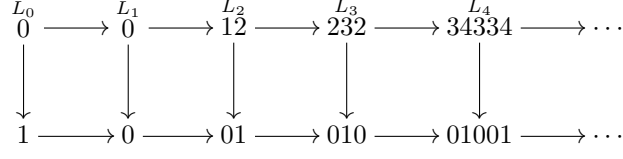


Figure 3: Link between  $L_n$  and  $\mathcal{F}_n$

It is clear that any starting word of *Collapse* corresponds to a subword of some  $L_k$  via repeated application of rule (i), so by answering questions about subwords of  $\mathcal{F}_k$ , we can, by the correspondence afforded by  $\Phi_k$ , answer questions about *Collapse*.

### 3 A Unifying Connection

In *Collapse*, the same starting word may have many different ways in which it might be reduced. In order to exploit the link to the Fibonacci words that we established in the previous section, we will provide a standard procedure, given in Proposition 13, of how to reduce starting words. To lead up to it, we need to establish a few properties of the rules of *Collapse*.

First, we define what it means for applications of rule (i) or (ii) to happen in descending order (conversely ascending order). Intuitively, descending order means applications of the rule first apply to the largest pairs of letters, then the second largest, etc. To be more precise, let  $\cdots \circ \rho_3 \circ \rho_2 \circ \rho_1$  be a sequence of applications of rule (i) (or (ii)). This sequence is in *descending order*, if for all  $i$  the largest letter  $\rho_i$  operates on is greater than or equal to the largest letter  $\rho_{i+1}$  operates on; the sequence is in *ascending order* if the largest letter  $\rho_i$  operates on is less than or equal to the largest letter  $\rho_{i+1}$  operates on.

Observe that the rules of *Collapse* have the following exchange properties.

Two consecutive applications of rule (i) in descending order can be replaced by two consecutive applications of rule (i) in ascending order.

Two consecutive applications of rule (ii) in ascending order can be replaced by two consecutive applications of rule (ii) in descending order.

Two consecutive applications of rule (iii) can be replaced by a single application of rule (iii).

An application of rule (ii) followed by an application of rule (i) can be replaced by an application of rule (i) followed by an application of rule (ii).

An application of rule (iii) followed by an application of rule (i) can be replaced by some number of applications of rule (i) followed by an application of rule (iii).

An application of rule (ii) followed by an application of rule (iii) can be replaced by some number of applications of rule (i) followed by an application of rule (iii) and then followed by some number of applications of rule (ii).

We will show the last exchange property in a moment. The proofs of the first five exchange properties as well as of the following Lemmas we will leave to the reader as an easy exercise, but we will give an example for why we might need multiple applications of rule (i) in the fifth property. For that, consider  $454 \xrightarrow{(iii)} (45)^* \xrightarrow{(i)} (565)^*$ . To reverse the order of the rules applied, we can proceed as follows:  $454 \xrightarrow{(i)} 5654 \xrightarrow{(i)} 56556 \xrightarrow{(iii)} (565)^*$ . Note that we needed to replace two occurrences of the letter 4 using rule (i) before being able to apply rule (iii).

**Lemma 9.** *For words  $S$  and  $T$ , suppose  $S \xrightarrow{(iii)} T^*$  and  $S'$  and  $T'$  arise from  $S$  and  $T$  by applying rule (i) to every occurrence of the letter  $k$  in each word. Then,  $S' \xrightarrow{(iii)} (T')^*$ .*

We can use Lemma 9 to show the last exchange property. Suppose we have  $W \xrightarrow{(ii)} S \xrightarrow{(iii)} T$ , where the application of rule (ii) replaces two consecutive letters  $k+1, k+2$  by the letter  $k$ . Let  $S'$  and  $T'$  be defined as in Lemma 9. Then we can find  $W \xrightarrow{(i)} \dots \xrightarrow{(i)} S' \xrightarrow{(iii)} (T')^* \xrightarrow{(ii)} \dots \xrightarrow{(ii)} T^*$ .

**Proposition 10.** *If the starting word  $X$  can be fully reduced by the rules of Collapse, then it can be reduced by first applying rule (i) a number of times in ascending order, then rule (iii) once, and then rule (ii) a number of times in descending order.*

*Proof.* Observing the properties stated at the beginning of this section, we note that all applications of rule (i) can be moved ahead of applications of rules (ii) and (iii) and all applications of rule (iii) ahead of applications of rule (ii) by inserting some additional applications of rule (i) if necessary. Finally all applications of rule (i) can be rearranged in ascending order, all applications of rule (ii) in descending order, and all applications of rule (iii) combined to a single application.  $\square$

**Lemma 11.** *If  $X \in \mathcal{W}_{\{k-1, k\}}$  and  $X$  reduces to the letter  $r$  using only rule (ii), then  $X = T_{k-1} \circ \dots \circ T_r(r)$ .*

**Lemma 12.** *Let  $S$  be a starting word of Collapse with largest letter  $k$ . By repeatedly using rule (i) on the smallest letters,  $S$  can be expanded to a subword of  $L_{k+1}$ .*

Note that  $L_k$  is not sufficient in Lemma 12, as for example 233 is not a subword of  $L_3$ , whereas its expansion 3433 is a subword of  $L_4$ .

We will now limit our quest to understand *Collapse* to understanding the reduction of subwords of  $L_k$ . This is justified by the following algorithm sketch.

**Proposition 13.** *If a starting word  $S$  can be reduced to  $(r)^*$ , then for suitable  $k$  and  $m = k - r$  it can be reduced by the following steps.*

- (1) *Use rule (i) in ascending order to expand  $S$  to  $L \subseteq L_k$ .*
- (2) *Use rule (iii) to rewrite  $L$  as  $(X)^*$ , where  $X = \Phi_k^{-1}(\mathcal{F}_m)$ .*
- (3) *Use rule (ii) in descending order to reduce  $(X)^*$  to  $(r)^*$ .*

*Proof.* By Proposition 10, the starting word  $S$  can be reduced to  $(r)^*$  in the form

$$S \xrightarrow{(i)} \dots \xrightarrow{(i)} S' \xrightarrow{(iii)} (S'')^* \xrightarrow{(ii)} \dots \xrightarrow{(ii)} (r)^*.$$

By Lemmas 9 and 12, we can expand  $S'$  to a word  $L \subseteq L_k$  and  $S''$  to a word  $X$  with  $L \xrightarrow{(iii)} X^*$ . This gives us the reduction

$$S \xrightarrow{(i)} \dots \xrightarrow{(i)} L \xrightarrow{(iii)} X^* \xrightarrow{(ii)} \dots \xrightarrow{(ii)} (r)^*,$$

where it may be assumed that all rule (i)'s are in ascending order and all rule (ii)'s are in descending order. Finally, since  $X \in \mathcal{W}_{\{k-1, k\}}$ , we obtain from Lemma 11 that

$$X = T_{k-1} \circ T_{k-2} \cdots \circ T_{k-m}(r),$$

which is equivalent (as shown by Figure 2) to

$$\Phi_k(X) = T^m(\Phi_{k-m}(r)) = T^m(\mathcal{F}_0) = \mathcal{F}_m. \quad \square$$

We observe that Proposition 13 tells us nothing about how to find  $k$  and  $m$  nor whether  $m$  even exists without the assumption that a starting word reduces to  $(r)^*$ . This leads to the following question.

**Question C.** Suppose  $w$  is a finite subword of  $\mathcal{F}$ . Does there exist an  $m$  so that  $|w| \geq |\mathcal{F}_m|$  and  $w \subset (\mathcal{F}_m)^*$ ?

If the answer for Question C is positive for all words  $w$ , then any admissible  $k$  in Proposition 13 would give us a possible reduction of a starting word, answering Question A in the affirmative.

## 4 An Exceptional Investigation

We recall from Section 2 that the infinite Fibonacci word  $\mathcal{F}$ , being a characteristic Sturmian word, contains  $r + 1$  distinct words of length  $r$ . If  $r = f_n$  is a Fibonacci number, then  $r$  of those words must be cyclic rotations of  $\mathcal{F}_n$ . The remaining word will be called the *Exceptional word*  $\mathcal{E}_n$  (in the general context of Sturmian words, these are called *singular* words). The first few Exceptional words are:

$$\begin{aligned}
\mathcal{E}_0 &= 0, \\
\mathcal{E}_1 &= 1, \\
\mathcal{E}_2 &= 00, \\
\mathcal{E}_3 &= 101, \\
\mathcal{E}_4 &= 00100, \\
\mathcal{E}_5 &= 10100101.
\end{aligned}$$

To analyze the Exceptional words, we introduce some arithmetic on finite words.

**Definition 14.** For a finite word  $w \in \mathcal{W}_{\{0,1\}}$ , the word arising from  $w$  by

- (i) rotating the first letter to the end is denoted by  $w^\frown$ ,
- (ii) rotating the last letter to the front is denoted by  $w^\circleftarrow$ ,
- (iii) changing the first letter (0 to 1 or vice versa) is denoted by  $w^\times$ .

As an example, if  $w = 10100$  then  $w^\frown = 01001$ ,  $w^\circleftarrow = 01010$ , and  $w^\times = 00100$ . The following lemmas show a few properties of the Exceptional words that we will be using later.

**Lemma 15** (Wen & Wen [5]). *For any  $n \geq 0$ :*

- (i)  $\mathcal{E}_n = (\mathcal{F}_n^\frown)^\times$ ,
- (ii)  $\mathcal{F}_n = (\mathcal{E}_n^\times)^\circleftarrow$ ,
- (iii)  $\mathcal{F}_{n+2} = (\mathcal{E}_{n+1}\mathcal{E}_n)^\circleftarrow$ ,
- (iv)  $\mathcal{E}_{n+3} = \mathcal{E}_{n+1}\mathcal{E}_n\mathcal{E}_{n+1}$ .

*Proof.* (i) We know that  $\mathcal{F}_{n+2} = \mathcal{F}_{n+1}\mathcal{F}_n$ . Consider the subword of length  $f_n$  starting with the last letter of  $\mathcal{F}_{n+1}$  and continuing with the first  $f_n - 1$  letters of  $\mathcal{F}_n$ . Since  $\mathcal{F}_n$  and  $\mathcal{F}_{n+1}$  end on different letters, this word is  $(\mathcal{F}_n^\frown)^\times$  and thus not a rotate of  $\mathcal{F}_n$ . Therefore it must be the Exceptional word  $\mathcal{E}_n$ .

(ii) This follows directly from part (i) and the definitions of the arithmetic notations.

(iii) We have

$$\mathcal{F}_{n+2} = \mathcal{F}_{n+1}\mathcal{F}_n = (\mathcal{E}_{n+1}^\times)^\circleftarrow(\mathcal{E}_n^\times)^\circleftarrow = (\mathcal{E}_{n+1}\mathcal{E}_n)^\circleftarrow,$$

where the last equality follows from the fact that the first letters of  $\mathcal{E}_{n+1}$  and  $\mathcal{E}_n$  are different (since the last letters of  $\mathcal{F}_n$  and  $\mathcal{F}_{n+1}$  are different).

(iv) Using the previous parts of this lemma, we have

$$\begin{aligned}
\mathcal{E}_{n+3} &= (\mathcal{F}_{n+3})^\times = ((\mathcal{F}_{n+2}\mathcal{F}_{n+1})^\wedge)^\times = ((\mathcal{F}_{n+1}\mathcal{F}_n\mathcal{F}_{n+1})^\wedge)^\times \\
&= (((\mathcal{E}_{n+1}^\times)^\wedge(\mathcal{E}_n^\times)^\wedge(\mathcal{E}_{n+1}^\times)^\wedge)^\times)^\times = (((\mathcal{E}_{n+1}\mathcal{E}_n\mathcal{E}_{n+1})^\times)^\wedge)^\times \\
&= \mathcal{E}_{n+1}\mathcal{E}_n\mathcal{E}_{n+1}. \quad \square
\end{aligned}$$

**Lemma 16.** *For any  $n \geq 0$ , any occurring copies of  $\mathcal{E}_n$  and  $\mathcal{E}_{n+1}$  in a word  $w$  are disjoint.*

*Proof.* We use induction on  $n$ . The statement is clearly true for  $n \leq 1$ , so assume  $n \geq 2$ . Suppose there exist copies of  $\mathcal{E}_n$  and  $\mathcal{E}_{n+1} = \mathcal{E}_{n-1}\mathcal{E}_{n-2}\mathcal{E}_{n-1}$  that are not disjoint. Then  $\mathcal{E}_n$  (being longer than  $\mathcal{E}_{n-2}$ ) is also not disjoint from at least one of the two copies of  $\mathcal{E}_{n-1}$ , a contradiction.  $\square$

**Lemma 17.** *If  $T$  is the replacement function on words in  $\mathcal{W}_{\{0,1\}}$  defined by  $0 \mapsto 01$  and  $1 \mapsto 0$  (as in Section 2), then*

- (i)  $T(\mathcal{E}_{2n}1) = 0\mathcal{E}_{2n+1}0$ ,
- (ii)  $T(\mathcal{E}_{2n-1}0) = \mathcal{E}_{2n}1$ ,
- (iii)  $T^2(\mathcal{E}_{2n}1) = 01\mathcal{E}_{2n+2}1$ ,
- (iv)  $T^2(\mathcal{E}_{2n-1}0) = 0\mathcal{E}_{2n+1}0$ .

*Proof.* Using Lemma 15 and the fact that  $T(\mathcal{F}_n) = \mathcal{F}_{n+1}$  for all  $n$ , we see that

$$T(\mathcal{E}_{2n}1) = T(0\mathcal{F}_{2n}) = 01\mathcal{F}_{2n+1} = 0\mathcal{E}_{2n+1}0,$$

and

$$T(\mathcal{E}_{2n-1}0) = T(1\mathcal{F}_{2n-1}) = 0\mathcal{F}_{2n} = \mathcal{E}_{2n}1.$$

The remaining equalities follow.  $\square$

Returning to Question C, we want to know, when a word  $w \subset \mathcal{F}$  is contained in a repeated Fibonacci word. While not required for Question C, we will include the following lemma for a later purpose.

**Lemma 18.** *Every finite word  $w \subset \mathcal{F}$  with  $|w| < f_n$  is a subword of  $(\mathcal{F}_n)^*$ .*

*Proof.* By Proposition 7,  $w$  appears infinitely often in  $\mathcal{F}$ . Therefore we can find a word  $w' \subset \mathcal{F}$  of length  $f_n$  that ends with  $w$ . This word has to be either  $\mathcal{E}_n$  or a rotate of  $\mathcal{F}_n$ . In the latter case we are done. If the former holds, then  $(w')^\times$  is a rotate of  $\mathcal{F}_n$  by Lemma 15, part (ii). However, since  $|w'| > |w|$ , we know that  $(w')^\times$  still ends with  $w$ , and so  $w \subset (\mathcal{F}_n)^*$ .  $\square$

**Theorem 19.** *Let  $w$  be a finite subword of  $\mathcal{F}$ . Then  $w \subset (\mathcal{F}_n)^*$  if and only if  $\mathcal{E}_n \not\subseteq w$ .*

*Proof.* Since  $(\mathcal{F}_n)^*$  does not contain  $\mathcal{E}_n$ , any word  $w$  containing  $\mathcal{E}_n$  cannot be contained in  $(\mathcal{F}_n)^*$ . Suppose that  $w$  does not contain  $\mathcal{E}_n$ . If  $|w| < f_n$ , then we are done by Lemma 18. Otherwise  $|w| \geq f_n$ . Since  $\mathcal{E}_n \not\subseteq w$ , any  $f_n$  consecutive letters of  $w$  must be a rotate of  $\mathcal{F}_n$ , and therefore it must have the same number of 1's. Comparing any consecutive subwords of length  $f_n$ , we see that the letters of  $w$  in positions  $k$  and  $k + f_n$  must be the same for all  $k$ . Therefore,  $w$  is formed by repeating the first  $f_n$  letters and is thus contained in  $(\mathcal{F}_n)^*$ .  $\square$

All that remains to be shown is that for every word  $w \in \mathcal{F}$ , there exists an  $n$  with  $|w| \geq f_n$  such that  $\mathcal{E}_n \not\subseteq w$ .

**Proposition 20.** *If  $w \in \mathcal{F}$  with  $f_n \leq |w| < f_{n+1}$ , then either  $\mathcal{E}_n \not\subseteq w$  or  $\mathcal{E}_{n-1} \not\subseteq w$ .*

*Proof.* Suppose  $w$  contains both  $\mathcal{E}_n$  and  $\mathcal{E}_{n-1}$ . By Lemma 16, part (v),  $\mathcal{E}_n$  and  $\mathcal{E}_{n-1}$  are disjoint and so  $|w| \geq |\mathcal{E}_n| + |\mathcal{E}_{n-1}| = f_{n+1}$ , a contradiction.  $\square$

Combining Theorem 19 and Proposition 20 gives us Theorem 21, which answers Question C and in turn Question A in the affirmative.

**Theorem 21** (Currie & Saari [1]). *If  $w \in \mathcal{F}$ , then there exists an  $\mathcal{F}_n$  with  $|\mathcal{F}_n| \leq |w|$  such that  $w \in (\mathcal{F}_n)^*$ .*  $\square$

It should be noted that the proof presented here is quite different from that in [1]. In particular, our use of Exceptional words in Theorem 19 will allow a complete description of reductions.

## 5 A Unique Solution

For any  $w \in \mathcal{F}$ , we know we can find an  $n$  with  $|w| \geq f_n$  and  $w \in (\mathcal{F}_n)^*$ , but that still leaves Question B, concerning uniqueness of reductions, open. We start by answering the following.

**Question D.** Suppose  $w$  is a finite subword of  $\mathcal{F}$ . For which  $n$  is  $w \in (\mathcal{F}_n)^*$ ?

In light of Theorem 19, we need to know which Exceptional words  $\mathcal{E}_n$  are contained in  $w$ .

**Proposition 22.** *Let  $w$  be a finite subword of  $\mathcal{F}$  with  $f_n \leq |w| < f_{n+1}$ . Then the Exceptional words contained in  $w$  are precisely  $\mathcal{E}_i$  for  $i \leq n - 3$  and either*

- (a)  $\mathcal{E}_n$  and  $\mathcal{E}_{n-2}$ ,
- (b)  $\mathcal{E}_{n-1}$ ,
- (c)  $\mathcal{E}_{n-1}$  and  $\mathcal{E}_{n-2}$ ,
- (d)  $\mathcal{E}_{n-2}$ .

*Proof.* Clearly  $w$  cannot contain  $\mathcal{E}_i$  for  $i \geq n+1$ . Notice that if  $w$  contains both  $\mathcal{E}_i$  and  $\mathcal{E}_{i-1}$  for some  $i$ , then  $w$  contains  $\mathcal{E}_j$  for all  $j \leq i$ , which follows from the recursion for the Exceptional words (Lemma 15, part (iv)).

By Lemma 16,  $w$  cannot contain both  $\mathcal{E}_n$  and  $\mathcal{E}_{n-1}$ . Suppose  $\mathcal{E}_n \subset w$ . Then by the recursion for Exceptional words,  $w$  contains  $\mathcal{E}_{n-2}$  and  $\mathcal{E}_{n-3}$ , so case (a) applies.

If  $\mathcal{E}_{n-1} \subseteq w$ , then both  $\mathcal{E}_{n-3}$  and  $\mathcal{E}_{n-4}$  are also contained in  $w$ . Thus, either case (b) or (c) applies.

Finally, assume  $\mathcal{E}_n, \mathcal{E}_{n-1} \not\subseteq w$ . By Lemma 18,  $w \subset (\mathcal{F}_{n+1})^*$ . Further,  $(\mathcal{F}_{n+1})^* = (\mathcal{E}_n \mathcal{E}_{n-1})^*$  and so  $w \subset \mathcal{E}_n \mathcal{E}_{n-1}$  or  $w \subset \mathcal{E}_{n-1} \mathcal{E}_n$ . We will only consider the case  $w \subset \mathcal{E}_n \mathcal{E}_{n-1}$  since the argument for both cases is similar. Expanding, we have

$$\mathcal{E}_n \mathcal{E}_{n-1} = \overbrace{\mathcal{E}_{n-2} \mathcal{E}_{n-3}}^{\text{length } f_n} \underbrace{\mathcal{E}_{n-2} \mathcal{E}_{n-3} \mathcal{E}_{n-4} \mathcal{E}_{n-3}}_{\text{length } f_n}.$$

Since  $|w| \geq f_n$ , we conclude that  $\mathcal{E}_{n-2} \subset w$ . Upon closer inspection, we can also conclude that  $\mathcal{E}_{n-3} \subset w$ , and thus case (d) applies.  $\square$

All of the cases in Proposition 22 can in fact appear. As an example, we consider the  $f_n + 1$  subwords of  $\mathcal{F}$  of length  $f_n$ . Out of these, one word (the Exceptional word  $\mathcal{E}_n$ ) satisfies case (a),  $f_{n-2} - 1$  words satisfy case (b), two words (the words  $\mathcal{E}_{n-1} \mathcal{E}_{n-2}$  and  $\mathcal{E}_{n-2} \mathcal{E}_{n-1}$ ) satisfy case (c), and  $f_{n-1} - 1$  words satisfy case (d). Reformulating Proposition 22 using Theorem 19, we obtain the following answer to Question D.

**Theorem 23.** *Let  $w$  be a finite subword of  $\mathcal{F}$  with  $f_n \leq |w| < f_{n+1}$  and let  $R = \{i \leq n \mid w \subset (\mathcal{F}_i)^*\}$ . Then exactly one of the following cases holds:*

- (a)  $R = \{n - 1\}$ ,
- (b)  $R = \{n - 2, n\}$ ,
- (c)  $R = \{n\}$ ,
- (d)  $R = \{n - 1, n\}$ .

$\square$

To finally answer Question B, we need to investigate, what happens to the set of reductions if we apply the replacement function  $T$  from Section 2 to a word  $w \subset \mathcal{F}$  (multiple times, possibly). Clearly if  $w \subset (\mathcal{F}_i)^*$ , then  $Tw \subset (\mathcal{F}_{i+1})^*$ . Therefore, if for example a word  $w$  satisfies case (a) in Theorem 23, then  $Tw$  has to satisfy either case (a) or (d). On the other hand, if a word  $w$  satisfies case (b), then  $Tw$  also has to satisfy case (b). The following proposition shows that all possible reductions can be obtained by applying  $T$  just twice.

**Proposition 24.** *Let  $w$  be a finite subword of  $\mathcal{F}$ . Then the words  $T^k w$  for  $k \geq 2$  all satisfy the same case in Theorem 23.*

*Proof.* By the preceding discussion,  $T^k w$  never has fewer reductions than  $w$ . Since no word reduces in more than two ways (Theorem 23), we only need to consider words  $w \in \mathcal{F}$  that fall into either case (a) or (c) in Theorem 23. Consider case (a), where  $w$  contains the Exceptional word  $\mathcal{E}_n$ . Suppose  $\mathcal{E}_n$  does not appear at the end of  $w$ . Then by Lemma 17,  $T^k w$  contains  $\mathcal{E}_{n+k}$  and therefore falls into case (a) for all  $k$ . On the other hand, suppose  $w$  ends with  $\mathcal{E}_n$  (which must be the only occurrence of  $\mathcal{E}_n$ ). Then, again by Lemma 17,  $T^2 w$  does not contain  $\mathcal{E}_{n+2}$  and thus falls into case (d).

The same arguments can be made if  $w$  falls into case (c). If both  $\mathcal{E}_{n-1}$  and  $\mathcal{E}_{n-2}$  appear before the end of  $w$ , then  $T^k w$  will fall into case (c) for all  $k$ . If  $w$  ends on  $\mathcal{E}_{n-1}$  then  $T^k$  falls into case (d) for all  $k \geq 2$ , and if  $w$  ends on  $\mathcal{E}_{n-2}$  and this is the only occurrence of  $\mathcal{E}_{n-2}$  in  $w$ , then  $T^k w$  falls into case (b) for all  $k \geq 2$ .  $\square$

We can now state an algorithm that obtains all possible reductions of a starting word in *Collapse*.

**Theorem 25.** *Given a starting word in Collapse, all possible reductions can be found using the following algorithm.*

- (1) Use rule (i) to expand the starting word to a subword  $s$  of  $L_k$ , where  $k-3$  is the largest letter in the starting word. Set  $n$  such that  $f_n \leq |s| < f_{n+1}$ .
- (2) Check whether  $s \subset (X_i)^*$  ( $i = 0, 1, 2$ ), where  $X_i = \Phi_k^{-1}(\mathcal{F}_{n-i})$ .
- (3) In each case where the question is answered positively, apply rule (iii) to  $s$  to obtain  $(X_i)^*$  and successively apply rule (ii) to reduce this to  $(r)^*$ , where  $r = (k - n + i)$ .

*Proof.* Comparing the statement with Proposition 13, we only need to show that the choices of  $k$  (three larger than the largest letter in the starting word) and  $m = n - i$  give all possible reductions.

If  $k-3$  is the largest letter in the starting word, then by Lemma 12 we can expand the starting word to a subword of  $L_{k-2}$ . Let  $w$  be the corresponding word in  $\mathcal{W}_{\{0,1\}}$ . By Proposition 24, we can get all possible reductions from  $T^2 w$ , which corresponds to a subword  $s$  of  $L_k$ . By Theorem 23, the choices of  $m = n - i$  ( $i = 0, 1, 2$ ), are sufficient.  $\square$

## 6 An Ultimate Conclusion

We have fully described an algorithm for winning *Collapse*, but that doesn't take away from its fun. Knowing you can always win can even taunt you as you attempt to do so in a minimal number of moves, without resorting to the expand-everything approach presented. But, *Collapse* is just a single instance of an uncountable collection of games based on Sturmian words and the Sturmian recursion given in Definition 4.

In all these games, which can be created by looking at the subscripts of standard Sturmian words, reducibility of starting words is always implied by the following.

**Proposition 26.** *Let  $\alpha \in [0, 1] \setminus \mathbb{Q}$  and let  $w \in \mathcal{S}_\alpha$ . There exists a rational  $p/q$  with  $q \leq |w|$  so that  $w \in \mathcal{S}_{p/q}$ .*

Though we will not do so here, Proposition 26 can be proven with beautiful geometric arguments that look completely different from the combinatoric arguments presented here.

We hope that *Collapse* and its analysis will inspire others to play games with the recursive and beautiful structures we all stumble upon every day.

## References

- [1] J. D. Currie and K. Saari, *Least periods of factors of infinite words*, RAIRO-Theor. Inform. Appl. 43(1), 165-178, 2009.
- [2] N. P. Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Lecture Notes in Mathematics **1794**, Springer, Berlin, 2002.
- [3] M. Lothaire, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications 90. Cambridge University Press. 2002
- [4] A. de Luca, *A Division Property of the Fibonacci Word*, Inform. Process. Lett. **12** (1981), no. 4, 193-195.
- [5] Z.-X. Wen and Z.-Y. Wen, *Some properties of the singular words of the Fibonacci word*, European J. Combin. **15** (1994), 587-598.