TEAM B FINAL DRAFT

HILLEL DEI, ALIZAYE MANIGO, DAVID MORGAN

Abstract. In this paper, we consider a special type of sequence obtained by applying a procedure Φ_n of the form

$$1 \mapsto 1\underbrace{00\cdots0}_{n \text{ zeros}}$$
$$0 \mapsto 1\underbrace{00\cdots0}_{n-1 \text{ zeros}}.$$

In particular, we explore the behavior of binary sequences generated by rules of this form. We prove the existence of such sequences, that they are quasiperiodic, and determine the density of ones and zeros in them. We also provide computational methods to support some of the results.

David Morgan proved all lemmas and theorems in section 2, wrote section 2 and helped write the abstract and section 1. Hillel Dei helped with the proof of asymptotic density in section 5 and wrote the sections on computational methods (5 and 6). Alizaye Manigo proved the theorems and lemmas in section 3 and 4, wrote sections 3 and 4, and helped write the abstract and section 1.

1. Introduction

Consider the following binary sequence:

$$T = 10010010110$$

On this sequence, we perform the following replacement operation: replace each 1 with 100, and each 0 with 10. After applying the operations, we obtain the following sequence T':

$$T' = 10010101001010100101001001010$$

It is natural to wonder if such a procedure can be applied "in reverse"; that is, is it possible to recover T given T'? Indeed, it is. We introduce the concept of an inverse operation, which replaces each 100 with 1 and each 10 with 0. We attempt to recover T by applying this operation to T':

$$\begin{array}{rcccccccccccc} T' & = & \underbrace{100} & \underbrace{10} & \underbrace{10} & \underbrace{100} & \underbrace{10} & \underbrace{10} & \underbrace{100} & \underbrace{10} & \underbrace{100} & \underbrace{100} & \underbrace{10} \\ T & = & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{array}$$

We see that it is possible to recover T by applying an inverse operation to T'. If a sequence has an inverse, then we refer to it as a *lucky* sequence. In order to provide a more formal definition for this concept, we will first define our general replacement procedure.

Definition 1.1 (Replacement operations)**.** For $n \geq 2$, the replacement operation Φ_n is given by

$$\Phi_n\colon 1 \to 1\underbrace{00\cdots0}_{n \text{ zeros}}$$
$$0 \to 1\underbrace{00\cdots0}_{n-1 \text{ zeros}}.$$

The inverse operation is Φ_n^{-1} given by

$$\Phi_n^{-1}\colon 1\underbrace{00\cdots0}_{n \text{ zeros}} \to 1$$
$$1\underbrace{00\cdots0}_{n-1 \text{ zeros}} \to 0.$$

We also let Φ_n^m denote m applications of Φ_n to a sequence; similarly, we let Φ_n^{-m} denote m applications of Φ_n^{-1} to a sequence. In the example with the sequence T, we set $n = 2$. Now that we have defined our procedures, we provide a definition for lucky:

Definition 1.2 (Lucky sequences)**.** A *lucky sequence* is any sequence S for which the inverse sequence $\Phi_n^{-1}(S)$ exists for a given $n \geq 2$.

We see that by our definition, T' is lucky, since its inverse $\Phi_2^{-1}(T')$ exists. In general, it may be possible that Φ_2 and Φ_2^{-1} can be applied any number of times to a binary sequence.

Let us continue to explore this operation Φ_2 by returning to our example sequence T and attempting to find $\Phi_2^{-1}(T)$. We group the terms of T as we did with T' above.

$$T = \underbrace{100}_{1}\underbrace{100}_{1}\underbrace{10}_{0}\underbrace{110101}_{??}.$$

We run into an issue: no matter how one tries to group the terms, it is not possible to "invert" two consecutive ones since Φ_2 cannot produce two consecutive ones. Thus, T is *not* lucky by Definition 1.2.

Now let us try to apply Φ_2 in the other direction. If we were to repeatedly apply Φ_2 to T a total of m times, then we could say that our resulting sequence is *very very ... very* lucky. Henceforth we will call such a sequence m-lucky:

Definition 1.3 (Very lucky sequences)**.** An *m-lucky sequence* is any sequence S for which the sequence $\Phi_n^{-m}(S)$ exists for a given $n \geq 2$ and some $m \in \mathbb{Z}^+$.

Definition 1.4 (Ancestors)**.** Suppose $\Phi_n^{-m}(S)$ exists for a binary sequence S for some m. We say this inverse is an *ancestor* of S.

Suppose we had a sequence that was m-lucky for all m. Such a sequence could be thought of as infinitely lucky—for any m, there exists a sequence W such that applying the lucky procedure to it m times results in our sequence. We refer to such a sequence as *happy*.

Definition 1.5 (Happy sequences)**.** A *happy sequence* is any sequence S for which an inverse sequence $\Phi_n^{-m}(S)$ exists for all $m \in \mathbb{Z}^+$.

In this paper, we are concerned with happy sequences infinite in both directions under the replacement rule Φ_2. In section 2, we provide a proof for the existence of happy sequences. In section 3, we explore the structure of happy sequences and reveal that such sequences are quasiperiodic. In section 4, we compute and examine a closed form for the density of ones and zeros in lucky sequences, and provide a conjecture about the density of happy sequences. Finally, in section 5, we introduce computational methods that can easily generate lucky sequences under the generalized replacement rule Φ_n.

2. Existence of happy sequences

In this section, we will prove that happy sequences exist for Φ_2. We will show existence by first considering a related infinite binary sequence.

Lemma 2.1. *Consider an infinite binary sequence $(a_m)_{m\geq 0}$. Then,*

$$a_0 \subset \Phi_2(a_1) \subset \Phi_2^2(a_2) \subset \cdots \subset \Phi_2^m(a_m) \subset \Phi_2^{m+1}(a_{m+1}) \subset \cdots \tag{1}$$

where $\Phi_2^i(a_i) \subset \Phi_2^{i+1}(a_{i+1})$ means $\Phi_2^i(a_i)$ is embedded in $\Phi_2^{i+1}(a_{i+1})$; that is, $\Phi_2^{i+1}(a_{i+1})$ is formed by appending or prepending ones or zeros to $\Phi_2^i(a_i)$.

Proof. We will prove the lemma through casework. Consider two consecutive terms in the sequence a_m and a_{m+1}. Since (a_m) is a binary sequence, we know that both a_m and a_{m+1} could be either one or zero:

- $a_{m+1} = 1$: If we were to apply Φ_2 to "1" a total of $m+1$ times, we would have
$$\Phi_2^{m+1}(1) = \Phi_2^m(100) = \Phi_2^m(\underbrace{1}_{a_m=1})\Phi_2^m(\underbrace{0}_{a_m=0})\Phi_2^m(0),$$
where the second equation is a result of the linearity of our operation Φ_2. In the above expression, we see that both $\Phi_2^m(0)$ and $\Phi_2^m(1)$ work to comprise $\Phi_2^{m+1}(1)$. No matter the choice for a_m, $\Phi_2^{m+1}(1)$ is formed by appending or prepending ones and zeros to $\Phi_2^m(a_m)$. If $a_m = 0$, then one can see that either $\Phi_2^m(1)$ or $\Phi_2^m(1)\Phi_2^m(0)$ are prepended to $\Phi_2^m(0)$. Likewise, if $a_m = 1$, then $\Phi_2^m(0)$ is appended to $\Phi_2^m(1)$.

- $a_{m+1} = 0$: By applying Φ_2 to "0" as we did to "1" above, we obtain
$$\Phi_2^{m+1}(0) = \Phi_2^m(10) = \Phi_2^m(\underbrace{1}_{a_m=1})\Phi_2^m(\underbrace{0}_{a_m=0}).$$
Once again, both choices for $\Phi_2^m(a_m)$ comprise the expression for $\Phi_2^{m+1}(0)$. As above, we have that $\Phi_2^m(1)$ prepends $\Phi_2^m(0)$ if $a_m = 0$ and $\Phi_2^m(0)$ appends $\Phi_2^m(1)$ if $a_m = 1$.

Since we've shown that $\Phi_2^m(a_m) \subset \Phi_2^{m+1}(a_{m+1})$ for every case of a_m and a_{m+1}, we know that any two consecutive terms in (a_m) must satisfy this relationship. Thus, the lemma is true. □

We claim that this relationship between terms in a binary sequence is enough to show that happy sequences exist. More formally,

Theorem 2.2. *Given the replacement rule Φ_2, there exists a happy sequence.*

Proof. Consider an infinite sequence of ones and zeros $(b_m)_{m\geq 0}$. By Lemma 2.1, we have the following relationship:
$$b_0 \subset \Phi_2(b_1) \subset \Phi_2^2(b_2) \subset \cdots \subset \Phi_2^m(b_m) \subset \Phi_2^{m+1}(b_{m+1}) \subset \cdots$$
We argue that examining the above relationship for large enough m will reveal a happy sequence. Suppose we examined $\Phi_2^m(b_m)$ for large m. To show this sequence is happy, we need to show two things:

- **It is possible to always find an ancestor for $\Phi_2^m(b_m)$.**

- **The sequence $\Phi_2^m(b_m)$ will extend infinitely in both directions as m tends to ∞.**

To prove the first point, consider an arbitrary binary sequence S. We know that
$$\Phi_2(\Phi_2^{m-1}(S)) = \Phi_2^m(S) \iff \Phi_2^{m-1}(S) = \Phi_2^{-1}(\Phi_2^m(S)) \tag{2}$$
for any integer $m \geq 1$. The right hand side of (2) implies that $\Phi_2^m(S)$ has an inverse, and thus it has an ancestor by Definition 1.4. If we set $S = b_m$, then we can say $\Phi_2^m(b_m)$ has an inverse and thus an ancestor.

To show the second point, we examine the special infinite binary sequences
$$\left.\begin{aligned} (c_m) &= 1 \\ (d_m) &= 0 \end{aligned}\right\} \forall m \in \mathbb{Z}^+.$$
We can use Lemma 2.1 to write the following for each sequence
$$\underline{1} \subset \underline{1}00 \subset \underline{1}001010 \subset \underline{1}0010101001010010 \subset \cdots$$
$$\underline{0} \subset 1\underline{0} \subset 1001\underline{0} \subset 10010101001\underline{0} \subset \cdots$$

One can see that the underlined bit in each line above does not move; rather, elements generated following an application of Φ_2 are prepended and appended to d_0 and c_0, respectively. Since an application of Φ_2 always replaces one bit with at least two, $\Phi_2^m(b_m)$ will continue expanding for larger m. More specifically, each "1" in (b_m) will continue expanding to the right after many applications of Φ_2, and each "0" in (b_m) will similarly expand to the left. Since (b_m) has infinitely many ones and zeros, $\Phi_2^m(b_m)$ will extend infinitely in both directions as m goes to ∞.

We have shown that $\Phi_2^m(b_m)$ satisfies both the desired properties for large enough m, meaning a happy sequence exists for Φ_2. □

We have seen so far that not only do happy sequences exist, they also have a particular form by virtue of the replacement rules which generate them. The following section examines this form by analyzing quasiperiodicity of happy sequences.

3. Quasiperiodicity of happy sequences

In this section, we aim to show that happy sequences follow a particular pattern; they are *quasiperiodic*, meaning every finite subsequence in a happy sequence repeats infinitely.

In order to show that happy sequences are quasiperiodic, we need to observe the behavior of the finite subsequences in a happy sequence. We will find the subsequences, and then show that they repeat infinitely.

We first investigate the ancestors of a happy sequence. Since an ancestor of a happy sequence H is any $\Phi_2^{-m}(H)$ for some m, applying Φ_2 to an ancestor of H enough times will regenerate H. If we can find any finite subsequences in the ancestors repeat infinitely, then we know the finite subsequences of H that repeat infinitely. We give the following lemma explicitly:

Lemma 3.1. *All of the ancestors and descendants of a happy sequence must be happy sequences.*

Proof. Consider the following family tree of a happy sequence H:

$$\cdots \longleftarrow H_{n+1} \longleftarrow H_n \longleftarrow \underbrace{\cdots}_{\Phi^{-1}\text{ applied many times}} \longleftarrow H_1 \longleftarrow H$$

where each sequence H_i for $i \in \mathbb{Z}^{\geq 0}$ is an ancestor of H, and each arrow indicates an application of Φ^{-1}.

By definition, there exists an ancestor of H given by $\Phi^{-m}(H)$ for all m, so H has an infinite amount of unique ancestors. Let us select a particular ancestor $\Phi^{-n}(H) = H_n$, the sequence resulting from n applications of Φ^{-1} to H. We observe that H_n also has an infinite amount of ancestors; it is a part of the same infinite family tree.

Then

$$\Phi^{-m}(H) \text{ exists for all } m \implies \Phi^{-m}(\Phi^{-n}(H)) \text{ exists for all } m, n,$$

and all ancestors of H satisfy the definition of a happy sequence. □

Using this lemma, we can show that some infinite subsequences repeat infinitely in a happy sequence—particularly, the subsequences 1 and 0. We state the following lemma:

Lemma 3.2. *The subsequence 1 repeats infinitely in a happy sequence, and the subsequence 0 repeats infinitely in a happy sequence.*

Proof. Suppose we have a happy sequence H. By proposition 3.1, there exists an ancestor of H given by $\Phi^{-1}(H) = H_1$ which is happy, and therefore of infinite length.

To obtain H from H_1, we must apply Φ_2 to H once. Recall that the procedure Φ_2 is defined as follows:

$$\begin{aligned} \Phi_2\colon 1 &\to 100 \\ 0 &\to 10. \end{aligned}$$

Then when we apply Φ to H_1, each digit in H_1 becomes a subsequence containing both 1 and 0 in H. Since H_1 has an infinite amount of digits, H has infinite repeats of 1 and infinite repeats of 0. □

We have shown that the subsequence 1 repeats infinitely in a happy sequence, as well as the subsequence 0. But we note that this means H_1 from the proof of lemma 3.2 has infinite repeats of 1 and infinite repeats of 0—then it follows that H, the image of H_1 under Φ_2, has infinite repeats of 100 and 10. It seems intuitive to suggest that a happy sequence will also have infinite repeats of the subsequences obtained from applying Φ_2 to 1 and to 0 many successive times. Indeed, we claim that a happy sequence has infinite repeats of $\Phi_2^m(1)$ and of $\Phi_2^m(0)$. As a consequence of lemma 3.2, we obtain the following corollary.

Corollary 3.2.1. *The subsequence $\Phi_2^m(1)$ repeats infinitely in a happy sequence.*

Proof. We follow similar logic to the proof of lemma 3.1.

Suppose we have a happy sequence H. By lemma 3.1, for some m, there exists an ancestor of H given by $\Phi_2^{-m}(H) = H_m$ which is happy, and therefore has infinitely repeating 1.

To obtain H from H_m, we must apply Φ_2 to H m times. We can denote this as Φ_2^m:

$$\begin{aligned} \Phi_2^m\colon 1 &\to \Phi_2^m(1) \\ 0 &\to \Phi_2^m(0). \end{aligned}$$

Applying Φ_2^m to H_m produces a $\Phi_2^m(1)$ in H for every 1 in H_m. Since H_m has an infinite amount of 1, H has infinite repeats of $\Phi_2^m(1)$.

We note that we can do this for *any* $m \in \mathbb{Z}^+$, so $\Phi_2^m(1)$ repeats infinitely in a given happy sequence for all $m \in \mathbb{Z}^+$.

We follow similar logic to claim $\Phi_2^m(0)$ repeats infinitely in H, but we note that the image of 0 under Φ_2^m is a subsequence of the image of 1 under Φ_2^m for any m—so $\Phi_2^m(1)$ repeating infinitely in H implies $\Phi_2^m(0)$ repeats infinitely in H. □

We have shown that a large amount of finite subsequences repeat infinitely in a happy sequence. Actually, all of these subsequences are subsequences of $\Phi^m(1)$ for some m—and these are all of the finite subsequences in a happy sequence. We are ready to state the main theorem of this section.

Theorem 3.3. *Every happy sequence is quasiperiodic.*

Proof. In order to prove the result, we show that we can define a set of all finite subsequences in a happy sequence, and that they repeat infinitely; quasiperiodicity results.

We will first show that we can define the set of all finite subsequences in a happy sequence.

A happy sequence has infinite repeats of $\Phi^m(1)$ for all m by 3.2.1, and must also have an inverse $\Phi^{-m}(H)$ for all m by definition. It follows, then, that a happy sequence will have no subsequences that are not subsequences of $\Phi^m(1)$ for some m. If there is a subsequence u in a happy sequence such that $u \not\subset \Phi^{-t}(H)$ for any t, then there will not exist an inverse $\Phi^{-m}(H)$ for some m. Then every finite subsequences in a happy sequence must be subsequences of $\Phi^m(1)$ for some m.

Now it remains to show that $\Phi^m(1)$ repeats infinitely in a happy sequence.

Recall that by corollary 3.2.1, $\Phi^m(1)$ repeats infinitely in a happy sequence for all m. Then since every finite subsequence in a happy sequence is a subsequence of $\Phi^m(1)$ for some arbitrary m, we have shown that all finite subsequences in a happy sequence repeat infinitely, and we are done. □

Now that we've explored the structure of happy sequences, in the next section, we will explore the asymptotic density of ones and zeros in them.

4. Asymptotic density of ones and zeros in a happy sequence

In this section, we provide a conjecture for the asymptotic density of ones and zeros in a happy sequence. We define asymptotic density as follows:

Definition 4.1 (Asymptotic density)**.** The asymptotic density of ones in a lucky sequence is given by

$$\lim_{m\to\infty}\frac{i_m}{a_m},$$

where i_m denotes the number of ones in the sequence, a_m denotes the length, and m denotes the number of applications of Φ_n.

First, we find the density for when our procedure is Φ_2, and then we find a general formula for Φ_n.

4.1. **Proof via generating functions for Φ_2.**

First, we will derive a recurrence for the length of the sequence obtained by applying Φ_2 m times to a given sequence s, as well as recurrences for the number of ones and zeros in this sequence. Then we will use these recurrences to produce generating functions, which we will then use to determine the asymptotic density of ones and zeros in a happy sequence.

Given some binary sequence, we can repeatedly apply the procedure to make it lucky. We let s_0 represent our starting sequence, and let s_m denote the sequence obtained by applying Φ_2^m to s_0 times. Additionally, let a_m denote the length of s_m, and define s_0 as the sequence we begin with; it may or may not already be lucky. We also let i_m denote the number of ones in s_m, and j_m denote the number of zeros in s_m.

Proposition 4.1. The length of s_m is given by $a_m = 3i_{m-1} + 2j_{m-1}$.

Proof. Since every 1 in the original sequence becomes 100 after applying Φ_2, we have one 1 and two 0s in s_m for every 1 in s_{m-1}. Each 0 becomes 10, so we have one 1 and one 0 in s_m for every 0 in s_{m-1}. Then the number of 1s in s_m is given by

$$i_m = i_{m-1} + j_{m-1}$$

and the number of 0s is given by

$$j_m = 2i_{m-1} + j_{m-1}.$$

Then the length of s_m is given by

$$\begin{aligned} a_m &= i_m + j_m \\ &= i_{m-1} + j_{m-1} + 2i_{m-1} + j_{m-1} \\ &= 3i_{m-1} + 2j_{m-1}. \end{aligned}$$

□

It is possible to define a_m purely in terms of the length of the previous sequences. We obtain the following corollary:

Corollary 4.0.1 (Recursive formula for length)**.** *The length of s_m is given by $a_m = 2a_{m-1} + a_{m-2}$.*

Proof. We rearrange the formula given in proposition 4.1:

$$\begin{aligned} a_m &= 3i_{m-1} + 2j_{m-1} \\ &= 2(i_{m-1} + j_{m-1}) + i_{m-1}, \end{aligned}$$

and substitute $a_{m-1} = i_{m-1} + j_{m-1}$ and $i_{m-1} = i_{m-2} + j_{m-2}$:

$$\begin{aligned} a_m &= 2a_{m-1} + i_{m-2} + j_{m-2} \\ &= 2a_{m-1} + a_{m-2}. \end{aligned}$$

□

Note that this formula means that for $m \geq 2$, we only need the length of the two most recent ancestors of s_m to determine its length; we do not need to know the specific makeup of ones and zeros in s_m or its ancestors. We also note that i_m can be defined in terms of a_m:

Corollary 4.0.2 (Recursive formula for number of ones)**.** *The number of ones in s_m is given by $i_m = a_{m-1}$.*

Proof. We have the recursions

$$a_m = i_m + j_m$$

and

$$i_m = i_{m-1} + j_{m-1},$$

but $i_{m-1} + j_{m-1}$ is equal to a_{m-1}, so we get

$$i_m = a_{m-1}.$$

□

Note that as a result of both corollaries 4.0.1 and 4.0.2, we can provide an upper bound on the density of ones in s_m:

Corollary 4.0.3 (Asymptotic density bound for a lucky sequence)**.** *The asymptotic density of ones in s_m is strictly less than 1/2 as m approaches infinity.*

Proof. The asymptotic density is given by

$$\lim_{m\to\infty} \frac{i_m}{a_m},$$

and we have the recursions for a_m and i_m from corollaries 4.0.1 and 4.0.2, respectively. Then the asymptotic density is

$$\lim_{m\to\infty} \frac{a_{m-1}}{2a_{m-1} + a_{m-2}}.$$

a_{m-2} is greater than 0, so we have an upper bound on our limit:

$$\lim_{m\to\infty} \frac{a_{m-1}}{2a_{m-1} + a_{m-2}} < \lim_{m\to\infty} \frac{a_{m-1}}{2a_{m-1}} < \frac{1}{2}$$

□

We will keep this result in mind as we explore asymptotic densities. If the density of ones in s_m is strictly less than 1/2, this should hold for happy sequences as well. Happy sequences must have infinite inverses; if the density of ones is 1/2 or more, this principle is violated.

Now that we have illustrated some basic properties of lucky sequences, we will use these to obtain non-recursive formulae for the length and number of ones in a lucky sequence.

First we obtain a non-recursive formula for the length of a lucky sequence.

Proposition 4.2. The length of s_m is given by

$$a_m = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^m} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^m}.$$

Proof. We will prove the result using generating functions.

Let $A(x) = \sum_{m\geq 0} a_m x^m$ be the generating function for a_m, and note that a_m is the coefficient of x_m in the mth term of the sum. We aim to find a formula for $A(x)$ which allows us to easily extract an explicit formula for a_m.

From corollary 4.0.1, we have a recurrence for length, $a_{m+2} = 2a_{m+1} + a_m$. Multiplying both sides of the recurrence by x^{m+2} and summing over all nonnegative integers m, we obtain

$$\begin{aligned}
\sum_{m\geq 0} a_{m+2}x^{m+2} &= \sum_{m\geq 0} 2a_{m+1}x^{m+2} + \sum_{m\geq 0} a_m x^{m+2} \\
A(x) - a_0 - a_1 x &= 2x[A(x) - a_0] + x^2 A(x) \\
(-1 + 2x + x^2)A(x) &= -a_0 + 2a_0 x - a_1 x \\
A(x) &= \frac{(2a_0 - a_1)x - a_0}{(-1 + 2x + x^2)}
\end{aligned}$$

Now, we continue to manipulate the formula for $A(x)$ until it resembles a geometric sum.

$$A(x) = \frac{-(3+2\sqrt{2})a_0 + (1+\sqrt{2})a_1}{2\sqrt{2}(-1-\sqrt{2}-x)} + \frac{(3-2\sqrt{2})a_0 - (-1+\sqrt{2})a_1}{2\sqrt{2}(-1+\sqrt{2}-x)}.$$

$A(x)$ is now in the form of a sum of a geometric series, so we can use the geometric series expansion to obtain

$$A(x) = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4} \sum_{m\geq 0} \frac{x^m}{(-1-\sqrt{2})^m} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4} \sum_{m\geq 0} \frac{x^m}{(-1+\sqrt{2})^m} \tag{3}$$

Recall that the coefficient of x^m in $A(x)$ is a_m. Hence, the coefficient of x^m in (3) is a_m; it is given by

$$a_m = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^m} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^m}. \tag{4}$$

□

Based on this formula, we need only know m and the lengths of any two sequential ancestors to find a_m. For example, if we have a sequence s_m obtained from applying the procedure to s_0 m times, but we know $|s_2| = a_2$ and $|s_3| = a_3$, and we wish to calculate a_m, the formula becomes

$$a_{m-2} = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^{m-2}} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m-2}},$$

because the sequence s_m obtained from applying the procedure to s_0 m times is the same as the sequence obtained from applying the procedure to s_2 $m-2$ times.

Now, we obtain a non-recursive formula for the number of ones in a lucky sequence.

Corollary 4.0.4. *The number of ones in s_m is given by*

$$i_m = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^{m-1}} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m-1}}, \textit{ for } m \geq 2.$$

Proof. Recall that $i_m = a_{m-1}$ from corollary 4.0.2. Let $I(x) = \sum_{m\geq 1} i_m x^m$ and $A(x) = \sum_{m\geq 0} a_m x^m$. We obtain

$$\begin{aligned}\sum_{m\geq 1} i_m x^m &= \sum_{m\geq 1} a_{m-1} x^m \\ I(x) &= x[A(x)] \\ I(x) &= x\left[\frac{-(3+2\sqrt{2})a_0 + (1+\sqrt{2})a_1}{2\sqrt{2}(-1-\sqrt{2}-x)} + \frac{(3-2\sqrt{2})a_0 - (-1+\sqrt{2})a_1}{2\sqrt{2}(-1+\sqrt{2}-x)}\right]\end{aligned}$$

Hence the coefficient of x^m is given by

$$i_m = \frac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^{m-1}} + \frac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m-1}}.$$

We could have also simply changed the exponents in the non-recursive formula for a_m (4) to $m-1$ to obtain this formula for i_m.

□

Now we are ready to determine the asymptotic density of ones and zeros in a lucky sequence.

Theorem 4.1. *The asymptotic density of ones in a lucky sequence is* $-1+\sqrt{2}$.

Proof. The density of ones in a lucky sequence is given by

$$\lim_{m\to\infty} \frac{i_m}{a_m}.$$

Since $i_m = a_{m-1}$, this limit becomes

$$\lim_{t\to\infty} \frac{a_{m-1}}{a_m}.$$

Using (4), the explicit formula for a_m from proposition 4.2, we obtain

$$\lim_{m\to\infty} \frac{a_{m-1}}{a_m} = \frac{\left[\dfrac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^{m-1}} + \dfrac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m-1}}\right]}{\left[\dfrac{(\sqrt{2}+2)a_0 - (\sqrt{2})a_1}{4(-1-\sqrt{2})^{m}} + \dfrac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m}}\right]}$$

As m approaches infinity, $(-1-\sqrt{2})^{-(m-1)}$ approaches 0. So then our limit becomes

$$\lim_{m\to\infty} \frac{a_{m-1}}{a_m} = \frac{\left[\dfrac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m-1}}\right]}{\left[\dfrac{(-\sqrt{2}+2)a_0 + (\sqrt{2})a_1}{4(-1+\sqrt{2})^{m}}\right]} = \frac{(-1+\sqrt{2})^m}{(-1+\sqrt{2})^{m-1}} = -1+\sqrt{2}$$

This means that as the number of times m that we apply the procedure to 0 approaches infinity, the density of ones in the resulting subsequence approaches $-1+\sqrt{2}$. We find that our choice of starting sequence does not impact the asymptotic density. That is, a_0 and a_1 cancel out. □

This aligns with the result of corollary 4.0.3. We obtain the following corollary as a consequence of theorem 4.1.

Corollary 4.1.1. *The asymptotic density of zeros in a lucky sequence is* $2-\sqrt{2}$.

Proof. The number of zeros in s_m is given by $j_m = a_m - i_m$. We have

$$\begin{aligned}
\lim_{m\to\infty} \frac{j_m}{a_m} &= \lim_{m\to\infty} \frac{a_m - i_m}{a_m} \\
&= \lim_{m\to\infty} \left(1 - \frac{i_m}{a_m}\right) \\
&= 1 - (-1+\sqrt{2}) \\
&= 2 - \sqrt{2}
\end{aligned}$$

□

Now that we have found the asymptotic density of ones and zeros in a lucky sequence, we propose a conjecture for the asymptotic density of ones in a happy sequence.

Conjecture 4.1. *The asymptotic density of ones in a happy sequence is* $-1+\sqrt{2}$.

Argument. We can use some intuition to support this result. We know that lucky sequences are composed of the subsequences $\Phi_2^t(1)$ and $\Phi_2^t(0)$ for all t, and we know that a happy sequence could not be composed of all $\Phi_2^t(1)$ or all $\Phi_2^t(0)$, else it is certainly unlucky for some m. But we will consider two extreme cases in which our sequence is solely composed of one of these.

If a lucky sequence is composed of infinitely repeating subsequences $\Phi_2^1(1) = 100$, it is easy to see that the density of ones is $1/3$. Similarly, if a lucky sequence is composed of infinitely repeating subsequences $\Phi_2^1(0) = 10$, the density of ones is $1/2$. These are the minimum and maximum possible densities of ones in a lucky sequence, so it follows that for a happy sequence, the asymptotic density of ones must be somewhere between these two values. As we showed in corollary 4.0.3, it is strictly less than $1/2$, and our reasoning suggests it is strictly greater than $1/3$. So we can bound the asymptotic density of ones in a happy sequence by these values. Letting D_2 denote the asymptotic density of a happy sequence with Φ_2, we have

$$\frac{1}{3} < D_2 < \frac{1}{2}$$

Next we consider a sequence composed only of $\Phi_2^2(1) = 1001010$. We can see that its asymptotic density of ones will be $3/7$. A sequence only composed of $\Phi_2^2(0) = 10010$ will have an asymptotic density of $2/5$. Again, a happy sequence should not have a density outside of this range, else it is unlucky, which certainly makes it unhappy. So we can tighten our bounds:

$$\frac{2}{5} < D_2 < \frac{3}{7}.$$

When $t = 3$, we have $\Phi_2^3(1) = 10010101001010010$ and $\Phi_2^3(0) = 100101010010$. The bound is

$$\frac{7}{17} < D_2 < \frac{5}{12},$$

which is now quite tight. It suggests that D_2 may equal $-1+\sqrt{2}$, the same as the asymptotic density for lucky sequences. Intuitively, happy sequences are a special type of lucky sequences, and so it seems to make sense that the behavior of both would be similar. If we think of a happy sequence as infinitely lucky, then perhaps an approximation of a happy sequence with a lucky sequence to which we apply Φ infinitely many times makes sense. Indeed, we could try finding the densities of $\Phi_2^t(1)$ and $\Phi_2^t(0)$ as t approaches infinity.

In the proof of theorem 4.1, it was noted that the limit of $\frac{i_t}{a_t}$ tended to $-1+\sqrt{2}$ as t tended to infinity, no matter our choice of starting sequence. So the asymptotic densities of both $\Phi_2^t(1)$ and $\Phi_2^t(0)$ evaluate to $-1+\sqrt{2}$. Earlier we claimed that D_2 was strictly in between the two values, not equal to either. However, intuitively, the sequences $\Phi_2^t(1)$ and $\Phi_2^t(0)$ may become "lucky" enough after some critical number of applications of Φ_2 that we find their densities are equal, and we can no longer claim D_2 is strictly between their densities. This seems plausible, since we showed that the densities get quite close to $-1+\sqrt{2}$ after just 3 applications of Φ_2.

□

4.2. **Asymptotic density for Φ_n with $n \geq 2$.** In the previous subsection, we considered the procedure with $n = 2$, which replaced 1 with 100 and 0 with 10. In this subsection, we consider what happens when $n \geq 2$.

We find that

$$i_m = i_{m-1} + j_{m-1} = a_{m-1}$$

still holds, since every digit in s_{m-1} still contributes a one to s_m. Now, since every one in s_{m-1} contributes n zeros to s_m, and every zero in in s_{m-1} contributes $n-1$ zeros to s_m, we get the following formula for the number of zeros:

$$\begin{aligned} j_m &= n i_{m-1} + (n-1) j_{m-1} \\ &= (n-1)a_{m-1} + i_{m-1}. \end{aligned}$$

Then for the length a_m, we obtain

$$\begin{aligned} a_m &= i_m + j_m \\ &= i_{m-1} + j_{m-1} + n i_{m-1} + (n-1) j_{m-1} \\ &= (n+1) i_{m-1} + (n) j_{m-1} \\ &= n a_{m-1} + i_{m-1} \\ &= n a_{m-1} + a_{m-2}. \end{aligned} \tag{5}$$

The result (5) makes sense, since every digit in s_{m-1} contributes at least n digits—the ones contribute $n+1$ digits and the zeros contribute n. So we can simply multiply the length of s_{m-1} by n, and then add in the number of ones in s_{m-1} to find a_m.

Using these formulae, we can find a non-recursive formula for the length of s_m by using generating functions.

Proposition 4.3. The length of s_m is given by $a_m =$

$$\frac{2a_0 + (n^2 - n\sqrt{n^2+4})a_0 - (n - \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n+\sqrt{n^2+4}}{2}\right)^{m+1}} + \frac{-2a_0 - (n^2 + n\sqrt{n^2+4})a_0 + (n + \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n-\sqrt{n^2+4}}{2}\right)^{m+1}}.$$

Proof. We prove the result using generating functions. We proceed using the same method as for the proof of theorem 4.1.

Let $A(x) = \sum_{m\geq 0} a_m x^m$ be the generating function for a_m.

We have $a_{m+2} = n a_{m+1} + a_m$. Multiplying both sides of the recurrence by x^{m+2} and summing over all nonnegative integers m, we obtain

$$\begin{aligned} \sum_{m\geq 0} a_{m+2} x^{m+2} &= \sum_{m\geq 0} n a_{m+1} x^{m+2} + \sum_{m\geq 0} a_m x^{m+2} \\ A(x) - a_0 - a_1 x &= nx[A(x) - a_0] + x^2 A(x) \\ (-1 + nx + x^2) A(x) &= -a_0 + n a_0 x - a_1 x \\ A(x) &= \frac{(n a_0 - a_1)x - a_0}{(-1 + nx + x^2)}. \end{aligned}$$

Using partial fraction decomposition, we get

$$A(x) = \frac{2a_0 + (n^2 - n\sqrt{n^2+4})a_0 - (n - \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n+\sqrt{n^2+4}}{2} - x\right)} + \frac{-2a_0 - (n^2 + n\sqrt{n^2+4})a_0 + (n + \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n-\sqrt{n^2+4}}{2} - x\right)}.$$

Using the geometric series expansion and simplifying, we obtain

$$A(x) = \frac{2a_0 + (n^2 - n\sqrt{n^2+4})a_0 - (n - \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n+\sqrt{n^2+4}}{2}\right)} \sum_{m\geq 0} \frac{x^m}{\left(\frac{-n+\sqrt{n^2+4}}{2}\right)^m} +$$
$$\frac{-2a_0 - (n^2 + n\sqrt{n^2+4})a_0 + (n + \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n-\sqrt{n^2+4}}{2}\right)} \sum_{m\geq 0} \frac{x^m}{\left(\frac{-n-\sqrt{n^2+4}}{2} - x\right)^m}$$

Hence the coefficient of x^m is given by

(6)
$$a_m = \frac{2a_0 + (n^2 - n\sqrt{n^2+4})a_0 - (n - \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n+\sqrt{n^2+4}}{2}\right)^{m+1}} + \frac{-2a_0 - (n^2 + n\sqrt{n^2+4})a_0 + (n + \sqrt{n^2+4})a_1}{2\sqrt{n^2+4}\left(\frac{-n-\sqrt{n^2+4}}{2}\right)^{m+1}}$$

□

Corollary 4.1.2. *The asymptotic density of ones in a lucky sequence generated by Φ_n with $n \in \mathbb{N} : n \geq 1$ is given by*

$$\frac{-n + \sqrt{n^2+4}}{2}.$$

Proof. Following the same steps as the proof for theorem 4.1, we get

$$\begin{aligned} \lim_{m\to\infty} \frac{i_m}{a_m} &= \frac{a_{m-1}}{a_m} \\ &= \frac{\left(\frac{-n+\sqrt{n^2+4}}{2}\right)^{m+1}}{\left(\frac{-n+\sqrt{n^2+4}}{2}\right)^{m}} \\ &= \frac{-n + \sqrt{n^2+4}}{2} \end{aligned}$$

□

We check that this formula matches the result that was obtained for Φ_2 in the last section by substituting $n = 2$:

$$\lim_{m\to\infty} \frac{i_m}{a_m} = \frac{-2 + \sqrt{2^2+4}}{2} = -1 + \sqrt{2}.$$

We obtain the expected result. By a similar argument to that presented in conjecture 4.1, we claim that this may be the value of the asymptotic density of a happy sequence. In the next section, we explore some computational methods which appear to support this claim.

5. Computational Methods to Generate Sequences

The goal of this section is to provide computational methods that validate results derived above, provide alternative perspectives and lend credibility to models developed. $1 \equiv a$ and $0 \equiv b$ for sake of readability. All notation used here is exactly identical to the prior sections, with n as the number of zero's in the replacement operation and m as number of times Φ has been performed. The linear algebra based proof of asymptotic densities provided below sets up the formulation of all the computation methods, so is included in this section.

5.1. **Observed recurrences.** First, we recognize that each a in s will generate 1 a and n b's in $\Phi(s)$. Next, we recognize that each b in s will generate 1 a and n - 1 $b's$ in $\Phi(s)$. We let A_i and B_i respectively denote the number of $a's$ and $b's$ in $\Phi^m(s)$.

Proposition 5.1. Cardinality of 0's: $A_{m+1} = A_m + B_m$.

Proof. To begin with the $(a's)$, these follow because when the operation Φ is done, each a in the updated sequence will remain an a, but each b will contribute another (additional) a, so the total number of $a's$ is dependent on the total count of $a's$ and $b's$ present at the previous step. □

Proposition 5.2. Cardinality of 1's: $B_{m+1} = nA_m + (n-1)B_m$.

Proof. For $b's$ on the other hand, each (a) in $\Phi(s)$ will contribute $(n)b's$ and each b in $\Phi(s)$ will contribute $(n-1)$ b's so then, the total number of $b's$ is updated accordingly. □

Proposition 5.3. Cardinality of a Sequence: $L_{m+1} = A_{m+1} + B_{m+1}$.

Proof. These follow from logical observation but may be proven more extensively by induction. □

Corollary 5.0.1. *Asymptotic Density of Term a:* $d_a = \lim_{m\to\infty} \dfrac{A_m}{L_m}$.

Corollary 5.0.2. *Asymptotic Density of Term b:* $d_b = \lim_{m\to\infty} \dfrac{B_m}{L_m}$.

5.2. **Identifying a transition matrix.** We have effectively constructed a pair of recurrence relationships for A and B after applying ϕ. It is to be noted that although we have identified a transition matrix, this is not technically a Stochastic or a Markov matrix (although it shares the property of capturing the asymptotic density of the elements it acts on in its eigenvectors). The system will not converge and although it has an asymptotic density, the cardinality of any finite subsequence grows with each instance of the ϕ operation.

Corollary 5.0.3. *Transition Matrix between Consecutive Sequences:* $\begin{pmatrix} A_{m+1} \\ B_{m+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ n & n-1 \end{pmatrix} \begin{pmatrix} A_m \\ B_m \end{pmatrix}$.

This may be simplified into: $X_{i+1} = MX_i$. The asymptotic densities d_a and d_b in s will require that we consider the behaviors of $\frac{A_m}{L_m}$ and $\frac{B_m}{l_m}$ as $m \to \infty$. As one would expect, a single eigenvalue would dominate the behaviors of A and B and $m \to \infty$. We may compute the eigenvalues and eigenvectors as:

$$\lambda_a = \frac{n - \sqrt{n^2+4}}{2}; v_a = \begin{pmatrix} \frac{2}{\sqrt{n^2+4}+(n-2)} \\ 1 \end{pmatrix} \tag{7}$$

$$\lambda_b = \frac{n + \sqrt{n^2+4}}{2}; v_b = \begin{pmatrix} \frac{-2}{\sqrt{n^2+4}-(n-2)} \\ 1 \end{pmatrix} \tag{8}$$

Proposition 5.4. Asymptotic Density of 1 $= \frac{-n+\sqrt{n^2+4}}{2}$

Proof. Since the transition matrix generates the future state, and is not a typical Stochastic matrix, we must invert this matrix to identify the current state preceding the transformation to a future state. Due to the self-repeating nature of happy sequences, the inversion operation done enough times will yield the original sequence iteration and following sequences. The inverted matrix in question has the eigenvalues:

$$\lambda_1 = \frac{-n + \sqrt{n^2+4}}{2} \tag{9}$$

$$\lambda_2 = \frac{-n - \sqrt{n^2+4}}{2} \tag{10}$$

Obviously, $\lambda_1 > \lambda_2$ as $n, m \to \infty$. Thus, the λ_1 term dominates any sequence for any number of operations. Let v_1 be the eigenvector associated with λ_1. From the well-known result that the eigenvectors (of the dominant eigenvalue) of a state matrix are the asymptotic density of the elements in future states, this leads to us being able to identify λ_1 as the asymptotic density of 1. This validates the result demonstrated by **Corollary 4.1.2**. This is shown below in the test case in **Figure 1**. □

Alternatively, one could prove this result with the dominant eigenvector approach, where the eigenvector associated with the larger eigenvalue, so λ_b in this case, is normalized to sum to 1, and the result is such that the first term will be the asymptotic density of one's and the second will be the asymptotic density of zero's. The densities computed are identical.

5.3. **MATLAB Script for replacement operations.** This following code does the operation on any starting sequence with a specified m, n and returns the output (i.e. the sequence after m^{th} transformation for a replacement rule with n zeroes). The density is then computed using the eigenvalue. Results of this script do approximately match results derived earlier, to the limit of computational precision.

In other words, the user inputs a starting sequence, and the n for a replacement operation and this program manually does the replacement operation on each element in the set and returns an output of the updated sequence for each instance where the operation is done. It will do this repeatedly until terminated for m steps.

Please comment out lines 48 to 62 and 67 to 73 to endlessly generate new sequences. A sample output is included below in a case where the starting sequence was 10 and $n = 2$. Alternatively, this code may be run in the "numpy" Python package at the user's preference, although syntax may differ slightly between the two languages.

Step 1: New sequence = 10010, Length = 5
Step 2: New sequence = 100101010010, Length = 12
Step 3: New sequence = 10010101001010010100101010010, Length = 29

5.4. **Function Definition.**

- **findHappySequence**: Main function to check if a given sequence is a happy sequence.
  - **Input**: Binary string input_sequence and integer n.
  - **Output**:
    - isHappy: Boolean flag indicating if the sequence is happy.
    - density_a, density_b: Asymptotic densities of '1's and '0's.
    - m_happy: Number of iterations before finding a happy sequence (0 if not found).
    - first5Sequences: Cell array containing the first 5 generated sequences.
    - sequenceLengths: Array containing lengths of generated sequences up to m iterations.

5.5. **Pseudocode for findHappySequence.**

(1) **Initialization**
- Define the transition matrix T based on the parameter n.
- Compute the eigenvalues and eigenvectors of T.
- Find the dominant eigenvalue and the corresponding eigenvector.
- Normalize this dominant eigenvector to calculate the asymptotic densities density_a and density_b.

(2) **Variable Initialization**
- Initialize variables such as isHappy, m_happy, first5Sequences, and sequenceLengths.

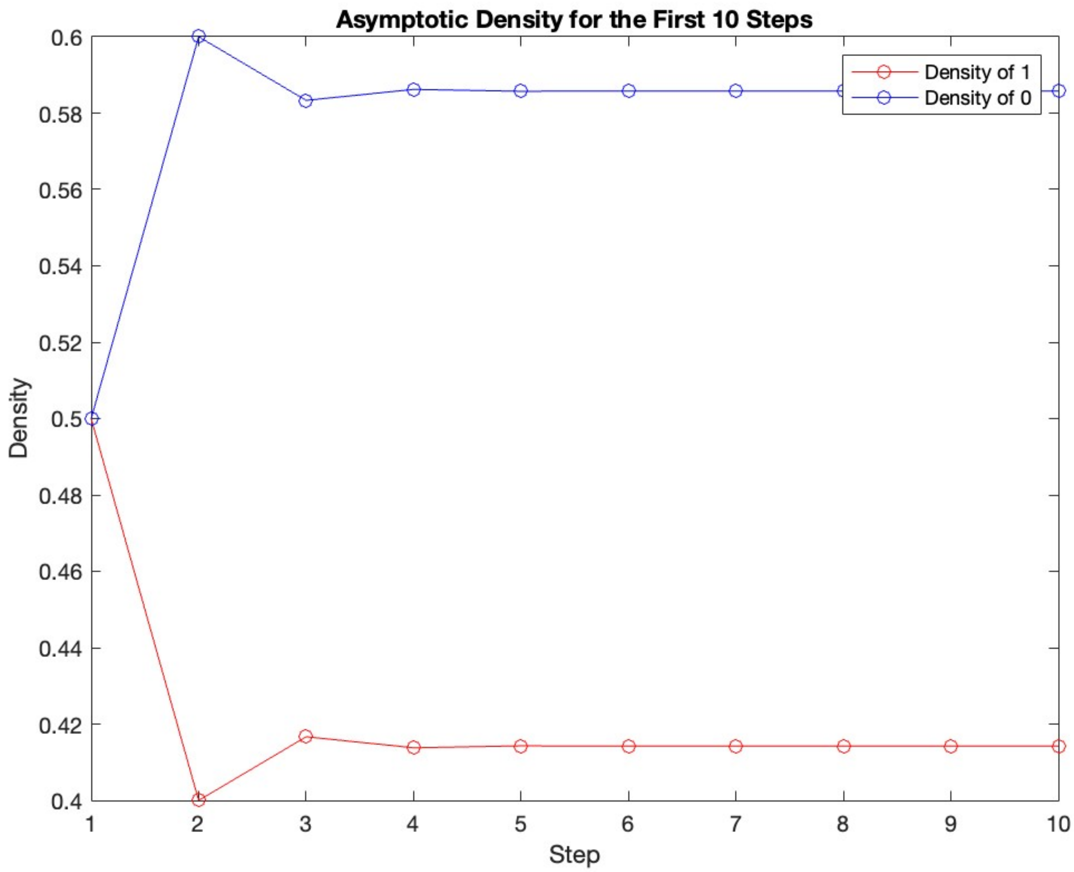

FIGURE 1. Starting sequence of (10) where $n = 2$ and $m = 10$. There is clear evidence of the asymptotic density settling to $\sqrt{2} - 1 \approx 0.414$ where $n = 2$. This is exactly what was expected from **Section 5**.

- Initialize arrays density_a_values and density_b_values to store density values for the first 10 steps.

(3) **Sequence Analysis**
- Perform iterations for $m = 1$ to 12 (or any arbitrary limit) to simulate an infinite scenario.
  - Apply the transformation ϕ to generate new_sequence.
  - Store the length of new_sequence in sequenceLengths.
  - Calculate and store densities of '1' and '0' for the first 10 steps in density_a_values and density_b_values.
  - If $m \leq 5$, store new_sequence in first5Sequences.
  - Check if input_sequence is a subsequence of new_sequence.
    * If true, set isHappy = true, m_happy = m, and exit the loop.

(4) **Plotting and Final Check**
- If $m\_happy = 0$, set isHappy = false.
- Plot the densities for the first 10 steps using density_a_values and density_b_values.

(5) **Output**
- Return all output variables isHappy, density_a, density_b, m_happy, first5Sequences, and sequenceLengths.

The program may be downloaded via GitHub from: https://github.com/HillelDei/18.821 as "findHappySequence.m". As mentioned earlier, **Figure 1** is an example output of running this code. The specific conditions are elaborated upon in its caption.

6. Computational Methods to Validate Sequences

The goal of this section is to go a step beyond simply generating transformed sequences with the Φ operation but rather to create sequences that are happy via computational methods. Two algorithms are given below that make use of the quasiperiodicity property. They reiterate the concept that a given happy sequence contains repeating subsequences. The former is a comprehensive, brute force method of finitely approximating a fundamentally infinite problem. However, as there are 2^n possible sequences at any given moment, each of which are individually verified for rigor, this means that this program is computationally expensive and time-consuming. The latter, inspired by the method of Penrose tilings, is far simpler and runs orders of magnitudes faster, however it suffers in terms of accuracy and each random sequence must be verified.

To provide guidance on the code's functionality for the first method, a few things must be considered: first and most importantly, the problem deals with infinite sequences. We, however, are only ever capable of modelling a finite subset of that. This leads to a few relaxations being introduced. The most important condition is this: for a sequence to be happy, it must be invertible. This means it must be constructed with only subseqeuences $\{1, 0^n\}$ or $\{1, 0^{n-1}\}$. This code receives an input n and alternates between these 2 subsequences to find a valid example. Furthermore, for a given subsequence of length q there will be $q+1$ permissible subsequences and everything else is forbidden. This can be proven by construction but for example, when $n = 2$ for subsequences of length 5, only 6 are valid, namely $00100, 00101, 01001, 10010, 10100, 10101$ and all others are invalid. We may computationally predict what is valid and what is not and use this to eliminate invalid sequences. These facts helps us quickly remove sequences that can not be inverted, and are incorporated into the scripts. They are available via https://github.com/HillelDei/18.821 and are respectively "find_invertible_happy_sequences.pdf" and "find_limited_invertible_sequences.m".

6.1. Introduction.

This document provides a formal description of a MATLAB code designed to find "happy sequences" based on certain transformation rules and constraints. A "happy sequence" is defined as a sequence that can be forward and backward transformed multiple times and can still regenerate itself. Additionally, the sequence must be "invertible," meaning it does not contain any forbidden subsequences.

6.2. Variables and Parameters.

- N: Number of zeros following a '1' in the first type of subsequence (input by the user)
- M: Number of transformation operations to apply (input by the user)
- `seq1`: A subsequence consisting of a '1' followed by N zeros
- `seq2`: A subsequence consisting of a '1' followed by $N - 1$ zeros
- `forbidden_subsequences`: A list of all subsequences that are not invertible
- `happy_sequences`: An array to store unique happy sequences
- `tested_sequences`: A cell array to store sequences that have been tested
- `unique_sequences`: A cell array to store unique happy sequences in a normalized form

Functions

6.3. Main Function (`find_invertible_happy_sequences`).

(1) Takes N and M as inputs
(2) Generates forbidden subsequences
(3) Iterates through 1000 trials to generate, test, and store happy sequences

6.4. **Forward Transformation** (`apply_forward_transformation`).

- Takes a sequence and returns its forward transformation

6.5. **Backward Transformation** (`apply_backward_transformation`).

- Takes a sequence and returns its backward transformation, if possible

6.6. **Random Sequence Generation** (`generate_random_sequence`).

- Generates a random sequence of length 200 based on `seq1` and `seq2`

6.7. **Invertibility Check** (`is_sequence_invertible`).

- Checks whether a sequence is invertible based on forbidden subsequences

6.8. **Forbidden Subsequence Generation** (`generate_forbidden_subsequences`).

- Generates a list of forbidden subsequences based on N

6.9. **Subsequence Invertibility Check** (`is_subseq_invertible`).

- Checks whether a subsequence is invertible based on the number of '1's and '0's

6.10. **Algorithm.**

(1) Take user inputs for N and M
(2) Generate a list of forbidden subsequences
(3) For each trial (up to 1000):
  (a) Generate a random initial sequence that doesn't contain any forbidden subsequences
  (b) Store this sequence in `tested_sequences`
  (c) Apply M forward and backward transformations
  (d) After each transformation, check for invertibility
  (e) If the sequence is valied, store it in `happy_sequences`
(4) Display all unique happy sequences

6.11. **Modifications to the Algorithm for the Second Method.** There are some simplifications made to the iteration presented below (also written in MATLAB), to optimize for speed. An example of an output for $N = 2, M = 4$ is: 1001010010101001010010101. This closely matches expected results, if one was to use the generator provided in **Section 5**. The user inputs values for N (which determines the transformation mechanism), the number of forward transformations, and the number of backward transformations. The script uses a stochastic approach to generate initial sequences of length 50, and it does this up to 3 times (as set by max_trials). For each generated sequence, it first applies the forward transformations and subsequently the backward transformations. After all transformations are applied, if the final sequence matches the original one, it's deemed unique and stored for later display. The transformations are guided by two rules: a 1 in the sequence results in appending 1 followed by N zeroes, while a 0 results in appending 1 followed by N-1 zeroes. The script concludes by visualizing and displaying the unique sequences that meet the regeneration criteria. Thus, instead of generating every possible sequence, this alternative method acts like a guess-and-check. This is represented visually as shown below.

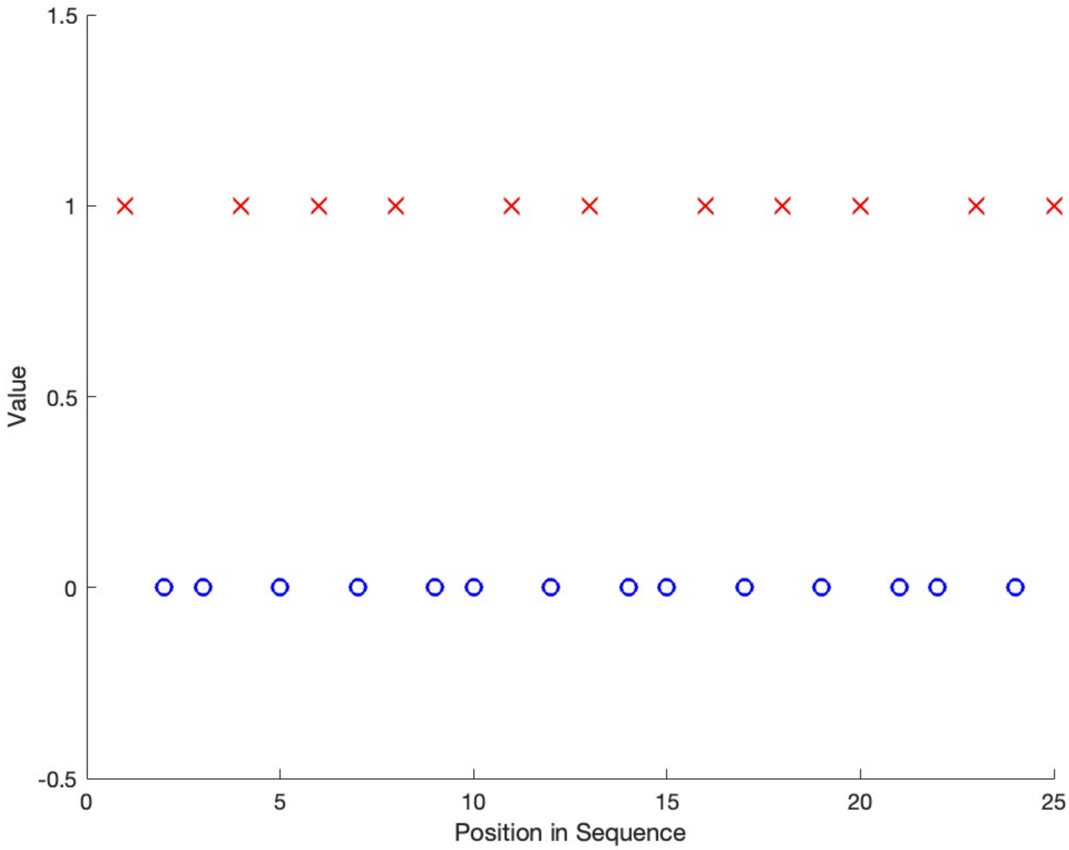

FIGURE 2. Pictorial representation of simplified algorithm. Red crosses represent '1's and blue circles represent '0's. A clear pattern emerges of ...1001010... with only the first 25 elements plotted for clarity.