AN INTRODUCTION TO NONSTANDARD LOOK AND SAY SEQUENCES

JONATHAN COMES

1. Introduction

Here is an example of a standard look and say sequence:

$$111 \to 31 \to 1311 \to 111321 \to 31131211 \to 132113111221 \to \cdots \tag{1.1}$$

Roughly speaking, the rule for generating the sequence is "say what you see". More precisely, the sequence above starts with the *seed* 111. When we look at 111 we see three 1's and thus the next term is 31. Looking at 31 we see one 3 followed by one 1, so the next term is 1311. Continuing, from 1311 we see one 1, one 3, and then two 1's, which gives the next term 111321. Eventually the terms in the sequence above appear to be growing, which brings us to the following:

Problem 1. *Determine the growth rate of a given look and say sequence.*

The solution to Problem 1 for the sequence (1.1)[1] is given by John Conway in [Con]. In order to describe Conway's solution let us consider the number of digits (i.e. the *length*) of each term in (1.1):

$$3, 2, 4, 6, 8, 12, \ldots$$

According to Conway, the ratios of those lengths approach what is now called *Conway's constant*:

$$1.303577269\ldots$$

In other words, on average each term in (1.1) is approximately 30% longer than the previous term. Conway's constant is remarkable in that it is the largest real root of the following irreducible polynomial:

$$\begin{aligned}
&\lambda^{71} - \lambda^{69} - 2\lambda^{68} - \lambda^{67} + 2\lambda^{66} + 2\lambda^{65} + \lambda^{64} - \lambda^{63} - \lambda^{62} - \lambda^{61} - \lambda^{60} - \lambda^{59} + 2\lambda^{58}\\
&+ 5\lambda^{57} + 3\lambda^{56} - 2\lambda^{55} - 10\lambda^{54} - 3\lambda^{53} - 2\lambda^{52} + 6\lambda^{51} + 6\lambda^{50} + \lambda^{49} + 9\lambda^{48} - 3\lambda^{47}\\
&- 7\lambda^{46} - 8\lambda^{45} - 8\lambda^{44} + 10\lambda^{43} + 6\lambda^{42} + 8\lambda^{41} - 5\lambda^{40} - 12\lambda^{39} + 7\lambda^{38} - 7\lambda^{37} + 7\lambda^{36}\\
&+ \lambda^{35} - 3\lambda^{34} + 10\lambda^{33} + \lambda^{32} - 6\lambda^{31} - 2\lambda^{30} - 10\lambda^{29} - 3\lambda^{28} + 2\lambda^{27} + 9\lambda^{26} - 3\lambda^{25}\\
&+ 14\lambda^{24} - 8\lambda^{23} - 7\lambda^{21} + 9\lambda^{20} + 3\lambda^{19} - 4\lambda^{18} - 10\lambda^{17} - 7\lambda^{16} + 12\lambda^{15} + 7\lambda^{14} + 2\lambda^{13}\\
&- 12\lambda^{12} - 4\lambda^{11} - 2\lambda^{10} + 5\lambda^{9} + \lambda^{7} - 7\lambda^{6} + 7\lambda^{5} - 4\lambda^{4} + 12\lambda^{3} - 6\lambda^{2} + 3\lambda - 6
\end{aligned}$$

The surprisingly high degree of the polynomial above indicates that the underlying mathematical structure of these look and say sequences may be more complicated than the simple "say what you see" rule might initially lead you to believe. Indeed,

Date: November 20, 2021.

[1] And every other standard look and say sequence!

Conway discovered a beautiful structure of these look and say sequences that is ultimately governed by a linear transformation of a 92-dimensional vector space!

The purpose of this article is definitely not to attempt to improve upon Conway's delightful explanation of his methods and results in [Con]. Instead, we hope to explain how Conway's methods and results can be generalized to other (nonstandard) look and say sequences. Many variations of look and say sequences have already been explored (see e.g. [EBGSN$^+$2, EBGSN$^+$1], [OM], [Mor], [SS]). We will be interested in look and say sequences that arise from various (nonstandard) number systems. For example, one could use Roman numerals[2] to generate the following look and say sequence:

$$\text{I} \to \text{II} \to \text{III} \to \text{IIII} \to \text{IVI} \to \text{IIIVII} \to \text{IIIIIVIII} \to \text{VIIVIIII} \to \cdots$$

1.1. **Outline.** In Section 2 we will completely describe the structure of all look and say sequences for a particular nonstandard number system we call the *negafibnary number system*, which is related to Fibonacci numbers. The structure in the negafibnary case is much simpler than the standard case worked out by Conway. Along the way we will compare the negafibnary results to the analogous results in [Con]. Moreover, we will provide exercises so that the reader can work out the details for a third case coming from the so-called *negabinary number system*.

2. Negafibnary look and say sequences

2.1. **The negafibnary number system.** In this subsection we will explain how to create a binary number system from Fibonacci numbers using something called a *Zeckendorf representation*. First, recall the definition of the Fibonacci sequence:

$$F_0 = 0, \quad F_1 = 1, \quad F_{n+1} = F_n + F_{n-1}.$$

Although Fibonacci numbers are usually considered with nonnegative index, we always have $F_{n-1} = F_{n+1} - F_n$ which allows use to extend Fibonacci numbers to negative indexes. Indeed, $F_{-1} = F_1 - F_0 = 1 - 0 = 1$, $F_{-2} = F_0 - F_{-1} = 0 - 1 = -1$, and so on. Here is the extended Fibonacci sequence:

$$\ldots, 34, -21, 13, -8, 5, -3, 2, -1, 1, 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \ldots$$

Note that $F_{-k} = (-1)^{k+1} F_k$ for all k.

Zeckendorf's Theorem states that every positive integer can be written uniquely as a sum of nonadjacent Fibonacci numbers (see [Zec]). For our purposes, we will use the following negative index analog of Zeckendorf's Theorem due to Bunder:

Theorem 2.1. [Bun] *Every nonzero integer can be written uniquely in the form* $\sum_{j=1}^{k} b_j F_{-j}$ *where each* $b_j \in \{0, 1\}$*,* $b_k = 1$*, and* $b_{j+1} = 0$ *whenever* $b_j = 1$*.*

If we interpret the b_j's in the theorem above as bits we obtain a nonstandard binary representation of the integers. More precisely, we will write the binary string $b_k \cdots b_2 b_1$ for the integer $n = \sum_{j=1}^{k} b_j F_{-j}$. We will call $b_k \cdots b_2 b_1$ the *negafibnary representation* of n. For example, 101 is the negafibnary representation of the integer $F_{-3} + 0 \cdot F_{-2} + F_{-1} = 2 + 0 + 1 = 3$. Similarly, 10010 is the negafibnary

[2]Conway once said that Roman numeral look and say sequences correspond to a degree 20 polynomial (see [Har]). I would love for someone to show me how to prove this.

representation of $F_{-5} + F_{-2} = 5 - 1 = 4$. Using this terminology, Theorem 2.1 says that every integer has a unique negafibnary representation (with no adjacent 1's). The following table shows the negafibnary representation for the first few nonnegative integers:

0	1	2	3	4	5	6	7	8	9
0	1	100	101	10010	10000	10001	10100	10101	1001010

2.2. **A negafibnary look and say sequence.** Using negafibnary representations gives us a new way to *say* what we see when generating a look and say sequences. For example, if we look at 1111 then we see four 1's. Since the negafibnary representation of four is 10010 we would say 100101. As in the standard case, repeatedly applying this say-what-you-see operation will generate a look and say sequence.

Consider the look and say sequence starting with the seed 0. First, we see one 0 so the next term is 10. From 10 we see one 1 and one 0 so the next term is 1110. Now, looking at 1110 we see three 1's followed by one 0; since the negafibnary representation of three is 101, the next term will be 101110. Continuing on in this manner gives us the following look and say sequence:

$$0 \to 10 \to 1110 \to 101110 \to 1110101110 \to 1011101110101110 \to \cdots \tag{2.1}$$

You might be able to spot a pattern by inspecting the terms above. Following Conway, in order to fully understand the structure of the sequence we should "split" it into smaller sequences. We explain how to do so in §2.4. First, we introduce another example of a nonstandard look and say sequence.

2.3. **The negabinary case.** Suppose $b_0, b_1, \ldots, b_k \in \{0, 1\}$ are bits with $b_k = 1$. We call $b_k \cdots b_1 b_0$ the *negabinary representation* of the integer $n = \sum_{j=0}^{k} b_j(-2)^j$. For example, $6 = (-2)^4 + (-2)^3 + (-2)^1$ so the negabinary representation of 6 is 11010. It turns out that every integer has a unique negabinary representation. As we develop the theory of negafibnary look and say sequences, we will explore the analogous structure of negabinary look and say sequences in exercises.

Exercise 2.2. Find the negabinary representation for each of $1, 2, 3, \ldots, 10$.

Exercise 2.3. Write the first several terms of the negabinary look and say sequence starting with seed 0.

2.4. **Splitting into elements.** Consider the following negafibnary look and say sequences with seeds $x = 0$, $y = 10$, and $z = 010$:

$$\begin{array}{ccccccccccc} x & = & 0 & \to & 10 & \to & 1110 & \to & 101110 & \to & \cdots \\ y & = & 10 & \to & 1110 & \to & 101110 & \to & 1110101110 & \to & \cdots \\ z & = & 010 & \to & 101110 & \to & 1110101110 & \to & 1011101110101110 & \to & \cdots \end{array}$$

Notice that each term of the look and say sequence with seed z can be obtained by concatenating the corresponding terms from seed x to the left of the terms from seed y. When this happens we say that z *splits* and write $z = x.y$. Note that the splitting allows us to completely determine the sequence with seed z from the sequences with smaller seeds x and y. This reduction is the key to Conway's method for analyzing look and say sequences.

When reducing look and say sequences via splitting, one must be careful that the splitting occurs for *all* of the terms in a look and say sequence. For example, consider the standard look and say sequences with seeds $x = 3$, $y = 2$, and $z = 32$:

$$\begin{array}{ccccccccccccc} x & = & 3 & \to & 13 & \to & 1113 & \to & 3113 & \to & 132113 & \to & \cdots \\ y & = & 2 & \to & 12 & \to & 1112 & \to & 3112 & \to & 132113 & \to & \cdots \\ z & = & 32 & \to & 1312 & \to & 11131112 & \to & 31133112 & \to & 1321232112 & \to & \cdots \end{array}$$

While the first few terms of sequence from seed z is obtained by concatenating those from x and y, eventually we find a term from z, namely 1321232112, which differs from the concatenation of the corresponding terms from x and y, namely 132113132113. Thus, in the standard case $32 \neq 3.2$.

Problem 2. *Determine precisely when and how look and say sequences split.*

The solution to Problem 2 for standard look and say sequences is Conway's Splitting Theorem (see [Con]). Before providing the solution in the negafibnary case, it will be convenient to introduce a bit of notation. Let us write xy for the concatenation of x and y (i.e. x placed to the left of y). For example, if $x = 10$ and $y = 11100$ we have $xy = 1011100$. Note that we are straying from the usual convention in that xy does not refer to multiplication. Now we are in a good position to solve Problem 2 in the negafibnary case:

Theorem 2.4. *(Negafibnary Splitting Theorem) Given any binary string z, the corresponding negafibnary look and say sequence splits as $z = x.y$ whenever $z = xy$, the rightmost bit in x is a 0, and the leftmost bit in y is a 1.*

Proof. Write $x = x_0 \to x_1 \to x_2 \to \cdots$, $y = y_0 \to y_1 \to y_2 \to \cdots$, and finally $z = z_0 \to z_1 \to z_2 \to \cdots$ for the relevant look and say sequences. We use induction to show that for every $k \geq 0$ we have (i) $z_k = x_k y_k$, (ii) the rightmost bit of x_k is a 0, and (iii) the leftmost bit of y_k is a 1. The base case is given by the hypothesis of the theorem. For the inductive step, assume (i)–(iii) hold for some $k \geq 0$. Now, since the rightmost bit of x_k is 0, the last thing you see in x_k will be some run of 0's, hence x_{k+1} will also end with a 0. Next, since the start of y_{k+1} is the negafibnary representation of the number of 1's on the very left of y_k, and every negafibnary representation of a positive integer starts with a 1, it follows that y_{k+1} starts with a 1. Finally, since x_k ends with a 0 and y_k starts with a 1, there is no interaction between x_k and y_k when performing the say-what-you-see operation on $z_k = x_k y_k$. It follows that $z_{k+1} = x_{k+1} y_{k+1}$. □

Exercise 2.5. Does Theorem 2.4 still hold if we replace negafibnary with negabinary? Why or why not?

Using Theorem 2.4 we can completely split any binary string in the negafibnary case. For example, $111011000 = 1110.11000$. Some binary strings can be split more than once: $101101001100 = 10.110.100.1100$. More generally, any binary string splits into strings consisting of some 1's followed by some 0's. To state this a bit more precisely, let us write x^n for the concatenation of the string x with itself n times. Then every binary string in the negafibnary case splits into strings of the form

$$1^m 0^n = \underbrace{1 \cdots 1}_{m} \underbrace{0 \cdots 0}_{n} \tag{2.2}$$

where m and n are nonnegative integers.

Strings that cannot be split are called *elements*. An arbitrary string is a *compound* of the elements it splits into. We have shown that the elements in the negafibnary case are precisely the binary strings (2.2). This terminology goes back to Conway's work in [Con]. Conway's splitting theorem for the standard case is more complicated than Theorem 2.4, and thus the elements in the standard case do not possess a simple form like (2.2). However, Conway found 92 *common elements* that appear in almost every look and say sequence; he named these common elements after the 92 naturally-occurring elements hydrogen, helium,. . ., uranium.

Since every string splits into elements, and the say-what-you-see operation on the elements do not interact with one another, the structure of any look and say sequence is completely determined by the elements that appear in the sequence. For example, consider the sequence (2.1). Every term after the initial seed is a compound of the two elements 10 and 1110. Thus we will be able to solve Problem 1 for sequence (2.1) by analyzing the say-what-you-see operation on just those two elements.

We say an element appears *frequently* in a look and say sequence if that element appears in infinitely many of the terms. For example, the only elements that appear frequently in (2.1) are 10 and 1110. Determining the frequent elements in a look and say sequence is a crucial step towards solving Problem 1.

Problem 3. *Given a look and say sequence equipped with a splitting theorem, determine all the elements that appear frequently.*

Exercise 2.6. Find all the elements that appear frequently in the negabinary look and say sequence from Exercise 2.3.

2.5. **Decay and decay matrices.** Continuing with the chemical terminology, if the say-what-you-see operation takes $x \to y$ then we say the compound x *decays* into y. For example, in the negafibnary case we have $0 \to 10$, so 0 decays into 10. Now let

$$e_1 = 10 \quad \text{and} \quad e_2 = 1110. \tag{2.3}$$

Then $e_1 = 10 \to 1110 = e_2$ and $e_2 = 1110 \to 101110 = e_1e_2$. Thus e_1 decays into e_2, and e_2 decays into e_1e_2.

Exercise 2.7. Determine the negabinary decay of each element from Exercise 2.6.

Suppose $e_1, \dots, e_k$ is a collection of elements such that each e_j decays into a compound of some of the e_i's. To these elements we associate a k-dimensional vector space of column vectors called the *compound space*. Given a compound of these elements, the corresponding *compound vector* is the column vector whose ith entry is the number of times e_i appears in the compound. For example, consider the last string written in the sequence (2.1). Keeping with (2.3), that string is a compound of 2 e_1's and 3 e_2's, so the corresponding compound vector is $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

Returning to the general setup with k elements, the *decay matrix* is the $k \times k$ matrix whose i, j-entry is the number of times e_i occurs in the decay of e_j. In other words, the jth column is the compound vector corresponding to the decay of e_j. For example, with (2.3) in the negafibnary case the decay matrix is

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}. \tag{2.4}$$

Indeed, the first column is $\begin{pmatrix}0\\1\end{pmatrix}$ since e_1 decays into zero e_1's and one e_2; the second column is $\begin{pmatrix}1\\1\end{pmatrix}$ because e_2 decays into one e_1 and one e_2.

Exercise 2.8. Use your answer to Exercise 2.7 to find a decay matrix for the negabinary look and say sequence from Exercise 2.3. To get started you will need to fix an order on the elements. Your matrix will depend on the order you choose.

The decay matrix is so-named because multiplying a compound vector by the decay matrix corresponds to decaying the compound via the say-what-you-see operation. For example, we have already seen $1110101110 \to 1011101110101110$ in the negafibnary case (see (2.1)). The first of these compounds corresponds the vector $\begin{pmatrix}1\\2\end{pmatrix}$. Multiplying by the decay matrix we get $\begin{pmatrix}0&1\\1&1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix}$, which we've already seen is the compound vector for the decay compound. Similarly, one can check that every step in the following sequence is obtained via left multiplication by the decay matrix (2.4):

$$\begin{pmatrix}1\\0\end{pmatrix} \mapsto \begin{pmatrix}0\\1\end{pmatrix} \mapsto \begin{pmatrix}1\\1\end{pmatrix} \mapsto \begin{pmatrix}1\\2\end{pmatrix} \mapsto \begin{pmatrix}2\\3\end{pmatrix} \mapsto \begin{pmatrix}3\\5\end{pmatrix} \mapsto \cdots \tag{2.5}$$

Note that the first 5 vectors in (2.5) are the compound vectors for the last 5 terms of (2.1). The first term of (2.1), 0, does not correspond to a compound vector in this compound space since 0 is not a compound of $e_1 = 10$ and $e_2 = 1110$. The last vector in (2.5) tells us that the next (unwritten) term in (2.1) will be a compound of 3 e_1's and 5 e_2's. Notice that all the coordinates in (2.5) are Fibonacci numbers!

Exercise 2.9. Find the sequence of compound vectors for the negabinary look and say sequence from Exercise 2.3. In other words, (2.1) is to (2.5) as your answer to Exercise 2.3 is to what? The coordinates in these compound vectors are the so-called *Padovan numbers*.

Although the decay matrix encodes a significant amount of decay information, one cannot completely recover a look and say sequence from the corresponding decay matrix and compound vectors. This is because the compound vector of a given compound does not encode the *order* in which the elements appear in a given compound. For example, the two different compounds e_1e_2 and e_2e_1 have the same compound vector $\begin{pmatrix}1\\1\end{pmatrix}$. However, if we are concerned with a problem that does not depend on the order of the elements, like Problem 1, then we can often find the solution within the structure of the decay matrix.

2.6. **Solving Problem 1 with an eigenvalue.** The key to solving Problem 1 is to find eigenvalues of the decay matrix D. In order to find eigenvalues we first find the characteristic polynomial[3] $\det(\lambda I - D)$. For example, the characteristic polynomial of the decay matrix (2.4) is

$$\det\begin{pmatrix}\lambda & -1\\ -1 & \lambda-1\end{pmatrix} = \lambda^2 - \lambda - 1. \tag{2.6}$$

[3]Some prefer to define the characteristic polynomial of D as $\det(D-\lambda I)$. I prefer $\det(\lambda I - D)$ since it will always be a monic polynomial (i.e. the leading coefficient will be 1). These two conventions at most differ by a multiple of -1. In particular, they always have the same roots

.

The eigenvalues of a matrix are the roots of its characteristic polynomial. The roots of (2.6) are $\lambda = \frac{1\pm\sqrt{5}}{2}$. In particular, the maximal real eigenvalue is the celebrated golden ratio

$$\varphi = \frac{1+\sqrt{5}}{2} = 1.61803\ldots \tag{2.7}$$

Exercise 2.10. Find the characteristic polynomial for the decay matrix from Exercise 2.8. Use technology to determine the maximal real root of the characteristic polynomial. This number is the so-called *plastic ratio* often denoted ρ.

We will show that the golden ratio is negafibnary analog of Conway's constant. In particular, the ratio of the lengths of successive terms of (2.1) approach the golden ratio, and thus solves Problem 1 for that look and say sequence. This procedure works more generally:

Theorem 2.11. *Given a look and say sequence whose terms are eventually compounds consisting of elements $e_1, \ldots, e_k$, the ratios of the lengths of the terms approach the maximal real eigenvalue*[4] *of the corresponding $k \times k$ decay matrix.*

In the standard case, Conway's 92 elements give rise to a 92×92 decay matrix. Conway's constant (1) is the maximal real root of this decay matrix. The irreducible degree 71 polynomial given at the beginning of the introduction is a factor of the degree 92 characteristic polynomial.

Instead of proving Theorem 2.11 in general, we will verify only the negafibnary case for the look and say sequence (2.1). However, with just a bit more linear algebra the argument given below can be generalized to give a full proof of Theorem 2.11.

Proof of Theorem 2.11 for (2.1). Let D denote the decay matrix (2.4) and write $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$ for the eigenvalues of D. One can show that $v_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$ are respective eigenvectors of D. In other words, $Dv_i = \lambda_i v_i$ for both $i = 1, 2$. Now, set $u_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and write $u_n = D^n u_0$ for each $n > 0$. In other words, the sequence $u_0 \mapsto u_1 \mapsto u_2 \mapsto \cdots$ is precisely (2.5). Since v_1 and v_2 span all of $\mathbb{R}^2$ we can find $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $u_0 = \alpha_1 v_1 + \alpha_2 v_2$. The interested reader can find the exact values of α_1 and α_2. For our purposes, it suffices to know that such scalars exist and $\alpha_1 \neq 0$ (because u_0 is not a scalar multiple of v_2). The following computation is the key to the proof:

$$\begin{aligned}
\frac{1}{\lambda_1^n} u_n &= \frac{1}{\lambda_1^n} D^n u_0 && \text{(definition of } u_n\text{)} && (2.8)\\
&= \frac{1}{\lambda_1^n} D^n(\alpha_1 v_1 + \alpha_2 v_2) && \text{(definition of } \alpha_1 \text{ and } \alpha_2\text{)} && (2.9)\\
&= \frac{1}{\lambda_1^n} (\alpha_1 \lambda_1^n v_1 + \alpha_2 \lambda_2^n v_2) && \text{(definition of eigenvector)} && (2.10)\\
&= \alpha_1 v_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^n v_2 && && (2.11)
\end{aligned}$$

[4]The fact that such an eigenvalue will always exist follows from the Perron-Frobenius Theorem.

$$\to \alpha_1 v_1 \quad \text{as } n \to \infty \qquad\qquad (\text{since } |\lambda_1| > |\lambda_2|) \tag{2.12}$$

To connect the computation above to the statement of the theorem we let ℓ denote the vector whose coordinates are the lengths of the elements $e_1 = 10$ and $e_2 = 1110$. More precisely, we have $\ell = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$. Then the length of any compound corresponding to compound vector u is given by the dot product $u \cdot \ell$. Thus, the ratio of lengths of successive terms of (2.1) is given by $\dfrac{u_{n+1} \cdot \ell}{u_n \cdot \ell}$. We can use the computation above to determine the limit of this ratio:

$$\frac{u_{n+1} \cdot \ell}{u_n \cdot \ell} = \lambda_1 \left(\frac{\frac{1}{\lambda_1^{n+1}} u_{n+1} \cdot \ell}{\frac{1}{\lambda_1^n} u_n \cdot \ell} \right) \to \lambda_1 \left(\frac{\alpha_1 v_1 \cdot \ell}{\alpha_1 v_1 \cdot \ell} \right) = \lambda_1 \quad \text{as } n \to \infty.$$

□

Exercise 2.12. Mimic the argument above to prove that the ratios of the lengths of terms in the negabinary look and say sequence from Exercise 2.3 approach the plastic ratio from Exercise 2.10.

2.7. **Abundances.** Since Problem 1 can be solved using the largest real eigenvalue of a decay matrix, it is natural to ask what the corresponding eigenvector tells us about the look and say sequence. We will see that eigenvectors can be used to solve the following problems:

Problem 4. *Determine the limiting relative abundance of each element in a look and say sequence.*

Problem 5. *Determine the limiting relative abundance of each digit in a look and say sequence.*

To make Problem 4 more precise we introduce some notation for abundances. Given a column vector u whose coordinates sum to $s \neq 0$, we write $\mathrm{ab}(u) = \frac{1}{s}u$ for the *abundance vector of* u. For example, consider the last term written in the look and say sequence (2.1), namely the compound $e_1e_2e_2e_1e_2$. The corresponding compound vector is $u = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ so the abundance vector is $\mathrm{ab}(u) = \frac{1}{2+3}u = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$. Note that the coordinates of $\mathrm{ab}(u)$ are exactly the relative abundances of the elements in the compound: 40% of the elements in the compound are e_1 and 60% are e_2. Thus, to solve Problem 4 is to find the limit of the abundance vectors for a given look and say sequence. The following theorem explains how to find such a limit.

Theorem 2.13. *Given a look and say sequence whose terms are eventually compounds consisting of elements $e_1, \ldots, e_k$, the corresponding sequence of abundance vectors approach $\mathrm{ab}(v)$ where v is any eigenvector of the decay matrix with maximal real eigenvalue.*

Before proving Theorem 2.13 we will see how it can be used to give explicit solutions to both Problem 4 and Problem 5 in the negafibnary case.

Consider the negafibnary look and say sequence (2.1). Taking the abundance of each vector in (2.5) gives us the sequence of abundance vectors:

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \mapsto \begin{pmatrix} 0.333\ldots \\ 0.666\ldots \end{pmatrix} \mapsto \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix} \mapsto \begin{pmatrix} 0.375 \\ 0.625 \end{pmatrix} \mapsto \cdots \tag{2.13}$$

Now, we have already seen that the maximal eigenvalue of the decay matrix (2.4) is the golden ratio $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and a corresponding eigenvector is $v_1 = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix}$. Using the notation $\varphi = \frac{1+\sqrt{5}}{2}$ we have

$$\mathrm{ab}(v_1) = \frac{1}{1+\varphi}\begin{pmatrix} 1 \\ \varphi \end{pmatrix} = \frac{1}{\varphi^2}\begin{pmatrix} 1 \\ \varphi \end{pmatrix} = \begin{pmatrix} \varphi^{-2} \\ \varphi^{-1} \end{pmatrix} = \begin{pmatrix} 0.381\ldots \\ 0.618\ldots \end{pmatrix}. \tag{2.14}$$

According to Theorem 2.13 the limit of the sequence (2.13) is (2.14). In particular, abundances of elements $e_1 = 10$ and $e_2 = 1110$ in (2.1) are approaching $\varphi^{-2} \approx 0.38$ and $\varphi^{-1} \approx 0.62$ respectively. Hence, the ratio of the number of e_2's to e_1's occurring in (2.1) is approaching the golden ratio. This solves Problem 4 in the negafibnary case for (2.1).

Exercise 2.14. Solve Problem 4 in the negabinary case for the look and say sequence from Exercise 2.3. In particular, explain how the abundances of the elements are related by the plastic ratio ρ (see Exercise 2.10).

In the standard case, the abundances of Conway's 92 common elements correspond to an eigenvector of a 92×92 decay matrix. All 92 abundances are listed in The Periodic Table found in [Con]. Conway was not concerned with Problem 5, but solving that problem is easy once we have a solution to Problem 4. Indeed, the abundance of each digit in a look and say sequence can be determined from the abundance of each element in the look and say sequence together with the abundance of each digit in each element. More precisely, if v is the limiting abundance vector for elements (the solution to Problem 4) and A is the matrix whose jth column lists the abundance of each digit in the jth element, then the product Av gives the limiting abundance of each digit in the look in say sequence.

For example, consider the negafibnary look and say sequence (2.1). Since $e_1 = 10$ we see the abundances of 0 and 1 in e_1 are both 0.5. Similarly, the abundances of 0 and 1 in $e_2 = 1110$ are 0.25 and 0.75 respectively. Packaging these abundances into a matrix A and multiplying by the vector v given in (2.14) gives

$$\begin{pmatrix} 0.5 & 0.25 \\ 0.5 & 0.75 \end{pmatrix}\begin{pmatrix} \varphi^{-2} \\ \varphi^{-1} \end{pmatrix} = \begin{pmatrix} 0.5\varphi^{-2} + 0.25\varphi^{-1} \\ 0.5\varphi^{-2} + 0.75\varphi^{-1} \end{pmatrix} = \begin{pmatrix} 0.25 + 0.25\varphi^{-2} \\ 0.5 + 0.25\varphi^{-1} \end{pmatrix} \approx \begin{pmatrix} 0.345 \\ 0.655 \end{pmatrix}.$$

This gives us a solution to Problem 5 for the negafibnary look and say sequence (2.1): eventually there are approximately 34.5% 0's and 65.5% 1's in each term.

Exercise 2.15. Solve Problem 5 for the negabinary look and say sequence from Exercise 2.7.

As promised, we conclude this section with a proof of Theorem 2.13. The proof below relies on our proof of Theorem 2.11, which was only given for the negafibnary (2.1). The motivated reader is encouraged to generalize the following proof to the general case.

Proof of Theorem 2.13 for (2.1). Let λ_i, v_i, α_i, and u_n be as in the proof of Theorem 2.11. To prove the theorem at hand, we will show $\mathrm{ab}(u_n) \to \mathrm{ab}(v_1)$ as $n \to \infty$. Note that the sum of the coordinates of any vector u is given by the dot product $u \cdot \mathbf{1}$ where $\mathbf{1}$ is the column vector with all 1's as coordinates. Hence $\mathrm{ab}(u) = \frac{1}{u \cdot \mathbf{1}} u$

for all u. Thus, using the result of computation (2.8)–(2.12) we have

$$\mathrm{ab}(u_n) = \frac{1}{u_n \cdot \mathbf{1}} u_n = \frac{1}{\lambda_1^{-n} u_n \cdot \mathbf{1}} \lambda_1^{-n} u_n \to \frac{1}{\alpha_1 v_1 \cdot \mathbf{1}} \alpha_1 v_1 = \frac{1}{v_1 \cdot \mathbf{1}} v_1 = \mathrm{ab}(v_1)$$

as $n \to \infty$, which completes the proof. □

2.8. **Completing the chemistry.** The elements that appear frequently (i.e. infinitely many times) in a look and say sequence depend on the seed. For example, in the negafibnary case we know the look and say sequence (2.1) with seed 0 has two frequently occurring elements: $e_1 = 10$ and $e_2 = 1110$. On the other hand, if the seed is instead 1 we get the following look and say sequence:

$$1 \to 11 \to 1001 \to 11100011 \to 101110101001 \to 1110101110111011100011 \to \cdots$$

In addition to the elements e_1 and e_2, the elements $1, 11, 100$, and 111000 also appear frequently in the sequence above. There are still other elements that appear frequently in other negafibnary look and say sequences, like 110 in the following:

$$110 \to 100110 \to 111000100110 \to 10111010111000100110 \to \cdots \tag{2.15}$$

In fact, there exist negafibnary look and say sequences with nine different frequent elements[5]. This motivates the following:

Problem 6. *For a given number system, describe all the elements that appear frequently in at least one look and say sequence.*

Exercise 2.16. Find the frequent elements in the negabinary look and say sequence with seed 1.

Exercise 2.17. Find a negabinary look and say sequence which has a frequently occurring element that is not among your answers for Exercises 2.6 and 2.16.

A complete solution to Problem 6 comes in two parts: (1) a list of elements along with their decay and abundances; and (2) a proof that *every* seed eventually decays into compounds of those elements. The solution in the standard case is provided by Conway in [Con] where he provides (1) a *periodic table* of his 92 common elements and (2) a *Cosmological Theorem* stating that every standard look and say sequence eventually consists of compounds of his 92 elements (along with a couple families of so-called *transuranic elements*).

For the negafibnary case we will show that nine elements suffice. These elements are denoted $e_1, e_2, \ldots, e_9$ according to the following:

The Negafibnary Periodic Table

n	e_n	Decay	Abundance
1	10	e_2	$\varphi^{-2} = 0.381\ldots$
2	1110	$e_1 e_2$	$\varphi^{-1} = 0.618\ldots$
3	11100000	$e_1 e_3$	0
4	111000	$e_1 e_2 e_1$	0
5	100	e_4	0
6	110	$e_5 e_6$	0
7	1100000	$e_5 e_7$	0
8	11	$e_5 e_9$	0
9	1	e_8	0

(2.16)

[5]For example, use the seed $1^5 0^5 1101$.

Note that the decay of each of the nine elements is listed in the table above. For example, the $n = 6$ row states that $e_6 = 110$ decays into $e_5e_6 = 100110$, which agrees with the start of (2.15).

Exercise 2.18. Verify the decay column of the periodic table (2.16) by performing the negafibnary say-what-you-see operation on each of the nine elements in the e_n column.

Following §2.5 we can encode the decay column of the periodic table into a 9×9 decay matrix:

$$\begin{pmatrix} 0 & 1 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{2.17}$$

Notice that the smaller decay matrix (2.4) is found in the upper left corner of the larger matrix above. The characteristic polynomial of the 9×9 matrix is

$$\lambda^9 - 4\lambda^8 + 4\lambda^7 + 3\lambda^6 - 7\lambda^5 + 2\lambda^4 + 2\lambda^3 - \lambda^2 = \lambda^2 (\lambda - 1)^4 (\lambda + 1) \left(\lambda^2 - \lambda - 1\right).$$

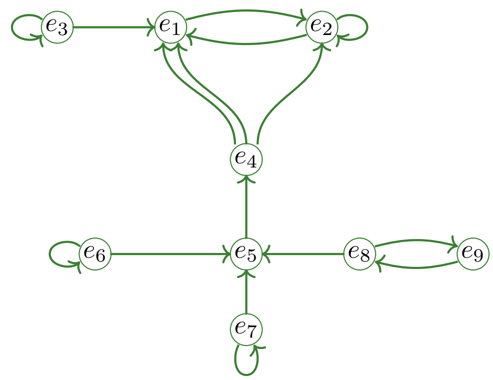
The roots of the polynomial above, i.e. the eigenvalues of (2.17), are $0, \pm 1, \frac{1 \pm \sqrt{5}}{2}$. In particular, the maximal real eigenvalue is $\varphi = \frac{1+\sqrt{5}}{2}$. It follows from Theorem 2.11 that any negafibnary look and say sequence whose frequent elements are those in the periodic table (2.16) grows at rate φ. Moreover, by Theorem 2.13 the relative abundance of the nine elements in such a look and say sequence will approach the abundance vector of any eigenvector with eigenvalue φ. This is exactly how the abundance column of the periodic table (2.16) is obtained.

Exercise 2.19. Create a negabinary periodic table with the following ten elements:

$$110, 10, 1110, 111110, 100, 111100, 11100, 11111100, 11, 1.$$

Use appropriate technology to determine abundances.

Elements that appear in a look and say sequence with nonzero relative abundance zero will be called *abundant*, elements with relative abundance 0 will be called *rare*. According to the periodic table (2.16), if the terms of a negafibnary look and say sequence are eventually compounds of $e_1, \ldots, e_9$ and the two elements e_1, e_2 appear somewhere, then all other elements will be rare in that look and say sequence. In particular, such sequences will always grow at rate φ. However, negafibnary look and say sequence in which e_1 and e_2 do not appear would have different abundant elements, and thus might have different growth rates. It turns out that no such look and say sequences exist in the negafibnary case. To see this for sequences whose terms are compounds of $e_1, \ldots, e_9$ consider the following *decay graph* for these nine elements:



(2.18)

In the graph above each arrow from e_j to e_i indicates an occurrence of e_i in the decay of e_j. For example, since e_4 decays into $e_1e_2e_1$ we have one arrow from e_4 to e_2 and two arrows from e_4 to e_1. In other words, the number of arrows from e_j to e_i in the decay graph is equal to the i, j-entry in the decay matrix (2.17).

Now, looking at the decay graph we see there is a path from each of the nine elements $e_1, \dots, e_9$ to both e_1 and e_2. It follows that e_1 and e_2 occur in any negafibnary look and say sequence whose terms are compounds of $e_1, \dots, e_9$. Hence, all such look and say sequences grow at rate φ.

Exercise 2.20. Draw the decay graph for the ten elements listed in Exercise 2.19 in the negabinary case. What are the possible growth rates of the negabinary look and say sequences whose terms are eventually compounds of those ten elements?

2.9. **Cosmological Theorems.** In this section we will complete our solution to Problem 6 by providing a Cosmological Theorem for negafibnary look and say sequences (see Theorem 2.24). As we have previously mentioned, Conway provided a Cosmological Theorem for standard look and say sequences in [Con]. However, Conway's proof of his Cosmological Theorem was not included in [Con] because it involved a "very subtle and complicated argument, which (almost) reduced the problem to tracking a few hundred cases". Proofs of Conway's Cosmological Theorem can be found in [EZ] and [Lit]. In the negafibnary case, we will provide a relatively simple proof. The steps in the proof are broken into lemmas, the first of which shows there is a bound on the number of consecutive 1's appearing in negafibnary look and say sequence:

Lemma 2.21. *Suppose $x_0 \to x_1 \to x_2 \to \cdots$ is a negafibnary look and say sequence. Whenever $j > 0$, x_j is a compound of elements of the form 1^m0^n with $1 \leq m \leq 3$ and $n \geq 0$.*

Proof. If $j > 0$, x_j starts with the negafibnary representation of some integer, and thus the first bit in x_j is a 1. Therefore x_j is a compound of elements of the form 1^m0^n with $m \geq 1$ and $n \geq 0$. To show each $m \leq 3$, suppose $B = 111$ appears somewhere within x_j. Since negafibnary representations cannot have consecutive 1's, the first (resp. last) 1 in B must be the end (resp. start) of some negafibnary representation, whence there cannot be another 1 in x_j to the left (resp. right) of B. Thus, there cannot be a run of more than 3 consecutive 1's in x_j. $\square$

In order to show the number of consecutive 0's in a negafibnary look and say sequence is bounded, we need a more precise understanding of consecutive 0's in negafibnary representations. Part (3) of the following lemma gives us precisely that:

Lemma 2.22. *Suppose n is an integer with k bits in its negafibnary representation.*

(1) If k is odd, then $F_{k-1} < n \leq F_{k+1}$.
(2) If k is even, then $-F_{k+1} < n \leq -F_{k-1}$.
(3) If $n > 0$ and there is a run of r consecutive 0's in the negafibnary representation of n, then $n \geq F_{r+1}$.

Proof. (1) Suppose $k = 2\ell + 1$ for some integer $\ell \geq 0$. The maximal negafibnary representation with k bits is obtained by using the maximal number of positive Fibonacci numbers, namely $1010\cdots101 = (10)^\ell 1$, which is the negafibnary representation of the integer

$$\sum_{i=0}^{\ell} F_{-2i-1} = \sum_{i=0}^{\ell} (F_{-2i} - F_{-2i-2}) = F_0 - F_{-2\ell-2} = 0 - F_{-k-1} = F_{k+1}.$$

Similarly, we use the maximal number of negative Fibonacci numbers to obtain the minimal k-bit negafibnary representation $100(10)^{\ell-1}$, which corresponds to

$$\begin{aligned} F_{-k} + \sum_{i=1}^{\ell-1} F_{-2i} &= F_{-k} + \sum_{i=1}^{\ell-1} (F_{-2i+1} - F_{-2i-1}) = F_{-k} + F_{-1} - F_{-2\ell+1} \\ &= F_k + 1 - F_{2\ell-1} = F_k + 1 - F_{k-2} = 1 + F_{k-1}. \end{aligned}$$

(2) The proof is similar to that of (1). Set $k = 2\ell$. Then the minimal k-bit negafibnary representation is $(10)^\ell$, which corresponds to

$$\sum_{i=1}^{\ell} F_{-2i} = \sum_{i=1}^{\ell} (F_{-2i+1} - F_{-2i-1}) = F_{-1} - F_{-2\ell-1} = 1 - F_{2\ell+1} = 1 - F_{k+1}.$$

The maximal k-bit negafibnary representation is $10(01)^{\ell-1}$, which corresponds to

$$\begin{aligned} F_{-k} + \sum_{i=1}^{\ell-1} F_{-2i+1} &= F_{-k} + \sum_{i=1}^{\ell-1} (F_{-2i+2} - F_{-2i}) = F_{-k} + F_0 - F_{-2\ell+2} \\ &= F_{-k} - F_{-k+2} = -F_{-k+1} = -F_{k-1}. \end{aligned}$$

(3) Let n denote the minimal positive integer whose negafibnary representation contains a run of at least r consecutive 0's, and let k denote the number of bits in that representation. It follows from (1) and (2) that k will be the smallest odd integer which allows for r consecutive 0's. In other words, $k = r + 1$ if r is even and $k = r + 2$ if r is odd. Moreover, n corresponds to the smallest such k-bit representation with r consecutive 0's, namely 10^r or 10^{r+1}. Therefore $n = F_{r+1}$ or $n = F_{r+2}$. In either case, $n \geq F_{r+1}$. □

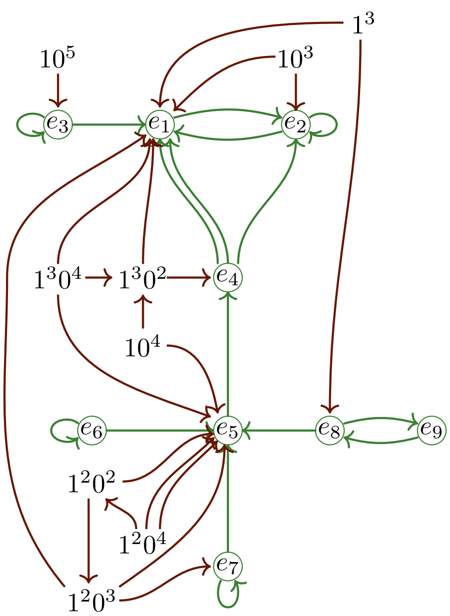
We are now in position to show that in any negafibnary look and say sequence, the terms will eventually have at most five consecutive 0's:

Lemma 2.23. *Suppose the element $1^{m'}0^{n'}$ appears in the negafibnary decay of $1^m 0^n$. If $1 \leq m \leq 3$ and $n > 5$, then $1 \leq m' \leq 3$ and $n' < n$.*

Proof. The inequality for m' follows from Lemma 2.21. To show the inequality for n', let a and b denote the negafibnary representations of m and n respectively so that $1^m0^n \to a1b0$. Since $1 \leq m \leq 3$ we know a is one of $1, 100$, or 101. Thus, if $1^{m'}0^{n'}$ appears in $a1$ then $n' \leq 2 < 5 < n$. On the other hand, if $1^{m'}0^{n'}$ appears in $b0$, then there must exist a run of at least $n'-1$ consecutive 0's in b. Whence $n \geq F_{n'}$ by part (3) of Lemma 2.22. Finally, since $n > 5$ we have $F_n > n \geq F_{n'}$, which implies $n > n'$. $\square$

Theorem 2.24. *(Negafibnary Cosmological Theorem) The terms of every negafibnary look and say sequence are eventually compounds of the nine elements listed in the negafibnary periodic table (2.16). Consequently, every negafibnary look and say sequence grows at the rate φ.*

Proof. By Lemma 2.21 it suffices to show every look and say sequence with seed of the form 1^m0^n with $1 \leq m \leq 3$ has terms that are eventually compounds of the nine elements listed in the periodic table (2.16). We will do so by inducting on n. For the base case we consider the 18 seeds 1^m0^n with $1 \leq m \leq 3$ and $0 \leq n \leq 5$. The following decay graph (which extends (2.18)) summarizes the decay of all 18 elements, completing the base case:



For the inductive step, fix $n > 5$ and assume all elements of the form $1^{m'}0^{n'}$ with $1 \leq m' \leq 3$ and $0 \leq n' < n$ eventually decay in into compounds of the elements in (2.16). By Lemma 2.23, each 1^m0^n with $1 \leq m \leq 3$ decays into a compound of such $1^{m'}0^{n'}$'s, and thus 1^m0^n eventually decays into the elements in (2.16). $\square$

Exercise 2.25. The goal of this exercise is to prove the following:

(Negabinary Cosmological Theorem) The terms of every negabinary look and say sequence are eventually compounds of the ten elements listed in Exercise 2.19.

The following steps will lead to such a proof:

(1) Assume $n > 0$ has ℓ bits in its negabinary representation. Show that ℓ is odd so that $\ell = 2k+1$ for some integer k. Moreover, show $n \geq \dfrac{4^k+2}{3}$.
(2) Assume m and n have negabinary representations a and b respectively so that $1^m0^n \to a1b0$. Now, let $2k_1+1$ and $2k_2+1$ denote the number of bits in a and b respectively, and set $k = \max\{k_1, k_2\}$. Show that the length of $a1b0$ is less than the length of 1^m0^n whenever $4k+4 < m+n$.
(3) Continuing with the notation of the previous part, use parts (1) and (2) to show that the length of $a1b0$ is less than the length of 1^m0^n whenever $4k+4 < \dfrac{4^k+2}{3}$.
(4) Show that $4k+4 < \dfrac{4^k+2}{3}$ if and only if $k \geq 3$.
(5) Use the previous parts to show that the length of $a1b0$ is less than the length of 1^m0^n whenever $m \geq 22$ or $n \geq 22$.
(6) Write a computer program[6] to show that each of the 462 elements of the form 1^m0^n with $1 \leq m < 22$ and $0 \leq n < 22$ eventually decays into compounds of the ten elements listed in Exercise 2.19.
(7) Use the previous two parts to give an inductive proof that every element of the form 1^m0^n with $m > 0$ and $n \geq 0$ eventually decays into compounds of the ten elements listed in Exercise 2.19. Explain why this proves The Negabinary Cosmological Theorem.

References

[Bun] M. W. Bunder. Zeckendorf representations using negative Fibonacci numbers. *The Fibonacci Quarterly*, 30:111–115, 1992.

[Con] J. H. Conway. The weird and wonderful chemistry of audioactive decay. *In: Cover T.M., Gopinath B. (eds) Open Problems in Communication and Computation. Springer, New York, NY*, 1987.

[EBGSN+1] Éric Brier, Rémi Géraud-Stewart, David Naccache, Alessandro Pacco, and Emanuele Troiani. The look-and-say the biggest sequence eventually cycles, 2020.

[EBGSN+2] Éric Brier, Rémi Géraud-Stewart, David Naccache, Alessandro Pacco, and Emanuele Troiani. Stuttering Conway sequences are still Conway sequences, 2020.

[EZ] S. B. Ekhad and D. Zeilberger. Proof of Conway's lost cosmological theorem. *Electronic Research Announcements of the American Mathematical Society*, 3:78–82, 1997.

[Har] B. Haran. Look-and-Say Numbers (feat John Conway) – Numberphile. `https://www.youtube.com/watch?v=ea7lJkEhytA`.

[Lit] R. A. Litherland. Conway's cosmological theorem.

[Mor] T. Morrill. Look, knave. *Bulletin of the Australian Mathematical Society*, 103(2):210–217, 2021.

[OM] Óscar Martín. Look-and-say biochemistry: Exponential rna and multistranded dna. *The American Mathematical Monthly*, 113(4):289–307, 2006.

[SS] J. Sauerberg and L. Shu. The long and the short on counting sequences. *The American Mathematical Monthly*, 104(4):306–317, 1997.

[Zec] E. Zeckendorf. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. R. Sci. Liège*, 41:179–182, 1972.

E-mail address: `jcomes@collegeofidaho.edu`

[6] You can write a short python program using the `look_and_say` module.