LOOK, KNAVE

THOMAS MORRILL

(Received 14 April 2020; accepted 20 May 2020; first published online 21 October 2020)

Dedicated to the life and works of John Horton Conway

Abstract

We examine a recursive sequence in which \( s_n \) is a literal description of what the binary expansion of the previous term \( s_{n-1} \) is not. By adapting a technique of Conway, we determine the limiting behaviour of \( \{s_n\} \) and dynamics of a related self-map of \( 2^\mathbb{N} \). Our main result is the existence and uniqueness of a pair of binary sequences, each the complement-description of the other. We also take every opportunity to make puns.

2020 Mathematics subject classification: primary 11B37; secondary 11B05.

Keywords and phrases: recursion, Look-Say sequence, Knights and Knaves.

1. Introduction

The Look-Say sequence is defined as follows. Let \( s_1 = 1 \). Given \( s_n \), the next term of the sequence is a literal description of the digits of the previous one [4]. The first few terms are

\[ 1, 11, 21, 1211, 111221, \ldots. \]

We use \( |s| \) to denote the length of a finite string \( s \).

**Theorem 1.1** (Conway [1, 5]). *Let \( s_n \) be the \( n \)th term of the Look-Say sequence. Then

\[ \lim_{n \to \infty} \frac{|s_{n+1}|}{|s_n|} = \lambda, \]

where

\[ \lambda = 1.3035 \ldots. \]

Shockingly, \( \lambda \) is an algebraic integer of degree 71 [5]. Theorem 1.1 follows from Conway’s cosmological theorem [1]. In short, the terms of any Look-Say-type sequence (not necessarily starting at \( s_1 = 1 \)) will eventually decompose into a concatenation of certain fundamental substrings identified by Conway as ‘elements’.

Supported by Australian Research Council Discovery Project DP160100932.

© 2020 Australian Mathematical Publishing Association Inc.
This problem has also been considered in terms of binary strings. Given a binary string $s_n$, the next term of the binary Look-Say sequence is a literal description of the bits of the previous term, where the counts are expressed in base two [6]. The first few terms are

$$1, 11, 101, 111011, \ldots$$

**Theorem 1.2 (Johnston [2]).** Let $s_n$ be the $n$th term of the binary Look-Say sequence. Then

$$\lim_{n \to \infty} \frac{|s_{n+1}|}{|s_n|} = \lambda,$$

where

$$\lambda = 1.465571 \ldots$$

We shake this up by introducing a new player, a Knave in the style of Smullyan. As opposed to the previous recursions, our $s_n$ is instead the literal description of what the bits of $s_{n-1}$ are not. Our main result concerns the limiting behaviour of the Look-Knave sequence.

**Theorem 1.3.** There is a unique pair of binary sequences $S_{\text{even}}$ and $S_{\text{odd}}$ such that $S_{\text{even}}$ is a literal description of the bitwise complement of $S_{\text{odd}}$ and vice versa.

The rest of the paper is organised as follows. In Section 2 we define the Look-Knave sequence and pose our problem. Then, in Section 3, we simplify the problem and prove Theorem 1.3. Finally, in Section 4, we offer avenues for future work.

2. The Knave

Recall Smullyan’s game of Knights and Knaves, a logic puzzle in which Knights always tell the truth and Knaves are always compelled to lie [3]. Our Knave is a very idiosyncratic liar. When looking at a string of $n$ 0s, the Knave correctly tells us they see $n$ bits of the same parity, but they will lie by saying that there are $n$ 1s. Likewise, while looking at $k$ 1s, the Knave will happily tell us there are $k$ 0s instead.

The Knave understands how to express natural numbers in base two and will write down their observations for us as such. Thus, when the Knave looks at the string

$$110,$$

they write down

$$10 \ 0 \ 1 \ 1$$

for the two 0s and one 1 they claim to have seen. Here, we have inserted white space to enhance the Knave’s handwriting.

Now, our Knave has not yet realised that they could have lied about their count by inverting the bits representing $n$ and $k$ above. I won’t tell them if you won’t.
TABLE 1. The first 10 entries of the Look-Knave sequence.

| $s_{2n+1}$       | $s_{2n+2}$       |
|------------------|------------------|
| 1                | 10               |
| 1011             | 1011100          |
| 10111010101      | 1011100011101110 |
| 10111101111011101 | 1011100011101011101 |
| 1011111011110111011101               | 1011100011101011100011100 |

Thus begins our new game. We will supply a binary string and command ‘Look, Knave’. Dutifully, the Knave will read the string and then record the observations on a fresh piece of paper for us. We return this paper to the Knave, who reads their own report and transcribes it in the only way a Knave can. The game continues.

Let us begin with the string $s_1 = 1$ and take $s_n$ to be the Knave’s description of $s_{n-1}$. This defines the Look-Knave sequence. For example, $s_3 = 1011$. We see that there is one bit which is not 0, followed by one bit which is not 1 and then two bits which are not 0. Thus, $s_4$ must be the string 10111000. In short, $s_n$ is a a binary string describing precisely what $s_{n-1}$ is not.

Looking at Table 1, it is tempting to conjecture that the subsequences $\{s_{2n+1}\}$ and $\{s_{2n+2}\}$ are approaching some bitwise limits. So, do there exist binary sequences $S_{\text{even}}$ and $S_{\text{odd}}$ such that $S_{\text{odd}}$ is the Knave’s description of $S_{\text{even}}$ and vice versa?

A binary sequence $S$ can be described by the Knave, so long as the tail end of $S$ is not all 0s or all 1s. Let $S \subset 2^N$ be the set of all such sequences. Then the Knave imposes a map $k : S \rightarrow S$.

It will be convenient to view finite strings as belonging to $2^N$. We say that a string whose final bit is 0 is followed by a tail of all 1s and vice versa. For example,

$$101 \leftrightarrow 101000 \ldots,$$
$$100 \leftrightarrow 100111 \ldots.$$

Our Knave does not have the patience for these infinite matters, so when we do compel them to act on $2^N$, the Knave will report

$$\ldots000$$

as

$$\ldots111$$

and vice versa. Thus, these tails will never interfere with the preceding string. We will (somewhat abusively) treat these either as sequences or strings, depending on which is more convenient.

Note that $k$ is not invertible; already

$$k(10) = k(00000) = 1011.$$
3. Metamorphosis

For a natural number \( n \), let \([n]\) denote the string which represents \( n \) in base two. We will call any string of \( n \) 0s or \( k \) 1s a ribbit, short for repeated bit. If we need to clarify what bit is repeated, we can say that 111 is a ribbit of three 1s or an odd ribbit. Likewise, \( 000 \) is a ribbit of three 0s and an even ribbit. Thus, any binary sequence \( S \in S \) decomposes into a sequence of ribbits of alternating parity.

Let \( S \in S \). Since the Knave must begin the report with a 1, we assume that \( S \) begins with an odd ribbit. Then \( S \) decomposes into ribbits as

\[
S = r_1 r_2 r_3 \ldots.
\]

Happily, this means that odd ribbits are indexed by odd subscripts and \textit{vice versa}.

We may write

\[
k(S) = [\vert r_1 \vert] \ 0 \ [\vert r_2 \vert] \ 1 \ [\vert r_3 \vert] \ 0 \ldots.
\]

It is unfortunate here that the leftmost 1 arising from \( r_{2\ell+1} \) will always form a ribbit with \([|r_{2\ell+2}|]\), as in

\[
k(101) = 101110.
\]

However, the decomposition of \( s_n \) into even and odd ribbits allows us to get the Knave’s reports piecemeal; keeping

\[
S = r_1 r_2 r_3 \ldots
\]

with \( r_1 \) odd, then

\[
k(S) = k(r_1) \ k(r_2 r_3) \ k(r_4 r_5) \ldots.
\]

Thus, we can determine the behaviour of \( k \) by examining all possible pairs of ribbits occurring in the decomposition of all \( s_n \). Fortunately, there are not many to check. We will call a ribbit \( r \) belonging to \( S \) maximal if it is not contained in a ribbit of larger size.

**Lemma 3.1.** Let \( \{s_n\} \) be the Look-Knave sequence. A maximal ribbit occurring in \( s_n \) cannot have length greater than five.

**Proof.** Suppose that \( n \) is the smallest index such that \( s_n \) contains a ribbit \( r \) of length six or greater, either

\[
s_n = \ldots \ 1 \ 0 \ldots \ 0 \ 1 \ldots
\]

or

\[
s_n = \ldots \ 0 \ 1 \ldots \ 1 \ 0 \ldots.
\]

What is \( s_n \) describing? Or, rather, what \textit{isn’t} \( s_n \) describing? If \( r \) is even, then \( s_{n-1} \) contains a ribbit of length at least 64; this ribbit can only occur if \( s_{n-1} \) has a ribbit \( r' \) such that the binary representation of \( |r'| \) has at least six 0s. This is a contradiction.
The case where \( r \) is odd is more complicated. We already see that such an \( r \) could arise from an \( r' \) in \( s_{n-1} \), where the binary representation of \(|r'|\) has at least five 1s, which is again impossible.

However, \( r \) could represent the concatenation of two separate descriptions of ribbits; the first even and the second odd. In this case,

\[
s_n = \ldots \underline{0}1\ldots1 1\ldots0\ldots0 \ldots ,
\]

where the first overbrace indicates the binary expansion of the length of an odd ribbit in \( s_{n-1} \), and the second overbrace indicates the binary expansion of the length of an even ribbit in \( s_{n-1} \). From our assumption on \( n \), we see that the only acceptable arrangement is

\[
s_n = \ldots \underline{111} 1 \underline{11} 0 \ldots .
\]

Unfortunately,

\[
\ldots111
\]

is the binary expansion of some \( n \geq 7 \) and we croak. \( \square \)

In fact, once we know the bound for maximal ribbits in general, we can tighten up the proof for some edge cases.

**Corollary 3.2.** A maximal even ribbit occurring in \( s_n \) cannot have length greater than three.

**Corollary 3.3.** If \( 11111 \) occurs in \( s_n \), it is not preceded by \( \underline{000} \).

**Proof.** If we consider \( s_n \) as

\[
s_n = \ldots \underline{00} 0 \underline{11} 1 \underline{11} 0 \ldots ,
\]

then \( s_{n-1} \) contains at least four 1s. Otherwise, if we consider \( s_n \) as

\[
s_n = \ldots \underline{011} 1 \underline{11} 0 \ldots ,
\]

then \( s_{n-1} \) contains at least 11 1s. Both possibilities contradict our previous results. \( \square \)

We may now examine the Knave’s behaviour on all possible ribbit pairs \((r, r')\) occurring in some \( s_n \), with \( r' \) possibly empty. This is shown in Table 2. Note that in all cases, \( k(r r') \) is no shorter than \( rr' \).

From our observation in Table 1, we want to determine if the sequences \( \{s_{2n+1}\} \) and \( \{s_{2n+2}\} \) converge in \( \mathcal{S} \). To this end, we will endow \( 2^{2n} \) with a simple metric. Two distinct binary sequences \( S, S' \) which first differ at the \( n \)th bit satisfy \( d(S, S') = 2^{-n} \). As expected, we set \( d(S, S) = 0 \). Note that \( \mathcal{S} \) is not complete under this metric, but \( 2^{2n} \) is.

For \( \ell \geq 1 \), let \( \beta_r \) be the string given by the first \( \ell \) bits in \( s_r \), which are then truncated to the last maximal ribbit. Here \( \beta \) stands for \( \beta\beta\beta\beta\beta\beta\beta \), of course. For example, \( \beta_3 \) is the string \( \underline{10} \), taken from \( s_3 = 1011 \).

**Lemma 3.4.** For \( \ell \geq 1 \), the strings \( s_{\ell+1} \) and \( s_{\ell+3} \) agree up to the \((|\ell| + 1)\)th bit.
TABLE 2. Elements of the Knave map.

| r r’ | k(r r’) |
|------|---------|
| 0    | 1       |
| 00   | 101     |
| 000  | 111     |
| 1    | 10      |
| 01   | 1110    |
| 001  | 10110   |
| 0001 | 11110   |
| 011  | 11100   |
| 0011 | 101100  |
| 00011| 111100  |
| 0111 | 111000  |
| 00111| 101110  |
| 000111| 111110 |
| 01111| 1110000 |
| 001111| 1011000|
| 0001111| 1111000|
| 011111| 111010  |
| 0011111| 10110010|

PROOF. We induct on \( \ell \). According to Table 2, we have \( |k(r r’)| \geq |r r’| \) for all elements of the Knave map. Because \( \beta_\ell \) begins with 10, we see that \( |k(\beta_\ell)| > |r_\ell| \). In the induction, we see that the first \( |\beta_\ell| \) bits of \( s_\ell \) and \( s_{\ell+2} \) determine at least the first \( |\beta_\ell| + 1 \) bits of \( s_{\ell+1} \) and \( s_{\ell+3} \). \( \Box \)

Note that \( \ell - 4 \leq |r_\ell| \leq \ell \).

COROLLARY 3.5. The sequences \( \{k^{2n}(1)\} \) and \( \{k^{2n}(10)\} \) converge in \( S \).

Thus, we can take \( S_{\text{even}} = \lim_{n \to \infty} k^{2n}(1) \) and \( S_{\text{odd}} = \lim_{n \to \infty} k^{2n}(1) \). It turns out that not only are \( S_{\text{even}} \) and \( S_{\text{odd}} \) fixed points of \( k^2 \), they attract all other orbits under \( k \) in \( S \).

THEOREM 3.6. Let \( S \in S \) be a binary sequence. Then either

\[
\lim_{n \to \infty} d(k^n(S), k^n(1)) = 0
\]

or

\[
\lim_{n \to \infty} d(k^n(S), k^n(10)) = 0.
\]

PROOF. We see that \( k(S) \) must begin with an odd ribbit. If \( k(S) \) begins with an odd ribbit of length \( \ell \geq 2 \), then \( k^2(S) \) begins with an odd ribbit of length strictly less than \( \ell \).
Otherwise, $k(S)$ begins with $10$ and so does $k^2(S)$. Thus, some iterate $k^n(S)$ begins with $10$. Then $k^{n+1}(S)$ begins with $101$, and $k^{n+2}(S') = 10r\ldots$, where $r$ is a maximal odd ribbit of length at least three. If $|r| \geq 5$, then $k^{n+3}(S) = 10r'\ldots$, where $r'$ is a maximal odd ribbit of length at most $2 + \log_2(r)$. Further iteration of the Knave map reduces to the case $|r| = 3, 4$.

If $|r| = 3, 4$, using the argument in Lemma 3.4, we see that the prefix of $k^{n+2}(S)$ determines a longer prefix of $k^{n+4}(S')$ and so on. Then

$$\lim_{n \to \infty} d(k^n(S), k^n(1)) = 0$$

or

$$\lim_{n \to \infty} d(k^n(S), k^n(10)) = 0,$$

depending on the parities of $|r|$ and $n$. □

**Corollary 3.7.** Let $S$ be any binary sequence in $S$. Then $\lim_{n \to \infty} k^{2n}(S)$ exists and is equal to one of $S_{even}$ or $S_{odd}$.

**Corollary 3.8.** The only fixed points of $k^2$ in $S$ are $S_{even}$ and $S_{odd}$.

**Corollary 3.9.** The only fixed points of $k^2$ in $2^{\mathbb{N}}$ are $S_{even}$ and $S_{odd}$, which are attracting, and $000\ldots$ and $111\ldots$, which are repelling.

### 4. Future study

We have left open the question of the asymptotic growth of $|s_n|$. Experimentally, we expect that

$$\lim_{n \to \infty} \frac{|s_{n+1}|}{|s_n|} = 1.12\ldots.$$  

Adapting Johnston’s argument to this problem would be an appropriate problem for a student.

Further, we conjecture that the binary strings $S$ are in fact the sections of a larger dynamical system via the diagonal entries of certain Kermitian matrices.

**Acknowledgement**

We are grateful to the referee for pointing out some errors in our argument and in improving our references.

**References**

[1] J. H. Conway, ‘The weird and wonderful chemistry of audioactive decay’, *Open Problems in Communication and Computation* (eds. T. M. Cover and B. Gopinath) (Springer, New York, NY, 1987), 173–188.

[2] N. Johnston, ‘The binary ‘look-and-say’ sequence’, http://www.njohnston.ca/2010/11/the-binary-look-and-say-sequence/, 2010.
[3] R. M. Smullyan, *What is the Name of This Book? The Riddle of Dracula and Other Logical Puzzles* (Prentice-Hall, Upper Saddle River, NJ, 1978).

[4] The On-Line Encyclopedia of Integer Sequences (2020), https://oeis.org/A005150.

[5] The On-Line Encyclopedia of Integer Sequences (2020), https://oeis.org/A014715.

[6] The On-Line Encyclopedia of Integer Sequences (2020), https://oeis.org/A001387.

THOMAS MORRILL, School of Science, The University of New South Wales Canberra, Australia
e-mail: t.morrill@adfa.edu.au; morrillt@trine.edu