SKOLEM GOES DOWN THE RABBIT HOLE

SAAD MNEIMNEH AND SAMAN FARHAT

Abstract. An infinite Skolem sequence $\{s_i : i \in \mathbb{N}\}$ is such that for every $n \in \mathbb{N}$ there exist exactly two integers $a_n < b_n$ that satisfy $s_{a_n} = s_{b_n} = n$. Furthermore, $b_n - a_n = n$. We show that the first infinite Skolem sequence in a lexicographic order is related to the binary Fibonacci sequence.

1. Introduction

A Skolem sequence is a sequence of $2k$ integers $s_1, \ldots, s_{2k}$ such that for every $n \in \{1, \ldots, k\}$ there exist two integers $a_n < b_n$ that satisfy $s_{a_n} = s_{b_n} = b_n - a_n = n$ [1]. We define an infinite Skolem sequence in an analogous way where every $n \in \mathbb{N}$ appears exactly twice. While not every finite k admits a Skolem sequence, it is trivial to construct an infinite Skolem sequence. In this article, we show that a special infinite Skolem sequence, the first in a lexicographic order, exhibits an interesting link to the binary Fibonacci sequence [2]. Namely, $a_{n+1} - a_n = 1 + f_n$ (equivalently $a_n = n + \sum_{i=1}^{n-1} f_i$), where f_n is the n^{th} bit of the binary Fibonacci sequence.

2. A Special Skolem Sequence

The first lexicographic infinite Skolem sequence can be trivially constructed using a greedy algorithm. For every n in increasing order, we find the smallest i for which s_i has not been assigned, then assign s_i and s_{i+n} the value n. The algorithm in Figure 1 constructs the first $k \geq 0$ terms.

$I \leftarrow \emptyset$, $i \leftarrow 1$, $n \leftarrow 1$
while $i \leq k$
    do $s_i \leftarrow n$
        $I \leftarrow I \cup \{i\}$
        if $i + n \leq k$
            then $s_{i+n} \leftarrow n$
                $I \leftarrow I \cup \{i + n\}$
        $i \leftarrow \min_{j \in \mathbb{N} - I} j$
        $n \leftarrow n + 1$

Figure 1. Algorithm for constructing $s_1, \ldots, s_k$.

The leading 15 terms of the first lexicographic infinite Skolem sequence are shown below.

$$1\ 1\ 2\ 3\ 2\ 4\ 3\ 5\ 6\ 4\ 7\ 8\ 5\ 9\ 6\ \dots$$

By construction, this infinite Skolem sequence satisfies (uniquely): $n < m$, $a_n < a_m$, and $b_n < b_m$ are equivalent.

Lemma 2.1. *If $n < m$, then $b_m - b_n > 1$.*

Proof. We know that $b_n < b_m$, so we only need to prove that $b_m - b_n \neq 1$. Assume that $b_m - b_n = 1$, then $a_n = b_n - n = b_m - (n+1) \geq b_m - m = a_m$, a contradiction since $a_n < a_m$. □

3. Gaps and Offsets

We now define for every $n \in \mathbb{N}$, an offset string T_n over the alphabet $\{1, 2\}$. First, when the algorithm of Figure 1 makes $s_i = s_{i+n} = n$, n defines a finite *gap* sequence $\{g_j : i < j < i + n\}$ of length $n - 1$ using $\mathbb{N} \cup \{-\}$ as follows: if $j \in I$, $g_j = s_j$, and if $j \notin I$, g_j is a gap symbol $-$. In this sequence, if the first gap symbol is g_j, then s_j will be assigned the value $n + 1$.

Next, we identify maximal gaps from left to right and concatenate their lengths to form the offset string T_n. Figure 2 shows the gap sequences and offset strings for $n = 1, \dots, 9$ (T_1 is the empty string ϵ).

Lemma 3.1. *The gap sequence defined by n cannot have $\{g_j, g_{j+1}\} \subset \mathbb{N}$ for any j.*

Proof. When the algorithm makes $s_i = s_{i+n} = n$, we know that $a_m < i$ for all $m < n$. Therefore, every $j \in [i + 1, i + n - 1]$ satisfies $j = b_m$ for some m. Having $\{g_j, g_{j+1}\} \subset \mathbb{N}$ in the gap sequence means there are two integers $m_1 < m_2$ such that $b_{m2} - b_{m1} = 1$, a contradiction to Lemma 2.1. □

Theorem 3.2. *For every n, the offset string T_n is obtained by starting with the empty string ϵ and applying the following substitution rules $n - 1$ times, where S represents a string over the alphabet $\{\mathbf{1}, 1, 2\}$ (when reading the offset string T_n we make no distinction between a $\mathbf{1}$ and a 1).*

$$\epsilon \to \mathbf{1}$$

$$\mathbf{1}S \to S1$$

$$1S \to S2$$

$$2S \to \mathbf{1}S2$$

Proof. The proof is by induction on n. The base cases can be easily verified. Assume that the statement of the theorem holds up to n and that $s_i = s_{i+n} = n$. We consider three exhaustive cases.

$$\begin{gathered}
1 \ \underbrace{\quad}_{T_1=\epsilon} \ 1 \\
1\ 1\ 2 \ \underbrace{-}_{T_2=1} \ 2 \\
1\ 1\ 2\ 3 \ \underbrace{2\ -}_{T_3=1} \ 3 \\
1\ 1\ 2\ 3\ 2\ 4 \ \underbrace{3\ -\ -}_{T_4=2} \ 4 \\
1\ 1\ 2\ 3\ 2\ 4\ 3\ 5 \ \underbrace{-\ 4\ -\ -}_{T_5=12} \ 5 \\
1\ 1\ 2\ 3\ 2\ 4\ 3\ 5\ 6 \ \underbrace{4\ -\ -\ 5\ -}_{T_6=21} \ 6 \\
1\ 1\ 2\ 3\ 2\ 4\ 3\ 5\ 6\ 4\ 7 \ \underbrace{-\ 5\ -\ 6\ -\ -}_{T_7=112} \ 7 \\
1\ 1\ 2\ 3\ 2\ 4\ 3\ 5\ 6\ 4\ 7\ 8 \ \underbrace{5\ -\ 6\ -\ -\ 7\ -}_{T_8=121} \ 8 \\
1\ 1\ 2\ 3\ 2\ 4\ 3\ 5\ 6\ 4\ 7\ 8\ 5\ 9 \ \underbrace{6\ -\ -\ 7\ -\ 8\ -\ -}_{T_9=212} \ 9
\end{gathered}$$

FIGURE 2. Gap sequences and offset strings for $n = 1, \ldots, 9$.

$\underline{T_n = \mathbf{1}S}$: By the inductive hypothesis, T_{n-1} starts with 2. Therefore, $s_{i+1} = s_{i+n+2} = n + 1$, and $T_{n+1} = S1$.

$\underline{T_n = 1S}$: The gap sequence defined by n must start with an integer followed by $-$; it cannot start with $-$ because this means T_{n-1} starts with 2, and hence $T_n = \mathbf{1}S$ by the inductive hypothesis. Therefore, $s_{i+2} = s_{i+n+3} = n + 1$, and $T_{n+1} = S2$.

$\underline{T_n = 2S}$: The gap sequence defined by n must start with an integer followed by $-\ -$ (see above case). Therefore, $s_{i+2} = s_{i+n+3} = n + 1$, and $T_{n+1} = \mathbf{1}S2$.

□

The importance of the offset string T_n is manifested in the following result. Define $s(T_n)$ to be the sum of all the digits that appear in T_n, then:

Lemma 3.3. $a_{n+1} - a_n = 1 + s(T_{n+1}) - s(T_n)$.

Proof. When the algorithm makes $s_i = s_{i+n} = n$, $s_1, \ldots, s_{i-1}$ have all been assigned, and every $m \leq n$ appears exactly twice in $s_1, \ldots, s_{bn}$ (see Figure 2). Therefore, $b_n = 2n + s(T_n)$, which implies that $a_n = n + s(T_n)$. Finally, $a_{n+1} - a_n = 1 + s(T_{n+1}) - s(T_n)$. □

4. Follow the Rabbit

We now make the link between the first lexicographic infinite Skolem sequence and the binary Fibonacci sequence. We will show that $s(T_{n+1}) - s(T_n) = f_n$, where f_n is the n^{th} bit of the binary Fibonacci sequence.

The crucial observation is that the rules in the statement of Theorem 3.1 can be combined to eliminate the use of **1** (numbers above arrows show $T_{n+1} - T_n$).

$$\epsilon \to \mathbf{1} \xrightarrow{0} 1$$
$$1S \xrightarrow{1} S2$$
$$2S \xrightarrow{1} \mathbf{1}S2 \xrightarrow{0} S21$$

Furthermore, by renaming 1 and 2 as 0 and 1 respectively, we will generate bits of the binary Fibonacci sequence, starting from the second bit $f_2 = 0$ (the underlined bits show the correspondence with $T_{n+1} - T_n$).

$$\epsilon \to \underline{0}$$
$$0S \to S\underline{1}$$
$$1S \to S\underline{10}$$

Observe that for $n \geq 2$, $s(T_{n+1}) - s(T_n)$ is preserved and is equal to the n^{th} generated bit. We also verify that for $n = 1$, $s(T_2) - s(T_1) = f_1 = 1$.

We finally establish the following result:

Theorem 4.1. *$a_{n+1} - a_n = 1 + f_n$ (or equivalently, $a_n = n + \sum_{i=1}^{n-1} f_i$).*

References

[1] http://mathworld.wolfram.com/SkolemSequence.html
[2] http://www.maths.surrey.ac.uk/hosted-sites/R.Knott/Fibonacci/fibrab.html

Department of Computer Science Hunter College of the City University of New York, 695 Park Avenue, New York, USA
E-mail address: saad@hunter.cuny.edu

Department of Computer Science, The Graduate Center of the City University of New York, 365 Fifth Avenue, New York, USA
E-mail address: sfarhat@gc.cuny.edu