Journal of Integer Sequences, Vol. 13 (2010),
Article 10.3.2

Some Remarks on Differentiable Sequences and Recursivity

Jean-Marc Fédou and Gabriele Fici[1]
Laboratoire d'Informatique, Signaux et Systèmes de Sophia-Antipolis
UMR6070 - CNRS et Université de Nice-Sophia Antipolis
2000, route des Lucioles
06903 Sophia-Antipolis Cedex
France
fedou@i3s.unice.fr
fici@i3s.unice.fr

Abstract

We investigate the recursive structure of differentiable sequences over the alphabet $\{1, 2\}$. We derive a recursive formula for the $(n + 1)$-th symbol of a differentiable sequence, which yields to a new recursive formula for the Kolakoski sequence. Finally, we show that the sequence of absolute differences of consecutive symbols of a differentiable sequence u is a morphic image of the run-length encoding of u.

1 Introduction

In 1965, W. Kolakoski [9] proposed the following problem:

"Describe a simple rule for constructing the sequence:

$$\mathbf{K} = 12211212212211211221211212211211212212211212212\cdots$$

What is the n-th term? Is the sequence periodic?"

This sequence, called now *the Kolakoski sequence*, is in fact the unique sequence starting with 1 and identical to its own run-length encoding.

The Kolakoski sequence has been investigated in many papers [1, 3–8, 10–13]. Although the non-periodicity of the sequence was shown immediately, the problem of finding a good

[1]The second author acknowledges the support of an Exchange Grant on the program "AutoMathA: Automata, from Mathematics to Applications" of the European Science Foundation.

formula for the n-th term is still open, and is related to other open problems. The most famous open problems on the Kolakoski sequence (up to now) are as follows: Is the sequence recurrent? Is the frequency of 1's and 2's asymptotically the same? Is the set of factors of the sequence closed under reversal and/or swap of the symbols?

The Kolakoski sequence $\mathbf{K}$ (sequence A000002 in Sloane's database) and the sequence $\mathbf{k}$ (sequence A078880) obtained from the former by deleting the first symbol are the unique fixed point of the run-length encoding operator Δ. A sequence u over the alphabet $\{1,2\}$ such that $\Delta(u)$ is still a sequence over $\{1,2\}$ is called a *differentiable sequence*. The problems stated above for the Kolakoski sequence remain unsolved for the wider class of sequences that are differentiable arbitrary many times, called *smooth sequences* [1, 2].

In this paper we study the recursive relationship between any differentiable sequence u and its run-length encoding $\Delta(u)$. We start by defining, for any differentiable sequence u, the sequences $\varphi_n(u)$ and $\gamma_n(u)$. The sequence $\varphi_n(u)$ is defined by $\varphi_n(u) = |\Delta(u_1 u_2 \cdots u_n)|$. In other words, $\varphi_n(u)$ is equal to 1 plus the number of symbol changes in $u_1 u_2 \cdots u_n$. The sequence $\gamma_n(u)$ is defined by $\gamma_n(u) = |u_{n+1} - u_n|$.

The sequences $\varphi_n(\mathbf{K})$, $\varphi_n(\mathbf{k})$ and $\gamma_n(\mathbf{K})$ are known (sequences A156253, A156351, and A156728 respectively).

Remark 1. We shall write φ_n, γ_n instead of $\varphi_n(u), \gamma_n(u)$ when no confusion arises.

In Theorem 3.1 we derive a recursive formula for γ_n

$$\gamma_n = 1 - (u'_{\varphi_n} - 1)\gamma_{n-1}$$

where $u' = \Delta(u)$. This formula yields to recursive formulas for u and $\varphi(u)$ (Corollaries 3.2 and 3.3)

$$u_{n+1} = 3 - u_n + (u'_{\varphi_n} - 1)(u_n - u_{n-1})$$

$$\varphi_{n+1} = \varphi_n + 1 - (u'_{\varphi_n} - 1)(\varphi_n - \varphi_{n-1})$$

When $u = K_1 K_2 \cdots$ is the Kolakoski sequence, our recursive formula gives

$$K_{n+1} = 3 - K_n + (K_{\varphi_n} - 1)(K_n - K_{n-1})$$

A different approach allows us to derive an alternative recursive formula for the $n+1$-th term of a differentiable sequence. Indeed, in Theorem 3.4, we prove that

$$u_{n+1} = u_n + (3 - 2u_n)\left(n + 1 - \sum_{i=1}^{\varphi_n} u'_i\right)$$

When u is the Kolakoski sequence, this latter formula is equivalent to one of Steinsky [13], obtained with different techniques.

As a last result, in Lemma 3.5, we show that for any differentiable sequence u, the sequence γ_n is a morphic image of the sequence $\Delta(u)$, under the morphism $\mu : 1 \mapsto 1, 2 \mapsto 01$.

2 Differentiable sequences

An *alphabet*, denoted by Σ, is a finite set of symbols. A *sequence* over Σ is a sequence of symbols from Σ. The *length* of a finite sequence u is denoted by $|u|$. A right-infinite sequence over Σ is a non-ending sequence of symbols from Σ. Formally, a right-infinite sequence is a function $f : \mathbb{N} \longmapsto \Sigma$. For an abuse of notation, we shall often write f_n for $f(n)$.

A *run* in a sequence u is a maximal block of consecutive identical symbols.

Let u be a sequence over Σ. Then u can be uniquely written as a concatenation of consecutive runs of the symbols of Σ, i.e. $u = x_1^{i_1} x_2^{i_2} x_3^{i_3} \cdots$, with $x_j \in \Sigma$, $x_j \neq x_{j+1}$ and $i_j > 0$. The *run-length encoding* of u, noted $\Delta(u)$, is the sequence of exponents i_j, i.e. $\Delta(u) = i_1 i_2 i_3 \cdots$.

Remark 2. From now on we set $\Sigma = \{1, 2\}$.

We say that a sequence u over Σ is *differentiable* if $\Delta(u)$ is still a sequence over Σ. Since $\Sigma = \{1, 2\}$ we have that u is differentiable if and only if neither 111 nor 222 appear in u.

In the sequel we note $\Delta(u) = u'_1 u'_2 \cdots$ for u a differentiable sequence.

Definition 2.1. A right-infinite sequence u over Σ is a *smooth sequence* if it is differentiable arbitrary many times over Σ.

The most famous examples of smooth sequences are the Kolakoski sequences:

$$\mathbf{k} = 22112122122112112212112122112112122122112122\cdots$$

and

$$\mathbf{K} = 1\mathbf{k} = 122112122122112112212112122112112122122112122\cdots$$

which are the fixed points of Δ.

The following lemma is a straightforward consequence of the definition of Δ.

Lemma 2.1. *Let uv be a differentiable sequence. Then $\Delta(uv) = \Delta(u)\Delta(v)$ if and only if the last symbol of u and the first symbol of v are different.*

Let $u = u_1 u_2 u_3 \cdots$ be a finite or infinite sequence over Σ. We define the two functions Δ_1^{-1} and Δ_2^{-1} by:

$$\Delta_1^{-1}(u) = 1^{u_1} 2^{u_2} 1^{u_3} \cdots$$

$$\Delta_2^{-1}(u) = 2^{u_1} 1^{u_2} 2^{u_3} \cdots$$

In such a way, $u = \Delta(\Delta_x^{-1}(u))$ for any $x \in \Sigma$.

Remark 3. Let $u = u_1 u_2 \cdots u_n$ be a sequence over Σ. Then for every $x \in \Sigma$

$$|\Delta_x^{-1}(u_1 u_2 \cdots u_n)| = \sum_{i=1}^{n} u_i$$

3 Recursivity

Let $u = u_1u_2\cdots$ be a differentiable sequence. We define, for every $n > 0$

$$\varphi_n(u) = |\Delta(u_1u_2\cdots u_n)|$$

The definition of Δ directly implies that $\varphi_n(u)$ is equal to 1 plus the number of symbol changes in $u_1u_2\cdots u_n$. With our notation:

$$\varphi_n(u) = \varphi_{n-1}(u) + |u_n - u_{n-1}| = 1 + \sum_{i=1}^{n-1} |u_{i+1} - u_i| \tag{1}$$

for every $n > 1$.

We also define, for every $n > 0$

$$\gamma_n(u) = |u_{n+1} - u_n|$$

Example 1. The sequences $\varphi_n(\mathbf{K}), \varphi_n(\mathbf{k})$ and $\gamma_n(\mathbf{K})$ are present in the Sloane's database as sequence A156253, A156351, and A156728 respectively. The first values of these sequences are reported in Table 1.

Let $u = u_1u_2\cdots$ be a (right infinite) differentiable sequence and let $u' = \Delta(u) = u'_1u'_2\cdots$ be its run-length encoding. For any $n > 0$ the two sequences $\Delta(u_1u_2\cdots u_n)$ and $u'_1u'_2\cdots u'_{\varphi_n}$ are equal if and only if $u_{n+1} \neq u_n$, as a consequence of Lemma 2.1. If instead $u_n = u_{n+1}$ then $u_n \neq u_{n-1}$ since u is differentiable, and hence the last symbol of $\Delta(u_1\cdots u_n)$ is equal to 1, while $u'_{\varphi_n} = 2$.

In other words, if $u_n = u_{n-1}$ then clearly $u_{n+1} \neq u_n$ since u is differentiable. If instead $u_n \neq u_{n-1}$ then $u_{n+1} = u_n$ when $u'_{\varphi_n} = 2$, while $u_{n+1} \neq u_n$ when $u'_{\varphi_n} = 1$. We thus have

Theorem 3.1. *Let $u = u_1u_2\cdots$ be a differentiable sequence. Then for every $n > 0$*

$$\gamma_n = 1 - (u'_{\varphi_n} - 1)\gamma_{n-1} \tag{2}$$

Remark 4. Since $\Sigma = \{1, 2\}$, one has that, $\forall x, y \in \Sigma$, $y = x + (3 - 2x)|y - x|$.

From the previous remark and from Equation 2 we have

Corollary 3.2. *Let $u = u_1u_2\cdots$ be a differentiable sequence. Then for every $n > 0$*

$$u_{n+1} = 3 - u_n + (u'_{\varphi_n} - 1)(u_n - u_{n-1}) \tag{3}$$

And from Equations 1 and 2 we derive

Corollary 3.3. *Let $u = u_1u_2\cdots$ be a differentiable sequence. Then for every $n > 0$*

$$\varphi_{n+1} = \varphi_n + 1 - (u'_{\varphi_n} - 1)(\varphi_n - \varphi_{n-1}) \tag{4}$$

Example 2. When u is the Kolakoski sequence $\mathbf{K}$, Equation 3 gives

$$K_{n+1} = 3 - K_n + (K_{\varphi_n} - 1)(K_n - K_{n-1}) \tag{5}$$

We now give another recursive formula for the $n+1$-th symbol of a differentiable sequence.

Theorem 3.4. *Let $u = u_1u_2\cdots$ be a differentiable sequence. Then for every $n > 0$*

$$u_{n+1} = u_n + (3 - 2u_n)\left(n + 1 - \sum_{i=1}^{\varphi_n} u'_i\right) \tag{6}$$

Proof. If $u_n = u_{n+1}$ then $\varphi_n = \varphi_{n+1}$, so $\Delta^{-1}_{u_1}(u'_1 \cdots u'_{\varphi_n}) = \Delta^{-1}_{u_1}(u'_1 \cdots u'_{\varphi_{n+1}}) = u_1u_2\cdots u_{n+1}$. Hence, by Remark 3, $\sum_{i=1}^{\varphi_n} u'_i = n + 1$.

If instead $u_n \neq u_{n+1}$ then, by Lemma 2.1, $\Delta(u_1\cdots u_n) = u'_1 \cdots u'_{\varphi_n}$, which implies that $\Delta^{-1}_{u_1}(u'_1 \cdots u'_{\varphi_n}) = u_1u_2\cdots u_n$. Hence, again by Remark 3, $\sum_{i=1}^{\varphi_n} u'_i = n$.

Thus

$$|u_{n+1} - u_n| = n + 1 - \sum_{i=1}^{\varphi_n} u'_i$$

The claim then follows from Remark 4.

□

Example 3. For $u = \mathbf{K}$, Equation 6 becomes

$$K_{n+1} = K_n + (3 - 2K_n)\left(n + 1 - \sum_{i=1}^{\varphi_n} K_i\right) \tag{7}$$

Equation 7 can be found in a paper of Steinsky [13], where $\varphi_n(\mathbf{K})$ is replaced by $\rho_n(\mathbf{K}) = \min\left\{j : \sum_{i=1}^{j} K_i \geq n\right\}$. But it is easy to see that for any differentiable sequence u one has $\varphi_n = \min\left\{j : \sum_{i=1}^{j} u'_i \geq n\right\}$.

We now show that, for any differentiable sequence u, the sequence $\gamma_n(u)$ is a morphic image of the sequence $\Delta(u)$.

Lemma 3.5. *Let μ be the morphism defined on Σ by*

$$\mu : \begin{cases} 1 & \longmapsto 1 \\ 2 & \longmapsto 01 \end{cases}$$

and let v_n be the sequence $\mu(\Delta(u))$. Then $v_n = \gamma_n$.

Proof. It is sufficient to prove that the sequences v_n and γ_n have the same partial sums. We have two cases:

Case 1. $u_n \neq u_{n+1}$. Then, by Lemma 2.1, $\Delta(u_1 \cdots u_n) = u'_1 \cdots u'_{\varphi_n}$. From the definition of μ, one has $\sum_{i=1}^{n} v_i = |u'_1 \cdots u'_{\varphi_n}| = \varphi_n$.

Case 2. $u_n = u_{n+1}$. This implies that $u'_{\varphi_n} = 2$ and so $\sum_{i=1}^{n} v_i = \sum_{i=1}^{n-1} v_i$. On the other hand, we must have $u_{n-1} \neq u_n$ and therefore, arguing as in Case 1, we obtain $\sum_{i=1}^{n-1} v_i = |u'_1 \cdots u'_{\varphi_{n-1}}| = \varphi_{n-1} = \varphi_n - 1$.

In summary, we have $\sum_{i=1}^{n} v_i = \varphi_n - 1 + \gamma_n$.

On the other hand, by Equation 1, we have $\sum_{i=1}^{n} \gamma_i = \sum_{i=1}^{n-1} \gamma_i + \gamma_n = \varphi_n - 1 + \gamma_n$.

□

4 Conclusion

The main purpose of this paper was to unify the description of various sequences described in the Sloane's database and related to the Kolakoski sequence. We showed that, indeed, all these sequences or recurrences can be easily deduced from more general equalities holding for any differentiable sequence. Unfortunately, it appears that all these results are finally only another way to write the definition of differentiability of a sequence over the alphabet $\{1, 2\}$. Thus, the challenge to find a formula for the n-th symbol of the Kolakoski sequence without the knowledge of the preceding symbols is still open.

Table 1: First values of the sequences $\mathbf{K}$, $\varphi(\mathbf{K})$, $\varphi(\mathbf{k})$, $\gamma(\mathbf{K})$, corresponding, respectively, to Sloane's database entries A000002, A156253, A156351, and A156728
.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
$\mathbf{K}$	1	2	2	1	1	2	1	2	2	1	2	2	1	1	2	1	1	2	2	1	2	1	1	2	1
$\varphi(\mathbf{K})$	1	2	2	3	3	4	5	6	6	7	8	8	9	9	10	11	11	12	12	13	14	15	15	16	17
$\varphi(\mathbf{k})$	1	1	2	2	3	4	5	5	6	7	7	8	8	9	10	10	11	11	12	13	14	14	15	16	17
$\gamma(\mathbf{K})$	1	0	1	0	1	1	1	0	1	1	0	1	0	1	1	0	1	0	1	1	1	0	1	1	1

References

[1] S. Brlek and A. Ladouceur. A note on differentiable palindromes. *Theor. Comput. Sci.* **302** (1–3) (2003), 167–178.

[2] S. Brlek, S. Dulucq, A. Ladouceur, and L. Vuillon. Combinatorial properties of smooth infinite words. *Theor. Comput. Sci.* **352** (1) (2006), 306–317.

[3] A. Carpi. Repetitions in the Kolakovski [sic] sequence. *Bull. European Assoc. Theor. Comput. Sci.*, No. 50, (1993), 194–196.

[4] A. Carpi. On repeated factors in C^∞-words. *Inform. Process. Lett.* **52** (1994), 289–294.

[5] V. Chvátal. Notes on the Kolakoski sequence. Technical Report 93-84, DIMACS, March 1994. Revised.

[6] K. Culik II, J. Karhumäki, and A. Lepistö. Alternating iteration of morphisms and the Kolakovski [sic] sequence. In G. Rozenberg and A. Salomaa, editors, *Lindenmayer Systems*, pp. 93–103. Springer-Verlag, 1992.

[7] F. M. Dekking. What is the long range order in the Kolakoski sequence? In R. V. Moody, editor, *The Mathematics of Long-Range Aperiodic Order*, Vol. 489 of *NATO ASI Ser., Ser. C., Math. Phys. Sci.*, pp. 115–125. Kluwer, 1997.

[8] C. Kimberling. Advanced problem 6281. *Amer. Math. Monthly* **86** (1979), 793.

[9] W. Kolakoski. Elementary problem 5304. *Amer. Math. Monthly* **72** (1965), 674. Solution in *Amer. Math. Monthly* **73** (1966), 681–682.

[10] A. Lepistö. Repetitions in Kolakoski sequence. In G. Rozenberg and A. Salomaa, editors, *Developments in Language Theory*, pp. 130–143. World Scientific, 1994.

[11] G. Paun. How much Thue is Kolakovski [sic]? *Bull. European Assoc. Theor. Comput. Sci.*, No. 49, (February 1993), 183–185.

[12] R. Steacy. Structure in the Kolakoski sequence. *Bull. European Assoc. Theor. Comput. Sci.*, No. 59, (1996), 173–182.

[13] B. Steinsky. A recursive formula for the Kolakoski sequence A000002. *J. Integer Seq.* **9** (3) (2006), Article 06.3.7.

2010 *Mathematics Subject Classification*: Primary 68R15; Secondary 11Y55, 11B83.
Keywords: Kolakoski sequence, integer sequences, differentiable sequences, smooth sequences, combinatorics of words.

(Concerned with sequences A000002, A078880, A156253, A156351, and A156728.)

Received December 8 2009; revised version received February 25 2010. Published in *Journal of Integer Sequences*, February 25 2010.

Return to Journal of Integer Sequences home page.