Journal of Integer Sequences, Vol. 9 (2006),
Article 06.3.7

A Recursive Formula for the Kolakoski Sequence A000002

Bertran Steinsky
Fürbergstrasse 56
5020 Salzburg
Austria
steinsky@finanz.math.tu-graz.ac.at

Abstract

We present a recursive formula for the nth term of the Kolakoski sequence. Using this formula, it is easy to find recursions for the number of ones in the first n terms and for the sum of the first n terms of the Kolakoski sequence.

1 Introduction

The Kolakoski sequence K_n [6, 7], which we study here, is the unique sequence starting with 1 which is identical to its own runlength sequence. K_n is Sloane's sequence A000002. Kimberling asks 5 questions about this sequence on his homepage [5]. The first one is, whether there exists a formula for the nth term of the Kolakoski sequence. For a survey of known properties of the Kolakoski sequence we refer to Dekking [4]. Cloitre wrote the formulas

$$K_N = \frac{3+(-1)^n}{2} \text{ and } K_{N+1} = \frac{3-(-1)^n}{2}, \text{ where } N = \sum_{i=1}^{n} K_i,$$

in the entry of Sloane's sequence A000002, where we also find block-substitution rules, which where posted by Lagarias. I.e., starting with 22 we have to apply $22 \to 2211$, $21 \to 221$, $12 \to 211$, and $11 \to 21$, as it was mentioned by Dekking [3, 4]. Culik *et al.* [2] proposed the double substitution rules $\sigma_1(1 \to 1, 2 \to 11)$ and $\sigma_2(1 \to 2, 2 \to 22)$, which are applied alternatingly to each letter of a word. These substitutions can also be found at Allouche *et al.* [1, p. 336]. Cloitre added the relationship

$$f_1(n) + f_2(n) = 1 + \sum_{i=0}^{n-1} f_2(i)$$

to Sloane's sequence A054349, where $f_1(n)$ denotes the number of ones and $f_2(n)$ denotes the number of twos in the nth string of Sloane's sequence A054349.

2 A Recursive Formula for the Kolakoski Sequence

We will now derive a recursive formula for K_n. Let $k_n = \min\left\{j : \sum_{i=1}^{j} K_i \geq n\right\}$.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
K_n	1	2	2	1	1	2	1	2	2	1	2	2	1	1	2
k_n	1	2	2	3	3	4	5	6	6	7	8	8	9	9	10

Table 1: K_n and k_n

Lemma 2.1.

$$k_n = k_{n-1} + n - \sum_{i=1}^{k_{n-1}} K_i, \;\; \textit{where } n \geq 2.$$

Proof. We first notice that

$$n - 1 \leq \sum_{i=1}^{k_{n-1}} K_i \leq n.$$

The left inequality holds by definition and the right one is valid, since if

$$\sum_{i=1}^{k_{n-1}} K_i \geq n + 1$$

we would have

$$\sum_{i=1}^{k_{n-1}-1} K_i \geq n - 1$$

which is a contradiction to the minimality of k_{n-1}. So, as the first case, we consider $\sum_{i=1}^{k_{n-1}} K_i = n - 1$ which implies $k_n = k_{n-1} + 1 = k_{n-1} + n - \sum_{i=1}^{k_{n-1}} K_i$. In the second case $\sum_{i=1}^{k_{n-1}} K_i = n$ leads to $k_n = k_{n-1} = k_{n-1} + n - \sum_{i=1}^{k_{n-1}} K_i$. □

We notice that Lemma 2.1 holds in general for every sequence, whose only values are 1 and 2.

Lemma 2.2.

$$k_n = k_{n-1} + |K_n - K_{n-1}| = 1 + \sum_{i=2}^{n} |K_i - K_{i-1}|, \;\; \textit{where } n \geq 2.$$

Proof. The following well known construction produces a sequence which is identical to K. Start with K_1 ones, continue with K_2 twos, followed by K_3 ones, and so on. In this construction, after k_{n-1} steps two cases can appear, as described in the proof of Lemma 2.1. The first possibility is that $\sum_{i=1}^{k_{n-1}} K_i = n-1$, which means that we have constructed $n-1$ terms of the sequence. Therefore, by construction, K_n must be different from K_{n-1} implying $k_n - k_{n-1} = |K_n - K_{n-1}|$. In the second case that $\sum_{i=1}^{k_{n-1}} K_i = n$, it is necessary that $K_{k_{n-1}} = 2$, for if otherwise $\sum_{i=1}^{k_{n-1}-1} K_i = n-1$, contradicting the minimality of k_{n-1}. So our construction has added 2 equal numbers at the k_{n-1}th step, such that $K_n = K_{n-1}$ and finally $k_n - k_{n-1} = |K_n - K_{n-1}|$. The second equality follows by induction. □

Corollary 2.1 is an implication of Lemma 2.2.

Corollary 2.1.

$$K_n \equiv k_n \pmod 2 \text{ or } K_n = \frac{(-1)^{k_n}+1}{2} + 1 \text{ respectively.}$$

Corollary 2.2 uses Lemma 2.1 and Corollary 2.1.

Corollary 2.2.

$$k_n = n - \frac{1}{2}\sum_{i=1}^{k_{n-1}} \left((-1)^{k_i}+1\right), \text{ where } n \geq 2.$$

Corollary 2.3 follows from Corollary 2.2.

Corollary 2.3.

$$k_n = k_{n-1} + 1 - \frac{1}{2}\left(k_{n-1} - k_{n-2}\right)\left((-1)^{k_{k_{n-1}}}+1\right), \text{ where } n \geq 3.$$

Theorem 2.1. *For $n \geq 3$ we have*

$$\begin{aligned} K_n &= K_{n-1} + (3-2K_{n-1})\left(n - \sum_{i=1}^{1+\sum_{j=2}^{n-1}|K_j - K_{j-1}|} K_i\right) & (1) \\ &= K_{n-1} + (3-2K_{n-1})\left(n - \sum_{i=1}^{1+\sum_{j=2}^{n-1}\frac{K_j - K_{j-1}}{3-2K_{j-1}}} K_i\right) & (2) \\ &= K_{n-1} + (3-2K_{n-1})\left(1 - \frac{1}{2}\frac{K_{n-1}-K_{n-2}}{3-2K_{n-2}}\left(1 + (-1)^{K_{1+\sum_{j=2}^{n-1}\frac{K_j-K_{j-1}}{3-2K_{j-1}}}}\right)\right). & (3) \end{aligned}$$

Proof. From Lemma 2.1 and Lemma 2.2 we obtain

$$|K_n - K_{n-1}| = n - \sum_{i=1}^{1+\sum_{j=2}^{n-1}|K_j-K_{j-1}|} K_i$$

and use the fact that

$$|K_n - K_{n-1}| = \frac{K_n - K_{n-1}}{3-2K_{n-1}}$$

to complete the proof of (1) and (2). The third equation (3) follows from Corollary 2.3 and Lemma 2.2. □

3 Concluding Remarks

Let $s_n = \sum_{i=1}^{n} K_i$, which is Sloane's sequence A054353, $o_n = |\{1 \leq j \leq n : K_j = 1\}|$, and $t_n = |\{1 \leq j \leq n : K_j = 2\}|$, which is Sloane's sequence A074286. With Theorem 2.1 and the equations

$$\begin{aligned} K_n &= s_n - s_{n-1}, \\ K_n &= -o_n + o_{n-1} + 2, \text{ and} \\ K_n &= t_n - t_{n-1} + 1 \end{aligned}$$

it is easy to produce recursive formulas for s_n, o_n, and t_n.

By Lemma 2.1, we obtain $k_n = n - t_{k_{n-1}}$, from which it follows that t_n/n converges if and only if the limit of k_n/n exists. The definition of k_n gives the equations $k_{s_n} = n$ and $k_{s_n+1} = n + 1$, which yield that the limit of k_n/n exists, if and only if s_n/n converges. Therefore, if we assume that one of the sequences t_n/n, o_n/n, k_n/n or s_n/n converges then all sequences have a limit. If $a = \lim_{n\to\infty} t_n/n$ then $\lim_{n\to\infty} o_n/n = 1 - a$, $\lim_{n\to\infty} s_n/n = 1 + a$, and $\lim_{n\to\infty} k_n/n = 1/(1 + a)$.

Using the recursion of Corollary 2.3, we computed k_n/n up to $n = 3 \cdot 10^8$. Figure 1 shows k_n/n for n from 10^8 to $3 \cdot 10^8$, where only each 1000th point is drawn, i.e., the subsequence $k_{1000n}/(1000n)$, for $n = 100000, \ldots, 300000$. The x-axis is positioned at 2/3.

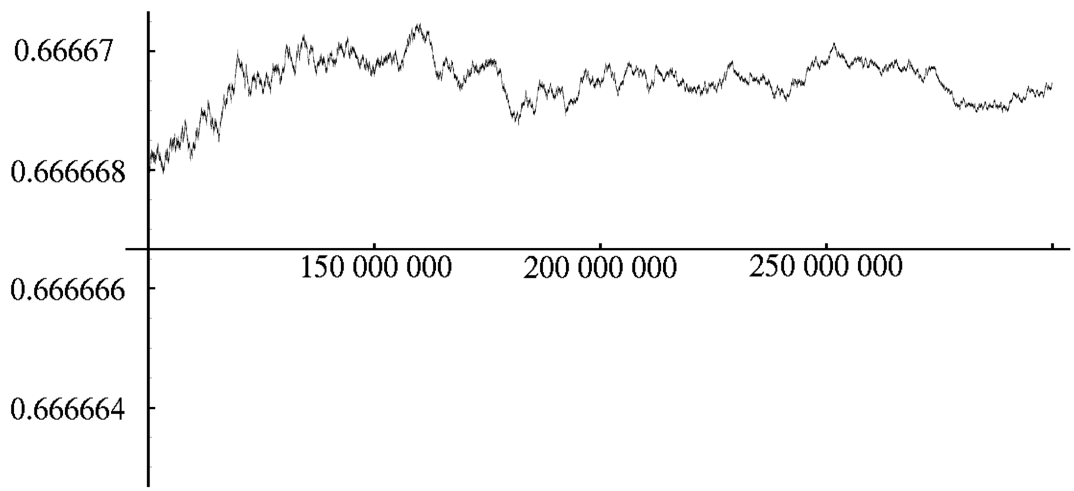

Figure 1: $\frac{k_n}{n}$ for n from 10^8 to $3 \cdot 10^8$.

If we assume that the limit of o_n/n exists and is equal to 1/2 then k_n/n must tend to 2/3. Thus, the graph in Figure 1 does not support the conjecture that o_n/n converges to 1/2.

4 Acknowledgement

We thank the referees for fruitful suggestions and Benoit Cloitre for helpful email discussion.

References

[1] J.-P. Allouche, J. Shallit, *Automatic Sequences*, Cambridge Univ. Press, Cambridge, 2003.

[2] K. Culik, J. Karhumäki, Iterative devices generating infinite words, *Lec. Notes in Comp. Sc.* **577** (1992), 531–544.

[3] F. M. Dekking, Regularity and irregularity of sequences generated by automata, *Sém. Th. Nombres Bordeaux '79-'80* (1980), 901–910.

[4] F. M. Dekking, What is the long range order in the Kolakoski sequence?, *The Mathematics of Long-Range Aperiodic Order*, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **489**, Kluwer Acad. Publ., Dordrecht (1997), 115–125.

[5] C. Kimberling, `http://faculty.evansville.edu/ck6/integer/index.html`

[6] W. Kolakoski, Problem 5304: Self Generating Runs, *Amer. Math. Monthly* **72** (1965), 674.

[7] W. Kolakoski, Problem 5304, *Amer. Math. Monthly* **73** (1966), 681–682.

2000 *Mathematics Subject Classification*: Primary 11B83; Secondary 11B85, 11Y55, 40A05.
Keywords: Kolakoski sequence, recursion, recursive formula.

(Concerned with sequences A000002, A054349, A054353, and A074286.)

Received January 13 2006; revised version received August 19 2006. Published in *Journal of Integer Sequences*, August 19 2006.

Return to Journal of Integer Sequences home page.