k-regular Sequences

Jeffrey Shallit
Department of Computer Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada
shallit@graceland.uwaterloo.ca
http://www.math.uwaterloo.ca/~shallit

Introduction

A sequence $(a(n))_{n\geq 0}$ over a finite alphabet Δ is said to be <u>k-automatic</u> if there exists a finite automaton with output

$$M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$$

such that

$$a(n) = \tau(\delta(q_0, (n)_k))$$

for all $n \geq 0$.

Here

- Q is a finite nonempty set of states;
- $\Sigma_k = \{0, 1, \ldots, k-1\}$;
- $\delta : Q \times \Sigma_k \to Q$ is the transition function;
- q_0 is the initial state;
- $(n)_k$ is the canonical base-k representation of n;
- $\tau : Q \to \Delta$ is the output mapping.

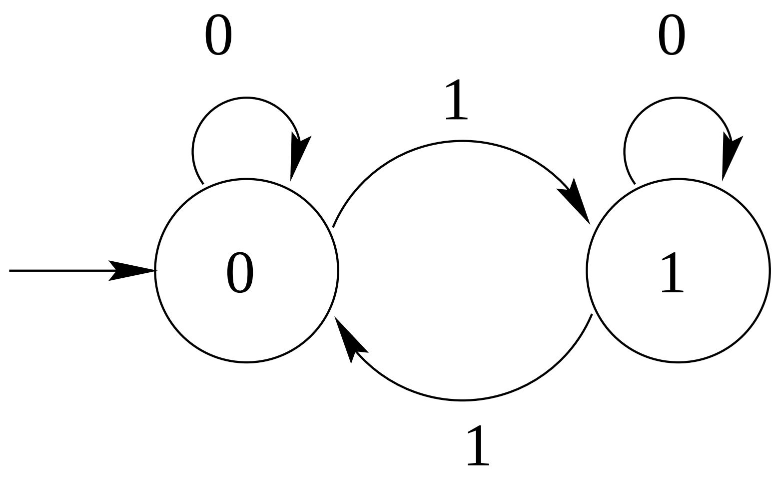
Example: The Thue-Morse Sequence

This sequence

$$(t(n))_{n \geq 0} = \mathtt{0110100110010110} \cdots$$

counts the number of 1's (mod 2) in the base-2 representation of n.

It is generated by the following finite automaton:



Automatic Sequences

- Automatic sequences were introduced by Cobham
- Popularized and further studied by Mendès France, Allouche, and others
- Extremely useful, with well-developed theory (e.g., theorem of Christol)
- However, they are somewhat restricted because of the restriction to a finite alphabet
- Want a generalization that preserves the flavor of automatic sequence, but over an infinite alphabet

The k-kernel

The $\underline{k\text{-kernel}}$ of a sequence $(a(n))_{n\geq 0}$ is the set of subsequences

$$\{(a(k^e n + r))_{n\geq 0} \ : \ e \geq 0, \ 0 \leq r < k^e\}.$$

Theorem. (Eilenberg) A sequence $(a(n))_{n\geq 0}$ is k-automatic if and only if the k-kernel is finite.

Example. Consider the Thue-Morse sequence $(t(n))_{n\geq 0}$. Then clearly

$$t(2^e n + r) \equiv t(n) + t(r) \pmod 2$$

so every sequence in the k-kernel is either $(t(n))_{n\geq 0}$ or $(t(2n+1))_{n\geq 0}$.

k-regular Sequences

To generalize automatic sequences, we use the k-kernel.

Instead of demanding that the k-kernel be finite, we instead ask that the set of sequences generated by the k-kernel be finitely generated.

Example 1. Consider the sequence $(s_2(n))_{n\geq 0}$, where $s_2(n)$ is the sum of the bits in the base-2 representation of n. Then

$$s_2(2^e n + r) = s_2(n) + s_2(r),$$

so every sequence in the k-kernel is a $\mathbb{Z}$-linear combination of the sequence $(s_2(n))_{n\geq 0}$ and the constant sequence 1.

Properties of k-regular Sequences

Theorem. A sequence is k-regular and takes finitely many values if and only if it is k-automatic.

Theorem. If $(a(n))_{n\geq 0}$ and $(b(n))_{n\geq 0}$ are k-regular sequences, then so are $(a(n)+b(n))_{n\geq 0}$, $(a(n)b(n))_{n\geq 0}$, and $(ca(n))_{n\geq 0}$ for any c.

Theorem. Let $c, d \geq 0$ be integers. If $(a(n))_{n\geq 0}$ is k-regular, then so is $(a(cn+d))_{n\geq 0}$.

Theorem. The sequence $(a(n))_{n\geq 0}$ is k-regular iff it is k^e-regular for any $e \geq 1$.

Examples of k-regular Sequences

Example 2. Families of Separating Subsets. Consider a set S containing n elements. If a family $F = \{A_1, A_2, \ldots, A_k\}$ of subsets of S has the property that for every pair (x, y) of distinct elements of S, we can find indices $1 \leq i, j \leq k$ such that

(i) $A_i \cap A_j = \emptyset$ and

(ii) $x \in A_i$ and $y \in A_j$,

then we call F a *separating family*. Let $f(n)$ denote the minimum possible cardinality of F.

For example, the letters of the alphabet can be separated by only 9 subsets:

$$\begin{array}{ll}
\{a,b,c,d,e,f,g,h,i\} & \{j,k,l,m,n,o,p,q,r\} \\
\{s,t,u,v,w,x,y,z\} & \{a,b,c,j,k,l,s,t,u\} \\
\{d,e,f,m,n,o,v,w,x\} & \{g,h,i,p,q,r,y,z\} \\
\{a,d,g,j,m,p,s,v,y\} & \{b,e,h,k,n,q,t,w,z\} \\
\{c,f,i,l,o,r,u,x\} &
\end{array}$$

Examples of k-regular Sequences

Cai Mao-Cheng showed that

$$f(n) = \min_{0 \le i \le 2} f_i(n),$$

where

$$f_i(n) = 2i + 3\lceil \log_3 n/2^i \rceil.$$

The first few terms of this sequence are given in the following table:

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$f(n)$	0	2	3	4	5	5	6	6	6	7	7	7	8	8

A priori, it is not clear that f is 3-regular, since the minimum of two k-regular sequences is not necessarily k-regular. However, in this case it is possible to prove the following characterization:

Examples of k-regular Sequences

Theorem. Let j be an integer such that $3^j < n \le 3^{j+1}$, i.e., $j = \lceil \log_3 n \rceil - 1$. Then

$$f(n) = \begin{cases} 3j+1, & \text{if } 3^j < n \le 4 \cdot 3^{j-1}; \\ 3j+2, & \text{if } 4 \cdot 3^{j-1} < n \le 2 \cdot 3^j; \\ 3j+3, & \text{if } 2 \cdot 3^j < n \le 3^{j+1}. \end{cases}$$

From this, it now easily follows that $f(n)$ is 3-regular.

Example 3. Mallows showed there there is a unique monotone sequence $(a(n))_{n \geq 0}$ of non-negative integers such that $a(a(n)) = 2n$ for $n \neq 1$. Here are the first few terms of this sequence:

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13
$a(n)$	0	1	3	4	6	7	8	10	12	13	14	15	16	18

It can be shown that $a(2^i + j) = 3 \cdot 2^{i-1} + j$ for $0 \leq j < 2^{i-1}$, and $a(3 \cdot 2^{i-1} + j) = 2^{i+1} + 2j$ for $0 \leq j < 2^{i-1}$.

We have

$$\begin{aligned} a(4n) &= 2a(2n) \\ a(4n+1) &= a(2n) + a(2n+1) \\ a(4n+3) &= -2a(n) + a(2n+1) + a(4n+2) \\ a(8n+2) &= 2a(2n) + a(4n+2) \\ a(8n+6) &= -4a(n) + a(2n+1) + a(4n+2) \end{aligned}$$

Hence this sequence is also 2-regular.

Example 4. A greedy partition of the natural numbers into sets avoiding arithmetic progressions.

Suppose we consider the integers $0, 1, 2, \ldots$ in turn, and place each new integer i into the set of lowest index S_k $(k \geq 0)$ so that S_k never contains three integers in arithmetic progression. For example, we put 0 and 1 in S_0, but placing 2 in S_0 would create an arithmetic progression of size 3 (namely, $\{0, 1, 2\}$), so we put 2 in S_1, etc.

Now define the sequence $(a_k)_{k \geq 0}$ as follows: $a_k = n$ if k is placed into set S_n. Here are the first few terms of this sequence:

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
a_k	0	0	1	0	0	1	1	2	2	0	0	1	0	0	1	1

Gerver, Propp, and Simpson showed that $a_{3k+r} = \lfloor (3a_k + r)/2 \rfloor$ for $k \geq 0$, $0 \leq r < 3$. It follows that $(a_k)_{k \geq 0}$ is 3-regular.

Example 5. Merge sort.

Consider sorting a list of n numbers as follows:

- sort the first half of the list recursively;
- sort the second half of the list recursively;
- merge the two sorted lists together.

The total number of comparisons needed is given by $T(1) = 0$ and

$$T(n) = T(\lfloor n/2 \rfloor) + T(\lceil n/2 \rceil) + n - 1$$

for $n \geq 2$.

It is now not hard to see that $T(n)$ is 2-regular, and in fact

$$T(n) = n\lceil \log_2 n \rceil - 2^{\lceil \log_2 n \rceil} + 1$$

for $n \geq 1$.

Inferring k-regular Sequences

Given a sequence $(s_n)_{n\geq 0}$, how can we determine if it is k-regular?

- construct a matrix in which the rows are elements of the k-kernel, and attempt to do row reduction
- as elements further out in the k-kernel are examined, the number of columns of the matrix that are known in all entries decreases
- if rows that are previously linearly independent suddenly become dependent with the elimination of terms further out in the sequence, then no relation can be accurately deduced; stop and retry after computing more terms
- if the subsequence $(s(k^j n + c))_{n\geq 0}$ is not linearly dependent on the previous sequences, try adding the subsequences $(s(k^j(kn+a)+c))_{n\geq 0}$ for $0 \leq a < k$

Inferring k-regular Sequences

- when no more linearly independent sequences can be found, you have found hypothetical relations for the sequence

Inferring k-regular Sequences

- (N. Strauss, 1988) Define
$$r(n) = \sum_{0 \le i < n} \binom{2i}{i},$$
- let $\nu_3(n)$ be the exponent of the highest power of 3 that divides n.
- The first few terms of $\nu_3(r(n))$ are:

$0, 1, 2, 0, 2, 3, 1, 2, 4, 0, 1, 2, 0, 3, 4, 2, 3, 5, 1, 2, \ldots$

- A 3-regular sequence recognizer easily produces the following conjectured relations (where $f(n) = \nu_3(r(n+1))$):
- $f(3n+2) = f(n) + 2$;
- $f(9n) = f(9n+3) = f(3n)$;
- $f(9n+1) = f(9n+4) = f(9n+7) = f(3n) + 1$.

- With a little more work, one arrives at the conjecture

$$\nu_3(r(n)) = \nu_3(n^2\binom{2n}{n}).$$

- proved by Allouche and JOS.
- A beautiful proof of this identity using 3-adic analysis was also given by Don Zagier.
- Zagier showed that if we set

$$F(n) = \frac{\sum_{0\le k\le n-1}\binom{2k}{k}}{n^2\binom{2n}{n}},$$

then $F(n)$ extends to a 3-adic analytic function from $\mathbb{Z}_3$ to $-1+3\mathbb{Z}_3$, and has the expansion:

$$F(-n) = -\frac{(2n-1)!}{(n!)^2}\sum_{0\le k\le n-1}\frac{(k!)^2}{(k-1)!}.$$

A "Mechanically-Produced" Conjecture

Let

$$a(n) = \sum_{0 \le k \le n} \binom{n}{k} \binom{n+k}{k}.$$

Let $b(n) = \nu_3(a(n))$. Then computer experiments suggest:

$$b(n) =$$

$$\begin{cases} b(\lfloor n/3 \rfloor) + (\lfloor n/3 \rfloor \bmod 2), & \text{if } n \equiv 0, 2 \pmod 3; \\ b(\lfloor n/9 \rfloor) + 1, & \text{if } n \equiv 1 \pmod 3. \end{cases}$$

This has been verified for $0 \le n \le 10,000$.

Open Problems on k-regular Sequences

1. Prove or disprove: $(\lfloor \frac{1}{2} + \log_2 n \rfloor)_{n \geq 1}$ is not a 2-regular sequence.

Comment. Suppose $a(n) = \lfloor \frac{1}{2} + \log_2 n \rfloor$ is 2-regular. Define $b(n) := a(n+1) - a(n)$ for $n \geq 1$. Then $(b(n))_{n \geq 0}$ would be 2-automatic, and is over the alphabet $\{0, 1\}$. The 1's in b are in positions $c_1 = 1$, $c_2 = 2$, $c_3 = 5$, $c_4 = 11$, $c_5 = 22$, $c_6 = 45$, $c_7 = 90$, etc. Then $c_{i+1} - 2c_i$ is the i'th bit in the binary expansion of $\sqrt{2}$.

2. Suppose S and T are k–regular sequences and $T(n) \neq 0$ for all n. Prove or disprove: if $S(n)/T(n)$ is always an integer, then $S(n)/T(n)$ is k–regular.

Comment. This is an analogue of van der Poorten's Hadamard quotient theorem.

Open Problems on k-regular Sequences

3. Prove or disprove: the 5-term analogue of the Gerver-Propp-Simpson sequence is not 5-regular.

Comment. Computer experiments show that if it is, the $\mathbb{Z}$-module generated by the 5-kernel must have large rank.

4. Prove or disprove: if a sequence $(a(n))_{n\geq 0}$ is simultaneously k- and l-regular, where k and l are multiplicatively independent, then $(a(n))_{n\geq 0}$ satisfies a linear recurrence.

Theorem. (Allouche, 1999) If $(a(n))_{n\geq 0}$ is simultaneously k- and l-regular, then it is kl-regular.

Open Problems on k-regular Sequences

5. Prove or disprove: if q is a polynomial taking integer values and p is a prime, then $(\nu_p(q(n)))_{n\geq 0}$ is either ultimately periodic or not p-regular.

Comment. If we understood, for example, the sequence $\nu_5(n^2+1)$, then we would understand the 5-adic expansion of $\sqrt{-1}$.

For Further Reading

1. J.-P. Allouche and J. O. Shallit. The ring of k–regular sequences. *Theoret. Comput. Sci.* **98** (1992), 163–187.
2. J. Gerver, J. Propp, and J. Simpson. Greedily partitioning the natural numbers into sets free of arithmetic progressions. *Proc. Amer. Math. Soc.* **102** (1988), 765–772.
3. R. Honsberger, Cai Mao-Cheng's solution to Katona's problem on families of separating subsets, in *Mathematical Gems III*, Mathematical Association of America, 1985, pp. 224–239.
4. G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bull. Soc. Math. France* **108** (1980), 401–419.
5. D. Zagier. Solution to advanced problem 6625. *Amer. Math. Monthly* **99** (1992), 66–69.