Theoretical Computer Science ••• (••••) •••–•••

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Note

Quasi-universal k-regular sequences

Juha Honkala

Department of Mathematics and Statistics, University of Turku, FI-20014 Turku, Finland

ARTICLE INFO

Article history:
Received 10 June 2021
Received in revised form 12 August 2021
Accepted 27 August 2021
Available online xxxx
Communicated by L.M. Kirousis

Keywords:
Automatic sequences
k-regular sequences
Recursively enumerable sets
Undecidability

ABSTRACT

We study the k-regular sequences introduced by Allouche and Shallit. We call a k-regular integer sequence s quasi-universal, if for every recursively enumerable set A of positive integers, the k-kernel of s contains a sequence t such that A equals the set of positive terms of t. We show that there are quasi-universal k-regular sequences for every integer $k \geq 2$. As a consequence, no nontrivial property is decidable for the set of positive terms of a k-regular sequence.

1. Introduction

Let $k \geq 2$ be an integer. Allouche and Shallit have introduced the k-regular sequences as a generalization of the k-automatic sequences (see [1]). In their seminal paper they showed that many known sequences are k-regular and proved that the theory of k-regular sequences is closely related to the theory of rational power series (see [3,4,8,15]). In a recent paper, Krenn and Shallit study decision problems concerning k-regular sequences (see [7]). Their main tool is the undecidability of Hilbert's tenth problem (see [9]).

In this paper we define and study quasi-universal k-regular integer sequences. We say that a k-regular integer sequence s is quasi-universal, if for every recursively enumerable set A of positive integers, the k-kernel of s contains a sequence t such that A equals the set of positive terms of t. The existence of quasi-universal k-regular sequences follows from the universal Diophantine representation of recursively enumerable sets of positive integers due to Matiyasevich (see [9]).

If $m \geq 2$ is an integer, every m-regular sequence can easily be coded as a recursively enumerable set of positive integers. Hence, if s is a quasi-universal k-regular sequence, then the subsequences of s in the kernel of s contain full information about all m-regular sequences.

In the construction of quasi-universal k-regular sequences we use the close connection between rational series and k-regular sequences. We use also ideas from [6], where the universal Diophantine representation of recursively enumerable sets is used to construct rational power series with high image complexity.

We will use quasi-universal k-regular sequences to prove that no nontrivial property is decidable for the set of positive terms of a k-regular sequence. This can be seen as a variant of Rice's theorem (stating that no nontrivial property is decidable for recursively enumerable sets) for k-regular sequences.

We will assume that the reader is familiar with the basics of language theory (see [2,12]), formal power series (see [3,4,8,15]) and recursively enumerable sets (see [9–11,14]). These references should be consulted for all unexplained notation and terminology.

E-mail address: juha.honkala@utu.fi.

https://doi.org/10.1016/j.tcs.2021.08.028

2. Definitions

As usual, $\mathbb{Z}$ is the set of integers and $\mathbb{Z}_+$ is the set of positive integers.

Let $k \geq 2$ be an integer. If n is a positive integer, there exist unique integers $j \geq 0$ and $c_j, \ldots, c_0 \in \{0, 1, \ldots, k-1\}$ such that

$$n = c_j k^j + c_{j-1} k^{j-1} + \cdots + c_0 \quad \text{and} \quad c_j \neq 0.$$

The word $c_j c_{j-1} \cdots c_0$ over the alphabet $\mathbf{k} = \{0, 1, \ldots, k-1\}$ is said to be the *representation* of n in base k. We write

$$\mathrm{rep}_k(n) = c_j c_{j-1} \cdots c_0.$$

Furthermore, $\mathrm{rep}_k(0) = \varepsilon$. Here ε is the empty word.

Let $u = a_n a_{n-1} \cdots a_0$, where $a_n, \ldots, a_0 \in \mathbf{k}$. The *numerical value* $\mathrm{val}_k(u)$ of u in base k is

$$\mathrm{val}_k(u) = a_n k^n + a_{n-1} k^{n-1} + \cdots + a_0.$$

Next we recall some notions concerning formal power series. Let X be a finite nonempty set of variables and let A be a semiring. The set of *formal power series* with noncommutative variables in X and coefficients in A is denoted by $A\langle\langle X\rangle\rangle$. If $r \in A\langle\langle X\rangle\rangle$, r is a mapping from the free monoid X^* generated by X into A. The image of a word $w \in X^*$ by r is denoted by $r(w)$ and r is written as

$$r = \sum_{w \in X^*} r(w) w.$$

Here $r(w)$ is called the *coefficient* of w in r. An alternative notation which is often used for $r(w)$ is (r, w).

If $r \in A\langle\langle X\rangle\rangle$, the *support* of r is the set

$$\mathrm{supp}(r) = \{w \in X^* \mid r(w) \neq 0\}.$$

If $r \in A\langle\langle X\rangle\rangle$, the *image* $\mathrm{Im}(r)$ of r is the set of its coefficients. Hence

$$\mathrm{Im}(r) = \{r(w) \mid w \in X^*\}.$$

If $r \in A\langle\langle X\rangle\rangle$ and $u \in X^*$, the series $ru^{-1} \in A\langle\langle X\rangle\rangle$ is defined by

$$(ru^{-1})(w) = (r, wu)$$

for all $w \in X^*$.

In what follows we will take $\mathbb{Z}$ as the basic semiring.

There are many equivalent ways to define the k-regular sequences and there are many equivalent ways to define the $\mathbb{Z}$-rational series. Here we choose among these equivalent definitions two which most clearly reveal the close connection between k-regular integer sequences and $\mathbb{Z}$-rational series. For more information we refer to [1–4,8,15].

For positive integers p and q, $\mathbb{Z}^{p \times q}$ is the set of $p \times q$ matrices having integer entries.

Definition 1. A series $r \in \mathbb{Z}\langle\langle X\rangle\rangle$ is called $\mathbb{Z}$*-rational* if there exist an integer $d \geq 1$, a monoid morphism

$$\mu : X^* \to \mathbb{Z}^{d \times d}$$

and two matrices $\alpha \in \mathbb{Z}^{1 \times d}$ and $\beta \in \mathbb{Z}^{d \times 1}$ such that for all $w \in X^*$,

$$r(w) = \alpha \mu(w) \beta.$$

Then the triple (α, μ, β) is called a *linear representation* of r.

Definition 2. Let $k \geq 2$ be an integer. A sequence $(s(n))_{n \geq 0}$ of integers is called k*-regular* if there exist an integer $d \geq 1$, a monoid morphism

$$\mu : \mathbf{k}^* \to \mathbb{Z}^{d \times d}$$

and two matrices $\alpha \in \mathbb{Z}^{1 \times d}$ and $\beta \in \mathbb{Z}^{d \times 1}$ such that for all $w \in \mathbf{k}^*$,

$$s(\mathrm{val}_k(w)) = \alpha \mu(w) \beta.$$

Then the triple (α, μ, β) is called a *linear representation* of s.

Let $r \in \mathbb{Z}\langle\langle \mathbf{k} \rangle\rangle$, where $k \geq 2$. We say that r satisfies the *left zero condition* if

$$r(0^i w) = r(w)$$

for all $i \geq 1$ and $w \in \mathbf{k}^*$. There is a one-to-one correspondence between the k-regular integer sequences and the $\mathbb{Z}$-rational series in $\mathbb{Z}\langle\langle \mathbf{k} \rangle\rangle$ satisfying the left zero condition.

Let $s = (s(n))_{n\geq 0}$ be a k-regular sequence. Then the *k-kernel* $\mathrm{Ker}_k(s)$ of s consists of the subsequences

$$(s(k^p n + q))_{n\geq 0}$$

of s, where $p \geq 0$ and $0 \leq q < k^p$. Let (α, μ, β) be a linear representation of s and let r be the $\mathbb{Z}$-rational series in $\mathbb{Z}\langle\langle \mathbf{k} \rangle\rangle$ which has the same linear representation as s. Then r satisfies the left zero condition. Assume that $u \in \mathbf{k}^*$ is a word of length p such that $\mathrm{val}_k(u) = q$. Then

$$(ru^{-1})(w) = r(wu) = \alpha\mu(wu)\beta = s(\mathrm{val}_k(wu)) = s(k^p\,\mathrm{val}_k(w) + q)$$

for all $w \in \mathbf{k}^*$. Hence ru^{-1} corresponds to $(s(k^p n + q))_{n\geq 0}$. The triple $(\alpha, \mu, \mu(u)\beta)$ is a linear representation of ru^{-1} and a linear representation of $(s(k^p n + q))_{n\geq 0}$.

3. Quasi-universal *k*-regular sequences

Recall that a set is *recursively enumerable* if there is a Turing machine accepting the set (see [9–11,14]).

Definition 3. Let $k \geq 2$ be an integer. A k-regular sequence s of integers is *quasi-universal* if for every recursively enumerable set A of positive integers, there is a sequence s_A in the k-kernel of s such that A equals the set of positive terms of s_A.

To prove the existence of quasi-universal k-regular sequences we use the universal Diophantine representation of recursively enumerable sets.

For the remaining part of this paper, fix an enumeration

$$A_1, A_2, A_3, \ldots$$

of the recursively enumerable subsets of $\mathbb{Z}_+$. The following deep theorem is due to Matiyasevich.

Theorem 1. *There is a positive integer t and a polynomial $p(z_1, \ldots, z_t)$ with integer coefficients such that for all positive integers i and m we have*

$$m \in A_i$$

if and only if there exist positive integers $m_1, \ldots, m_{t-2}$ such that

$$p(m_1, \ldots, m_{t-2}, m, i) = 0.$$

Furthermore, we can compute such a polynomial $p(z_1, \ldots, z_t)$.

For the proof of Theorem 1 see the monograph [9].

The following lemma is well known. It follows easily from the basic properties of rational series (see [8,15]).

Lemma 2. *Let $X = \{x, y\}$ be a two-letter alphabet and let $q(z_1, \ldots, z_t)$ be a polynomial with integer coefficients. Then there is a $\mathbb{Z}$-rational series $r_1 \in \mathbb{Z}\langle\langle X \rangle\rangle$ such that*

$$r_1(x^{m_1} y x^{m_2} y \cdots y x^{m_{t-1}} y x^{m_t}) = q(m_1, \ldots, m_t) \tag{1}$$

for all positive integers $m_1, \ldots, m_t$ and

$$\mathrm{supp}(r_1) \subseteq (x^+ y)^{t-1} x^+. \tag{2}$$

Furthermore, we can compute a linear representation of r_1.

The following theorem is a variant of a result in [6] (see also [5]).

Theorem 3. *Let $X = \{x, y\}$ be a two-letter alphabet. There exists a $\mathbb{Z}$-rational series $r \in \mathbb{Z}\langle\langle X\rangle\rangle$ such that for all positive integers i we have*

$$\mathrm{Im}(r(yx^i)^{-1}) = A_i \cup \{z \in \mathbb{Z} \mid z \leq 0\} \tag{3}$$

and

$$\mathrm{supp}(r) \subseteq (x^+ y)^{t-1} x^+ \ \cup \ x^+ y^t x^+ \tag{4}$$

where $t \geq 2$ is an integer. Furthermore, we can compute a linear representation of r.

Proof. Let $p(z_1, \ldots, z_t)$ be the universal polynomial of Theorem 1. Define the polynomial $q(z_1, \ldots, z_t)$ by

$$q(z_1, \ldots, z_t) = (1 + z_{t-1})(1 - p(z_1, \ldots, z_t)^2) - 1.$$

By Lemma 2 there exists a $\mathbb{Z}$-rational series $r_1 \in \mathbb{Z}\langle\langle X\rangle\rangle$ such that (1) holds for all positive integers $m_1, \ldots, m_t$ and (2) holds. Let

$$r_2 = \sum_{j,\ell=1}^{\infty} (-j) x^j y^t x^\ell$$

and

$$r = r_1 + r_2.$$

Since r_2 is $\mathbb{Z}$-rational, so is r. Since we can compute linear representations of r_1 and r_2, we can compute a linear representation of r. By (2) and the definition of r_2 we have (4).

Next, fix a positive integer i. Then

$$r(yx^i)^{-1} = r_1(yx^i)^{-1} + r_2(yx^i)^{-1}.$$

Since

$$\mathrm{supp}(r_1(yx^i)^{-1}) \cap \mathrm{supp}(r_2(yx^i)^{-1}) = \emptyset,$$

we have

$$\mathrm{Im}(r(yx^i)^{-1}) = \mathrm{Im}(r_1(yx^i)^{-1}) \cup \mathrm{Im}(r_2(yx^i)^{-1}).$$

Here

$$\mathrm{Im}(r_2(yx^i)^{-1}) = \{z \in \mathbb{Z} \mid z \leq 0\}.$$

To conclude the proof it remains to show that a positive integer m belongs to $\mathrm{Im}(r_1(yx^i)^{-1})$ if and only if $m \in A_i$.

First, assume that $m \in A_i$. Then there exist positive integers $m_1, \ldots, m_{t-2}$ such that $p(m_1, \ldots, m_{t-2}, m, i) = 0$. This implies that $q(m_1, \ldots, m_{t-2}, m, i) = m$. Hence

$$r_1(x^{m_1} yx^{m_2} y \cdots yx^{m_{t-2}} yx^m yx^i) = m.$$

Consequently,

$$(r_1(yx^i)^{-1})(x^{m_1} yx^{m_2} y \cdots yx^{m_{t-2}} yx^m) = m.$$

This shows that $m \in \mathrm{Im}(r_1(yx^i)^{-1})$.

Conversely, assume that a positive integer m belongs to $\mathrm{Im}(r_1(yx^i)^{-1})$. Then there exist positive integers $m_1, \ldots, m_{t-1}$ such that

$$r_1(x^{m_1} yx^{m_2} y \cdots yx^{m_{t-2}} yx^{m_{t-1}} yx^i) = m.$$

This implies that

$$q(m_1, \ldots, m_{t-2}, m_{t-1}, i) = m. \tag{5}$$

Since m is positive, the definition of q implies that

$$p(m_1, \ldots, m_{t-2}, m_{t-1}, i) = 0. \tag{6}$$

Hence we have

$$m_{t-1} \in A_i. \tag{7}$$

Equation (6) implies that

$$q(m_1, \ldots, m_{t-2}, m_{t-1}, i) = m_{t-1}. \tag{8}$$

Now (5) and (8) imply that $m = m_{t-1}$. Hence (7) implies $m \in A_i$. $\square$

It is now easy to construct quasi-universal k-regular sequences.

Theorem 4. *Let $k \geq 2$ be an integer. There is a k-regular sequence $(s(n))_{n\geq 0}$ of integers such that for all positive integers i we have*

$$\{s(k^{i+1}n + k^{i-1} + \cdots + k + 1) \mid n \geq 0\} = A_i \cup \{z \in \mathbb{Z} \mid z \leq 0\}.$$

Furthermore, we can compute a linear representation of s.

Proof. By Theorem 3 there is a $\mathbb{Z}$-rational series $r \in \mathbb{Z}\langle\langle \mathbf{k} \rangle\rangle$ such that for all positive integers i we have

$$\mathrm{Im}(r(01^i)^{-1}) = A_i \cup \{z \in \mathbb{Z} \mid z \leq 0\}.$$

We may assume that r satisfies the left zero condition. Furthermore, we can compute a linear representation of r.

Now, let $s = (s(n))_{n\geq 0}$ be the k-regular sequence which has the same linear representation as r. In other words, $s(\mathrm{val}_k(w)) = r(w)$ for all $w \in \mathbf{k}^*$.

Fix a positive integer i. Then

$$\{s(k^{i+1}n + k^{i-1} + \cdots + k + 1) \mid n \geq 0\} = \{s(\mathrm{val}_k(w01^i)) \mid w \in \mathbf{k}^*\}$$
$$= \{r(w01^i) \mid w \in \mathbf{k}^*\} = \mathrm{Im}(r(01^i)^{-1}) = A_i \cup \{z \in \mathbb{Z} \mid z \leq 0\}. \quad \square$$

4. Examples and undecidability results

The set of positive terms of a k-regular sequence is always recursively enumerable. Theorem 4 shows that nothing more can be said in general. We continue with some specific examples. The claims in these examples follow immediately by Theorem 4.

Example 1. There exists a k-regular sequence s_1 such that the positive terms of s_1 are the prime numbers.

Example 2. There exists a k-regular sequence s_2 such that the positive terms of s_2 are the prime powers.

Example 3. There exists a k-regular sequence s_3 such that the positive terms of s_3 are the numbers $3, 31, 314, 3141, \ldots$ obtained from the decimal expansion of π.

Example 4. There exists a k-regular sequence s_4 such that the positive terms of s_4 are the numbers 3, 3^3, $3^{3^3}, \ldots$.

In the following example we discuss the properties of the sequence s constructed in Theorem 4 in an informal way.

Example 5. Consider a k-regular sequence t. Let us say that the sets $\mathrm{Im}(t_j)$, where t_j are the sequences in the k-kernel of t are memory places of t. Then our sequence s has in its memory places all recursively enumerable sets of positive integers. If we know the index of the recursively enumerable set we have access to the memory place which contains the elements of the set. The memory place also contains all nonpositive integers. Now, for every integer $m \geq 2$, every m-regular integer sequence can easily be coded as a recursively enumerable set of positive integers. Hence, our k-regular sequence s has in its memory places all m-regular integer sequences, in a coded form. It should be noted that while it is easy to compute any term of a k-regular sequence, the sequence s does not generate the m-regular sequences in a nice way. In this sense it is very far from a universal k-regular sequence if we interpret the term universal as it is understood in connection of Turing machines.

The sequence s constructed in the previous section can be used to prove that many decision problems concerning k-regular sequences are undecidable. Krenn and Shallit have recently given a long list of undecidable problems for k-regular sequences (see [7]). As already pointed out in [1], k-regular sequences are closely related to rational series. Hence many undecidability results known for rational series imply undecidability results for k-regular sequences (see [8,15]). For example, Ruohonen has proved that image equivalence is undecidable for $\mathbb{N}$-rational series (see [13]). The proofs of these older results are sometimes longer than the proofs given by Krenn and Shallit if these older results prove undecidability under more restrictive conditions depending on the problem under consideration, but they are based on the undecidability of Hilbert's tenth problem.

The following theorem gives an undecidability result of some generality for k-regular sequences.

Theorem 5. *Let P be a property which at least one but not all recursively enumerable subsets of $\mathbb{Z}_+$ have. Let s be the k-regular sequence of Theorem 4. There is no algorithm to decide, given a sequence t in the k-kernel of s, whether the set of positive terms of t has property P.*

Proof. Let i be a positive integer. Then we can compute a sequence t_i in the k-kernel of s such that the set of positive terms of t_i is A_i, where A_i is the ith recursively enumerable subset of $\mathbb{Z}_+$ in our fixed list of all recursively enumerable subsets of $\mathbb{Z}_+$. The claim now follows by Rice's theorem stating that no nontrivial property of recursively enumerable sets is decidable (see [10,11,14]). □

The following theorem lists some specific corollaries of Theorem 5.

Theorem 6. *Let s be the k-regular sequence of Theorem 4. There is no algorithm to decide any of the following problems for the sequences in the k-kernel of s:*

(a) is a given positive integer a term of the sequence,
(b) is the image of the sequence equal to $\mathbb{Z}$,
(c) are there infinitely many positive integers in the image,
(d) is every positive term a prime (resp. a prime power),
(e) is every positive term a perfect square,
(f) are the positive terms of the sequence the numbers $3, 31, 314, \ldots$ obtained from the decimal expansion of π,
(g) are the positive terms of the sequence the numbers $3, 3^3, 3^{3^3}, \ldots$.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The author would like to thank the anonymous referee for useful remarks.

References

[1] J.-P. Allouche, J. Shallit, The ring of k-regular sequences, Theor. Comput. Sci. 98 (1992) 163–197.
[2] J.-P. Allouche, J. Shallit, Automatic Sequences. Theory, Applications, Generalizations, Cambridge University Press, 2003.
[3] J. Berstel, C. Reutenauer, Rational Series and Their Languages, Springer, 1988.
[4] J. Berstel, C. Reutenauer, Noncommutative Rational Series with Applications, Cambridge University Press, 2011.
[5] J. Honkala, Products of matrices and recursively enumerable sets, J. Comput. Syst. Sci. 81 (2015) 468–472.
[6] J. Honkala, Rational series with high image complexity, RAIRO Theor. Inform. Appl. 51 (2017) 1–6.
[7] D. Krenn, J. Shallit, Decidability and k-regular sequences, Available at https://arxiv.org/abs/2005.09507, 2020.
[8] W. Kuich, A. Salomaa, Semirings, Automata, Languages, Springer, 1986.
[9] Y. Matiyasevich, Hilbert's Tenth Problem, MIT Press, 1993.
[10] H. Rogers, Theory of Recursive Functions and Effective Computability, McGraw-Hill, 1967.
[11] G. Rozenberg, A. Salomaa, Cornerstones of Undecidability, Prentice Hall, 1994.
[12] G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, vols. 1-3, Springer, 1997.
[13] K. Ruohonen, On equality of multiplicity sets of regular languages, Theor. Comput. Sci. 36 (1985) 113–117.
[14] A. Salomaa, Computation and Automata, Cambridge University Press, 1985.
[15] A. Salomaa, M. Soittola, Automata-Theoretic Aspects of Formal Power Series, Springer, 1978.