UNCOVERING HIDDEN AUTOMATIC SEQUENCES

ELŻBIETA KRAWCZYK

It has been observed in several contexts that certain substitutive sequences defined using substitutions of non-constant length could in fact also be obtained from substitutions of constant length. While it is easy to construct such examples artificially, they also occur naturally, and the corresponding constant-length substitution is often by no means obvious. Such discoveries of 'hidden automatic sequences' (a name we borrow from [2]) are often insightful since automatic sequences are considerably better understood and can be treated using more specialized tools (e.g. finite automata). A particularly striking example is the Lysënok morphism related to the presentation of the first Grigorchuk group [7], where spectral properties of the system generated by the Lysënok morphism are used to deduce spectral properties of the Schreier graph of the Grigorchuk group. In the opposite direction, a problem of showing that a given substitutive sequence is not automatic has also appeared in several contexts, e.g. in the study of gaps between factors in the famous Thue–Morse sequence [9] or in the mathematical description of the drawing of the classical Indian kolam [1]. In each case, some *ad hoc* methods are employed to prove or disprove the automaticity of the substitutive sequence under consideration.

The problem of how to recognize that a substitutive sequence is automatic has been raised recently in [2] by Allouche, Dekking and Queffélec; however, it also appeared in the classical book on automatic sequences by Allouche and Shallit [3, Section 7.11, Problem 3] and has been studied much earlier by Dekking [5].

Problem A. For a given substitutive sequence, decide whether it is automatic.

Let $k \geqslant 2$. A sequence is purely substitutive (resp. purely k-automatic) if it is a fixed point of some substitution (resp. substitution of constant length k). A sequence is *substitutive* (resp. *k-automatic*) if it can be obtained as the image of a purely substitutive (resp. purely k-automatic) sequence under a coding. We recall that for a substitution $\varphi\colon \mathscr{A} \to \mathscr{A}^*$, its *incidence matrix* is defined as $M_\varphi = (|\varphi(b)|_a)_{a,b}$, where $|\varphi(b)|_a$ denotes the number of occurrences of the letter a in $\varphi(b)$. A substitution φ is called *primitive* if its incidence matrix M_φ is primitive.

A necessary condition for a substitutive sequence to be automatic comes from a version of Cobham's theorem for substitutions proven by Durand [6]: it implies that a substitutive sequence, which is not ultimately periodic, can be k-automatic only if the dominant eigenvalue of the incidence matrix of the underlying substitution is multiplicatively dependent with k.[1] It is well-known, however, that this condition is not necessary: there are primitive substitutions whose dominant eigenvalue is an integer and whose nonperiodic fixed points are not automatic. In the opposite direction, a useful sufficient condition for a fixed point of a substitution to be automatic has been obtained by Dekking in 1976. It says that if the length vector ${}^t(|\varphi(a)|)_{a\in\mathscr{A}}$ is a left eigenvector of M_φ, then any fixed point of φ is automatic [5], see also [2]. A recent paper by Allouche, Shallit and Yassawi [4] provides a handy toolkit of methods of showing that a (substitutive) sequence is not automatic; nevertheless, no general necessary and sufficient condition is known.

In this talk we will give a solution to Problem A for uniformly recurrent substitutive sequences. Recall that for a letter $a \in \mathscr{A}$ and a sequence x, a word $w \in \mathscr{L}(x)$ is called *a return word* to a (in x) if w starts with a, w has exactly one occurrence of a, and $wa \in \mathscr{L}(x)$. For a purely substitutive sequence x given by φ, we let R_a denote the set of return words to a in x, and we let $\tau\colon \mathrm{R}_a \to \mathrm{R}_a^*$ denote the return substitution of φ to a (we do not define it formally here, but the reader will get the idea by looking at Example C below). Theorem B shows that for a uniformly

[1] Two real numbers $\alpha, \beta > 1$ are called *multiplicatively dependent* if $\alpha^n = \beta^m$ for some integers $m, n \geqslant 1$.

recurrent substitutive sequence x, Dekking's criterion applied to the return substitution of an underlying purely substitutive sequence essentially gives a necessary and sufficient condition for x to be automatic.

Theorem B. *Let $\varphi\colon \mathscr{A} \to \mathscr{A}^*$ be a primitive substitution, let $\pi\colon \mathscr{A} \to \mathscr{B}$ be a coding, let x be a fixed point of φ, and let $a = x_0$. Let $y = \pi(x)$ and assume that y is not periodic. Let $\tau\colon \mathrm{R}_a \to \mathrm{R}_a^*$ be the return substitution to a, let M_τ denote the incidence matrix of τ, and let s denote the size of the largest Jordan block of M_τ corresponding to the eigenvalue 0. The following conditions are equivalent:*

(*i*) *y is automatic;*
(*ii*) *${}^t(|\varphi^s(w)|)_{w\in \mathrm{R_a}}$ is a left eigenvector of M_τ.*

Example C. Let $\varphi\colon \mathscr{A} \to \mathscr{A}^*$ be a primitive substitution given by

$$a \mapsto aca,\ b \mapsto bca, c \mapsto cbcac,$$

and let $x = acac\ldots$ be a fixed point of φ starting with a. The set of return words to a in x is given by $\mathrm{R}_a = \{ac, acbc\}$. To see this, note that ac is the first return word to a occurring in x. The word $\varphi(ac) = ac|acbc|ac$ is a concatenation of 3 return words to a in which $acbc$ is the only new word. Applying φ to it, we see that $\varphi(acbc) = ac|acbc|acbc|acbc|ac$ is a concatenation of 5 return words and no new return words appear in this factorisation. Hence R_a consists exactly of these two words. Relabelling, $1 = ac$, $2 = acbc$, we get that the return substitution $\tau\colon \{1,2\} \to \{1,2\}^*$ is given by

$$1 \mapsto 121,\ 2 \mapsto 12221.$$

It is easy to check that the fixed point x of φ is not periodic. The incidence matrix of the return substitution $\tau\colon \{1,2\} \to \{1,2\}^*$ is given by

$$M_\tau = \begin{pmatrix} 2 & 2 \\ 1 & 3 \end{pmatrix},$$

and has eigenvalues 4 and 1; in particular, s=0. Since ${}^t(|w|)_{w\in\mathrm{R_a}} = (2,4)$ is a left eigenvector of M_τ corresponding to the eigenvalue 4, by Theorem B, the fixed point x of φ is automatic.

Time permitting, we will also discus the progress made on Problem A in the case of general substitutive sequences. The talk is based on a joint work with Clemens Müllner [8].

References

1. Gabrielle Allouche, Jean-Paul Allouche, and Jeffrey Shallit, *Kolam indiens, dessins sur le sable aux îles Vanuatu, courbe de Sierpiński et morphismes de monoïde*, Ann. Inst. Fourier (Grenoble) **56** (2006), no. 7, 2115–2130.
2. Jean-Paul Allouche, Michel Dekking, and Martine Queffélec, *Hidden automatic sequences*, Comb. Theory **1** (2021), Paper No. 20, 15.
3. Jean-Paul Allouche and Jeffrey Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003, Theory, applications, generalizations.
4. Jean-Paul Allouche, Jeffrey Shallit, and Reem Yassawi, *How to prove that a sequence is not automatic*, Expo. Math. **40** (2022), no. 1, 1–22.
5. F. Michel Dekking, *The spectrum of dynamical systems arising from substitutions of constant length*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **41** (1977/78), no. 3, 221–239.
6. Fabien Durand, *Cobham's theorem for substitutions*, J. Eur. Math. Soc. (JEMS) **13** (2011), no. 6, 1799–1814.
7. Rostislav Grigorchuk, Daniel Lenz, and Tatiana Nagnibeda, *Spectra of Schreier graphs of Grigorchuk's group and Schroedinger operators with aperiodic order*, Math. Ann. **370** (2018), no. 3-4, 1607–1637.
8. Elźbieta Krawczyk and Clemens Müllner, *Automaticity of uniformly recurrent substitutive sequences*, (2021), Preprint. arXiv:2111.13134 [math.NT].
9. Lukas Spiegelhofer, *Gaps in the Thue–Morse Word*, J. Aust. Math. Soc. **114** (2023), no. 1, 110–144.

Faculty of Mathematics and Computer Science, Institute of Mathematics, Jagiellonian University, Stanisława Łojasiewicza 6, 30-348 Kraków

Email address: ela.krawczyk7@gmail.com