On a generalization of Thue sequences

Jaka Kranjc Borut Lužar

Faculty of Information Studies
Novo mesto, Slovenia

jaka.kranjc@fis.unm.si borut.luzar@gmail.com

Martina Mockovčiaková

European Centre of Excellence NTIS -
New Technologies for the Information Society
Department of Mathematics
University of West Bohemia
Pilsen, Czech Republic

mmockov@ntis.zcu.cz

Roman Soták

Faculty of Science
Pavol Jozef Šafárik University
Košice, Slovakia

roman.sotak@upjs.sk

Submitted: Nov 7, 2014; Accepted: May 10, 2015; Published: May 22, 2015
Mathematics Subject Classifications: 68R15, 05C15

Abstract

A sequence is *Thue* or *nonrepetitive* if it does not contain a repetition of any length. We consider a generalization of this notion. A j-subsequence of a sequence S is a subsequence in which two consecutive terms are at indices of difference j in S. A k*-Thue sequence* is a sequence in which every j-subsequence, for $1 \leqslant j \leqslant k$, is also Thue. It was conjectured that $k+2$ symbols are enough to construct an arbitrarily long k-Thue sequence and shown that the conjecture holds for $k \in \{2,3,5\}$. In this paper we present a construction of k-Thue sequences on $2k$ symbols, which improves the previous bound of $2k+10\sqrt{k}$. Additionally, we define cyclic k-Thue sequences and present a construction of such sequences of arbitrary lengths when $k=2$ using four symbols, with three exceptions. As a corollary, we obtain tight bounds for total Thue colorings of cycles. We conclude the paper with some open problems.

Keywords: Thue sequence; k-Thue sequence; total Thue chromatic number

1 Introduction

A subsequence of consecutive terms of a sequence S is called a *block*. A *repetition* in a sequence S is a subsequence $\xi_1 \dots \xi_t \xi_{t+1} \dots \xi_{2t}$ of consecutive terms of S such that $\xi_i = \xi_{t+i}$ for every $i = 1, \dots, t$. A repetition thus consists of two identical *repetition blocks* and hence its length is always even. A sequence is called *nonrepetitive* or *Thue* if it does not contain a repetition of any length. Throughout the paper, the terms of sequences are not separated by commas as this is a usual notation in this area of combinatorics.

Repetitions and other regularities in sequences of symbols (words) take an important part in combinatorics on words (we refer the reader to [4] for a short introduction to this topic). The origins of their studies date back in 1906, when Thue [28] (see [3] for a translation) showed that using only three symbols one can construct an arbitrarily long sequence without a repetition. His famous work attracted a considerable attention and later many applications have been found in various fields of science [21].

Due to many unexpected applications that sequences without repetitions found in science, naturally a number of generalizations was presented. The most basic one is forbidding an appearance of k equal consecutive blocks, where already Thue showed that for $k = 3$, only two symbols suffice to construct an infinite sequence [28]. A more restrictive generalization are the sequences avoiding abelian powers. An *abelian k-th power* is a sequence of k consecutive blocks, where in each block every symbol s appears the same number $n(s)$ of times. Again, the most interesting case is when k equals 2. It was shown that 25 [13], five [22], and finally four [20] symbols suffice to construct an infinite sequence without abelian 2-repetitions. For $k = 3$ and $k = 4$, Dekking [12] showed that three and two symbols suffice, respectively. See e.g. [6] for more details on generalizations of Thue sequences.

In this paper we consider the following generalization of Thue sequences introduced by Currie and Simpson [11]: a (possibly infinite) sequence S is *k-Thue* (or *nonrepetitive up to mod k*) if every j-subsequence of S is Thue, for $1 \leqslant j \leqslant k$. Here, a *$j$-subsequence* of S is a subsequence $\xi_i \xi_{i+j} \xi_{i+2j} \dots$, for any i. Notice that a 1-Thue sequence is simply a Thue sequence. Currie and Simpson [11] introduced this notion in connection with *nonrepetitive tilings*, i.e. assignments of symbols to the lattice points of the plane such that all lines in prescribed directions are nonrepetitive.

As an example, consider a sequence $a\,b\,d\,c\,b\,c$, which is Thue, but not 2-Thue, since the 2-subsequence $b\,c\,c$ is not Thue. On the other hand, $a\,b\,c\,a\,d\,b$ is 2-Thue, but not 3-Thue.

A natural question arises what is the minimum number of symbols required to construct an arbitrarily long k-Thue sequence. In [11], the authors showed that four symbols are enough to create 2-Thue sequences and five symbols suffice for 3-Thue sequences of arbitrary lengths. The lower bound on the number of symbols needed to construct a k-Thue sequence of an arbitrary length is obvious; for a positive integer k, at least $k + 2$ symbols are required to construct such sequences.

In 2002 Grytczuk [16] conjectured that, in fact, the upper bound is equal to the lower bound for any k.

Conjecture 1 ([16])**.** For any k, $k + 2$ symbols suffices to construct a k-Thue sequence.

This conjecture is hence true for $k = 2$ and 3 and moreover, in [9], Currie and Moodie confirmed it also for $k = 5$. The case with $k = 4$ was considered by Currie and Pierce [10] using an application of the fixing block method. However, the upper bound for general k is still far from the conjectured. The bound of $e^{33}k$ established in [16] was substantially improved to $2k + O(\sqrt{k})$ in [17], which is currently the best known upper bound.

Theorem 2 ([17]). *For an arbitrary $k \geqslant 1$, there is an arbitrarily long k-Thue sequence using at most $2k + 10\sqrt{k}$ symbols.*

Motivated by the paper of Grytczuk et al. we present a construction of arbitrarily long k-Thue sequences using at most $2k$ symbols which improves the previous bound. We also introduce cyclic k-Thue sequences and establish the number of symbols required to construct a cyclic 2-Thue sequence of arbitrary length.

The paper is organized as follows: in Section 2, we improve the upper bound on the number of symbols needed to construct k-Thue sequences, in Section 3, we deal with cyclic 2-Thue sequences, and in Section 4 we briefly describe nonrepetitive colorings of graphs and present a result establishing tight bounds for total Thue colorings of cycles. Finally, in Section 5, we propose a conjecture, questions and possibilities for further work.

2 Upper bound on k-Thue sequences

In this section we improve the upper bound for the number of symbols needed to construct k-Thue sequences, with $k \geqslant 3$, given in [17].

Theorem 3. *For an arbitrary $k \geqslant 3$, there is an arbitrarily long k-Thue sequence using at most $2k$ symbols.*

Notice that our proof is constructive and provides a k-Thue sequence of a given length, while the proofs of upper bounds known so far just show an existence of such sequences.

Proof. Consider an arbitrary Thue sequence on three symbols α, β and γ. We replace each of the three symbols by $2k$ symbols, $a_1, \dots, a_k$ called *a-symbols*, and $b_1, \dots, b_k$ called *b-symbols*, in the following way: α by $a_1 \dots a_k\, b_1 \dots b_k$, β by $a_1 \dots a_k\, b_2 \dots b_k b_1$, and γ by $a_1 \dots a_k\, b_3 \dots b_k b_1 b_2$, and denote the obtained sequence τ_k. We refer to the blocks replacing α, β and γ as the *α-block*, *β-block* and *γ-block*, respectively, and denote these blocks *replacement blocks*. In Figure 1 the scheme of symbol adjacencies in τ_1 and τ_k is depicted.

We will use the three claims below, to show that the sequence τ_k is k-Thue. We denote the i-th term of τ_k by ξ_i.

Claim 4. *Let S be a j-subsequence of τ_k, for $1 \leqslant j \leqslant k$. If a term ξ of S is a b-symbol preceded or followed by an a-symbol, we can uniquely determine the replacement block in which ξ is contained.*

Proof. By construction of τ_k, it is clear that the distance (the number of terms) between an a-symbol a_s and a b-symbol b_t, for some fixed s and t, with $1 \leqslant s, t \leqslant k$, within a replacement block is different in each of the replacement blocks, and so we can uniquely determine to which replacement block b_t belongs. Moreover, this holds also if a_s is in a subsequent replacement block of the replacement block of b_t. □

Claim 5. *Let R be a repetition in some j-subsequence of τ_k starting with the symbol a_s, where $j < s \leqslant k$. Then there is a repetition in τ_k starting with the symbol a_{s-j}.*

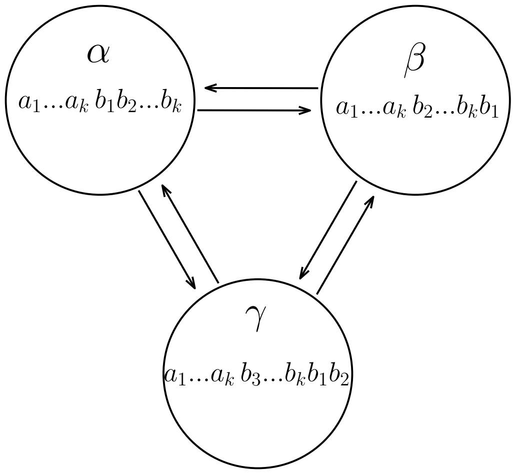

Figure 1: The scheme of symbol adjacencies in τ_1 and τ_k.

Proof. Let $\xi_i \xi_{i+j} \dots \xi_{i+(2r-1)j}$ be a repetition of length $2r$ in a j-subsequence of τ_k. Since every a-symbol is always at the same position within a replacement block, from $\xi_i = \xi_{i+rj}$ it follows that $\xi_{i-j} = \xi_{i+(r-1)j}$ and so the sequence $\xi_{i-j}\xi_i\xi_{i+j} \dots \xi_{i+(2r-2)j}$ is also a repetition in τ_k. $\square$

Claim 6. *If there is a repetition starting with a b-symbol in some j-subsequence of τ_k, there is also a repetition starting with an a-symbol.*

Proof. Let $\xi_i \xi_{i+j} \dots \xi_{i+(2r-1)j}$ be a repetition of length $2r$ in a j-subsequence of τ_k, where ξ_i is a b-symbol b_s for some $1 \leqslant s \leqslant k$. We consider two cases regarding the term $\xi_{i+(r-1)j}$. Suppose first that $\xi_{i+(r-1)j}$ is an a-symbol a_t, $1 \leqslant t \leqslant k$. Let p be the least positive integer such that ξ_{i+pj} is an a-symbol. By Claim 4, the terms $\xi_{i+(p-1)j}$ and ξ_{i+pj} uniquely determine the corresponding replacement block, i.e. the block in which also ξ_i is contained. Notice that the terms $\xi_{i+(r+p-1)j} = \xi_{i+(p-1)j}$ and $\xi_{i+(r+p)j} = \xi_{i+pj}$ determine the same replacement block in which similarly $\xi_{i+rj} = b_s$ is contained. Since $\xi_{i+(r-1)j} = a_t$, it follows that also $\xi_{i-j} = a_t$ and hence $\xi_{i-j}\xi_i\xi_{i+j} \dots \xi_{i+(2r-2)j}$ is also a repetition.

Thus, we may assume that $\xi_{i+(r-1)j}$ is a b-symbol b_u, $1 \leqslant u \leqslant k$. Let p and q be the least positive integers such that ξ_{i+pj} and $\xi_{i+(r-1-q)j}$ are a-symbols. Then, by Claim 4, the terms $\xi_{i+(r-1-q)j}$ and $\xi_{i+(r-q)j}$ uniquely determine the corresponding replacement block and so $\xi_{i+(r-q)j} \dots \xi_{i+(r+p-1)j} = \xi_{i+(2r-q)j} \dots \xi_{i+(2r+p-1)j}$. This means that $\xi_{i+pj}\xi_{i+(p+1)j} \dots \xi_{i+(2r+p-1)j}$ is also a repetition. $\square$

Suppose now that there is some repetition R in a j-subsequence of τ_k, with $1 \leqslant j \leqslant k$. Notice that since each replacement block has length $2k$, consecutive a-symbols (resp. b-symbols) in R correspond to one replacement block. By Claim 6, we may assume that the first term of R is an a-symbol ξ and by Claim 5, we have that $\xi = a_i$ for $1 \leqslant i \leqslant j$. It follows that the last term of the first repetition block of R (and also of R) is a b-symbol. Thus, by Claim 4, we infer that the replacement blocks of both repetition blocks in R are identical, which means that there is a repetition in τ_1 also, a contradiction. Hence, τ_k is indeed k-Thue. $\square$

3 Cyclic 2-Thue sequences

In this section we introduce the notion of cyclic Thue sequences and their generalization, cyclic k-Thue sequences, where for $k = 2$ we establish tight bounds on the number of symbols needed to construct sequences of arbitrary lengths. Cyclic Thue sequences were first investigated by Currie [7], who was motivated by the question of Alon et al. [1] asking which cycles can be Thue colored by three colors.

Let $S = \xi_1 \dots \xi_\ell$ be an arbitrary sequence. A sequence $\overline{S}$ is a *conjugate* of S if there is some integer i such that $\overline{S} = \xi_i \dots \xi_\ell \xi_1 \dots \xi_{i-1}$. A sequence is *cyclic Thue* if all its conjugates are Thue (we adopted this definition from [7]).

The definition of cyclic k-Thue sequences is analogous. A *cyclic j-subsequence* is a subsequence of S in which every two consecutive terms appear at indices i and $i + j$ (mod ℓ), for some integer i, and every term from S appears at most once. Hence, for ℓ even the length of any cyclic 2-subsequence is at most $\frac{\ell}{2}$, while for ℓ odd, its length is at most ℓ. A sequence S is *cyclic k-Thue* if every cyclic j-subsequence of S is Thue, for $1 \leqslant j \leqslant k$. One can imagine the terms of a cyclic sequence on a circle such that the last term is followed by the first term (see Fig. 3 for an example).

Observe that a cyclic sequence consisting of at least four terms requires at least four symbols in order to be 2-Thue. Moreover, for cyclic 2-Thue sequences of lengths 5, 7, and 11 at least five symbols are needed.

Lemma 7. *Every cyclic* 2*-Thue sequence of length* 5*,* 7*, or* 11 *is constructed by at least five symbols.*

Proof. Let $S = \xi_1 \dots \xi_\ell$ be a cyclic 2-Thue sequence of length $\ell \in \{5, 7, 11\}$. We consider each of the three cases separately assuming that S contains four distinct symbols a, b, c, and d, obtaining a contradiction.

$\ell = 5$. Clearly, every symbol appears at most $\lfloor \ell/3 \rfloor$ times in a cyclic 2-Thue sequence of length ℓ, since the terms at distance at most 2 have distinct symbols assigned. Hence, all terms must be distinct in S, a contradiction.

$\ell = 7$. Every symbol appears at most twice in a seven-term cyclic 2-Thue sequence. Without loss of generality, we may assume that a appears once, while the other three symbols appear twice in S. We may assume that $\xi_1 = c$, $\xi_2 = a$, and $\xi_3 = b$. Then, $\xi_4 = \xi_7 = d$, $\xi_5 = c$, and $\xi_6 = b$. But then there is a repetition $\xi_6 \xi_1 \xi_3 \xi_5$ in a cyclic 2-subsequence of S, a contradiction. On the other hand, a sequence *abcabde* constructed by five distinct symbols is cyclic 2-Thue.

$\ell = 11$. Any symbol of S occurs at most three times. Hence, three symbols appear three times and one, say d, appears twice. Assume that $\xi_1 = d$. Since the length of a 2-Thue sequence on three symbols is at most 5, the second occurrence of d is determined up to symmetry. Thus, assume d is at index 7. Moreover, a 2-Thue sequence of length 5 on three symbols has a unique ordering up to isomorphism, so we may assume that

$\xi_1 \dots \xi_7 = dabcabd$. The symbol c appears once at index 4 and the other two occurrences of this symbols are at index 8 and 11. Therefore, $S = dabcabdcxyc$. There are two possibilities: $xy = ab$ or $xy = ba$. In the former case a repetition $cbcb$ appears in a 2-subsequence of S, and in the latter case there is a repetition $adbadb$ in a cyclic 2-subsequence, a contradiction. Again, there exists a cyclic 2-Thue sequence $abcabdcbace$ on five symbols. $\square$

Before we prove the main theorem of this section, we present an alternative construction of 2-Thue sequences using a similar approach as in Section 2. We use such constructed sequences to construct cyclic 2-Thue sequences. Consider an arbitrary Thue sequence on an alphabet $\{\alpha, \beta, \gamma\}$ and insert a symbol δ between any two consecutive symbols α and γ (or γ and α) to obtain a new sequence τ_1. By Lemma 8 below, τ_1 is also Thue.

First, we present some additional notation used in the sequel. By $S(p_1, p_2)$ we denote a subsequence of a sequence $S = \xi_1 \dots \xi_n$ starting (at index p_1) with the term ξ_{p_1} and ending (at index p_2) with the term ξ_{p_2} for some indices p_1 and p_2.

Lemma 8 ([19]). *Let $A = a_1 \dots a_m$ be a nonrepetitive sequence with $a_i \in \mathbb{A}$ for all $i = 1, 2, \dots, m$. Let $B^i = b_1^i b_2^i, \dots, b_{m_i}^i, 0 \leqslant i \leqslant r+1$, be nonrepetitive sequences with $b_j^i \in \mathbb{B}$ for all $i = 0, 1, \dots, r+1$ and $j = 1, 2, \dots, m_i$. If $\mathbb{A} \cap \mathbb{B} = \emptyset$ then $S = B^0 A(1, n_1) B^1 A(n_1 + 1, n_2) \dots B^r A(n_r + 1, m) B^{r+1}$ with $1 \leqslant n_1 < n_2 < \cdots < n_r < m$ is a nonrepetitive sequence.*

Notice that in our case we have $\mathbb{A} = \{\alpha, \beta, \gamma\}$, $\mathbb{B} = \{\delta\}$, $B^i = \delta$, and $B^0 = B^{r+1} = \emptyset$.

Using the property that there is no pair of adjacent terms α and γ nor a pair of adjacent terms β and δ, we obtain the following lemma.

Lemma 9. *For every odd integer k there exists a subsequence of τ_1 of length k starting and ending with the same symbol.*

Proof. Let k be an arbitrary odd integer. Suppose, for a contradiction, that every subsequence of τ_1 of length k starts and ends with a different symbol. Let $T = \xi_{i+1} \dots \xi_{i+k}$ be a subsequence of τ_1 of length k. Without loss of generality we may assume that T starts with α. Let $j \in \{i+1, \dots, i+k\}$. Observe that for every j, with $j \equiv i+1 \pmod 2$, we have $\xi_j, \xi_{j+k-1} \in \{\alpha, \gamma\}$, and for j, where $j \equiv i \pmod 2$, we have $\xi_j, \xi_{j+k-1} \in \{\beta, \delta\}$. By the initial assumption, we also have that $\xi_j \neq \xi_{j+k-1}$ for every j. However, $\xi_j = \xi_{j+2k-2}$ for every j and hence there is a repetition of length $4k-4$ in τ_1, a contradiction. $\square$

Now, replace each of the symbols α, β, γ, and δ in τ_1 with the four-term blocks $abcd$, $abdc$, $bacd$, and $badc$, respectively. We denote the obtained sequence by τ_2 (see Figure 2 for the scheme of symbol adjacencies in τ_1 and τ_2).

Below, we list two basic properties of τ_2. Both of them are obvious by the construction of τ_2.

Property 1. *In τ_2 the symbols a and b appear only at positions* 1 *or* 2 $\pmod 4$ *and the symbols c and d appear only at positions* 0 *or* 3 $\pmod 4$*.*

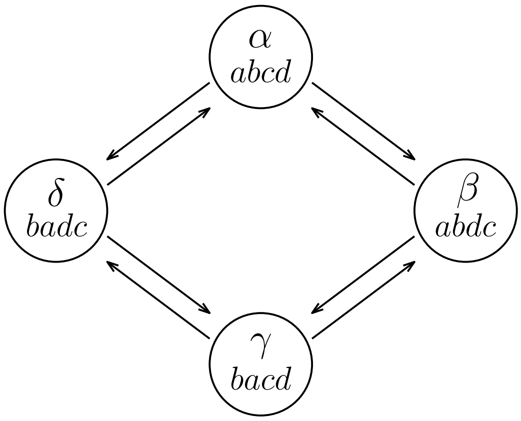

Figure 2: The scheme of symbol adjacencies in τ_1 and τ_2.

Property 2. *In τ_2 the symbol pairs cd and dc appear alternately, i.e. $\xi_i\xi_{i+1} = cd$ for $i \equiv 3 \pmod 8$ and $\xi_i\xi_{i+1} = dc$ for $i \equiv 7 \pmod 8$.*

Now we are ready to show that the sequence τ_2 is 2-Thue.

Theorem 10. *The sequence τ_2 is 2-Thue.*

Proof. First, we show that τ_2 is without a repetition by considering two cases. In the former, we show that there is no repetition of length at most 8 in τ_2 (this case is discussed in Claim 11).

Claim 11. *There is no repetition of length at most 8 in τ_2.*

Proof. By Property 1, the same symbol appears after at least two other symbols, so repetitions of lengths 2 or 4 are not possible. Moreover, Property 1 implies that every block of length 5 in τ_2 contains four distinct symbols, hence there is no repetition of length 6 in τ_2. Obviously, there is no repetition of length 8 in τ_2 due to Property 2. □

In the latter case, we show that there is no repetition of length greater than 8. The proof is divided into several subcases regarding the length of an eventual repetition and the position of its starting term. We start with the following claim.

Claim 12. *There is no repetition of length $2t \equiv 2 \pmod 4$ in τ_2.*

Proof. Since the symbols in τ_2 come in pairs, a with b and c with d, every block of a repetition must be of even size. □

Let k and t be positive integers and suppose there is a repetition

$$R = \xi_{k+1} \dots \xi_{k+t}\xi_{k+t+1} \dots \xi_{k+2t}$$

in τ_2 of length $2t$ starting at the position $k+1$ such that $\xi_{k+i} = \xi_{k+t+i}$ for $i = 1, \dots, t$. We distinguish several cases regarding the position of the starting term of R in τ_2. First, consider the values of t. By Claim 12, t is not odd, and moreover $t \not\equiv 2 \pmod 4$, since the second block would start with a symbol from the second pair. Hence $t \equiv 0 \pmod 4$. Now, we consider four subcases regarding the values of k.

- $k \equiv 0 \pmod 4$. Since $t \equiv 0 \pmod 4$ an appearance of a repetition R in τ_2 implies that there is also a repetition in τ_1, a contradiction.

- $k \equiv 1 \pmod 4$. By Property 1, the term $\xi_{k+1} = \xi_{k+t+1}$ of R is a or b and the previous term $\xi_k = \xi_{k+t}$ is the other symbol of a or b. Therefore there is a repetition R' in τ_2 of length $2t$ starting with the term ξ_k. According to the former case we obtain a contradiction.

- $k \equiv 2 \pmod 4$. By Property 2, we have that $\xi_{k+1}\xi_{k+2} = \xi_{k+2t+1}\xi_{k+2t+2}$, since $2t \equiv 0 \pmod 8$. Hence, the subsequence $\xi_{k+3} \dots \xi_{k+2t+2}$ is also a repetition of length $2t$ in τ_2, a contradiction.

- $k \equiv 3 \pmod 4$. As in the previous case, we obtain a repetition of length $2t$ starting with the term ξ_{k+2}, a contradiction.

It follows that τ_2 is a Thue sequence.

To show that the sequence τ_2 is 2-Thue it remains to prove that every 2-subsequence is nonrepetitive. Consider a subsequence τ_{odd} (τ_{even}) formed by the terms at odd (even) positions in τ_2. Clearly, there is no repetition of length 2 and 4 in τ_{odd}. Now suppose that there is a repetition $R_{odd} = \{\zeta_n\}_{k+1}^{k+t}$ in τ_{odd} of length $t, t \geqslant 6$. Notice that the terms at even positions are uniquely determined by Property 1. Therefore if there is a repetition in τ_{odd}, there is also a repetition $R_{even} = \{\zeta_n\}_{k+2}^{k+t+1}$ in τ_{even}. Consequently, these facts lead to a contradiction, since having a unique determination of the terms at odd and also at even positions, implies that there is also a repetition in the whole sequence τ_2 of length $2t$ starting at the position $k+1$. This establishes the theorem. $\square$

Now, we are ready to prove the main theorem of this section.

Theorem 13. *For every $\ell \geqslant 4, \ell \notin \{5, 7, 11\}$, there exists a cyclic 2-Thue sequence of length ℓ constructed by four symbols.*

Proof. In the proof, we proceed as follows. First, we present a construction of sequences of specified lengths obtained from subsequences of the sequence τ_2 described above and show that such sequences are cyclic 2-Thue. Then we show that we can construct a cyclic 2-Thue sequence of any length.

By Lemma 9, there exists a subsequence $S = \xi_1 \dots \xi_\ell$ of τ_2 of length ℓ starting and ending with equal blocks $abdc$ for every $\ell = 8t + 4$, with $t \geqslant 1$. Moreover we can assume that $\xi_5\xi_6\xi_7\xi_8 = bacd$ and by Property 2 we have $\xi_{\ell-5}\xi_{\ell-4} = cd$. Let S^c be the sequence obtained by replacing the first term a in S with the symbol c, and the last term c with the symbol a (see Figure 3). Now, we show that S^c is cyclic 2-Thue. Suppose the contrary, that there is a repetition R in some cyclic j-subsequence of S^c, for $j \in \{1, 2\}$. It is easy to verify that the subsequence $S^c(\ell - 3, 4)$ is 2-Thue. Furthermore, observe that R is not

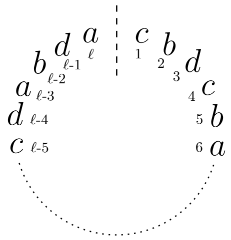

Figure 3: A cyclic 2-Thue sequence S^c obtained by replacing the first and the last term of S.

within any j-subsequence of $S^c(2, \ell-1)$, since its terms are the same as the terms at the same positions in S, which is 2-Thue. Hence, R contains at least one of the terms at indices 1 and ℓ and some of the terms of $S^c(5, \ell-4)$.

When $j = 1$, at least one of the subsequences $S^c(1, 5)$ and $S^c(\ell-4, \ell)$ is contained in R. In the former case, R contains a block cbd, which is unique in S^c by Property 1; in the latter case, there is a unique block bda, a contradiction. If $j = 2$, then the repetition R contains the terms of at least one pair of indices $\ell-2$ and ℓ, $\ell-1$ and 1, ℓ and 2, or 1 and 3. Similarly as above, these terms represent unique blocks ba, dc, ab or cd, a contradiction implying that S^c is cyclic 2-Thue. Notice that in both cases we assume that the unique blocks appear in one repetition block of R, since otherwise, if a unique block, say ab, is divided and appears in both repetition blocks of R, then another block, in this case ba, appears entirely in one repetition block of R.

In what follows, we show that we can modify the above sequences such that we obtain cyclic 2-Thue sequences of arbitrary lengths. Consider a cyclic sequence S^c of length $\ell = 8t + 4$ for some t constructed as above. We extend the sequence S^c by appending a block of symbols A_k from Fig. 4, obtaining a sequence S^c_k, for $k \in \{4, 9, 10, 11, 13, 14, 15\}$. Denote the indices of the terms of S^c_k corresponding to the terms of A_k by $\ell+1, \ldots, \ell+k$.

k	A_k
0	$\emptyset$
4	$cdba$
9	$cbadbcdba$
10	$cbadcabcda$
11	$cbadcabcdba$
13	$cdbacbadbcdba$
14	$cbadbcadcabcda$
15	$cbadbcadcabcdba$

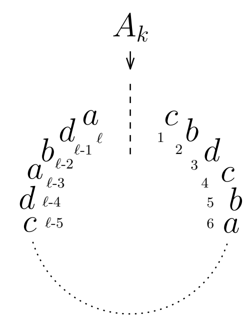

Figure 4: A list of blocks of symbols to be appended at the end of S^c.

We show that a sequence S^c_k is cyclic 2-Thue, for every k. Similarly as in the case above, suppose that there is a repetition R_k in some cyclic j-subsequence of S^c_k, for $j \in \{1, 2\}$.

As above, there is no repetition in any j-subsequence of $S_k^c(2,\ell-1)$ and, since the blocks $cdabdaA_kcbdcba$ are 2-Thue for every k, R_k is not contained in any j-subsequence of $S_k^c(\ell-5,6)$ (which is of length $k+12$). Now, we consider two cases regarding the value of j.

Suppose that R_k is a subsequence of some 1-subsequence of S_k^c. Then, by the observation above, it must contain the terms of $S_k^c(1,3)$ (i.e. cbd) or of $S_k^c(\ell-2,\ell)$ (i.e. bda). First, assume that at least one of the two blocks is contained in one repetition block of R_k. By Property 1 and the choice of A_k's, both blocks are unique in all the sequences S_k^c, a contradiction.

Now, we may assume that none of the blocks cbd and bda is contained in one repetition block of R_k. Hence only one of them is completely contained in R_k. This implies that one repetition block is entirely within $S(\ell-1,2)$, while the second is entirely within $S(2,\ell-1)$. By Properties 1 and 2, the former and the latter block are not the same, a contradiction implying that all the sequences S_k^c are cyclic Thue.

We may thus assume that R_k is a subsequence of some cyclic 2-subsequence of S_k^c. By the same argumentation as above, at least one of the four triples of indices $(1,3,5),(\ell-4,\ell-2,\ell),(\ell-3,\ell-1,\ell+1)$ or $(\ell+k,2,4)$ appears in R_k. Recall that when k is odd (then also $\ell+k$ is odd), a cyclic 2-subsequence might be of length $\ell+k$, whereas in cyclic sequences of even lengths there are two disjoint maximal cyclic 2-subsequences.

As in the previous case, the triples of indices represent unique blocks cdb, dba, adc and abc, which means that R_k contains entirely exactly one of them, and clearly not in one repetition block. This implies that one repetition block is entirely within $S(\ell-2,3)$, while the second is entirely within $S(2,\ell-1)$. Notice that from this it follows that in R_k only the terms with the same parities of indices in S_k^c appear. Consider now how the four blocks can be divided between the two repetition blocks of R_k.

Each of the four blocks can be divided in two ways; we consider them separately. First, if the block cdb (resp. dba, adc, abc) is divided so that cd (resp. ba, dc, ab) appears in one repetition block, the same block cannot appear in the other repetition block by Property 1. Hence, we may assume that cdb (resp. dba, adc, abc) is divided so that db (resp. db, ad, bc) appears in one repetition block. Notice that db does not appear in any 2-subsequence of $S(\ell-2,3)$, while ad appears twice but at the indices of different parities. Only bc appears also at the indices of the same parities (in A_{13}), but it is followed by different symbols as bc at the indices 2 and 4. Hence, we may conclude that R_k does not exist and that S_k^c is indeed cyclic 2-Thue for every k.

By Lemma 7, there are no cyclic 2-Thue sequences on four symbols of lengths 5, 7 and 11. The sequences of the remaining lengths ℓ that cannot be obtained by our construction, i.e. $\ell = 8t+4+k$, for $t \geqslant 1$ and $k \in \{0,4,9,10,11,13,14,15\}$, are presented in Table 1. This completes the proof. $\square$

ℓ	S_ℓ
4	*abcd*
6	*cdbcda*
8	*cbadbcda*
9	*abcdbacbd*
10	*cbadcabcda*
13	*abcabdcbadbcd*
14	*abcabdacbadbcd*
15	*abcabdacbdcbacd*
17	*abcabdacbdcbadbcd*
18	*abcabdabcdacbadbcd*
19	*abcabdacbadbcabdcbd*

Table 1: Cyclic 2-Thue sequences of lengths that cannot be obtained by the above procedure.

4 On nonrepetitive colorings of graphs

The notion of repetitive sequences has been strongly adopted by the graph theory. It seems that Currie [5, 8] was the first to consider nonrepetitive sequences in graphs. In 2002, Alon et al. [1] generalized the problem to edge-colorings of graphs. A coloring of the edges of a graph G is called *nonrepetitive* if the sequence of colors on any path in G is nonrepetitive. The authors denote the minimal number of colors needed for such a coloring the *Thue number* of G.

That work initiated the study of a number of colorings of graphs with analogous constraints; together with already mentioned edge colorings and classical colorings of vertices, both colorings were defined also with restriction to faces and named *facial nonrepetitive* edge or vertex coloring regarding nonrepetitive sequences of colors of consecutive vertices or edges on the boundary of any face in plane graphs (see [19] and [2]). For more results on graph colorings we refer the reader to an early survey of Grytczuk [15]. Additionally, a list version of this problem was studied for paths [18], trees [14], and plane graphs [23, 24, 26] where in the latter two the entropy compression method was used.

Recently, Schreyer and Škrabuľáková [25] introduced a new type of a nonrepetitive coloring of graphs, a sort of total colorings: a weak and a (strong) total Thue coloring. A *weak total Thue coloring* of a graph G is a coloring of vertices and edges such that the color sequences of consecutive vertices and edges of every path in G are nonrepetitive. For a *(strong) total Thue coloring* of G an additional condition must be satisfied: the induced vertex-coloring and edge-coloring of G are also nonrepetitive. The minimum number of colors required in these colorings are called the *weak total Thue chromatic number* and the *(strong) total Thue chromatic number* of a graph G, respectively, and denoted by $\pi_{\tau_w}(G)$ and $\pi_\tau(G)$, respectively. Among other results, they showed that the total Thue chromatic number of a graph is at most $15\Delta^2(G)$ and the bound increases to $18\Delta^2(G)$ in the list version of the problem.

Note that an immediate corollary of the result of Currie and Simpson [11] is the fact that the total Thue chromatic number of any path of length at least 4 is equal to 4. In [25], the bounds for cycles were also presented.

Proposition 14 ([25])**.** *For every cycle C on at least 4 vertices it holds $4 \leqslant \pi_\tau(C) \leqslant 6$.*

As a Corollary of Theorem 13, we obtain also the tight bounds for total Thue coloring of cycles, improving Proposition 14.

Corollary 15. *For every cycle C_n it holds $\pi_\tau(C_n) = 4$.*

Notice that we need a sequence of length $2n$ to induce a coloring of the vertices and edges of a cycle on n vertices.

5 Discussion

In this paper we presented a construction of k-Thue sequences using $2k$ symbols. We believe that our method could be improved to obtain better or even tight upper bounds. However, Conjecture 1 is still wide open for $k = 4$ and $k \geqslant 6$. While there was some work already done on the upper bounds for constructing linear sequences, the results in Section 3 seem to be the first in this area. For cyclic k-Thue sequences we sometimes need an additional symbol already in the cases with $k = 1$ and $k = 2$. Even more, for a cyclic k-Thue sequence of length $2k+1$, we always need $2k+1$ symbols. On the other hand, from the construction used in proof of Theorem 3 one can deduce that for the lengths $\ell = 2kn$, cyclic k-Thue sequences on $2k$ symbols exist: instead of an arbitrary Thue sequence τ_1, we just take a cyclic Thue sequence on three symbols of length n. However, if the length is big enough, we conjecture the following.

Conjecture 16. For every k there exists an integer ℓ_k such that at most $k + 2$ symbols are needed to construct a cyclic k-Thue sequence of length $\ell \geqslant \ell_k$.

If the conjecture is true, what are the lengths, for which $k + 2$ symbols do not suffice?

Another interesting problem in this area is study of growth properties of Thue sequences (see [27] for a survey). An analogous question can be asked for k-Thue sequences.

Question 17. What is the number of k-Thue sequences of length n on $k + 2$ symbols?

In a k-Thue sequence we forbid repetitions in all j-subsequences, for $1 \leqslant j \leqslant k$. One can naturally define the following weaker version. A *weak k-Thue sequence* is a Thue sequence in which every k-subsequence is also Thue. Intuitively, the condition is at most as restrictive as the condition for 2-Thue sequences, i.e. there is at least as many weak k-Thue sequences as 2-Thue sequences on four symbols (note that every weak 2-Thue sequence is also 2-Thue). Thus one may expect that, for k big enough, three symbols suffice to construct a weak k-Thue sequence. Computer search shows that there are no weak k-Thue sequences of arbitrary length on three symbols for $k \in \{2, \dots, 12, 14, 15, 16, 20, 22\}$.

Question 18. Is there a positive integer k large enough such that three symbols suffice to construct an infinite weak k-Thue sequence?

Grytczuk et al. [17] considered the following problem concerning infinite sets of forbidden differences. Let K be a fixed (possibly infinite) set of positive integers. A K-nonrepetitive coloring is a coloring of the integers in which every arithmetic progression with common difference in K forms a nonrepetitive sequence. They denoted by $\pi(K)$ the minimum number of colors needed for K-nonrepetitive coloring of $\mathbb{Z}$. Clearly, by Theorem 3, $\pi(K)$ is at most $2k$ for finite sets K, where k is the maximum element of K. For infinite sets the authors gave an Erdős type conjecture that $\pi(K)$ is finite also for every lacunary set K, where a set $K = \{k_1 < k_2 < \dots\}$ is lacunary if there is a real number $\delta > 0$ such that $\frac{k_{i+1}}{k_i} > 1 + \delta$ for all indices i.

Notice that the notion of weak k-Thue sequences is just a special case of the above; namely, taken $K = \{1, k\}$ the problems are equivalent.

Acknowledgements

This work is partially supported by ARRS Program P1-0383 and by Creative Core FISNM-3330-13-500033. Additionally, the third author was financed by the project NEXLIZ - CZ.1.07/2.3.00/30.0038, which is co-financed by the European Social Fund and the state budget of the Czech Republic, and the fourth author was supported in part by Slovak VEGA Grant No. 1/0652/12, by the Slovak Science and Technology Assistance Agency under the contract APVV-0023-10, and the Project implementation: University Science Park TECHNICOM for Innovation Applications Supported by Knowledge Technology, ITMS: 26220220182, supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] N. Alon, J. Grytczuk, M. Haluszczak and O. Riordan. Non-repetitive colorings of graphs, *Random Struct. Algor.*, 21:336–346, 2002.

[2] J. Barát, J. Czap. Facial nonrepetitive vertex coloring of plane graphs, *J. Graph Theory*, 77:115–121, 2013.

[3] J. Berstel. Axel Thue's papers on repetitions in words: a translation, *Publications du LCIM*, volume 20, Université du Quebec à Monréal, 1994.

[4] J. Berstel, D. Perrin. The origins of combinatorics on words, *Europ. J. Combin.*, 28:996–1022, 2007.

[5] J. D. Currie. Non-repetitive Walks in Graphs and Digraphs, Ph.D. Thesis, University of Calgary, Alberta, Canada, 1987.

[6] J. D. Currie. Pattern avoidance: themes and variantions, *Theoret. Comput. Sci.*, 339:7–18, 2005.

[7] J. D. Currie. There are ternary circular square-free words of length n for $n \geqslant 18$, *Electron. J. Combin.*, 9:#N10, 2002.

[8] J. D. Currie. Which graphs allow infinite nonrepetitive walks?, *Discrete Math.*, 87:249–260, 1991.

[9] J. D. Currie, E. Moodie. A word on 7 letters which is non-repetitive up to mod 5, *Acta Inform.*, 39:451–468, 2003.

[10] J. D. Currie, C. W. Pierce. The fixing block method in combinatorics on words, *Combinatorica*, 23:571–584, 2003.

[11] J. D. Currie, J. Simpson. Non-repetitive tilings, *Electron. J. Combin.*, 9:#R28, 2002.

[12] F. M. Dekking. Strongly non-repetitive sequences and progression-free sets, *J. Combin. Theory Ser. A*, 27:181–185, 1979.

[13] A. A. Evdokimov. Strongly asymmetric sequences generated by a finite number of symbols, *Dokl. Akad. Nauk SSSR*, 179:1268–1271, 1968.

[14] F. Fiorenzi, P. Ochem, P. Ossona de Mendez and X. Zhu. Thue choosability of trees, *Discrete Appl. Math.*, 159:2045–2049, 2011.

[15] J. Grytczuk. Nonrepetitive coloring of graphs—A survey, *Int. J. Math. Math. Sci.*, Art. ID 74639:10, 2007.

[16] J. Grytczuk. Thue-like sequences and rainbow arithmetic progressions, *Electron. J. Combin.*, 9:#R44, 2002.

[17] J. Grytczuk, J. Kozik, M. Witkowski. Nonrepetitive sequences on arithmetic progressions, *Electron. J. Combin.*, 18:#R209, 2011.

[18] J. Grytczuk, J. Przybyło and X. Zhu. Nonrepetitive list colourings of paths, *Random. Struct. Algor.*, 38:162–173, 2011.

[19] F. Havet, S. Jendroľ, R. Soták and E. Škrabuľáková. Facial non-repetitive edge-coloring of plane graphs, *J. Graph Theory*, 66:38–48, 2010.

[20] V. Keränen. Abelian squares are avoidable on 4 letters, In W. Kuich (Ed.), *ICALP 1992*, volume 623 of *Lecture Notes in Comput. Sci.*, pages 41–52. Springer, 1992.

[21] M. Lothaire. *Applied Combinatorics on Words*, Cambridge University Press, 2005

[22] P. A. B. Pleasant. Non-repetitive sequences, *Proc. Camb. Philos. Soc.*, 68:267–274, 1970.

[23] J. Przybyło. *On the facial Thue choice index via entropy conpression*, J. Graph Theory **77** (2014), 180–189.

[24] J. Przybyło, J. Schreyer and E. Škrabuľáková. On the facial Thue choice number of plane graphs via entropy compression method, arXiv:1308.5128.

[25] J. Schreyer, E. Škrabuľáková. Total Thue colorings of graphs, *Europ. J. Math.*, 1:186–197, 2015.

[26] J. Schreyer, E. Škrabuľáková. On the facial Thue choice index of plane graphs, *Discrete Math.*, 312:1713–1721, 2012.

[27] A. M. Shur. Growth properties of power-free languages, *Comput. Sci. Review*, 6:187–208, 2012.

[28] A. Thue. Über unendliche Zeichenreichen, *Norske Vid. Selsk. Skr., I Mat. Nat. Kl., Christiana*, 7:1–22, 1906.