Non-repetitive 3-coloring of subdivided graphs

Andrzej Pezarski
Theoretical Computer Science Department
Faculty of Mathematics and Computer
Science, Jagiellonian University
Krakow, Poland
pezarski@tcs.uj.edu.pl

Michał Zmarz
Theoretical Computer Science Department
Faculty of Mathematics and Computer
Science, Jagiellonian University
Krakow, Poland
mrozik@tcs.uj.edu.pl

Submitted: Dec 22, 2008; Accepted: May 12, 2009; Published: May 20, 2009
Mathematics Subject Classification: 05C15

Abstract

We show that every graph can be subdivided in a way that the resulting graph can be colored without repetitions on paths using only 3 colors. This extends the result of Thue asserting the existence of arbitrarily long nonrepetitive strings over a 3-letter alphabet.

1 Introduction

A vertex coloring f of an undirected graph G is *nonrepetitive* if there is no simple path $v_1, \ldots, v_{2r}$ in G such that $f(v_i) = f(v_{r+i})$ for all $i = 1, \ldots, r$. The minimum number of colors needed in such coloring is called *Thue chromatic number* and is denoted by $\pi(G)$. Of course, every nonrepetitive coloring is also a proper coloring, so $\chi(G) \leqslant \pi(G)$.

A subdivision of an undirected graph G is a graph resulting from a (possibly empty) sequence of the edge subdivisions applied to G. A subdivision of an edge $e = \{u, v\}$ of the graph G is obtained by the addition of a new vertex w into G and replacement of e with two new edges $\{u, w\}$ and $\{w, v\}$.

For a graph G we want to find

$$\pi_S(G) = \min\{\pi(F) : F \text{ is a subdivision of } G\}$$

and to effectively determine a subdivision of G that witnesses $\pi_S(G)$.

In 1906 Axel Thue published his paper [Thu06] in which he presented an infinite, *square-free* sequence of digits `0`, `1`, `2`, that is in which no subword is repeated twice in a row. This can be expressed as $\pi(\mathcal{P}) \leqslant 3$ for every path $\mathcal{P}$. It implies easily that $\pi(\mathcal{C}_n) \leqslant 4$ for every cycle $\mathcal{C}_n$. Recently Currie proved [Cur02] that $\pi(\mathcal{C}_n) = 3$ for every $n \geqslant 18$. By combining these two results we obtain $\pi_S(G) \leqslant 3$ for every graph G with maximum degree at most 2. It is also known that $\pi(T) \leqslant 4$ for every tree T, see [BGK+07]. Moreover $\pi_S(T) \leqslant 3$, as proved in [BGK+07].

An effort has been made to bound $\pi_S(G)$ without restricting G to any specific class of graphs. In fact, it is easy to show that $\pi_S(G) \leqslant 5$ for every graph G [Gry07]. This bound was improved to 4 by Barát and Wood [BW08] and independently by Marx and Shaefer [MS09].

In fact, Grytczuk [Gry07] and Barát and Wood [BW08] conjectured that $\pi_S(G) \leqslant 3$ for every graph G. The aim of this paper is to confirm this conjecture. Our proof is based on the original construction of Thue and is similar to the approach by Marx and Schaefer. We give it in Section 3 after a necessary preparation in Section 2.

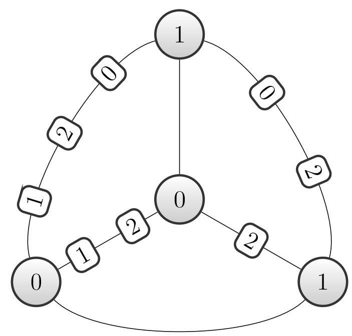

Figure 1: Non-repetitive 3-coloring of a subdivided clique $\mathcal{K}_4$

For other interesting problems on Thue colorings of graphs, see [AGHR02].

2 Preliminaries

First, we introduce some notation that will be used later. Upper left arrow denotes reversion of a string, that is

$$\overleftarrow{a_1 a_2 \ldots a_n} = a_n a_{n-1} \ldots a_1.$$

Upper bar denotes *complement modulo* 2, that is

$$\overline{a_1 a_2 \ldots a_n} = (2 - a_1)(2 - a_2) \ldots (2 - a_n),$$

where $a_i \in \{0, 1, 2\}$. Square brackets are used to extract the substring from a string, i.e., $a_1 a_2 \ldots a_n[i] = a_i$ and $a_1 a_2 \ldots a_n[i \mathinner{..} j] = a_i \ldots a_j$.

The sequence $\gamma = (\gamma_0, \gamma_1, \ldots)$ of Thue-Morse binary strings is defined as follows:

$$\gamma_0 = \mathsf{0}$$

and for $i > 0$, γ_i arises from γ_{i-1} by applying the substitutions

$$\mathsf{0} \to \mathsf{01} \qquad \mathsf{1} \to \mathsf{10}$$

to all digits in γ_{i-1}. The first few strings γ_i are listed in Table 1.

By counting the number of 1's between consecutive 0's in γ_i we obtain consecutive digits of a new string, which we denote β_i. More formally, if c_i^j is the position in γ_i at which the j-th 0 occurs then $\beta_i[j] = c_i^{j+1} - c_i^j - 1$.

The strings β_i were first defined by Thue [Thu06] who showed that each β_i is nonrepetitive and consist only of digits 0, 1, 2. As before, the first few values of β_i are listed in Table 1.

Table 1: γ and β sequences

i	γ_i	β_i
0	0	
1	01	
2	0110	2
3	01101001	210
4	0110100110010110	2102012
5	01101001100101101001011001101001	210201210120210

Lemma 2.1. *No β_i contains* 010 *or* 212 *as a substring.*

Proof. For 010 to appear as a substring of some β_i, 00100 must be a subword of γ_i. As γ_i is constructed from γ_{i-1} by substituting a single digit by a pair of different digits, there is no possibility that two leading and simultaniously two last zeros in 00100 can occur.

Analogously, 212 in β_i gives rise to 011010110 in γ_i which cannot happen by the very same token applied to the first two and the last two 1's. □

Lemma 2.2. *For every natural number m, there is a family of strings $\{\alpha_i : i = 1, 2, ..., m\}$ over $\{0, 1, 2\}$ satisfying the following properties:*

1. *$\alpha_i \neq \alpha_j$, for $i \neq j$*
2. *$\alpha_i \neq \overleftarrow{\alpha_j}$, for all i, j*
3. *every string α_i is nonrepetitive*
4. *all strings have the same length, i.e. $|\alpha_i| = |\alpha_j|$ and moreover $|\alpha_j| = O(m)$*
5. *all strings begin with* 1020 *and end with* 2021 *but no string contains* 010 *nor* 212 *as a subword.*

Proof. Recall, that

$$\beta_5 = 210201210120210, \quad \overline{\beta_5} = 012021012102012.$$

Observe, that

$$\begin{aligned}\beta_6 &= \beta_5 \cdot \mathtt{2} \cdot \overline{\beta_5}\\ \beta_7 &= \beta_5 \cdot \mathtt{2} \cdot \overline{\beta_5} \cdot \mathtt{1} \cdot \overline{\beta_5} \cdot \mathtt{0} \cdot \beta_5\\ \beta_8 &= \beta_5 \cdot \mathtt{2} \cdot \overline{\beta_5} \cdot \mathtt{1} \cdot \overline{\beta_5} \cdot \mathtt{0} \cdot \beta_5 \cdot \mathtt{2} \cdot \overline{\beta_5} \cdot \mathtt{0} \cdot \beta_5 \cdot \mathtt{1} \cdot \beta_5 \cdot \mathtt{2} \cdot \overline{\beta_5}\end{aligned}$$

and more generally, for $i \geqslant 2$ we have

$$\beta_i = \begin{cases} \beta_{i-1} \cdot \mathtt{2} \cdot \overline{\beta_{i-1}}, & \text{if } i \text{ is even,}\\ \beta_{i-1} \cdot \mathtt{1} \cdot \overline{\beta_{i-1}}, & \text{if } i \text{ is odd.}\end{cases}$$

Let $k = 6 + \lceil \log_2 m \rceil$. The string β_k contains at least m (disjoint) instances of β_5 as subwords. Now, let p_i denote the position of the i-th occurrence of β_5 in β_k, which means that $\beta_k [p_i \mathinner{..} p_i + 14] = \beta_5$.

We know that β_{k+2} has length $|\beta_{k+2}| = 4 |\beta_k| + 3$ and contains two instances of β_k shifted by $d = 3 |\beta_k| + 3$ positions. In particular, for $i = 1, 2, ..., m$ we have

$$\beta_{k+2} [p_i + d \mathinner{..} p_i + 14 + d] = \beta_{k+2} [p_i \mathinner{..} p_i + 14] = \beta_5$$

for $i = 1, 2, ..., m$.

Finally, we define our strings by

$$\alpha_i = \beta_{k+2} [p_i + 1 \mathinner{..} p_i + 13 + d] \qquad i = 1, 2, ..., m$$

Clearly, this definition satisfies (3) – (5). To see (1), assume to the contrary that $\alpha_i = \alpha_j$ for some $i < j$. Since the α-strings have length $d + 13 > \frac{|\beta_{k+2}|}{2}$, any two of them have to overlap so that this overlap is realized by

$$\alpha_i = u \cdot v \quad \alpha_j = v \cdot w,$$

where $u \cdot v \cdot w$ is a substring of β_{k+2}. But

$$u \cdot v \cdot w = u \cdot \alpha_j = u \cdot \alpha_i = u \cdot u \cdot v$$

has a repetition, which cannot appear in β_{k+2}.

Finally, (2) follows directly from (5). □

3 Main result

Theorem 3.1. *Every graph has a subdivision that can be nonrepetitively colored using 3 colors.*

Proof. For a graph $G = (V, E)$, let $m = 2 |E|$ and $\alpha_1, \alpha_2, ..., \alpha_m$ be the α-strings satisfying the conditions of Lemma 2.2. Subdivide every edge of G by placing $2 |\alpha_i| + 9$ new vertices on it and denote the resulting graph by $G' = (V', E')$.

Now we color G'. Assign the color 1 to all vertices from V. Color newly added vertices on the i-th edge from E in such a way that reading them in order of their placement gives a string

$$e_i = \mathtt{20} \cdot \alpha_{2i-1} \cdot \mathtt{20102} \cdot \overleftarrow{\alpha_{2i}} \cdot \mathtt{02}.$$

The numbering and orientation of edges from E does not matter.

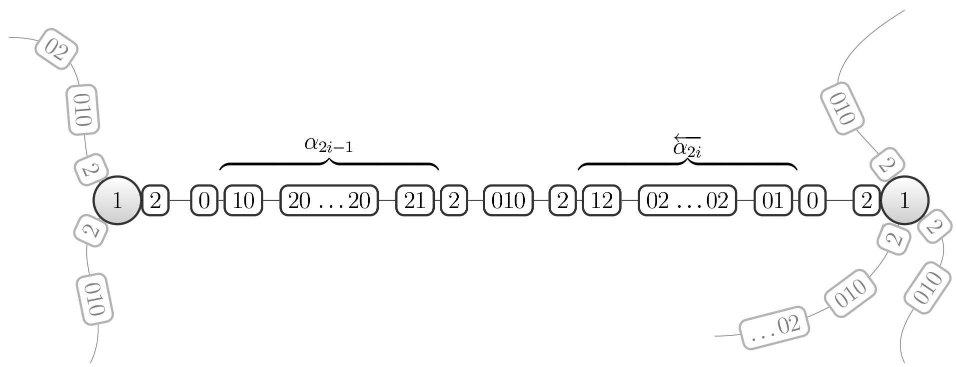

Figure 2: Coloring of the subdivision

Take any simple path $\mathcal{P}$ from the subdivided graph G' and let P be the word consisting the of the labels of the vertices on $\mathcal{P}$. Our aim is to show that P cannot be of the form $P = w \cdot w$ for any word w.

The string P may contain one or more occurences of `010` or `212`. Those substrings of P will play a crucial role in our proof by serving as *tags.* By Lemma 2.2.(5) α-strings contain no tags. Therefore, every two consecutive occurences of tags in P are either adjacent to each other or are separated by an occurence of $\alpha_i[3\ ..\ |\alpha_i| - 2]$ or $\overleftarrow{\alpha_i[3\ ..\ |\alpha_i| - 2]}$, for some i.

Recall that every α_i was used only once to color vertices from G' and that $\mathcal{P}$ is a simple path. This, together with Lemma 2.2.(5) leads to the following conclusion:

Fact 3.2. *Every α_i may occur at most once as a subword in P or in $\overleftarrow{P}$.*

Now we consider how long $P = w \cdot w$ might be. If $\mathcal{P}$ contains at least two vertices from V then P has to contain (at least) two instances of some α-strings, with one of them being reversed. One of those two instances must be located entirely in w, so that for some i the word $P = w \cdot w$ contains α_i twice or $\overleftarrow{\alpha_i}$ twice. In particular, $\mathcal{P}$ contains two different fragments colored by α_i (or $\overleftarrow{\alpha_i}$), contrary to Fact 3.2. Therefore $\mathcal{P}$ contains at most one vertex from V.

Hence P has to be a subword of the word:

$$M = \mathtt{20} \cdot \alpha_a \cdot \mathtt{20102} \cdot \overleftarrow{\alpha_b} \cdot \mathtt{02120} \cdot \alpha_c \cdot \mathtt{20102} \cdot \overleftarrow{\alpha_d} \cdot \mathtt{02}$$

that results from painting two consecutive edges from E by $\mathtt{20} \cdot \alpha_a \cdot \mathtt{20102} \cdot \overleftarrow{\alpha_b} \cdot \mathtt{02}$ and $\mathtt{20} \cdot \alpha_c \cdot \mathtt{20102} \cdot \overleftarrow{\alpha_d} \cdot \mathtt{02}$ with $|a - b| = 1$ and $|c - d| = 1$. Looking more closely how α-strings start and end we see that

$$
\begin{gathered}
M = \mathtt{2} \cdot \mathtt{0}\overbrace{\mathtt{10} \cdot \mathtt{20} \cdot \ldots \cdot \mathtt{20} \cdot \mathtt{21}}^{\alpha_a}\mathtt{2} \cdot \mathtt{010} \cdot \mathtt{2}\overbrace{\mathtt{12} \cdot \mathtt{02} \cdot \ldots \cdot \mathtt{02} \cdot \mathtt{01}}^{\overleftarrow{\alpha_b}}\mathtt{0} \cdot \\
\mathtt{212} \cdot \\
\mathtt{0}\underbrace{\mathtt{10} \cdot \mathtt{20} \cdot \ldots \cdot \mathtt{20} \cdot \mathtt{21}}_{\alpha_c}\mathtt{2} \cdot \mathtt{010} \cdot \mathtt{2}\underbrace{\mathtt{12} \cdot \mathtt{02} \cdot \ldots \cdot \mathtt{02} \cdot \mathtt{01}}_{\overleftarrow{\alpha_d}}\mathtt{0} \cdot \mathtt{2}.
\end{gathered}
\tag{$*$}
$$

The rest of the argument splits into cases depending on the number of tags appearing in P.

If $P = w \cdot w$ does not contain any tag then it must be entirely contained in one of: 201, α_a, 1201, 1021, $\overleftarrow{\alpha_b}$, 1021, 1201, α_c, 1201, 1021, $\overleftarrow{\alpha_d}$, 102. Since none of them contain a repetition, we are done.

Now suppose that P contains exactly one tag. Of course this tag must start in the first and finish in the second half of $P = w \cdot w$, having something in common with both w's in P. There are no tags in α-strings and each tag from M is adjacent to some other tag, with the exception of the first tag that occurs very closely (at distance 1) to the beginning of M and the last tag that occurs very closely to the end of M. Therefore, P cannot be too long, namely $|P| \leqslant 8$. By inspecting possible positions of our tag in P and how P lies within M, we are assured that P cannot be a repetition. For this inspection we simply analyze the digits presented in $(*)$.

Finally, in the rest of the proof we assume that there are at least two tags in P. To analyze how the tags can occur in P we fix their positions in M, which are simply defined to be the positions of their first digits. Let $s = |\alpha_a| = |\alpha_b| = |\alpha_c| = |\alpha_d|$. Note that M contains 11 tags and their positions are: $t_1 = 2$, $t_2 = s+1$, $t_3 = s+4$, $t_4 = s+7$, $t_5 = 2s+6$, $t_6 = 2s+9$, $t_7 = 2s+12$, $t_8 = 3s+11$, $t_9 = 3s+14$, $t_{10} = 3s+17$, $t_{11} = 4s+16$. Note that for odd i, t_i denotes a position of a 010 tag, while t_i with even i denotes a position of a 212 tag.

Since $P = w \cdot w$ contains at least two tags, we know that at least one of them is fully contained in w. We focus on the first tag T that occurs in P, so that after embedding P into M we know that $M[t_i \mathinner{..} t_i + 2] = T$ and $M[t_i + |w| \mathinner{..} t_i + |w| + 2] = T$, for some i. The minimum distance between tags in M of the same type is 6, therefore $|w| \geqslant 6$.

Suppose $|w| = 6$. There are 3 pairs of tags in M that are of the same type and in each pair tags are placed at distance 6. Their positions are: (t_2, t_4), (t_5, t_7), (t_8, t_{10}), so that for some $i \in \{2, 5, 8\}$, $w \cdot w$ contains $M[t_i \mathinner{..} t_{i+2} + 2]$ and at least two additional digits $M[t_i - 2 \mathinner{..} t_i - 1]$ or $M[t_{i+2} + 3 \mathinner{..} t_{i+2} + 4]$. However these two additional digits do not match digits in the middle tag $M[t_i \mathinner{..} t_{i+1} + 2]$. Again, it is sufficient to analyze only digits presented in $(*)$.

Hence $|w| > 6$ and referring to the structure of M we observe that the length of w must be even greater than s. Since $|w| > s > 12$, w has to contain at least 5 digits either to the left or to the right of tag T. Again by inspecting M described by $(*)$ we see that the only pair of matching tags that occur in the first and the second w of $P = w \cdot w$ consist of tags $M[t_i \mathinner{..} t_i + 2]$ and $M[t_{i+6} \mathinner{..} t_{i+6} + 2]$. For each such pair $|w| = 2s + 10$, so that $|P| = 4s + 20$. But this is more than $|M| = 4s + 19$. This means that P cannot be nested inside M and ends the proof. $\square$

4 Conclusion

It is easy to see that if a graph contains $\mathcal{P}_4$ – a path consisting of four vertices, then it cannot be nonrepetitively colored by 2 colors. It is also obvious that every graph that contains a cycle requires at least 3 colors in such coloring.

This leads to the formula that gives the exact value of $\pi_S(G)$ for any undirected graph $G = (V, E)$

$$\pi_S(G) = \begin{cases} 1, & \text{if } E = \emptyset, \\ 2, & \text{if } E \neq \emptyset \text{ and } G \text{ is a star forest,} \\ 3, & \text{otherwise.} \end{cases}$$

By a *star forest* we mean any acyclic graph that does not contain $\mathcal{P}_4$.

Acknowledgement

We would like to thank Jarosław Grytczuk and Paweł Idziak for their useful comments, suggestions and help in writing this article.

References

[AGHR02] Noga Alon, Jarosław Grytczuk, Mariusz Hałuszczak, and Oliver Riordan. Nonrepetitive colorings of graphs. *Random Struct. Algorithms*, 21(3-4):336–346, 2002.

[BGK+07] B. Brešar, J. Grytczuk, S. Klavžar, S. Niwczyk, and I. Peterin. Nonrepetitive colorings of trees. *Discrete Mathematics*, 307(2):163 – 172, 2007.

[BW08] János Barát and David R. Wood. Notes on Nonrepetitive Graph Colouring. *Electron. J. Combin.*, 15(1):#R99, 2008.

[Cur02] James D. Currie. There are ternary circular square-free words of length n for $n \geq 18$. *Electron. J. Combin.*, 9(1):#N10, 2002.

[Gry07] Jarosław Grytczuk. Nonrepetitive colorings of graphs—a survey. *Int. J. Math. Math. Sci.*, 2007:74639, 2007.

[MS09] Dániel Marx and Marcus Schaefer. The Complexity of Nonrepetitive Coloring. *Discrete Applied Mathematics*, 157(1):13 – 18, 2009.

[Thu06] Axel Thue. Über unendliche Zeichenreihen. *Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. Christiana 1 - 22*, 1906.