The Complexity of Nonrepetitive Coloring

Dániel Marx*
Department of Computer Science and Information Theory
Budapest University of Technology and Econonomics
Budapest H-1521, Hungary
dmarx@cs.bme.hu

Marcus Schaefer
Department of Computer Science
DePaul University
Chicago, IL 60604
mschaefer@cs.depaul.edu

Abstract

A coloring of a graph is *nonrepetitive* if the graph contains no path that has a color pattern of the form xx (where x is a sequence of colors). We show that determining whether a particular coloring of a graph is nonrepetitive is coNP-hard, even if the number of colors is limited to four. The problem becomes fixed-parameter tractable, if we only exclude colorings xx up to a fixed length k of x.

1 Squares and Nonrepetitive Colorings

In 1906 Axel Thue published his paper "Über unendliche Zeichenreihen" which showed the remarkable result that there is an infinite word over the alphabet $\Sigma = \{0, 1, 2\}$ that does not contain a *square*, namely a subword of the form xx:

$$010210120102120210120102101202 12\ldots$$ [1]

Remarkable, because over a binary alphabet there are only six square-free words: 0, 1, 01, 10, 010, 101. Remarkable also, because it is a rare

*Research partially supported by the Magyary Zoltán Felsőoktatási Közalapítvány and the Hungarian National Research Fund (Grant Number OTKA 67651).

[1]For the construction see, for example, Lothaire's book [11].

instance of a pattern avoidance theorem: a counter-example to Ramsey theory published when Ramsey was three years old. Thue's result points in two directions: the study of patterns in words and the study of repetition. Combinatorics on words has become an active research field, not least through its importance to computer science [11, 12, 13]. In this paper we want to follow the second direction studying repetition in structures more general than words. There are recent surveys by Grytczuk [8] and Currie [5] on avoiding repetition in various areas of mathematics including graph theory, geometry, and number theory.

One natural variant of words is circular words, that is, words whose last letter is adjacent to its first letter. Currie [5] showed that there are square-free circular words of every length $n \geq 18$ on the alphabet $\{0, 1, 2\}$. Currie's result can be rephrased as saying that the cycle C_n on $n \geq 18$ vertices can be colored using 3 colors so that no subpath of C_n has a coloring of the form xx. We call such a coloring *nonrepetitive.* The coloring point of view was introduced by Alon, Grytczuk, Hałuszczak, and Riordan in a 2002 paper [1], which also contained the definition of the *Thue chromatic number* of a graph, $\pi(G)$, as the smallest number of colors needed in a nonrepetitive coloring of G. In this terminology, Currie proved that $\pi(C_n) = 3$ for $n \geq 18$.

Remark One can distinguish between vertex and edge Thue numbers depending on whether one studies nonrepetitive vertex or edge colorings of graphs. The original paper [1] introduced both variants but emphasized the edge-coloring variant. Subsequent papers seem to have given more attention to the vertex-coloring variant. Here we use the term Thue chromatic number to suggest vertex coloring.

Many problems related to the Thue chromatic number are still open. For example, it is not yet known whether $\pi(G)$ is bounded by some constant for all planar graphs G, a particularly intriguing problem. Kündgen and Pelsmajer [10] showed that graphs of treewidth at most k have Thue chromatic number at most 4^k, settling the special case of outerplanar graphs. Moreover, $\pi(G) \leq 36\Delta^2$, as was shown by Alon, Grytczuk, Hałuszczak, and Riordan [1]. It is also known that every graph has a subdivision whose Thue chromatic number is at most 4 (shown by Grytczuk [9] for 5 and Barát and Wood for 4 [9, 3]).

We look at the Thue chromatic number from the point of view of computational complexity. Deciding whether $\pi(G) \leq k$ is an $\exists\forall$-question: is there a coloring such that no subpath of the graph has a square coloring. Deciding a question of this form belongs to the complexity class $\Sigma_2^{\mathrm{p}} = \mathrm{NP}^{\mathrm{NP}}$, the second level of the polynomial-time hierarchy (see [15] for more information on the polynomial-time hierarchy). We conjecture that the Thue chromatic number problem is complete for that class. As a first result towards settling this conjecture we show in Section 2 that determining whether a *given* coloring of a graph is nonrepetitive is coNP-complete (in other words, deciding whether a coloring is repetitive is NP-complete). Indeed, the problem remains coNP-complete even when restricted to four colors, as we show in Section 3. As an illustration of our technique, we obtain a new proof of the Grytczuk-Barát-Wood result that every graph has a subdivision with Thue chromatic number at most 4.

Deciding whether a given two-coloring of a graph is nonrepetitive (as well as deciding whether a given graph can be nonrepetitively two-colored) is easy, since a two-coloring is nonrepetitive if and only if it is a proper coloring and the graph does not contain a path of length at least 4. This raises the question of how hard it is to determine whether a coloring of a graph with three colors is nonrepetitive. This problem looks difficult; for example, by Currie's result, we can take a word w that is square-free as a circular word of any length $n \geq 18$. Then a path of length $2n$ with coloring ww is not square-free, but we have to look at a block of length n to find this out.

This example suggests studying nonrepetitiveness with restricted block-lengths. Let $\pi_k(G)$ be the smallest number of colors in a coloring of G which does not contain a path of length at most $2k$ with a repetitive coloring. This is a natural parameterization of the problem, $\pi_1(G)$ equals the chromatic number of G, and $\pi_2(G)$ is the *star-chromatic* number of G, introduced by Vince [16].

We complement the result that deciding the nonrepetitiveness of a coloring is coNP-hard by showing how to decide in time $k^{O(k)}n^5 \log n$ whether a coloring of a graph on n vertices contains a path of length at most $2k$ with a repetitive coloring. Using the terminology of parameterized complexity [6, 7], for bounded block-lengths, nonrepetitiveness of a coloring is *fixed-parameter tractable:* the exponent of the polynomial running time does not depend on the parameter k.

Complementing our results, Fedor Manin [14] has recently announced that determining whether a given *edge* coloring of a graph is nonrepetitive is coNP-hard and that deciding the edge Thue number of a graph is Σ_2^{p}-complete. He also established that the edge version of $\pi_k(G)$ is NP-complete for fixed k. Manin's and our results do not seem to imply each other.

2 Nonrepetitiveness of a Coloring

A word x is a *square* if $x = ww$ for some non-empty word w. A word is *nonrepetitive* if it does not contain a square as a subword. A *repetitive sequence* in a graph with a vertex-coloring is a path in the graph whose coloring, as read along the path, is a square. A graph coloring is *nonrepetitive* if it does not contain a repetitive sequence.

Theorem 2.1 *Deciding whether a coloring of a graph is nonrepetitive is* coNP*-complete.*

Proof We reduce from the Hamiltonian Path problem. Let $G = (V, E)$ be a graph with $V = \{v_1, \dots, v_n\}$. We construct a graph H and a coloring that is nonrepetitive if and only if G *does not* have a Hamiltonian path. The graph H consists of two parts. In the first part, for each v_i take a $K_{2,n}$ and color the two element partition using colors a and b, and the n-element partition using colors $c_{i,j}$ (for $1 \leq j \leq n$). Next, for every $i \neq j$ we introduce a new vertex colored $d_{i,j}$ and connect it to the b vertex of the $K_{2,n}$ belonging to v_i and the a vertex belonging to v_j. Also, we connect all the vertices colored b to a new vertex colored c. We construct the second part of H as follows: for each $1 \leq i, j \leq n$, we take a path $P_{i,j}$ on three vertices, coloring the vertices on $P_{i,j}$ by $a, c_{i,j}, b$. We connect the vertex colored by c to the a vertices of the paths $P_{i,1}$ $(1 \leq i \leq n)$. For every $P_{i,j}$ $(1 \leq i \leq n,\ 1 \leq j < n)$ and every edge $v_i v_{i'} \in E$ we add a new vertex of color $d_{i,i'}$ and connect it to the b vertex of $P_{i,j}$ and the a vertex of $P_{i',j+1}$. Finally, we connect all the b-vertices of $P_{i,n}$ to a new vertex colored c $(1 \leq i \leq n)$.

This finishes the construction of H and its coloring (for an example see Figure 1, where G is the diamond, i.e. $K_4 - e$). We claim that G contains a Hamiltonian path if and only if the coloring of H we constructed is repetitive. This implies that deciding the nonrepetitiveness of a graph coloring is coNP-complete.

To prove the claim, let us first assume that G has a Hamiltonian path $v_{\pi(1)}, \dots, v_{\pi(n)}$. Consider the following path through H: we start at the $K_{2,n}$ associated with $v_{\pi(1)}$, traversing it so we see colors $a, c_{\pi(1),1}, b$. We continue via the vertex colored $d_{\pi(1),\pi(2)}$ to the $K_{2,n}$ associated with $v_{\pi(2)}$, traversing it as $a, c_{\pi(2),2}, b$, etc. until we reach the b vertex in the $K_{2,n}$ belonging to $v_{\pi(n)}$. We then continue to the vertex colored c, and traverse the second half of H as follows: $P_{\pi(1),1}$, vertex colored $d_{\pi(1),\pi(2)}$, $P_{\pi(2),2}$, vertex colored $d_{\pi(2),\pi(3)}$, etc. finishing with $P_{\pi(n),n}$ and the vertex colored c. Since $v_{\pi(1)}, \dots, v_{\pi(n)}$ is a Hamiltonian path, this traversal of H is possible, and, comparing the

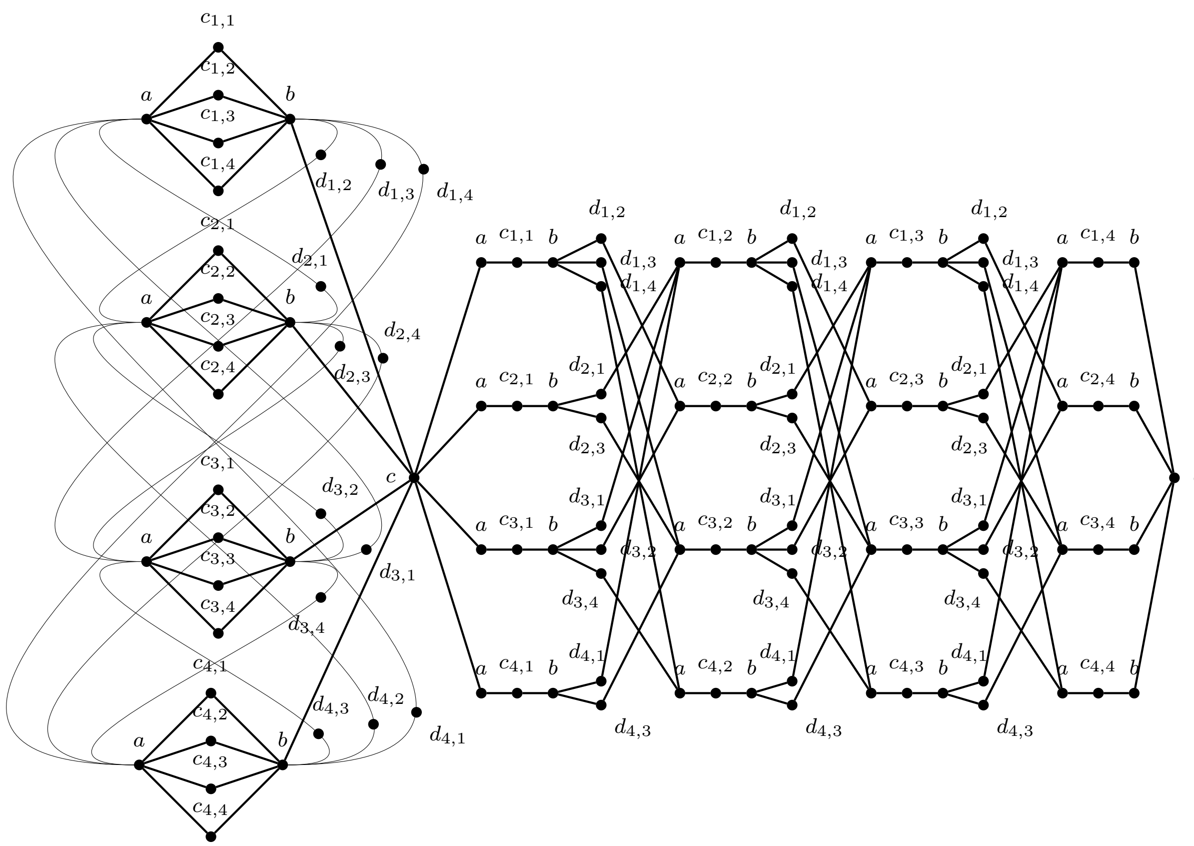

Figure 1: The graph H corresponding to the graph $(\{1,2,3,4\}, \{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{3,4\}\})$.

colors in the two halves of H, we see that they are the same, and, therefore, the coloring is repetitive.

For the reverse direction, assume that H contains a path P such that the colors along P are of the form ww for some word w. Let us first suppose that w does not contain the color c. Then P is entirely contained within the first or the second half of H. In either case we can argue that no repetition is possible, since all the colors except a and b are unique and vertices with colors a and b are not adjacent. We can therefore assume that w contains c. Consequently, P must contain both vertices z, z' colored c (let z be the vertex connecting the two halves). Without loss of generality, we can assume that P starts in the first half of H, and thus there are paths Q, Q', and Q'' such that $P = QzQ'z'Q''$. The first vertex of Q' has color a, while all neighbors of z' have color b, which means that Q'' is empty, and, therefore, $P = QzQ'z'$.

Now any path from z to z' must, for every j, pass through some vertex colored $c_{i,j}$, $1 \leq i \leq n$. In particular Q' must do so and therefore contain at least n vertices colored $c_{i,j}$ for some i, j. Therefore Q must also pass through n vertices colored $c_{i,j}$. Now Q cannot pass through two vertices colored $c_{i,j}$ and $c_{i,j'}$ since they belong to the same $K_{2,n}$ and this would Q to have $c_{i,j}$ or $c_{i,j'}$ as an endpoint, which is not possible as Q' has endpoints colored a and b. Since Q has to pass through n vertices colored $c_{i,j}$ this implies that for every i there is exactly one j such that Q passes through a vertex colored $c_{i,j}$. Since we also argued that for every j there is an i so that Q' and, thereby, Q pass through a vertex colored $c_{i,j}$, there is a permutation π such that $c_{\pi(j),j}$ occurs on Q.

By the construction of the second half of H, $v_{\pi(1)}, \ldots, v_{\pi(n)}$ is a Hamiltonian path of G. $\square$

We note that the proof used an unbounded number of colors to achieve the coding. This can be remedied as we will see in the next section.

3 The Case of 4 Colors

We reduce the number of colors in the construction by replacing colors with long nonrepetitive sequences on a fixed set of colors. As an illustration, we first prove a simple graph-theoretic result.

Proposition 3.1 (Grytczuk, Barát and Wood) *Every graph has a subdivision which can be nonrepetitively colored with at most* 4 *colors.*

Remark Grytczuk [9] proved that every graph has a subdivision which can be colored with at most 5 colors; Barát and Wood improved his result to 4 colors [3]. Our construction is closer in spirit to Grytczuk's original proof.

The following lemma constructs a family of nonrepetitive sequences with useful properties. We write x^R for the reverse of the sequence x.

Lemma 3.2 *We can in polynomial time construct m nonrepetitive sequences of length $O(m)$ on colors* 1*,* 2 *and* 3 *so that*

(i) *for any two sequences x and y, if we split each sequence into two halves of equal length, $x = x_1x_2$ and $y = y_1y_2$, then $x_i \neq y_i$ and $x_i \neq y^R_{3-i}$ (for $i = 1, 2$),* (*ii*) *all sequences begin* 31 *and end* 13*, and*

(*iii*) *all sequences have the same length.*

To see that the lemma is true, take a nonrepetitive sequence x of length $1764m+13$ and permute the colors so it starts with 31. We claim that every subword of 14 letters has to contain the sequences 13 and 31, a claim we will verify later. So if we let x_i be the subword of x that starts with the i-th 31 in x, and ends with the first 13 at least $1176m-1$ positions later, we know that $1176m \leq |x_i| \leq 1176m+13$ and we have fulfilled condition (*ii*). In this fashion we can pick $42m$ sequences x_i from x ($1 \leq i \leq 42m$), since x_{42m} ends no later than position $42m \cdot 14 + 1173m + 13 = 1764m + 13$. Note that any two of these sequences y and z overlap in at least $588m+13$ positions in x, because x_{42m} must contain position $14 \cdot (42m-1)+1 = 588m-13$ and x_1 ends no earlier than position $1176m$, so there is a string of length $588m+13$ common to *all* x_i. Since y and z have length at most $1176m+13$ the overlap of length at least $588m+13$ between them forces their first halves, as well as their second halves to overlap. Therefore, the first halves of y and z must differ from each other, as must the second halves (otherwise, x would contain a square). Among the $42m$ sequences, we can pick $3m$ sequences of the same length, fulfilling condition (*iii*). While it is possible that for two of these sequences y and z, the first half of y equals the reverse of the second half of z, it is not possible that the first half of y equals the reverse of the second half of two other sequences z and z', since in that case the second halves of z and z' would be identical, which we excluded. Similarly, the second half of y can be equal to the reverse of at most one other sequence. Hence we can pick $m = 3m/3$ sequences fulfilling condition (*i*).

We are left with the proof of the claim that any nonrepetitive sequence of length 14 contains the subsequence 13, and, consequently, every other two-digit subsequence. So let x be a nonrepetitive 14-digit string over the alphabet $\{1,2,3\}$. A 1 must occur within the first four digits of x. If that 1 is followed by a 3 we are done, so we know that there is a sequence 12 starting within the first four positions of x. Suppose that sequence continued with a 1, i.e. we see 121. Then the next digit cannot be 2 again, so we have 1213, and, therefore, a 13 within the first seven digits of x. In other words, we know that there is a sequence 123 starting within the first four positions of x. There are two cases: suppose the next digit after 123 is 1, i.e. we have 1231, the next digit has to be 2 (otherwise we have a 13), followed by 1 (since the word is nonrepetitive): 123121. The next digit cannot be a 2, since the word is nonrepetitive, so it has to be a 3 and we are done, since we have found a 13 within the first nine positions of x. In the second case, we have 1232. To avoid repetition, this sequence needs to continue 12321. If the next digit is a 3, we are done, so we can assume we see 123212, which cannot be followed by 1 (repetition), so we have 1232123, which cannot be followed by 2 (repetition), giving us 12321231 followed by 2 (otherwise we have a 13), followed by 1 (repetition), yielding 1232123121. Finally, this string cannot be followed by a 2, so we see 12321231213, which means a 13 within x.

Proof of Proposition 3.1 It is enough to prove the theorem for the case $G = K_n$. Let $(x_i)_{i=1}^m$ be a family of $m = \binom{n}{2}$ nonrepetitive sequences as described in Lemma 3.2. Replace the i-th edge of G with a path of length $|x_i| + 7$ and color it $210x_i012$. Also, give each vertex of G color 0. We claim that this coloring of a subdivision G' of G is nonrepetitive.

Suppose, to the contrary, that G' contains a path P with a coloring of the form ww. P has to contain the color 0, since otherwise ww would be a subword of some x_i which is not possible (as the x_i's are nonrepetitive). There are two types of vertices colored 0: the vertices of G, all of whose neighbors are colored 2, and the vertices introduced in the subdivision, all of whose neighbors are colored 1 and 3. Hence, for a repetition, P must contain two vertices colored 0 of the same type, and that is only possible if P contains a whole path Q between two vertices of G. It is not possible that the coloring of Q is a subword of w, since the colorings of the paths (and their reverses) are unique. Hence, Q must contain the border between the two halves of P. In other words, ww has to contain the following string:

$$0210v0120,$$

where $v = x_i$ for some i (if $v = x_i^R$ we reverse P), and the boundary of ww occurs within v. Assuming that the boundary occurs in the second half of v (the other case being similar), the first half of $0210v0120$ must occur a second time along P; but then the first half of v must occur as the prefix or the reverse of the suffix of some other x_j. This possibility, however, is precluded by choosing sequences x_i fulfilling Lemma 3.2 (i). □

Corollary 3.3 *Deciding whether a coloring of a graph is nonrepetitive is* coNP*-complete even for colorings with at most* 4 *colors.*

Proof We will show how to replace the colors in the graph H constructed in the proof of Theorem 2.1 with just 4 colors. Using Lemma 3.2 we obtain sequences x_i, one for each of the colors c, $c_{k,j}$, and $d_{k,j}$. If a vertex has color $c_{k,j}$ or $d_{k,j}$, and it has been assigned sequence x_i, replace the vertex with a path of length $|x_i| + 7$ and color it $210x_i012$. For the two vertices colored c, we proceed similarly, but in this case the vertex is replaced with a path colored $130x_i031$; call the two paths replacing the c vertices C and C' (where C is the path connecting the two halves of G). Finally, recolor vertices with colors a or b to have color 0. This construction uses colors $0, 1, 2, 3$ only.

We claim that the coloring of the resulting graph will be nonrepetitive if and only if the original graph G did not have a Hamiltonian path.

The proof of one direction remains unchanged: a Hamiltonian path in G still corresponds to a repetitive coloring, since we just replaced colors by color sequences.

Suppose then that G contains a path P colored ww. As we argued earlier, P has to contain the color 0, since otherwise ww would be a subword of some x_i, which is nonrepetitive.

We have four types of vertices colored 0: those with neighbors $1, 3$, those with neighbors $1, 2$, those with two neighbors colored 2 and those with two neighbors colored 3. Let us look at the last type first.

Suppose P does not contain the sequence 303 (which occurs exactly four times: twice on each of the paths replacing c). In that case P cannot traverse C (or C'), and is therefore caught within one of the two halves of G. We claim that this is impossible.

First of all, observe that P does have to contain at least one vertex from C or C', since otherwise we argue as in the proof of Theorem 2.1 that the two halves of the graph obtained by removing C and C' do not contain a square. (That part of the proof of Theorem 2.1 did not use the fact that a and b are different colors.)

Suppose next that P contains exactly one vertex from C and/or exactly one vertex from C'. Such a vertex must be one of the end-vertices of C or C' colored 1. Then P must contain one of the sequences 201 or 102; by changing direction of P if necessary, we can ensure that P contains the sequence 201. We distinguish two cases by whether P lies in the left or right half of G.

If P lies in the left half of G, we use the fact that all occurrences of 201 in that half share the same vertex colored 1 (the end-vertex of C which belongs to P), so 201 can occur at most once along P. Hence, the middle $|$ of P has to occur either at $2|01$ or $20|1$. In the first case the next color after 01 along P must be 0: $2|010$, so P itself must start with 010, however, 010 does not occur anywhere else in the left half of G without using the endpoint of C that is already in use, so this is impossible. In the second case, the next two colors of P after 201 are 02: $20|102$, that is P has to start with the sequence

102. Again, the only other occurrences of 102 in the left half of G use the endpoints of C which is already in use in P, so this is not possible either.

If P lies in the right half, the argument is similar: the occurrence of 201 can be either as 201 or 2|01 or 20|1 with respect to the midpoint of P. In the first case (i.e. the middle of P does not occur within 201), the 201 must be matched as a whole which is only possible by including vertices from C'. Hence the path P needs to contain vertices from both C and C', implying that ww contains a string of the form $210x_i012$. As we argued in Proposition 3.1 this is impossible by the construction of the x_i.

Consequently, P must contain at least two vertices from either C or C'; since we assumed that it does not contain 303, P must begin in C with 310 or 0310 (the last 0 corresponding to an a vertex) and/or end in C' with 013 or 0130 (the initial 0 corresponding to a b vertex). However any two occurrences of any of these sequences in the same half share a vertex, so this is not possible.

We conclude that P must contain the sequence 303. This sequence occurs exactly four times, twice in C and C'. The two occurrences in the same path C or C' cannot match with each other, since one begins $3031x_i$, and the other 30310 (and the x_i's do not contain zeroes). Hence a 303 from C must match with a 303 from C'. But then either all of C or all of C', and therefore both must belong to P.

From this point on, we can argue as in the original proof. □

4 Bounded-Length Sequences

Checking whether a coloring of a graph is nonrepetitive for block-lengths up to some fixed value k can be done in polynomial time: we have to check all the $O(n^{2k})$ paths of length at most $2k$. Here we present an algorithm that is significantly more efficient than brute force: we show that the problem is *fixed-parameter tractable*, i.e., it can be solved in time $O(f(k)n^c)$. This means that the exponent of n does not increase as k increases.

Theorem 4.1 *Given a vertex-colored graph $G(V, E)$, it can be checked in time $k^{O(k)} \cdot |V|^5 \log |V|$ whether G has a repetitive sequence of length $2k$.*

Proof The algorithm is based on color-coding, introduced by Alon et al. [2]. Assign a random label from $\{1, \ldots, 2k\}$ to each vertex of G independently with uniform distribution. Assume that we have a polynomial-time algorithm for checking whether there is a repetitive sequence $v_1, \ldots, v_{2k}$ where vertex v_i has label i (below we will present such an algorithm). If the graph has a repetitive sequence, then the sequence receives the labels 1, $\ldots$, $2k$ with probability $1/(2k)^{2k}$, hence the algorithm finds such a repetitive sequence with probability $1/(2k)^{2k}$. If the graph has no repetitive sequence, then of course no such sequence is found by the algorithm. Therefore, the algorithm produces a correct answer with probability at least $1/(2k)^{2k}$, which can be increased to a constant by repeating the algorithm $(2k)^{2k}$ times. Randomized algorithms based on color-coding can be derandomized using standard techniques, see [2] and [6, Section 8.3].

We still need to show how to check whether there is a repetitive sequence $v_1, \ldots, v_{2k}$ where vertex v_i has label i. For a given labeling $\lambda : V \to \{1, \ldots, 2k\}$ of the vertices, we proceed as follows. For a given vertex x, the algorithm below checks whether there is a repetitive sequence $v_1, \ldots, v_{2k}$ where $\lambda(v_i) = i$ and $v_k = x$. Therefore, the algorithm has to be repeated for every possible choice of x, i.e., $|V|$ times.

We build a directed graph $D(U, A)$ where the U is a subset of $V \times V$. For $v, v' \in V$, the pair (v, v') is a vertex of D only if

- v and v' have the same color in G,
- $\lambda(v') = \lambda(v) + k$,
- if $\lambda(v) = k$, then $v = x$, and
- if $\lambda(v') = k + 1$, then v' is a neighbor of x in G.

There is an arc from (v, v') to (u, u') in D if and only if

- u is a neighbor of v,
- u' is a neighbor of v', and
- $\lambda(u) = \lambda(v) + 1$.

Note that, by the properties of the vertices in D, the last requirement also implies $\lambda(u') = \lambda(v') + 1$.

It is easy to see that D is acyclic, hence the length of the longest directed path can be determined in time $O(|A|)$ using standard techniques. We claim that D has a directed path on k vertices if and only if G has a repetitive sequence on $2k$ vertices. Indeed, if (v_1, v'_1), (v_2, v'_2), $\ldots$, (v_k, v'_k) is a directed path in D, then $v_1, \ldots, v_k, v'_1, \ldots, v'_k$ is a path in G. Notice that the i-th vertex of the path in G has label i, thus a vertex cannot appear twice in the sequence. Furthermore, v_i and v'_i have the same color in G, hence the path is repetitive. The converse statement is also easy to see: if $v_1, \ldots, v_{2k}$

is a repetitive sequence such that $\lambda(v_i) = i$ and $v_k = x$, then the vertices (v_1, v_{k+1}), (v_2, v_{k+2}), $\ldots$, (v_k, v_{2k}) exist in D and they form a directed path.

The directed graph D contains at most $|V|^2$ vertices and hence at most $|V|^4$ edges. Finding the longest path in the acyclic graph D can be done in linear time. The algorithm has to be repeated for every possible vertex x, thus the running time is $|V|^5$ for a given labeling. The derandomization adds a factor $O(\log |V|)$ to the running time. $\square$

The case $k = 2$ is of special interest. Graphs that do not have repetitive sequences of length at most 4 are often called *star-free* or *apathic*. For star-free coloring, the complexity of the coloring problem is settled:

Proposition 4.2 (Coleman, Moré [4]) *Deciding whether a graph has a star-free coloring with three colors is* NP*-complete, even if the graph is bipartite.*

The proof is quite simple: replace each edge of a graph G with three paths of length 2. Then the original graph is 3-colorable, if and only if the resulting (bipartite) graph has a star-free 3-coloring. The result was proved by Coleman and Moré in the context of computing sparse Hessian matrices.

Acknowledgments

We would like to thank Michael Pelsmajer for carefully reviewing and discussing the proofs and making several helpful suggestions.

References

[1] Noga Alon, Jarosław Grytczuk, Mariusz Hałuszczak, and Oliver Riordan. Nonrepetitive colorings of graphs. *Random Structures Algorithms*, 21(3-4):336–346, 2002. Random structures and algorithms (Poznan, 2001).

[2] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *J. Assoc. Comput. Mach.*, 42(4):844–856, 1995.

[3] János Barát and David R. Wood. Notes on nonrepetitive graph colouring, 2005. `http://www.citebase.org/abstract?id=oai:arXiv.org:math/0509608` (accessed February 14th, 2008).

[4] Thomas F. Coleman and Jorge J. Moré. Estimation of sparse Hessian matrices and graph coloring problems. *Math. Programming*, 28(3):243–270, 1984.

[5] James D. Currie. Pattern avoidance: themes and variations. *Theoret. Comput. Sci.*, 339(1):7–18, 2005.

[6] Rod G. Downey and Michael R. Fellows. *Parameterized complexity*. Monographs in Computer Science. Springer-Verlag, New York, 1999.

[7] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory*. Springer-Verlag, Berlin, 2006.

[8] Jarosław Grytczuk. Thue type problems for graphs, points, and numbers. Discrete Math. (to appear).

[9] Jarosław Grytczuk. Nonrepetitive graph coloring. In *Graph theory in Paris*, Trends Math., pages 209–218. Birkhäuser, Basel, 2007.

[10] Andre Kündgen and Michael J. Pelsmajer. Nonrepetitive colorings of graphs of bounded treewidth. Discrete Math., available online 29 September 2007. (http://www.sciencedirect.com/science/article/B6V00-4PSJT4D-2/2/40732ef45641d8565e3f9f3506b5b005).

[11] M. Lothaire. *Combinatorics on words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997.

[12] M. Lothaire. *Algebraic combinatorics on words*, volume 90 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2002.

[13] M. Lothaire. *Applied combinatorics on words*, volume 105 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2005.

[14] Fedor Manin. The complexity of nonrepetitive edge coloring of graphs. `http://arxiv.org/abs/0709.4497` (accessed February 15th, 2008), 2008.

[15] Marcus Schaefer and Chris Umans. Completeness in the polynomial-time hierarchy: Part I: A compendium. *SIGACTN: SIGACT News (ACM Special Interest Group on Automata and Computability Theory)*, 33, 2002.

[16] Andrew Vince. Star chromatic number. *J. Graph Theory*, 12(4):551–559, 1988.