SEMI-TRANSITIVE ORIENTATIONS AND WORD-REPRESENTABLE GRAPHS*

MAGNÚS M. HALLDÓRSSON†, SERGEY KITAEV‡, AND ARTEM PYATKIN§

Abstract. A graph $G = (V, E)$ is a *word-representable graph* if there exists a word W over the alphabet V such that letters x and y alternate in W if and only if $(x, y) \in E$ for each $x \neq y$.

In this paper we give an effective characterization of word-representable graphs in terms of orientations. Namely, we show that a graph is word-representable if and only if it admits a *semi-transitive orientation* defined in the paper. This allows us to prove a number of results about word-representable graphs, in particular showing that the recognition problem is in NP, and that word-representable graphs include all 3-colorable graphs.

We also explore bounds on the size of the word representing the graph. The representation number of G is the minimum k such that G is a representable by a word, where each letter occurs k times; such a k exists for any word-representable graph. We show that the representation number of a word-representable graph on n vertices is at most $2n$, while there exist graphs for which it is $n/2$.

Keywords: graphs, words, orientations, word-representability, complexity, circle graphs, comparability graphs.

1. Introduction

A graph $G = (V, E)$ is *word-representable* if there exists a word W over the alphabet V such that for each pair of distinct letters x and y, $(x, y) \in E$ if and only if the occurrences of the letters alternate in W. As an example, the words *abcdabcd*, *abcddcba*, and *abdacdbc* represent the 4-clique, K_4; 4-independent set, $\overline{K_4}$; and the 4-cycle, C_4, labeled by a, b, c, d in clockwise direction, respectively.

If each letter appears exactly k times in the word, the word is said to be *k-uniform* and the graph is said to be *k-word-representable.* It is known that any word-representable graph is k-word-representable, for some k [12].

* A preliminary version of this work appeared in the International Workshop on Graph-Theoretic Concepts in Computer Science (WG), 2008.

† School of Computer Science, Reykjavik University, 101 Reykjavik, Iceland. Email: mmh@ru.is. Supported by grant-of-excellence 120032011 from the Iceland Research Fund.

‡ Department of Computer and Information Sciences, University of Strathclyde, Glasgow G1 1XH, UK. Email: sergey.kitaev@cis.strath.ac.uk.

§ Sobolev Institute of Mathematics; Novosibirsk State University; Novosibirsk, Russia. Email: artem@math.nsc.ru.

The class of word-representable graphs is rich, and properly contains several important graph classes to be discussed next.

Circle graphs. Circle graphs are those whose vertices can be represented as chords on a circle in such a way that two nodes in the graph are adjacent if and only if the corresponding chords overlap. Assigning a letter to each chord and listing the letters in the order they appear along the circle, one obtains a word where each letter appears twice and two nodes are adjacent if and only if the letter occurrences alternate [4]. Therefore, circle graphs are the same as 2-word-representable graphs.

Comparability graphs. A comparability graph is one that admits a transitive orientation of the edges, i.e. an assignment of directions to the edges such that the adjacency relation of the resulting digraph is transitive: the existence of arcs xy and yz yields that xz is an arc. Such a digraph induces a poset on the set of vertices V. Note that each poset is an intersection of several linear orders and each linear order corresponds to some permutation P_i of V. These permutations can be concatenated to a word of the form $P_1P_2\cdots P_k$. Then two letters alternate in this word if and only if they are in the same order in each permutation (linear order), and this means that they are comparable in the poset and, thus, the corresponding letters are adjacent in the graph. So, comparability graphs are a subclass of word-representable graphs that is known as the class of *permutationally representable graphs* in the literature.

Cover graphs. The (*Hasse*) *diagram* of a partial order $P = (V, <)$ is the directed graph on V with an arc from x to y if $x < y$ and there is no z with $x < z < y$ (in which case x "covers" y). A graph is a *cover graph* if it can be oriented as a diagram of a partial order. Limouzy [14] observed that cover graphs are exactly the triangle-free word-representable graphs.

3-colorable graphs. A corollary of our main structural result in this paper is that the class of word-representable graphs contains all 3-colorable graphs.

Various computational hardness results follow from these inclusions. Most importantly, since it is an NP-hard problem to recognize cover graphs [1, 16], the same holds for word-representable graphs. Also, the NP-hardness of various optimization problems, such as Independent Set, Dominating Set, Graph Coloring, and Clique Partition, follows from the case of 3-colorable graphs.

Our results. The main result of the paper is an alternative characterization of word-representable graphs in terms of orientations.

A directed graph (digraph) $G = (V, E)$ is *semi-transitive* if it is acyclic and for any directed path $v_1v_2 \cdots v_k$, either $v_1v_k \notin E$ or $v_iv_j \in E$ for all $1 \leq i < j \leq k$. Clearly, comparability graphs (i.e. those admitting transitive orientations) are semi-transitive. The main result of this paper is Theorem 3 saying that a graph is word-representable if and only if it admits a semi-transitive orientation.

The proof of the main result shows that any word-representable graph on $n \geq 3$ vertices is $(2n - 4)$-word-representable. This bound implies that the problem of recognizing word-representable graphs is contained in NP. Previously, no polynomial upper bound was known on the *representation number*, the smallest value k such that the given graph is k-word-representable. Our bound on the representation number is tight up to a constant factor, as we construct graphs with representation number $n/2$. We also show that deciding if a word-representable graph is k-word-representable is NP-complete for $3 \leq k \leq n/2$.

One corollary of the structural result is that all 3-colorable graphs are word-representable. This provides a generic reason for word-representability for some of the classes of graphs previously known to be word-representable, e.g. for outerplanar graphs and prisms. On the other hand, there are non-3-colorable graphs that are word-representable (for example, any complete graph on at least four vertices).

A motivating application. Consider a scenario with n recurring tasks with requirements on the alternation of certain pairs of tasks. This captures typical situations in periodic scheduling, where there are recurring *precedence* requirements.

When tasks occur only once, the pairwise requirements form precedence constraints, which are modeled by partial orders. When the orientation of the constraints is omitted, the resulting pairwise constraints form comparability graphs. The focus of this paper is to study the class of undirected graphs induced by the alternation relationship of recurring tasks.

Consider, e.g. the following five tasks that may be involved in operation of a given machine: 1) Initialize controller, 2) Drain excess fluid, 3) Obtain permission from supervisor, 4) Ignite motor, 5) Check oil level. Tasks 1 & 2, 2 & 3, 3 &4, 4 & 5, and 5 & 1 are expected to alternate between all repetitions of the events. This is shown in Fig. 1(b), where each pair of alternating tasks is connected by an edge. One possible task execution sequence that obeys these recurrence constraints – and no other – is shown in Fig. 1(a). Later in the paper, we will introduce an orientation of such graphs that we call a *semi-transitive orientation*; such an orientation for our example is shown in Fig. 1(c).

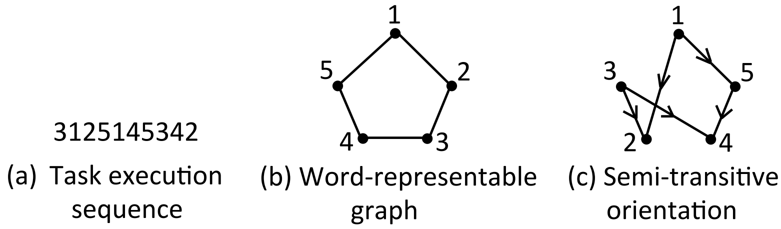

FIGURE 1. The word in (a) corresponds to the word-representable graph in (b). A semi-transitive orientation of the graph is given in (c).

Execution sequences of recurring tasks can be viewed as words over an alphabet V, where V is the set of tasks.

Related work. The notion of directed word-representable graphs was introduced in [13] to obtain asymptotic bounds on the free spectrum of the widely-studied *Perkins semigroup*, which has played central role in semigroup theory since 1960, particularly as a source of examples and counterexamples. In [12], numerous properties of word-representable graphs were derived and several types of word-representable and non-word-representable graphs pinpointed. Some open questions from [12] were resolved recently in [7], including the representability of the Petersen graph.

Circle graphs were generalized to polygon-circle graphs (see [17]), which are the intersection graphs of polygons inscribed in a circle. If we view each polygon as a letter and read the incidences of the polygons on the circle in order, we see that two polygons intersect if and only if there *exists* a pair of occurrences of the two polygons that alternate. This differs from word-representable graphs where *all* occurrences of the two letters must alternate in order for the nodes to be adjacent.

Cyclic (or periodic) scheduling problems have been studied extensively in the operations research literature [9, 15], as well as in the AI literature [5]. These are typically formulated with more general constraints, where, e.g. the 10th occurrence of task A must be preceded by the 5th occurrence of task B. The focus of this work is then on obtaining effective periodic schedules, while maintaining a small cycle time. We are, however, not aware of work on characterizing the graphs formed by the cyclic precedence constraints. A different periodic scheduling application related to word-representable graphs was considered by Graham and Zang [6], whose work involves a counting problem related to the cyclic movements of a robot arm. More generally, given a set of jobs to be performed periodically, certain pairs (a, b) must be done alternately, e.g. since the product of job a is used as a resource for job b. Any valid execution sequence corresponds to a word over the alphabet formed by the jobs. The word-representable graph given by such a word must then contain the constraint pairs as a subgraph.

The preliminary version of this work, that appeared in [8], claimed an upper bound of n on the representation number. This was based on a lemma (Lemma 2 in [8]) that turned out to be false. We present a corrected proof of the main result that gives a slightly weaker bound of $2n-4$ on the representation number.

Organization of the paper. The paper is organized as follows. In Section 2, we give definitions of objects of interest and review some of the known results. In Section 3, we give a characterization of word-representable graphs in terms of orientations and discuss some important corollaries of this fact. In Section 4, we examine the representation number, and show that it is always at most $2n-4$, but can be as much as $n/2$. We explore, in Section 5, which classes of graphs are word-representable, and show, in particular, that 3-colorable graphs are such graphs, but numerous other properties are independent from the property of being word-representable. Finally, we conclude with two open problems in Section 6.

2. Definitions and Properties

Let W be a finite word. If W involves letters $x_1, x_2, \ldots, x_n$ then we write $A(W) = \{x_1, x_2, \ldots, x_n\}$. A word is *$k$-uniform* if each letter appears in it exactly k times. A 1-uniform word is also called a *permutation*. Denote by W_1W_2 the concatenation of words W_1 and W_2. We say that letters x_i and x_j *alternate* in W if the word induced by these two letters contains neither x_ix_i nor x_jx_j as a factor. If a word W contains k copies of a letter x, then we denote these k appearances of x from left to right by $x^1, x^2, \ldots, x^k$. We write $x_i^j < x_k^\ell$ if x_i^j occurs in W before x_k^ℓ, i.e. x_i^j is to the left of x_k^ℓ in W.

We say that a word W *represents* a graph $G = (V, E)$ if there is a bijection $\phi : A(W) \to V$ such that $(\phi(x_i), \phi(x_j)) \in E$ if and only if x_i and x_j alternate in W. We call a graph G *word-representable* if there exists a word W that represents G. It is convenient to identify the vertices of a word-representable graph and the corresponding letters of a word representing it. If G can be represented by a k-uniform word, then we say that G is a *k-word representable* graph and the word *k-represents* G.

The *representation number* of a word-representable graph G is the minimum k such that G is a k-word-representable graph. It follows from [12] that the representation number is well-defined for any word-representable graph. We call a graph *permutationally representable* if it can be represented by a word of the form $P_1P_2\cdots P_k$, where each P_i is a permutation over the same alphabet given by the graph vertices.

A digraph $D = (V, E)$ is *transitive* if the adjacency relation is transitive, i.e. for any vertices $x, y, z \in V$, the existence of the arcs $xy, yz \in E$ yields that $xz \in E$. A *comparability graph* is an undirected graph that admits an orientation of the edges that yields a transitive digraph.

The following properties of word-representable graphs and facts from [12] are useful. A graph G is word-representable if and only if it is k-word representable for some k. If $W = AB$ is k-uniform word representing a graph G, then the word $W' = BA$ also k-represents G.

The wheel W_5 is the smallest non-word-representable graph. Some examples of non-word-representable graphs on 6 and 7 vertices are given in Fig. 2.

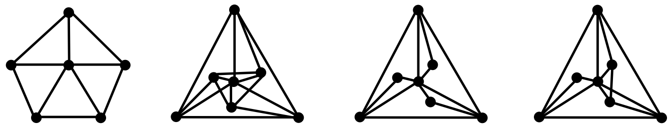

FIGURE 2. Small non-word-representable graphs.

3. Characterizing Word-Representable Graphs in Terms of Orientability

In this section we present a characterization, which implies that word-representability corresponds to a property of a digraph obtained by directing the edges in a certain way. It is known that a graph is permutationally representable if and only if it has a transitive orientation (i.e. is a comparability graph) [13]. We prove a similar fact on word-representable graphs, namely, that a graph is word-representable if and only if it has a certain *semi-transitive orientation* that we shall define.

Let $G = (V, E)$ be a graph. An acyclic orientation of G is *semi-transitive* if for any directed path $v_1 \to v_2 \to \cdots \to v_k$ either

- there is no arc $v_1 \to v_k$, or
- the arc $v_1 \to v_k$ is present and there are arcs $v_i \to v_j$ for all $1 \leq i < j \leq k$. That is, in this case, the (acyclic) subgraph induced by the vertices $v_1, \ldots, v_k$ is a transitive clique (with the unique source v_1 and the unique sink v_k).

We call such an orientation a *semi-transitive orientation.* For example, the orientation of the graph in Figure 3 is semi-transitive.

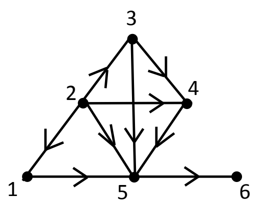

FIGURE 3. An example of a semi-transitive orientation.

A graph $G = (V, E)$ is *semi-transitive* if it admits a semi-transitive orientation.

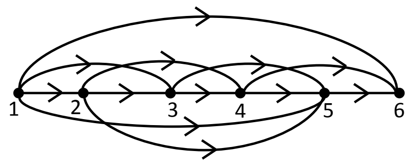

FIGURE 4. An example of a shortcut.

We can alternatively define semi-transitive orientations in terms of induced subgraphs. A *semi-cycle* is the directed acyclic graph obtained by reversing the direction of one arc of a directed cycle. An acyclic digraph is a *shortcut* if it is induced by the vertices of a semi-cycle and contains a pair of non-adjacent vertices. Thus, a digraph on the vertex set $\{v_1, \ldots, v_k\}$ is a shortcut if it contains a directed path $v_1 \to v_2 \to \cdots \to v_k$, the arc $v_1 \to v_k$, and it is missing an arc $v_i \to v_j$ for some $1 \leq i < j \leq k$; in particular, we must have $k \geq 4$, so that any shortcut is on at least four vertices. See Figure 4 for an example of a shortcut (there, the arcs $1 \to 4$, $2 \to 6$, and $3 \to 6$ are missing).

Definition 1. An orientation of a graph is semi-transitive, if it is acyclic and contains no shortcuts.

For a word W, denote by $P(W)$ its *initial permutation*, i.e. the permutation obtained by removal from W all but the first appearances of each letter. Let $D = (V, E)$ be an acyclic orientation of a graph G. For vertices u and v, let $u \rightsquigarrow v$ denote that there exists a directed path from u to v in D. By definition, $u \rightsquigarrow u$. We say that a permutation P of the set V is a *topological sort* of D if for every distinct $u, v \in V$ such that $u \rightsquigarrow v$, the letter u precedes v in P.

We say that a word W over the alphabet V representing a graph H *covers* a set A of non-edges of the digraph D if the following four conditions hold:
1) W is k-uniform for some k;
2) $P(W)$ is a topological sort of D;
3) G is a subgraph of H;
4) each non-edge in A is also a non-edge in H.
We need the following easy to prove lemma.

Lemma 1. *Let words W_1 and W_2 cover, respectively, sets A and B of non-edges of an acyclic digraph $D = (V, E)$. Then the word $W = W_1W_2$ covers the set of non-edges $A \cup B$.*

Proof. Let W_1 be k-uniform and W_2 be ℓ-uniform. Then, clearly, W is $(k+l)$-uniform. $P(W)$ is a topological sort of D since $P(W) = P(W_1)$. If $uv \in A$ then u and v do not alternate in W_1. If $uv \in B$ then u and v do not alternate in W_2. So, for every $uv \in A \cup B$ the letters u and v cannot alternate in W.

Finally, let $uv \in E(G)$. Then without loss of generality, it corresponds to an arc $u \to v$ in the digraph D. Then, by conditions 1)–3), the subwords of W_1 and W_2 induced by the letters u and v are alternating, starting with u and ending with v. But then the same is true for the subword of W induced by these letters, i.e. u and v alternate in W and so, $uv \in E(H)$. □

Now we can prove our main technical lemma.

Lemma 2. *Let $D = (V, E)$ be a semi-transitively oriented graph and $v \in V$. Then the non-edges incident with v can be covered by a 2-uniform word.*

Proof. Let $I(v) = \{u : u \to v\}$ be the set of all in-neighbors of v, and $O(v) = \{u : v \to u\}$ be the set of all out-neighbors of v. Also, let $A(v) = \{u \in V : u \rightsquigarrow v\} \setminus I$ be the set of v's non-neighboring vertices that can reach v, and $B(v) = \{u \in V : v \rightsquigarrow u\} \setminus O$ be the set of v's non-neighboring vertices that can be reached from v. Finally, let $T(v) = V \setminus (\{v\} \cup I(v) \cup O(v) \cup A(v) \cup B(v))$ be the set of remaining vertices. Note that the sets $I(v)$, $O(v)$, $A(v)$, $B(v)$ and $T(v)$ are pairwise disjoint and some of them can be empty.

Denote by A, B, I, O and T topological sorts of the corresponding digraphs induced by the sets $A(v)$, $B(v)$, $I(v)$, $O(v)$ and $T(v)$, respectively.

We now consider the 2-uniform word W given by

$$W = A\ I\ T\ A\ v\ O\ I\ v\ B\ T\ O\ B.$$

We claim that W covers all non-edges incident with v, i.e. the non-edges of type vu for $u \in T(v) \cup A(v) \cup B(v)$.

Condition 1) holds automatically. Clearly, W represents the graph H that is the union of the cliques $T \cup A \cup I$, $T \cup B \cup O$, and $I \cup O \cup \{v\}$. Since v is not adjacent to each $u \in T \cup A \cup B$ in H, condition 4) is true. Let us check condition 3). Indeed there are no edges connecting v with $T(v) \cup A(v) \cup B(v)$ in G by the definition. Note also that no arcs can go from $A(v)$ to $O(v) \cup B(v)$, or from $I(v)$ to $B(v)$ in D, since that would induce a shortcut. Also, there are no arcs in the other direction, since the digraph D is acyclic. So, G has no edges connecting $A(v)$ with $O(v) \cup B(v)$, or $I(v)$ with $B(v)$. Since all other edges exist in H, we have that G is a subgraph of H, i.e. condition 3) holds. Finally, let us check condition 2). We have $P(W) = AITvOB$. Let u and w be distinct vertices such that $u \rightsquigarrow w$. If u and w are in the same set A, I, T, O, or B then u precedes w in $P(W)$ since the corresponding set is a topological sort. If neither u nor w is in T then u precedes w in $P(W)$ because otherwise the digraph D would contain a directed cycle. If $u \in T$ then w cannot be in $A \cup I$ since otherwise there would be a directed path from u to v and thus u must be also in $A \cup I$ by the definition. So, u precedes w in $P(W)$. Finally, if $w \in T$ then u cannot be in $\{v\} \cup O \cup B$ since otherwise there would be a directed path from v to u and from u to w and thus w must be in $O \cup B$ by the definition. So, $u \in A \cup I$ and hence u precedes w in $P(W)$. Therefore, $P(W)$ is a topological sort of D and condition 2) is true. □

Now we are ready to prove the main result.

Theorem 3. *A graph G is word-representable if and only if it is semi-transitively orientable. Moreover, each non-complete word-representable graph is $2(n-\kappa)$-word-representable where κ is the size of the maximum clique in G.*

Proof. For the forward direction, given a word W, we direct an edge of G from x to y if the first occurrence of x is before that of y in the word. Let us show that such an orientation D of G_W is semi-transitive. Indeed, assume that $x_0x_t \in E(D)$ and there is a directed path $x_0x_1 \cdots x_t$ in D. Then in the word W we have $x_0^i < x_1^i < \cdots < x_t^i$ for every i. Since $x_0x_t \in E(D)$ we have $x_t^i < x_0^{i+1}$. But then for every $j < k$ and i there must be $x_j^i < x_k^i < x_j^{i+1}$, i.e. $x_ix_j \in E(D)$. So, D is semi-transitive.

The other direction follows directly from Lemmas 1 and 2. Indeed, let K be a maximum clique of G. Denote by D a semi-transitive orientation of the graph G. Let W_v be the 2-uniform word that covers all the non-edges in G incident with a vertex $v \in V \setminus K$. Concatenating $n-\kappa$ such words W_v induces a word W that covers all non-edges in G and preserves all edges. This follows from the fact that every non-edge has at least one endpoint outside K. Thus, G is represented by W. Moreover, this word is $2(n-\kappa)$-uniform, proving the second part of the theorem. □

Since each complete graph is 1-word-representable and each edgeless graph (having maximum clicks of size 1) is 2-word-representanle, we have the following statement.

Corollary 1. *Each word-representable graph G on $n \geq 3$ vertices is $2(n-2)$-word-representable.*

4. The Representation Number of Graphs

We focus now on the following question: Given a word-representable graph, how large is its representation number? In [12], certain classes of graphs were proved to be 2- or 3-word representable, and an example was given of a graph (the triangular prism) with the representation number of 3. More on graphs with representation number 3 can be found in [11]. On the other hand, no examples were known of graphs with representation numbers larger than 3, nor were there any non-trivial upper bounds known.

Theorem 3 implies that the graph property of word-representability is polynomially verifiable, i.e. the recognition problem is in NP. Limouzy [14] observed that triangle-free representable graphs are precisely the cover graphs, i.e. graphs that can be oriented as the diagrams of a partial order. Determining whether a graph is a cover graph is NP-hard [1, 16], and thus it is also hard to determine if a given (triangle-free) graph is word-representable. Thus, we obtain the following exact classification.

Corollary 2. *The recognition problem for word-representable graphs is NP-complete.*

We now show that there are graphs with representation number of $n/2$, matching the upper bound within a factor of 4.

The *crown graph* $H_{k,k}$ is the graph obtained from the complete bipartite graph $K_{k,k}$ by removing a perfect matching. Denote by G_k the graph obtained from a crown graph $H_{k,k}$ by adding a universal vertex (adjacent to all vertices in $H_{k,k}$).

Theorem 4. *The graph G_k has representation number $k = \lfloor n/2 \rfloor$.*

The proof is based on three statements.

Lemma 5. *Let H be a graph and G be the graph obtained from H by adding an all-adjacent vertex. Then G is a k-word-representable graph if and only if H is a permutational k-word-representable graph.*

Proof. Let 0 be the letter corresponding to the all-adjacent vertex. Then every other letter of the word W representing G must appear exactly once between two consecutive zeroes. We may assume also that W starts with 0. Then the word $W \setminus \{0\}$, formed by deleting all occurrences of 0 from W, is a permutational k-representation of H. Conversely, if W' is a word permutationally k-representing H, then we insert 0 in front of each permutation to get a (permutational) k-representation of G. □

Recall that the *order dimension* of a poset is the minimum number of linear orders such that their intersection induces this poset.

Lemma 6. *A comparability graph is a permutational k-word-representable graph if and only if the poset induced by this graph has dimension at most k.*

Proof. Let H be a comparability graph and W be a word permutationally k-representing it. Each permutation in W can be considered as a linear order where $a < b$ if and only if a meets before b in the permutation. We want to show that the comparability graph of the poset induced by the intersection of these linear orders coincides with H.

Two vertices a and b are adjacent in H if and only if their letters alternate in the word. So, they must be in the same order in each permutation, i.e. either $a < b$ in every linear order or $b < a$ in every linear order. But this means that a and b are comparable in the poset induced by the intersection of the linear orders, i.e. a and b are adjacent in its comparability graph. □

The next statement most probably is known but we give its proof here for the sake of completeness.

Lemma 7. *Let P be the poset over the $2k$ elements $\{a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_k\}$ such that $a_i < b_j$ for every $i \neq j$ and all other elements are not comparable. Then, P has dimension k.*

Proof. Assume that this poset is the intersection of t linear orders. Since a_i and b_i are not comparable for each i, their must be a linear order where

$b_i < a_i$. If we have in some linear order both $b_i < a_i$ and $b_j < a_j$ for $i \neq j$, then either $a_i < a_j$ or $a_j < a_i$ in it. In the first case we have that $b_i < a_j$, in the second that $b_j < a_i$. But each of these inequalities contradicts the definition of the poset. Therefore, $t \geq k$.

In order to show that $t = k$ we can consider a linear order $a_1 < a_2 < \ldots < a_{k-1} < b_k < a_k < b_{k-1} < \ldots < b_2 < b_1$ together with all linear orders obtained from this order by the simultaneous exchange of a_k and b_k with a_m and b_m respectively ($m = 1, 2, \ldots, k-1$). It can be verified that the intersection of these k linear orders coincides with our poset. □

Now we can prove Theorem 4. Since the crown graph $H_{k,k}$ is a comparability graph of the poset P, we deduce from Lemmas 7 and 6 that $H_{k,k}$ is a permutational k-word-representable graph but not a permutational $(k-1)$-word-representable graph. Then by Lemma 5 we have that G_k is a k-word-representable graph but not a $(k-1)$-word-representable graph. Theorem 4 is proved. □

The above arguments help us also in deciding the complexity of determining the representation number. From Lemmas 5 and 6, we see that it is as hard as determining the dimension k of a poset. Yannakakis [18] showed that the latter is NP-hard, for any $3 \leq k \leq \lceil n/2 \rceil$. We therefore obtain the following result.

Proposition 8. *Deciding whether a given graph is a k-word-representable graph, for any given $3 \leq k \leq \lceil n/2 \rceil$, is NP-complete.*

It was shown that it is also NP-hard to approximate the dimension of a poset within $n^{1/2-\epsilon}$-factor [10], and this has recently been strengthened to and $n^{1-\epsilon}$-factor [2]. We therefore obtain the same hardness for the representation number.

Proposition 9. *Approximating the representation number within $n^{1-\epsilon}$-factor is NP-hard, for any $\epsilon > 0$. That is, for every constant $\epsilon > 0$, it is an NP-complete problem to decide whether a given word-representable graph has representation number at most n^{ϵ}, or it has representation number greater than n^{ϵ}.*

5. Subclasses of Word-Representable Graphs

When faced with a new graph class, the most basic questions involve the kind of properties it satisfies: which known classes are properly contained (and which not), which graphs are otherwise contained (and which not), what operations preserve word representability (or non-representability), and which properties hold for these graphs.

Previously, it was known that the class of word-representable graphs includes comparability graphs, outerplanar graphs, subdivision graphs, and prisms. The purpose of this section is to clarify this situation significantly, including resolving some conjectures. We start with exploring the relation of colorability and representability.

Corollary 3. *3-colorable graphs are word-representable.*

Proof. Given a 3-coloring of a graph G, direct the edges from the first color class through the second to the third class. It is easy to see that we obtain a semi-transitive digraph. Thus, by Theorem. 3, the graph is word-representable. □

This implies a number of earlier results on word-representability, including that of outerplanar graphs, subdivision graphs, and prisms [12]. The theorem also shows that 2-degenerate graphs (graphs in which every subgraph has a vertex of degree at most 2) and sub-cubic graphs (graphs of maximum degree 3, via Brooks theorem) are word-representable.

Note, however, that the bound of $2n-4$ for the representability number can be improved for 3-colorable graphs.

Theorem 10. *3-colorable graphs are $2\lfloor 2n/3 \rfloor$-word-representable.*

Proof. Let G be a 3-colorable graph.

Suppose that the vertices of G are partitioned into three independent sets A, B and C. We may assume that the set B has the largest cardinality among these three sets. Direct the edges of G from A to $B \cup C$ and from B to C and denote the obtained orientation by D.

Let $v \in A$. Denote by N_v^B (resp., N_v^C) an arbitrary permutation over the set of neighbors of v in B (resp., in C) and by $\overline{N_v^B}$ (resp., $\overline{N_v^C}$) — arbitrary permutations over the remaining vertices of B (resp., of C). Also, let A_v be an arbitrary permutation over the set $A \setminus \{v\}$. For a permutation P denote by $R(P)$ the permutation written in the reverse order.

Consider the 2-uniform word

$$W_v^A = A_v\ \overline{N_v^B}\ v\ N_v^B\ N_v^C\ v\ \overline{N_v^C}\ R(A_v)\ R(N_v^B)\ R(\overline{N_v^B})\ R(\overline{N_v^C})\ R(N_v^C).$$

Note that W_v^A covers all non-edges lying inside B and C and also all non-edges incident with v. Indeed, the graph H induced by W_v^A is obtained from the complete tripartite graph with the partition A, B, C by removal of all edges connecting v with $\overline{N_v^B} \cup \overline{N_v^C}$. This justifies 3) and 4). Since $P(W_v^A) = A_v\ \overline{N_v^B}\ v\ N_v^B\ N_v^C\ \overline{N_v^C}$ and no directed path can go from v to $\overline{N_v^B}$, it is a topological sort of D, justifying 2).

Similarly, for $v \in C$ consider the 2-uniform word

$$W_v^C = N_v^A\ \overline{N_v^A}\ \overline{N_v^B}\ N_v^B\ C_v\ R(\overline{N_v^A})\ v\ R(N_v^A)\ R(N_v^B)\ v\ R(\overline{N_v^B})\ R(C_v),$$

where C_v, $\overline{N_v^A}$, N_v^A, $\overline{N_v^B}$ and N_v^B are defined similarly to the respective sets above. Using the similar arguments, one can show that W_v^C covers all non-edges lying inside A and B and all non-edges incident with v.

Concatenating all these words, we obtain a $2k$-uniform word W representing G, where $k = |A \cup C|$. Since B has the largest cardinality, $k \leq \lfloor 2n/3 \rfloor$. □

Corollary 3 does not extend to higher chromatic numbers. The examples in Fig. 2 show that 4-colorable graphs can be non-word-representable. We can, however, obtain a result in terms of the *girth* of the graph, which is the length of its shortest cycle.

Proposition 11. *Let G be a graph whose girth is greater than its chromatic number. Then, G is word-representable.*

Proof. Suppose the graph is colored with $\chi(G)$ natural numbers. Orient the edges of the graph from smaller to larger colors. There is no directed path with more than $\chi(G)-1$ arcs, but since G contains no cycle of $\chi(G)$ or fewer edges, there can be no shortcut. Hence, the digraph is semi-transitive. □

6. Conclusions

Two of open problems stated in the preliminary version of this paper [8] were solved. One of the solved problems is on NP-hardness of the problem of recognition whether a given graph is word-representable or not, and it is discussed in Section 4 (see Corollary 2). The other problem was solved in [3], where it was shown that there exist non-word-representable graphs of maximum degree 4. We end up the paper with stating revised versions of the remaining two open problems.

(1) What is the maximum representation number of a graph? We know that it lies between $n/2$ and $2n-4$.
(2) Is there an algorithm that forms an $f(k)$-representation of a k-word-representable graph, for some function f? Namely, can the representation number be approximated as a function of itself? By Prop. 9, this function must grow faster than any fixed polynomial.

References

[1] G. Brightwell. On the complexity of diagram testing. *Order* **10**:297–303, 1993.

[2] P. Chalermsook, B. Laekhanukit and D. Nanongkai. Graph products revisited: Tight approximation hardness of induced matching, poset dimension and more. Proc. *24th annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2013.

[3] A. Collins, S. Kitaev and V. Lozin: New results on word-representable graphs, *Discrete Applied Mathematics*, to appear.

[4] B. Courcelle. Circle graphs and Monadic Second-order logic. *J. Applied Logic* 6(3): 416-442, 2008.

[5] D. L. Draper, A. K. Jonsson, D. P. Clements and D. E. Joslin. Cyclic Scheduling. In *Proc. IJCAI*, 1999.

[6] R. Graham and N. Zang. Enumerating split-pair arrangements, *J. Combin. Theory, Series A* **115**, Issue 2 (2008), 293–303.

[7] M. Halldórsson, S. Kitaev and A. Pyatkin. Graphs capturing alternations in words. In *Proc. 14th Developments in Language Theory (DLT)*, LNCS 6224, pages 436–437, August 2010.

[8] M. Halldórsson, S. Kitaev and A. Pyatkin. Alternation graphs, *Lecture Notes in Computer Science* **6986** (2011) 191–202. Proceedings of the 37*th International Workshop on Graph-Theoretic Concepts in Computer Science*, *WG* 2011, Teplá Monastery, Czech Republic, June 21–24, 2011.

[9] C. Hanen and A. Munier. Cyclic scheduling on parallel processors: An overview. In *P. Chretienne, E. G. Coffman, Jr., J. K. Lenstra, and Z. Liu, editors, Scheduling Theory and its Applications*, chapter 9. John Wiley & Sons, 1995.

[10] R. Hegde and K. Jain. The hardness of approximating poset dimension. *Electronic Notes in Discrete Mathematics* **29** (2007), 435–443.

[11] S. Kitaev: On graphs with representation number 3, *Journal of Automata, Languages and Combinatorics*, to appear.

[12] S. Kitaev and A. Pyatkin. On representable graphs. *Automata, Languages and Combinatorics* **13** (2008) 1, 45–54.

[13] S. Kitaev and S. Seif. Word problem of the Perkins semigroup via directed acyclic graphs. *Order*, DOI 10.1007/s11083-008-9083-7 (2008).

[14] V. Limouzy. Private communication, 2014.

[15] M. Middendorf and V. G. Timkovsky. On scheduling cycle shops: Classification, complexity and approximation. *Journal of Scheduling*, 5:135–169, 2002.

[16] J. Nešetřil and V. Rödl. More on the complexity of cover graphs. *Comment. Math. Univ. Carolin* **36**(2):269–278, 1995.

[17] M. Pergel. Recognition of polygon-circle graphs and graphs of interval filament is NP-complete. In *Proc. 33rd Intl. Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, Lecture Notes in Computer Science 4769, Springer, 2007.

[18] M. Yannakakis. The complexity of the partial order dimension problem. *SIAM J. Algebraic Discrete Methods* 3(3) (1982), 351–358.