A Comprehensive Introduction to the Theory of Word-Representable Graphs

Sergey Kitaev

May 30, 2017

Abstract

Letters x and y alternate in a word w if after deleting in w all letters but the copies of x and y we either obtain a word $xyxy \cdots$ (of even or odd length) or a word $yxyx \cdots$ (of even or odd length). A graph $G = (V, E)$ is word-representable if and only if there exists a word w over the alphabet V such that letters x and y alternate in w if and only if $xy \in E$.

Word-representable graphs generalize several important classes of graphs such as circle graphs, 3-colorable graphs and comparability graphs. This paper offers a comprehensive introduction to the theory of word-representable graphs including the most recent developments in the area.

Contents

1 Introduction

The theory of word-representable graphs is a young but very promising research area. It was introduced by the author in 2004 based on the

joint research with Steven Seif [20] on the celebrated *Perkins semi-group*, which has played a central role in semigroup theory since 1960, particularly as a source of examples and counterexamples. However, the first systematic study of word-representable graphs was not undertaken until the appearance in 2008 of the paper [18] by the author and Artem Pyatkin, which started the development of the theory. One of the most significant contributors to the area is Magnús M. Halldórsson.

Up to date, nearly 20 papers have been written on the subject, and the core of the book [17] by the author and Vadim Lozin is devoted to the theory of word-representable graphs. It should also be mentioned that the software produced by Marc Glen [7] is often of great help in dealing with word-representation of graphs.

We refer the Reader to [17], where relevance of word-representable graphs to various fields is explained, thus providing a motivation to study the graphs. These fields are algebra, graph theory, computer science, combinatorics on words, and scheduling. In particular, word-representable graphs are important from graph-theoretical point of view, since they generalize several fundamental classes of graphs (e.g. *circle graphs*, 3-*colorable graphs* and *comparability graphs*).

A graph $G = (V, E)$ is *word-representable* if and only if there exists a word w over the alphabet V such that letters x and y, $x \neq y$, alternate in w if and only if $xy \in E$ (see Section 2 for the definition of alternating letters). Natural questions to ask about word-representable graphs are:

- Are all graphs word-representable?

- If not, how do we characterize word-representable graphs?

- How many word-representable graphs are there?

- What is graph's representation number for a given graph? Essentially, what is the minimal length of a word-representant?

- How hard is it to decide whether a graph is word-representable or not? (complexity)

- Which graph operations preserve (non-)word-representability?

- Which graphs are word-representable in your favourite class of graphs?

This paper offers a comprehensive introduction to the theory of word-representable graphs. Even though the paper is based on the book [17] following some of its structure, our exposition goes far beyond book's content and it reflects the most recent developments in the area. Having said that, there is a relevant topic on a generalization of the theory of word-representable graphs [12] that is discussed in [17, Chapter 6], but we do not discuss it at all.

In this paper we do not include the majority of proofs due to space limitations (while still giving some proofs, or ideas of proofs whenever

possible). Also, all graphs we deal with are simple (no loops or multiple edges are allowed), and unless otherwise specified, our graphs are unoriented.

2 Word-Representable Graphs. The Basics

Suppose that w is a word over some alphabet and x and y are two distinct letters in w. We say that x and y *alternate* in w if after deleting in w *all* letters *but* the copies of x and y we either obtain a word $xyxy\cdots$ (of even or odd length) or a word $yxyx\cdots$ (of even or odd length). For example, in the word 23125413241362, the letters 2 and 3 alternate. So do the letters 5 and 6, while the letters 1 and 3 do *not* alternate.

Definition 1. *A graph $G = (V, E)$ is* word-representable *if and only if there exists a word w over the alphabet V such that letters x and y, $x \neq y$, alternate in w if and only if $xy \in E$. (By definition, w must* contain *each letter in V.) We say that w represents G, and that w is a* word-representant.

Definition 1 works for both vertex-labeled and unlabeled graphs because any labeling of a graph G is equivalent to any other labeling of G with respect to word-representability (indeed, the letters of a word w representing G can always be renamed). For example, the graph to the left in Figure 1 is word-representable because its labeled version to the right in Figure 1 can be represented by 1213423. For another example, each *complete graph* K_n can be represented by any permutation π of $\{1, 2, \ldots, n\}$, or by π concatenated any number of times. Also, the *empty graph* E_n (also known as *edgeless graph*, or *null graph*) on vertices $\{1, 2, \ldots, n\}$ can be represented by $12 \cdots (n-1)nn(n-1) \cdots 21$, or by any other permutation concatenated with the same permutation written in the reverse order.
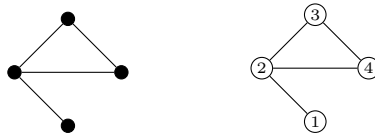

Figure 1: An example of a word-representable graph

Remark 1. *The class of word-representable graphs is* hereditary. *That is, removing a vertex v in a word-representable graph G results in a word-representable graph G'. Indeed, if w represents G then w with v removed represents G'. This observation is crucial, e.g. in finding*

4

asymptotics for the number of word-representable graphs [4], which is the only known enumerative result on word-representable graphs to be stated next.

Theorem 2 ([4])**.** *The number of non-isomorphic word-representable graphs on n vertices is given by $2^{\frac{n^2}{3}+o(n^2)}$.*

2.1 k-Representability and Graph's Representation Number

A word w is k-*uniform* if each letter in w occurs k times. For example, 243321442311 is a 3-uniform word, while 23154 is a 1-uniform word (a permutation).

Definition 2. *A graph G is k-word-representable, or k-representable for brevity, if there exists a k-uniform word w representing it. We say that w k-represents G.*

The following result establishes equivalence of Definitions 1 and 2.

Theorem 3 ([18])**.** *A graph is word-representable if and only if it is k-representable for some k.*

Proof. Clearly, k-representability implies word-representability. For the other direction, we demonstrate on an example how to extend a word-representant to a uniform word representing the same graph. We refer to [18] for a precise description of the extending algorithm, and an argument justifying it.

Consider the word $w = 3412132154$ representing a graph G on five vertices. Ignore the letter 1 occurring the maximum number of times (in general, there could be several such letters all of which need to be ignored) and consider the *initial permutation* $p(w)$ of w formed by the remaining letters, that is, $p(w)$ records the order of occurrences of the leftmost copies of the letters. For our example, $p(w) = 3425$. Then the word $p(w)w = 34253412132154$ also represents G, but it contains more occurrences of the letters occurring not maximum number of time in w. This process can be repeated a number of times until each letter occurs the same number of times. In our example, we need to apply the process one more time by appending 5, the initial permutation of $p(w)w$, to the left of $p(w)w$ to obtain a uniform representation of G: 534253412132154. □

Following the same arguments as in Theorem 3, one can prove the following result showing that there are *infinitely many* representations for any word-representable graph.

Theorem 4 ([18])**.** *If a graph is k-representable then it is also $(k+1)$-representable.*

By Theorem 3, the following notion is well-defined.

Definition 3. Graph's representation number *is the* least k *such that the graph is* k-*representable. For non-word-representable graphs (whose existence will be discussed below), we let* $k = \infty$. *Also, we let* $\mathcal{R}(G)$ *denote* G's *representation number and* $\mathcal{R}_k = \{G : \mathcal{R}(G) = k\}$.

Clearly, $\mathcal{R}_1 = \{G : G$ is a complete graph$\}$. Next, we discuss $\mathcal{R}_2$.

2.2 Graphs with Representation Number 2

We begin with discussing five particular classes of graphs having representation number 2, namely, *empty graphs*, *trees*, *forests*, *cycle graphs* and *ladder graphs*. Then we state a result saying that graphs with representation number 2 are exactly the class of *circle graphs*.

2.2.1 Empty Graphs

No empty graph E_n for $n \geq 2$ can be represented by a single copy of each letter, so $\mathcal{R}(E_n) \geq 2$. On the other hand, as discussed above, E_n can be represented by concatenation of two permutations, and thus $\mathcal{R}(E_n) = 2$.

2.2.2 Trees and Forests

A simple inductive argument shows that *any tree* T can be represented using two copies of each letter, and thus, if the number of vertices in T is at least 3, $\mathcal{R}(T) = 2$. Indeed, as the base case we have the edge labeled by 1 and 2 that can be 2-represented by 1212. Now, suppose that any tree on at most $n - 1$ vertices can be 2-represented for $n \geq 3$, and consider a tree T with n vertices and with a leaf x connected to a vertex y. Removing the leaf x, we obtain a tree T' that can be 2-represented by a word $w_1 y w_2 y w_3$ where w_1, w_2 and w_3 are possibly empty words not containing y. It is now easy to see that the word $w_1 y w_2 x y x w_3$ 2-represents T (obtained from T' by inserting back the leaf x). Note that the word $w_1 x y x w_2 y w_3$ also represents T.

Representing each tree in a forest by using two letters (trees on one vertex x and two vertices x and y can be represented by xx and $xyxy$, respectively) and concatenating the obtained word-representants, we see that for any forest F having at least two trees, $\mathcal{R}(F) = 2$. Indeed, having two letters in a word-represent for each tree guarantees that no pair of trees will be connected by an edge.

2.2.3 Cycle Graphs

Another class of 2-representable graphs is *cycle graphs*. Note that a cyclic shift of a word-representant may not represent the same graph, as

is the case with, say, the word 112. However, if a word-representant is uniform, a cyclic shift does represent the same graph, which is recorded in the following proposition.

Proposition 5 ([18]). *Let $w = uv$ be a k-uniform word representing a graph G, where u and v are two, possibly empty, words. Then the word $w' = vu$ also represents G.*

Now, to represent a cycle graph C_n on n vertices, one can first represent the *path graph* P_n on n vertices using the technique to represent trees, then make a 1-letter cyclic shift still representing P_n by Proposition 5, and swap the first two letters. This idea is demonstrated for the graph in Figure 2 as follows. The steps in representing the path graph P_6 obtained by removing the edge 16 from C_6 are

$$1212 \rightarrow 121323 \rightarrow 12132434 \rightarrow 1213243545 \rightarrow 121324354656.$$

The 1-letter cyclic shift gives the word 612132435465 still representing P_6 by Proposition 5, and swapping the first two letters gives the sought representation of C_6: 162132435465.
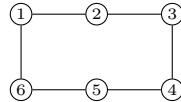

Figure 2: Cycle graph C_6

2.2.4 Ladder Graphs

The ladder graph L_n with $2n$ vertices, labeled $1, \ldots, n, 1', \ldots, n'$, and $3n - 2$ edges is constructed following the pattern for $n = 4$ presented in Figure 3. An inductive argument given in [15] shows that for $n \geq 2$, $\mathcal{R}(L_n) = 2$. Table 1 gives 2-representations of L_n for $n = 1, 2, 3, 4$.
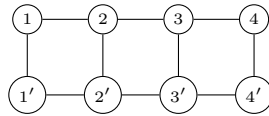

Figure 3: The ladder graph L_4

2.2.5 Circle Graphs

A *circle graph* is the *intersection graph* of a set of chords of a circle. That is, it is an unoriented graph whose vertices can be associated with chords of a circle such that two vertices are adjacent if and only if the

n	2-representation of the ladder graph L_n
1	$11'\mathbf{1}1'$
2	$1'21\mathbf{2'2}1'2'1$
3	$12'1'32\mathbf{3'3}2'3'121'$
4	$1'213'2'43\mathbf{4'4}3'4'231'2'1$

Table 1: 2-representations of the ladder graph L_n for $n = 1, 2, 3, 4$

corresponding chords cross each other. See Figure 4 for an example of a circle graph on four vertices and its associated chords.

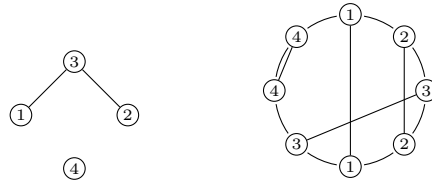The following theorem provides a complete characterization of $\mathcal{R}_2$.

Figure 4: A circle graph on four vertices and its associated chords

Theorem 6. *We have*

$$\mathcal{R}_2 = \{G : G \text{ is a circle graph different from a complete graph}\}.$$

Proof. Given a circle graph G, consider its representation on a circle by intersecting chords. Starting from any chord's endpoint, go through all the endpoints in clock-wise direction recording chords' labels. The obtained word w is 2-uniform and it has the property that a pair of letter x and y alternate in w if and only if the pair of chords labeled by x and y intersect, which happens if and only if the vertex x is connected to the vertex y in G. For the graph in Figure 4, the chords' labels can be read starting from the lower 1 as 13441232, which is a 2-uniform word representing the graph. Thus, G is a circle graph if and only if $G \in \mathcal{R}_2$ with the only exception if G is a complete graph, in which case $G \in \mathcal{R}_1$. $\square$

2.3 Graphs with Representation Number 3

Unlike the case of graphs with representation number 2, no characterization of graphs with representation number 3 is know. However, there is a number of interesting results on this class of graphs to be discussed next.
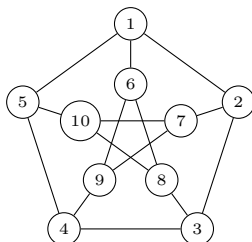
Figure 5: The Petersen graph

2.3.1 The Petersen Graph

In 2010, Alexander Konovalov and Steven Linton not only showed that the Petersen graph in Figure 5 is not 2-representable, but also provided two *non-equivalent* (up to renaming letters or a cyclic shift) 3-representations of it:

- 1387296(10)7493541283(10)7685(10)194562 and
- 134(10)58679(10)273412835(10)6819726495.

The fact that the Petersen graph does not belong to $\mathcal{R}_2$ is also justified by the following theorem.

Theorem 7 ([9]). *Petersen's graph is* not *2-representable.*

Proof. Suppose that the graph is 2-representable and w is a 2-uniform word representing it. Let x be a letter in w such that there is a minimal number of letters between the two occurrences of x. Since Petersen's graph is regular of degree 3, it is not difficult to see that there must be exactly three letters, which are all different, between the xs (having more letters between xs would lead to having two equal letters there, contradicting the choice of x).

By symmetry, we can assume that $x = 1$, and by Proposition 5 we can assume that w starts with 1. So, the letters 2, 5 and 6 are between the two 1s, and because of symmetry, the fact that Petersen's graph is *edge-transitive* (that is, each of its edges can be made "internal"), and taking into account that the vertices 2, 5 and 6 are pairwise non-adjacent, we can assume that $w = 12561w_1 6w_2 5w_3 2w_4$ where the w_is are some, possibly empty words for $i \in \{1, 2, 3, 4\}$. To alternate with 6 but not to alternate with 5, the letter 8 must occur in w_1 and w_2. Also, to alternate with 2 but not to alternate with 5, the letter 3 must occur in w_3 and w_4. But then 8833 is a subsequence in w, and thus 8 and 3 must be non-adjacent in the graph, a contradiction. $\square$
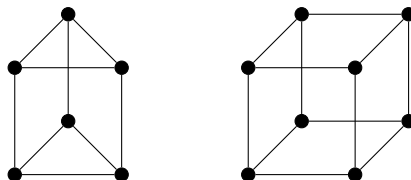
Figure 6: Prisms Pr_3 and Pr_4

2.3.2 Prisms

A *prism* Pr_n is a graph consisting of two cycles $12\cdots n$ and $1'2'\cdots n'$, where $n \geq 3$, connected by the edges ii' for $i = 1,\ldots,n$. In particular, the 3-dimensional cube to the right in Figure 6 is the prism Pr_4.

Theorem 8 ([18]). *Every prism Pr_n is 3-representable.*

The fact that the triangular prism Pr_3 is not 2-representable was shown in [18]. The following more general result holds.

Theorem 9 ([15]). *None of prisms Pr_n is 2-representable.*

Theorems 8 and 9 show that $\text{Pr}_n \in \mathcal{R}_3$ for any $n \geq 3$.

2.3.3 Colorability of Graphs in $\mathcal{R}_3$

Theorem 13 below shows that $\mathcal{R}_3$ does not even include 2-colorable graphs, and thus any class of c-colorable graphs for $c \geq 3$. Indeed, any c-colorable non-3-representable graph can be extended to a $(c+1)$-colorable graph by adding an apex (all-adjacent vertex), which is still non-3-representable using the hereditary nature of word-representability (see Remark 1).

A natural question to ask here is: Is $\mathcal{R}_3$ properly included in a class of c-colorable graphs for a constant c? A simple argument of replacing a vertex in the 3-representable triangular prism Pr_3 by a complete graph of certain size led to the following theorem.

Theorem 10 ([15]). *The class $\mathcal{R}_3$ is not included in a class of c-colorable graphs for some constant c.*

2.3.4 Subdivisions of Graphs

The following theorem gives a useful tool for constructing 3-representable graphs, that is, graphs with representation number at most 3.

Theorem 11 ([18]). *Let $G = (V, E)$ be a 3-representable graph and $x, y \in V$. Denote by H the graph obtained from G by adding to it a path of length at least 3 connecting x and y. Then H is also 3-representable.*

Definition 4. *A* subdivision *of a graph G is a graph obtained from G by replacing each edge xy in G by a* simple *path (that is, a path without self-intersection) from x to y. A subdivision is called a k-*subdivision *if each of these paths is of length at least k.*

Definition 5. *An* edge contraction *is an operation which removes an edge from a graph while gluing the two vertices it used to connect. An unoriented graph G is a* minor *of another unoriented graph H if a graph isomorphic to G can be obtained from H by contracting some edges, deleting some edges, and deleting some isolated vertices.*

Theorem 12 ([18])**.** *For every graph G there are infinitely many* 3-*representable graphs H that contain G as a minor. Such a graph H can be obtained from G by subdividing* each *edge into* any *number of, but* at least *three edges.*

Note that H in Theorem 12 does not have to be a k-subdivision for some k, that it, edges of G can be subdivided into different number (at least 3) of edges. In either case, the 3-subdivision of any graph G is always 3-representable. Also, it follows from Theorem 11 and the proof of Theorem 12 in [18] that a graph obtained from an edgeless graph by inserting simple paths of length at least 3 between (some) pairs of vertices of the graph is 3-representable.

Finally, note that subdividing *each* edge in any graph into *exactly* two edges gives a bipartite graph, which is word-representable by Theorem 15 (see the discussion in Section 2.4.1 on why a bipartite graph is word-representable).

2.4   Graphs with High Representation Number

In Theorem 31 below we will see that the upper bound on a shortest word-representant for a graph G on n vertices is essentially $2n^2$, that is, one needs at most $2n$ copies of each letter to represent G. Next, we consider two classes of graphs that require essentially $n/2$ copies of each letter to be represented, and these are the longest known shortest word-representants.

2.4.1   Crown Graphs

Definition 6. *A* crown graph *(also known as a* cocktail party graph*)* $H_{n,n}$ *is obtained from the complete bipartite graph $K_{n,n}$ by removing a* perfect matching. *That is, $H_{n,n}$ is obtained from $K_{n,n}$ by removing n edges such that* each *vertex was incident to* exactly one *removed edge.*

See Figure 7 for examples of crown graphs.

By Theorem 15 below, $H_{n,n}$ can be represented by a concatenation of permutations, because $H_{n,n}$ is a comparability graph defined in Section 3 (to see this, just orient all edges from one part to the other).
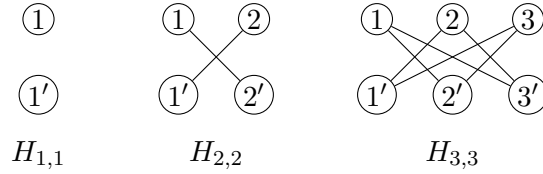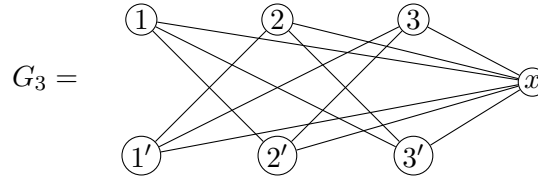
Figure 7: Crown graphs

Figure 8: Graph G_3

In fact, $H_{n,n}$ is known to require n permutations to be represented. However, can we provide a shorter representation for $H_{n,n}$? It turns out that we can, to be discussed next, but such representations are still long (linear in n).

Note that $H_{1,1} \in \mathcal{R}_2$ by Section 2.2.1. Further, $H_{2,2} \neq K_4$, the complete graph on 4 vertices, and thus $H_{2,2} \in \mathcal{R}_2$ because it can be 2-represented by $121'2'212'1'$. Also, $H_{3,3} = C_6 \in \mathcal{R}_2$ by Section 2.2.3. Finally, $H_{4,4} = \mathrm{Pr}_4 \in \mathcal{R}_3$ by Section 2.3.2. The following theorem gives the representation number $\mathcal{R}(H_{n,n})$ in the remaining cases.

Theorem 13 ([6])**.** *If $n \geq 5$ then the representation number of $H_{n,n}$ is $\lceil n/2 \rceil$ (that is, one needs $\lceil n/2 \rceil$ copies of each letter to represent $H_{n,n}$, but not fewer).*

2.4.2 Crown Graphs with an Apex

The graph G_n is obtained from a crown graph $H_{n,n}$ by adding an apex (all-adjacent vertex). See Figure 8 for the graph G_3.

It turns out that G_n is the *worst* known word-representable graph in the sense that it requires the maximum number of copies of each letter to be represented, as recorded in the following theorem.

Theorem 14 ([18])**.** *The representation number of G_n is $\lfloor n/2 \rfloor$.*

It is unknown whether there exist graphs on n vertices with representation number between $\lfloor n/2 \rfloor$ and essentially $2n$ (given by Theorem 32).
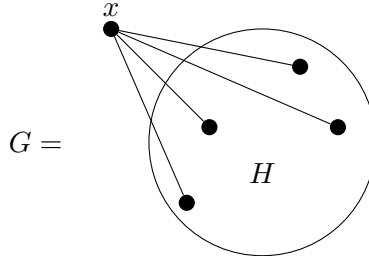
Figure 9: G is obtained from H by adding an apex

3 Permutationally Representable Graphs and their Significance

An orientation of a graph is *transitive* if presence of edges $u \to v$ and $v \to z$ implies presence of the edge $u \to z$. An unoriented graph is a *comparability graph* if it admits a transitive orientation. It is well known [17, Section 3.5.1], and is not difficult to show that the smallest non-comparability graph is the cycle graph C_5.

Definition 7. *A graph $G = (V, E)$ is* permutationally representable *if it can be represented by a word of the form $p_1 \cdots p_k$ where p_i is a permutation. We say that G is* permutationally *k-representable.*

For example, the graph in Figure 1 is permutationally representable, which is justified by the concatenation of two permutations 21342341.

The following theorem is an easy corollary of the fact that any partially ordered set can be represented as intersection of linear orders.

Theorem 15 ([20]). *A graph is permutationally representable if and only if it is a comparability graph.*

Next, consider a schematic representation of the graph G in Figure 9 obtained from a graph H by adding an all-adjacent vertex (apex). The following theorem holds.

Theorem 16 ([18]). *The graph G is word-representable if and only if the graph H is permutationally representable.*

A *wheel graph W_n* is the graph obtained from a cycle graph C_n by adding an apex. It is easy to see that none of cycle graphs C_{2n+1}, for $n \geq 2$, is a comparability graph, and thus none of wheel graphs W_{2n+1}, for $n \geq 2$ is word-representable. In fact, W_5 is the smallest example of a non-word-representable graph (the only one on 6 vertices). Section 6 discusses other examples of non-word-representable graphs.

As a direct corollary to Theorem 16, we have the following important result revealing the structure of neighbourhoods of vertices in a word-representable graph.
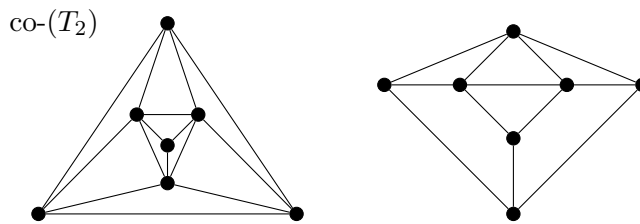
13

Figure 10: Non-word-representable graphs in which each neighbourhood is permutationally representable

Theorem 17 ([18])**.** *If a graph G is word-representable then the neighbourhood of each vertex in G is permutationally representable (is a comparability graph by Theorem 15).*

The converse to Theorem 17 is *not* true as demonstrated by the counterexamples in Figure 10 taken from [9] and [4], respectively.

A *clique* in an unoriented graph is a subset of pairwise adjacent vertices. A *maximum clique* is a clique of the maximum size. Given a graph G, the *Maximum Clique problem* is to find a maximum clique in G. It is well known that the Maximum Clique problem is NP-complete. However, this problem is polynomially solvable for word-representable graphs, which is a corollary of Theorem 17 and is discussed next.

Theorem 18 ([10, 11])**.** *The Maximum Clique problem is polynomially solvable on word-representable graphs.*

Proof. Each neighbourhood of a word-representable graph G is a comparability graph by Theorem 17. It is known that the Maximum Clique problem is solvable on comparability graphs in polynomial time. Thus the problem is solvable on G in polynomial time, since any maximum clique belongs to the neighbourhood of a vertex including the vertex itself. ☐

4 Graphs Representable by Pattern Avoiding Words

It is a very popular area of research to study patterns in words and permutations[1]. The book [14] provides a comprehensive introduction

---

[1]The patterns considered in this section are ordered, and their study comes from Algebraic Combinatorics. There are a few results on word-representable graphs and (unordered) patterns studied in Combinatorics on Words, namely on squares and cubes in words, that are not presented in this paper, but can be found in [17, Section 7.1.3]. One of the results says that for any word-representable graph, there exists a cube-free word representing it.
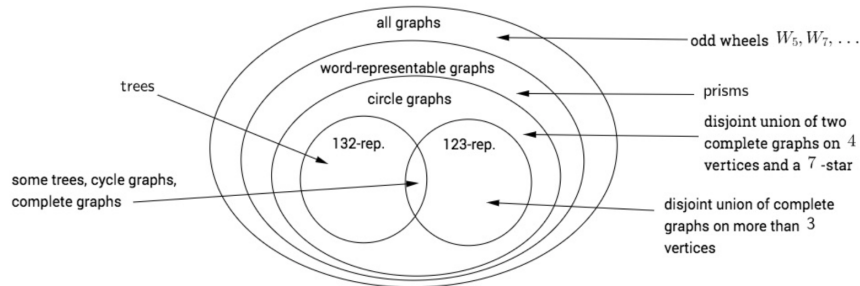
Figure 11: Relations between graph classes taken from [21]

to the field. A *pattern* is a word containing each letter in $\{1, 2, \ldots, k\}$ at least once for some k. A pattern $\tau = \tau_1 \tau_2 \cdots \tau_m$ *occurs* in a word $w = w_1 w_2 \cdots w_n$ if there exist $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ such that $\tau_1 \tau_2 \cdots \tau_m$ is *order-isomorphic* to $w_{i_1} w_{i_2} \cdots w_{i_m}$. We say that w *avoids* τ if w contains no occurrences of τ. For example, the word 42316 contains several occurrences of the pattern 213 (all ending with 6), e.g. the subsequences 426, 416 and 316.

As a particular case of a more general program of research suggested by the author during his plenary talk at the international Permutation Patterns Conference at the East Tennessee State University, Johnson City in 2014, one can consider the following direction (see [17, Section 7.8]). Given a set of words avoiding a pattern, or a set of patterns, which class of graphs do these words represent?

As a trivial example, consider the class of graphs defined by words avoiding the pattern 21. Clearly, any 21-avoiding word is of the form

$$w = 11 \cdots 122 \cdots 2 \cdots nn \cdots n.$$

If a letter x occurs *at least twice* in w then the respective vertex is isolated. The letters occurring *exactly once* form a *clique* (are connected to each other). Thus, 21-avoiding words describe graphs formed by a clique and an independent set.

Two papers, [5] and [21], are dedicated to this research direction and will be summarised in this section. So far, apart from Theorem 19 and Corollary 20 below, only 132-avoiding and 123-avoiding words were studied from word-representability point of view. The results of these studies are summarized in Figure 11, which is taken from [21]. In that figure, and more generally in this section, we slightly abuse the notation and call graphs representable by τ-avoiding words τ-*representable*.

We note that unlike the case of word-representability without extra restrictions, labeling of graphs *does matter* in the case of pattern avoiding representations. For example, the 132-avoiding word 4321234 represents the graph to the left in Figure 12, while *no* 132-avoiding
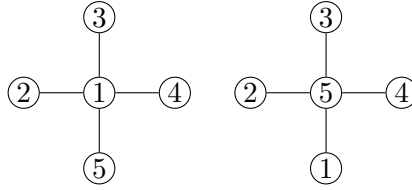
15

Figure 12: 132-representable (left) and non-132-representable (right) labelings of the same graph

word represents the other graph in that figure. Indeed, *no two* letters out of 1, 2 and 3 can occur *once* in a word-representant or else the respective vertices would *not* form an *independent set.* Say, w.l.o.g. that 1 and 2 occur at *least twice.* But then we can find 1 and 2 on *both sides* of an occurrence of the letter 4, and the patten 132 is *inevitable.*

The following theorem has a great potential to be applicable to the study of τ-representable graphs for τ of length 4 or more.

Theorem 19 ([21]). *Let G be a word-representable graph, which can be represented by a word avoiding a pattern τ of length $k + 1$. Let x be a vertex in G such that its degree $d(x) \geq k$. Then, any word w representing G that avoids τ must contain no more than k copies of x.*

Proof. If there are *at least $k + 1$* occurrences of x in w, we obtain a factor (i.e. consecutive subword) $xw_1x \cdots w_kx$, where k neighbours of x in G occur in each w_i. But then w contains *all* patterns of length $k + 1$, in particular, τ. Contradiction. □

Corollary 20 ([21]). *Let w be a word-representant for a graph which avoids a pattern of length $k + 1$. If some vertex y adjacent to x has degree at least k, then x occurs at most $k + 1$ times in w.*

4.1   132-Representable Graphs

It was shown in [5] that the minimum (with respect to the number of vertices) non-word-representable graph, the wheel graph W_5, is actually a minimum non-132-representable graph (we do not know if there exit other non-132-representable graphs on 6 vertices).

Theorem 21 ([5]). *If a graph G is 132-representable, then there exists a 132-avoiding word w representing G such that any letter in w occurs at most twice.*

Theorems 21 and 6 give the following result.

Theorem 22 ([5]). *Every 132-representable graph is a circle graph.*

Thus, by Theorems 6, 9 and 22, none of prisms Pr_n, $n \geq 3$, is 132-representable. A natural question is if there are circle graphs that are not 132-representable.

Theorem 23 ([21]). *Not all circle graphs are 132-representable. E.g. disjoint union of two complete graphs K_4 is a circle graph, but it is not 132-representable.*

Theorem 24 ([5]). *Any tree is 132-representable.*

Note that in the case of pattern avoiding representations of graphs, Theorem 3 does not necessarily work, because extending a representation to a uniform representation may introduce an occurrences of the pattern(s) in question. For example, while any complete graph K_n can be represented by the 132-avoiding word $n(n-1)\cdots 1$, it was shown in [21] that for $n \geq 3$ no 2-uniform 132-avoiding representation of K_n exists. In either case, [21] shows that any tree can actually be represented by a 2-uniform word thus refining the statement of Theorem 24. For another result on uniform 132-representation see Theorem 30 below.

Theorem 25 ([5]). *Any cycle graph is 132-representable.*

Proof. The cycle graph C_n labeled by $1, 2, \ldots, n$ in clockwise direction can be represented by the 132-avoiding word

$$(n-1)n(n-2)(n-1)(n-3)(n-2)\cdots 45342312.$$

$\square$

Theorem 26 ([5]). *For $n \geq 1$, a complete graph K_n is 132-representable. Moreover, for $n \geq 3$, there are*

$$2 + C_{n-2} + \sum_{i=0}^{n} C_i$$

different 132-representants for K_n, where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number. Finally, K_1 can be represented by a word of the form $11\cdots 1$ and K_2 by a word of the form $1212\cdots$ (of even or odd length) or $2121\cdots$ (of even or odd length).

As a corollary to the proof of Theorem 26, [5] shows that for $n \geq 3$, the length of any 132-representant of K_n is either n, or $n+1$, or $n+2$, or $n+3$.

4.2   123-Representable Graphs

An analogue of Theorem 22 holds for 123-representable graphs.

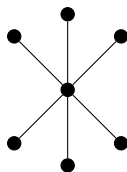**Theorem 27** ([21]). *Any 123-representable graph is a circle graph.*

Figure 13: Star graph $K_{1,6}$

Theorem 28 ([21])**.** *Any cycle graph is* 123*-representable.*

Proof. The cycle graph C_n labeled by $1, 2, \ldots, n$ in clockwise direction can be represented by the 123-avoiding word

$$n(n-1)n(n-2)(n-1)(n-3)(n-1)\cdots 23121.$$

$\square$

Theorem 29 ([21])**.** *The star graph* $K_{1,6}$ *in Figure 13 is not* 123*-representable.*

It is easy to see that $K_{1,6}$ is a circle graph, and thus not all circle graphs are 123-representable by Theorem 29. Also, by Theorem 29, not all trees are 123-representable.

Based on Theorems 23 and 29, it is easy to come up with a circle graph on 14 vertices that is neither 123- nor 132-representable (see [21]).

As opposed to the situation with 132-representation discussed in Section 4.1, any complete graph K_n can be represented by the 123-avoiding 2-uniform word $n(n-1)\cdots 1n(n-1)\cdots 1$ as observed in [21]. Also, it was shown in [21] that any path graph P_n can be 123-represented by a 2-uniform word. We conclude with a general type theorem on uniform representation applicable to both 123- and 132-representations.

Theorem 30 ([21])**.** *Let a pattern $\tau \in \{123, 132\}$ and $G_1, G_2, \ldots, G_k$ be τ-representable connected components of a graph G. Then G is τ-representable if and only if at most one of the connected components cannot be τ-represented by a 2-uniform word.*

5 Semi-transitive Orientations as a Key Tool in the Theory of Word-Representable Graphs

Recall the definition of a transitive orientation at the beginning of Section 3.

18

A *shortcut* is an *acyclic non-transitively oriented* graph obtained from a directed cycle graph forming a directed cycle on at least four vertices by changing the orientation of one of the edges, and possibly by adding more directed edges connecting some of the vertices (while keeping the graph be acyclic and non-transitive). Thus, any shortcut

- is *acyclic* (that it, there are *no directed cycles*);
- has *at least* 4 vertices;
- has *exactly one* source (the vertex with no edges coming in), *exactly one* sink (the vertex with no edges coming out), and a *directed path* from the source to the sink that goes through *every* vertex in the graph;
- has an edge connecting the source to the sink that we refer to as the *shortcutting edge*;
- is *not* transitive (that it, there exist vertices u, v and z such that $u \to v$ and $v \to z$ are edges, but there is *no* edge $u \to z$).

Definition 8. *An orientation of a graph is* semi-transitive *if it is* acyclic *and* shortcut-free.

It is easy to see from definitions that *any* transitive orientation is necessary semi-transitive. The converse is *not* true, e.g. the following schematic semi-transitively oriented graph is *not* transitively oriented:
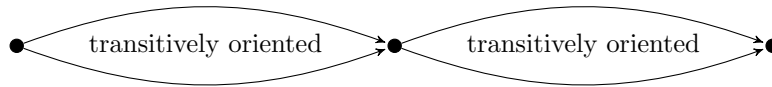
Thus semi-transitive orientations generalize transitive orientations.

A way to check if a given oriented graph G is semi-transitively oriented is as follows. First check that G is acyclic; if not, the orientation is not semi-transitive. Next, for a directed edge from a vertex x to a vertex y, consider *each* directed path P having at least three edges without repeated vertices from x to y, and check that the subgraph of G induced by P is transitive. If such non-transitive subgraph is found, the orientation is not semi-transitive. This procedure needs to be applied to each edge in G, and if no non-transitivity is discovered, G's orientation is semi-transitive.

As we will see in Theorem 31, finding a semi-transitive orientation is equivalent to recognising whether a given graph is word-representable, and this is an NP-hard problem (see Theorem 42). Thus, there is no efficient way to construct a semi-transitive orientation in general, and such a construction would rely on an exhaustive search orienting edges one by one, and thus branching the process. Having said that, there are several situations in which branching is not required. For example, the orientation of the partially oriented triangle below can be completed uniquely to avoid a cycle:
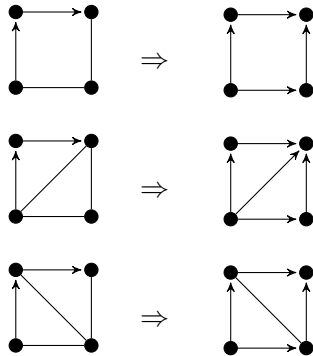
Figure 14: Completing orientations of quadrilaterals

For another example, the branching process can normally be shorten e.g. by completing the orientation of quadrilaterals as shown in Figure 14, which is *unique* to avoid cycles and shortcuts (the diagonal in the last case may require branching).

The main characterization theorem to date for word-representable graphs is the following result.

Theorem 31 ([11]). *A graph G is word-representable if and only if G admits a semi-transitive orientation.*

Proof. The backwards direction is rather complicated and is omitted. An algorithm was created in [11] to turn a semi-transitive orientation of a graph into a word-representant.

The idea of the proof for the forward direction is as follows (see [11] for details). Given a word, say, $w = 2421341$, orient the graph represented by w by letting $x \to y$ be an edge if the *leftmost x* is *to the left* of the *leftmost y* in w, to obtain a semi-transitive orientation:

□
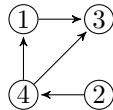
Any *complete graph* is 1-representable. The algorithm in [11] to turn semi-transitive orientations into word-representants gave the following result.

Theorem 32 ([11]). *Each non-complete word-representable graph G is $2(n - \kappa(G))$- representable, where $\kappa(G)$ is the size of the maximum clique in G.*

As an immediate corollary of Theorem 32, we have that the *recognition problem* of word-representability is in *NP*. Indeed, any word-representant is of length at most $O(n^2)$, and we need $O(n^2)$ passes through such a word to check alternation properties of all pairs of letters. There is an alternative proof of this complexity observation by Magnús M. Halldórsson in terms of *semi-transitive orientations*. In presenting his proof, we follow [17, Remark 4.2.3].

Checking that a given directed graph G is acyclic is a polynomially solvable problem. Indeed, it is well known that the entry (i, j) of the kth power of the adjacency matrix of G records the number of walks of length k in G from the vertex i to the vertex j. Thus, if G has n vertices, then we need to make sure that the diagonal entries are all 0 in all powers, up to the nth power, of the adjacency matrix of G. Therefore, it remains to show that it is polynomially solvable to check that G is shortcut-free. Let $u \to v$ be an edge in G. Consider the induced subgraph $H_{u \to v}$ consisting of vertices "in between" u and v, that is, the vertex set of $H_{u \to v}$ is

$$\{x \mid \text{there exist directed paths from } u \text{ to } x \text{ and from } x \text{ to } v\}.$$

It is not so difficult to prove that $u \to v$ is not a shortcut (that is, is not a shortcutting edge) if and only if $H_{u \to v}$ is transitive. Now, we can use the well known fact that finding out whether there exists a directed path from one vertex to another in a directed graph is polynomially solvable, and thus it is polynomially solvable to determine $H_{u \to v}$ (one needs to go through n vertices and check the existence of two paths for each vertex). Finally, checking transitivity is also polynomially solvable, which is not difficult to see.

The following theorem shows that word-representable graphs generalize the class of 3-colorable graphs.

Theorem 33 ([11])**.** *Any 3-colorable graph is word-representable.*

Proof. Coloring a 3-colorable graph in three colors, say, colors 1, 2 and 3, and orienting the edges based on the colors of their endpoints as $1 \to 2 \to 3$, we obtain a semi-transitive orientation. Indeed, obviously there are no cycles, and because the longest directed path involves only three vertices, there are no shortcuts. Theorem 31 can now be applied to complete the proof. $\square$

Theorem 33 can be applied to see, for example, that the Petersen graph is word-representable, which we already know from Section 2.3.1. More corollaries to Theorem 33 can be found below.

6 Non-word-representable Graphs

From the discussion in Section 3 we already know that the wheel graphs W_{2n+1}, for $n \geq 2$, are not word-representable, and that W_5 is the minimum (by the number of vertices) non-word-representable graph. But then, taking into account the hereditary nature of word-representability (see Remark 1), we have a family $\mathcal{W}$ of non-word-representable graphs characterised by containment of W_{2n+1} ($n \geq 2$) as an induced subgraph.

Note that *each* graph in $\mathcal{W}$ necessarily contains a vertex of degree 5 or more, and also a triangle as an induced subgraph. Natural questions are if there are non-word-representable graphs of *maximum degree* 4, and also if there are *triangle-free* non-word-representable graphs. Both questions were answered in affirmative. The graph to the right in Figure 10, which was found in [4], addresses the first question, while the second question is addressed by the following construction presented in [10].

Let M be a 4-chromatic graph with girth at least 10 (such graphs exist by a result of Paul Erdős; see [17, Section 4.4] for details). The *girth* of a graph is the length of a shortest cycle contained in the graph. If the graph does not contain any cycles (that is, it is an acyclic graph), its girth is defined to be infinity. Now, for every path of length 3 in M add to M an edge connecting path's end vertices. Then the obtained graph is triangle-free and non-word-representable [10].

6.1 Enumeration of Non-word-representable Graphs

According to experiments run by Herman Z.Q. Chen, there are 1, 25 and 929 non-isomorphic non-word-representable connected graphs on six, seven and eight vertices, respectively. These numbers were confirmed and extended to 68,545 for nine vertices, and 4,880,093 for 10 vertices, using a constraint programming (CP)-based method by Özgür Akgün, Ian Gent and Christopher Jefferson.

Figure 15 created by Chen presents the 25 non-isomorphic non-word-representable graphs on seven vertices. Note that the only non-word-representable graph on six vertices is the wheel W_5. Further note that the case of seven vertices gives just 10 minimal non-isomorphic non-word-representable graphs, since 15 of the graphs in Figure 15 contain W_5 as an induced subgraphs (these graphs are the first 11 graphs, plus the 15th, 16th, 18th and 19th graphs).

6.2 Non-word-representable Line Graphs

The *line graph* of a graph $G = (V, E)$ is the graph with vertex set E in which two vertices are adjacent if and only if the corresponding edges
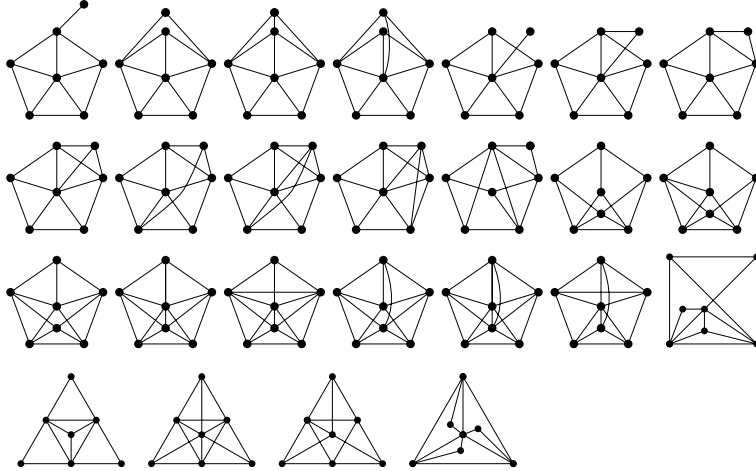
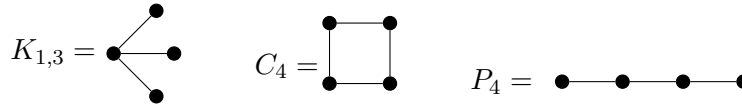Figure 15: 25 non-isomorphic non-word-representable graphs on seven vertices

Figure 16: The claw graph $K_{1,3}$, the cycle graph C_4, and the path graph P_4

of G share a vertex. The line graph of G is denoted $L(G)$. Line graphs give a tool to construct non-word-representable graphs as follows from the theorems below.

Theorem 34 ([19]). *Let $n \geq 4$. For any wheel graph W_n, the line graph $L(W_n)$ is non-word-representable.*

Theorem 35 ([19]). *Let $n \geq 5$. For any complete graph K_n, the line graph $L(K_n)$ is non-word-representable.*

The following theorem is especially interesting as it shows how to turn essentially any graph into non-word-representable graph.

Theorem 36 ([19]). *If a connected graph G is not a path graph, a cycle graph, or the claw graph $K_{1,3}$, then the line graph $L^n(G)$ obtained by application of the procedure of taking the line graph to G n times, is non-word-representable for $n \geq 4$.*

7 Word-Representability and Operations on Graphs

In this section we consider some of the most basic operations on graphs, namely, taking the complement, edge subdivision, edge contraction, connecting two graphs by an edge and gluing two graphs in a clique, replacing a vertex with a module, Cartesian product, rooted product and taking line graph.

We do not consider edge-addition/deletion trivially not preserving (non)-word-representability, although there are situations when these operations may preserve word-representability. For example, it is shown in [13] that edge-deletion preserves word-representability on K_4-free word-representable graphs.

Finally, we do not discuss the operations $Y \to \Delta$ (replacing an induced subgraph $K_{1,3}$, the claw, on vertices v_0, v_1, v_2, v_3, where v_0 is the apex, by the triangle on vertices v_1, v_2, v_3 and removing v_0) and $\Delta \to Y$ (removing the edges of a triangle on vertices v_1, v_2, v_3 and adding a vertex v_0 connected to v_1, v_2, v_3) recently studied in [13] in the context of word-representability of graphs.

7.1 Taking the Complement

Starting with a word-representable graph and taking its complement, we may either obtain a word-representable graph or not. Indeed, for example, both any graph on at most five vertices and its complement are word-representable. On the other hand, let G be the graph formed by the 5-cycle (2,4,6,3,5) and an isolated vertex 1. The 5-cycle can be represented by the word 2542643653 (see Section 2.2.3 for a technique to represent cycle graphs) and thus the graph G can be represented by the word 112542643653. However, taking the complement of G, we obtain the wheel graph W_5, which is not word-representable.

Similarly, starting with a non-word-representable graph and taking its complement, we can either obtain a word-representable graph or not. Indeed, the complement of the non-word-representable wheel W_5 is word-representable, as is discussed above. On the other hand, the graph G having two connected components, one W_5 and the other one the 5-cycle C_5, is non-word-representable because of the induced subgraph W_5, while the complement of G also contains an induced subgraph W_5 (formed by the vertices of C_5 in G and any of the remaining vertices) and thus is also non-word-representable.

7.2 Edge Subdivision and Edge Contraction

Subdivision of graphs (see Definition 4) is based on subdivision of individual edges, and it is considered in Section 2.3.4 from 3-representability

point of view.

If we change "3-representable" by "word-representable" in Theorem 12 we would obtain a weaker, but clearly still true statement, which is not hard to prove directly via semi-transitive orientations. Indeed, each path of length at least 3 added instead of an edge e can be oriented in a "blocking" way, so that there would be no directed path between e's endpoints. Thus, edge subdivision does not preserve the property of being non-word-representable. The following theorem shows that edge subdivision may be preserved on some subclasses of word-representable graphs, but not on the others.

Theorem 37 ([13]). *Edge subdivision preserves word-representability on K_4-free word-representable graphs, and it does not necessarily preserve word-representability on K_5-free word-representable graphs.*

Recall the definition of edge contraction in Definition 5. By Theorem 12, contracting an edge in a word-representable graph may result in a non-word-representable graph, while in many cases, e.g. in the case of path graphs, word-representability is preserved under this operation.

On the other hand, when starting from a non-word-representable graph, a graph obtained from it by edge contraction can also be either word-representable or non-word-representable. For example, contracting any edge incident with the bottommost vertex in the non-word-representable graph to the right in Figure 10, we obtain a graph on six vertices that is different from W_5 and is thus word-representable. Finally, any non-word-representable graph can be glued in a vertex with a path graph P (the resulting graph will be non-word-representable), so that contracting any edge in the subgraph formed by P results in a non-word-representable graph.

7.3   Connecting two Graphs by an Edge and Gluing two Graphs in a Clique

In what follows, by *gluing* two graphs in a clique we mean the following operation. Suppose $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$ are cliques of size k in graphs G_1 and G_2, respectively. Then gluing G_1 and G_2 in a clique of size k means identifying each a_i with one b_j, for $i, j \in \{1, \ldots, k\}$ so that the neighbourhood of the obtained vertex $c_{i,j}$ is the union of the neighbourhoods of a_i and b_j.

By the hereditary nature of word-representability (see Remark 1), if at least one of two graphs is non-word-representable, then gluing the graphs in a clique, or connecting two graphs by an edge (with the endpoints belonging to different graphs) will result in a non-word-representable graph.

25

On the other hand, suppose that graphs G_1 and G_2 are word-representable. Then gluing the graphs in a vertex, or connecting the graphs by an edge will result in a word-representable graph. The latter statement is easy to see using the notion of semi-transitive orientation. Indeed, by Theorem 31 both G_1 and G_2 can be oriented semi-transitively, and gluing the oriented graphs in a vertex, or connecting the graphs by an edge oriented arbitrarily, will not result in any cycles or shortcuts created. In fact, it was shown in [15] that if G_1 is k_1-representable (such a k_1 must exist by Theorem 3) and G_2 is k_2-representable, then essentially always the graph obtained either by gluing G_1 and G_2 in a vertex or by connecting the graphs by an edge is $\max(k_1, k_2)$-representable.

Even though glueing two word-representable graphs in a vertex (clique of size 1) always results in a word-representable graph, this is not necessarily true for glueing graphs in an edge (clique of size 2) or in a triangle (clique of size 3). We refer to [17, Section 5.4.3] for the respective examples. Glueing two graphs in cliques of size 4 or more in the context of word-representability remains an unexplored direction.

7.4   Replacing a Vertex with a Module

A subset X of the set of vertices V of a graph G is a *module* if all members of X have the same set of neighbours among vertices not in X (that is, among vertices in $V \setminus X$). For example, Figure 17 shows replacement of the vertex 1 in the triangular prism by the module K_3 formed by the vertices a, b and c. Thus, $\{a, b, c\}$ is a module of the graph on the right in Figure 17.
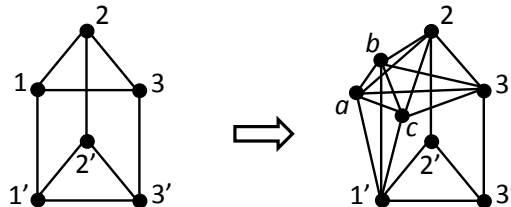

Figure 17: Replacing a vertex by a module

Theorem 38 ([15]). *Suppose that $G = (V, E)$ is a word-representable graph and $x \in V$. Let G' be obtained from G by replacing x with a module M, where M is any comparability graph (in particular, any clique). Then G' is also word-representable. Moreover, if $\mathcal{R}(G) = k_1$ and $\mathcal{R}(M) = k_2$ then $\mathcal{R}(G') = k$, where $k = \max\{k_1, k_2\}$.*

7.5 Cartesian Product of two Graphs

The *Cartesian product* $G\square H$ of graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$ is a graph such that

- the vertex set of $G\square H$ is the Cartesian product $V(G) \times V(H)$; and

- any two vertices (u, u') and (v, v') are adjacent in $G\square H$ if and only if either

  - $u = v$ and u' is adjacent to v' in H, or
  - $u' = v'$ and u is adjacent to v in G.

See Figure 18 for an example of the Cartesian product of two graphs.
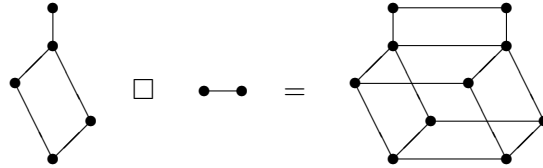

Figure 18: Cartesian product of two graphs

A proof of the following theorem was given by Bruce Sagan in 2014. The proof relies on semi-transitive orientations and it can be found in [17, Section 5.4.5].

Theorem 39 (Sagan). *Let G and H be two word-representable graphs. Then the Cartesian product $G\square H$ is also word-representable.*

7.6 Rooted Product of Graphs

The *rooted product* of a graph G and a rooted graph H (i.e. one vertex of H is distinguished), $G \circ H$, is defined as follows: take $|V(G)|$ copies of H, and for every vertex v_i of G, identify v_i with the root vertex of the ith copy of H. See Figure 19 for an example of the rooted product of two graphs.
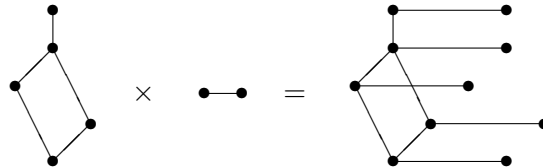

Figure 19: Rooted product of two graphs

The next theorem is an analogue of Theorem 39 for the rooted product of two graphs.

Theorem 40 ([17])**.** *Let G and H be two word-representable graphs. Then the rooted product $G \circ H$ is also word-representable.*

Proof. Identifying a vertex v_i in G with the root vertex of the ith copy of H in the definition of the rooted product gives a word-representable graph by the discussion in Section 7.3. Thus, identifying the root vertices, one by one, we will keep obtaining word-representable graphs, which gives us at the end word-representability of $G \circ H$. □

7.7 Taking the Line Graph Operation

Taking the line graph operation has already been considered in Section 6.2. Based on the results presented in that section, we can see that this operation can turn a word-representable graph into either a word-representable graph or non-word-representable graph. Also, there are examples of when the line graph of a non-word-representable graph is non-word-representable. However, it remains an open problem whether a non-word-representable graph can be turned into a word-representable graph by applying the line graph operation.

8 Computational Complexity Results and Word-Representability of Planar Graphs

In this section we will present known complexity results and also discuss word-representability of planar graphs.

8.1 A Summary of Known Complexity Results

Even though the Maximum Clique problem is polynomially solvable on word-representable graphs (see Theorem 18), many classical optimization problems are NP-hard on these graphs. The latter follows from the problems being NP-hard on 3-colorable graphs and Theorem 33.

The justification of the known complexity results presented in Table 2, as well as the definitions of the problems can be found in [17, Section 4.2]. However, below we discuss a proof of the fact that recognizing word-representability is an NP-complete problem. We refer to [17, Section 4.2] for any missed references to the results we use.

Suppose that P is a poset and x and y are two of its elements. We say that x *covers* y if $x > y$ and there is no element z in P such that $x > z > y$.

The *cover graph* G_P of a poset P has P's elements as its vertices, and $\{x, y\}$ is an edge in G_P if and only if either x covers y, or vice

problem	complexity
deciding whether a given graph is word-representable	NP-complete
approximating the graph representation number within a factor of $n^{1-\epsilon}$ for any $\epsilon > 0$	NP-hard
Clique Covering	NP-hard
deciding whether a given graph is k-word-representable for any fixed k, $3 \leq k \leq \lceil n/2 \rceil$	NP-complete
Dominating Set	NP-hard
Vertex Colouring	NP-hard
Maximum Clique	in P
Maximum Independent Set	NP-hard

Table 2: Known complexities for problems on word-representable graphs

versa. The *diagram* of P, sometimes called a *Hasse diagram* or *order diagram*, is a drawing of the cover graph of G in the plane with x being higher than y whenever x covers y in P. The three-dimensional cube in Figure 6 is an example of a cover graph.

Vincent Limouzy observed in 2014 that semi-transitive orientations of triangle-free graphs are exactly the 2-*good orientations* considered in [22] by Pretzel (we refer to that paper for the definition of a k-good orientation). Thus, by Proposition 1 in [22] we have the following reformulation of Pretzel's result in our language.

Theorem 41 (Limouzy). *The class of triangle-free word-representable graphs is exactly the class of cover graphs of posets.*

It was further observed by Limouzy, that it is an NP-complete problem to recognize the class of cover graphs of posets. This implies the following theorem, which is a key complexity result on word-representable graphs.

Theorem 42 (Limouzy). *It is an NP-complete problem to recognize whether a given graph is word-representable.*

8.2   Word-Representability of Planar Graphs

Recall that not all planar graphs are word-representable. Indeed, for example, wheel graphs W_{2n+1}, or graphs in Figure 10, are not word-representable.

Theorem 43 ([10]). *Triangle-free planar graphs are word-representable.*

Proof. By Grötzch's theorem [23], *every* triangle-free planar graph is 3-colorable, and Theorem 33 can be applied. □
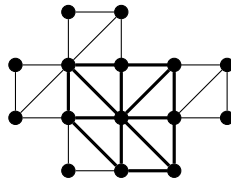
29

Figure 20: A triangulation of a poyomino

It remains a challenging open problem to classify word-representable planar graphs. Towards solving the problem, various *triangulations* and certain *subdivisions* of planar graphs were considered to be discussed next. Key tools to study word-representability of planar graphs are 3-*colorability* and *semi-transitive orientations*.

8.2.1 Word-Representability of Polyomino Triangulations

A *polyomino* is a plane geometric figure formed by joining one or more equal squares edge to edge. Letting corners of squares in a polyomino be vertices, we can treat polyominoes as graphs. In particular, well known *grid graphs* are obtained from polyominoes in this way. Of particular interest to us are *convex polyominoes*. A polyomino is said to be *column convex* if its intersection with any vertical line is convex (in other words, each column has no holes). Similarly, a polyomino is said to be *row convex* if its intersection with any horizontal line is convex. A polyomino is said to be *convex* if it is row and column convex.

When dealing with word-representability of triangulations of convex polyominoes (such as in Figure 20), one should watch for *odd wheel graphs* as induced subgraphs (such as the part of the graph in bold in Figure 20). Absence of such subgraphs will imply 3-colorability and thus word-representability, which is the basis of the proof of the following theorem.

Theorem 44 ([1]). *A triangulation of a* convex *polyomino is word-representable if and only if it is* 3-*colorable. There are non-3-colorable word-representable* non-convex *polyomino triangulations.*

The case of rectangular polyomino triangulations with a single domino tile (such as in Figure 21) is considered in the next theorem.

Theorem 45 ([8]). *A triangulation of a rectangular polyomino with a single domino tile is word-representable if and only if it is* 3-*colorable.*
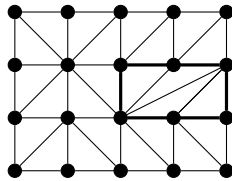
30

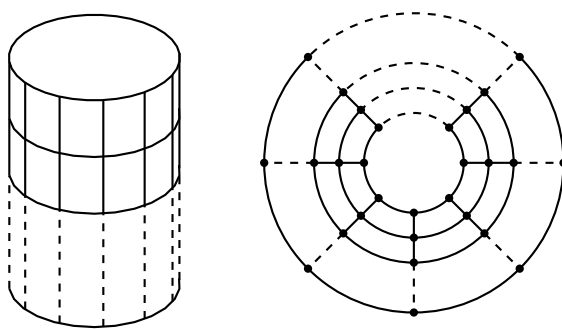Figure 21: A triangulation of a rectangular polyomino with a single domino tile

Figure 22: Grid-covered cylinder

8.2.2 Word-Representability of Near-Triangulations

A *near-triangulation* is a planar graph in which each *inner bounded face* is a triangle (where the *outer face* may possibly not be a triangle).

The following theorem is a far-reaching generalization of Theorems 44 and 45.

Theorem 46 ([6]). *A K_4-free near-triangulation is 3-colorable if and only if it is word-representable.*

Characterization of word-representable near-triangulations (containing K_4) is still an open problem.

8.2.3 Triangulations of Grid-covered Cylinder Graphs

A *grid-covered cylinder*, *GCC* for brevity, is a 3-dimensional figure formed by drawing vertical lines and horizontal circles on the surface of a cylinder, each of which are parallel to the generating line and the upper face of the cylinder, respectively. A GCC can be thought of as the object obtained by gluing the left and right sides of a rectangular grid. See the left picture in Figure 22 for a schematic way to draw a GCC. The vertical lines and horizontal circles are called the *grid lines.*
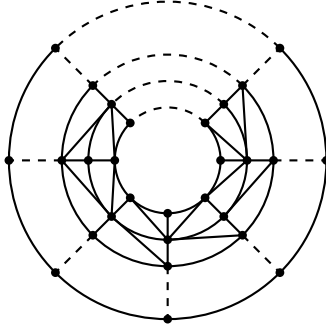
Figure 23: A triangulation of a GCCG

The part of a GCC between two consecutive vertical lines defines a *sector*.

Any GCC defines a graph, called *grid-covered cylinder graph*, or *GCCG*, whose set of vertices is given by intersection of the grid lines, and whose edges are parts of grid lines between the respective vertices. A typical triangulation of a GCCG is presented schematically in Figure 23.

Word-representability of triangulations of any GCCG is completely characterized by the following two theorems, which take into consideration the number of sectors in a GCCG.

Theorem 47 ([3]). *A triangulation of a GCCG with* more than three *sectors is word-representable if and only if it contains no W_5 or W_7 as an induced subgraph.*

Theorem 48 ([3]). *A triangulation of a GCCG with three sectors is word-representable if and only if it contains no graph in Figure 24 as an induced subgraph.*

8.2.4 Subdivisions of Triangular Grid Graphs

The *triangular tiling graph T^∞* is the *Archimedean tiling* 3^6 (see Figure 25). By a *triangular grid graph G* we mean a graph obtained from T^∞ as follows. Specify a finite number of triangles, called *cells*, in T^∞. The edges of G are then all the edges surrounding the specified cells, while the vertices of G are the endpoints of the edges (defined by intersecting lines in T^∞). We say that the specified cells, along with any other cell whose all edges are from G, *belong* to G.

The operation of *face subdivision of a cell* is putting a new vertex inside the cell and making it to be adjacent to every vertex of the cell.
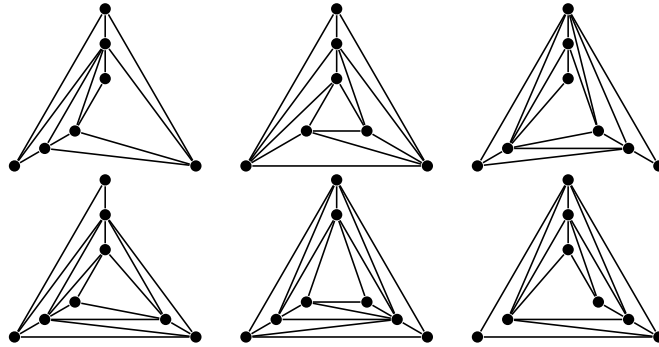
Figure 24: All minimal non-word-representable induced subgraphs in triangulations of GCCG's with three sectors
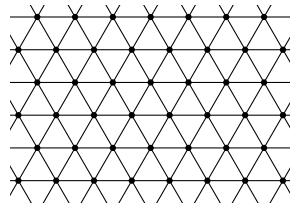

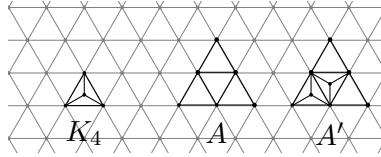Figure 25: A fragment of the graph T^∞

Figure 26: Examples of face subdivisions: K_4 is the face subdivision of a cell, and A' is a face subdivision of A
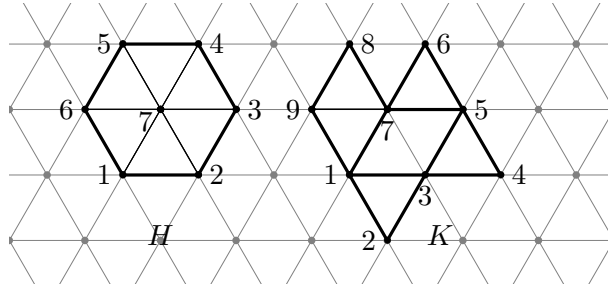

Figure 27: Graphs H and K, where boundary edges are in bold

Equivalently, face subdivision of a cell is replacing the cell (which is the complete graph K_3) by a plane version of the complete graph K_4. A *face subdivision of a set S of cells* of a triangular grid graph G is a graph obtained from G by subdividing each cell in S. The set S of subdivided cells is called a *subdivided set*. For example, Figure 26 shows K_4, the face subdivision of a cell, and A', a face subdivision of A.

If a face subdivision of G results in a word-representable graph, then the face subdivision is called a *word-representable face subdivision*. Also, we say that a word-representable face subdivision of a triangular grid graph G is *maximal* if subdividing any other cell results in a non-word-representable graph.

An edge of a triangular grid graph G shared with a cell in T^∞ that does not belong to G is called a *boundary edge*. A cell in G that is incident to at least one boundary edge is called a *boundary cell*. A non-boundary cell in G is called an *interior cell*. For example, the boundary edges in the graphs H and K in Figure 27 are in bold.

A face subdivision of a triangular grid graph that involves face subdivision of just boundary cells is called a *boundary face subdivision*. The following theorem was proved using the notion of a *smart orientation* (see [2] for details).
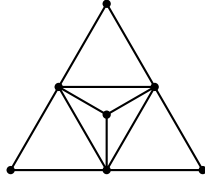
Figure 28: The graph A''

Theorem 49 ([2]). *A face subdivision of a triangular grid graph G is word-representable if and only if it has no induced subgraph isomorphic to A'' in Figure 28, that is, G has no subdivided interior cell.*

Theorem 49 can be applied to the *two-dimensional Sierpiński gasket graph $SG(n)$* to find its maximum word-representable subdivision (see [2] for details).

9 Directions for Further Research

In this section we list some of open problems and directions for further research related to word-representable graphs. The first question though the Reader should ask himself/herself is "Which graphs in their favourite class of graphs are word-representable?".

- Characterize (non-)word-representable planar graphs.
- Characterize word-representable near-triangulations (containing K_4).
- Describe graphs representable by words avoiding a pattern τ, where the notion of a "pattern" can be specified in any suitable way, e.g. it could be a *classical pattern*, a *vincular pattern*, or a *bivincular pattern* (see [14] for definitions).
- Is it true that out of all bipartite graphs on the same number of vertices, *crown graphs* require the *longest* word-representants?
- Are there any graphs on n vertices whose representation requires *more* than $\lfloor n/2 \rfloor$ copies of each letter?
- Is the *line graph* of a non-word-representable graph *always* non-word-representable?
- Characterize word-representable graphs in terms of *forbidden subgraphs*.
- Translate a known to you problem on *graphs* to *words* representing these graphs (assuming such words exist), and find an *efficient algorithm* to solve the *obtained problem*, and thus the *original problem*.

35

The last two problems are of fundamental importance.

References

[1] P. Akrobotu, S. Kitaev, Z. Masárová. On word-representability of polyomino triangulations. *Siberian Adv. Math.* **25** (2015) 1, 1–10.

[2] T. Z. Q. Chen, S. Kitaev, B. Y. Sun. Word-representability of face subdivisions of triangular grid graphs. *Graphs and Combinatorics* **32(5)** (2016) 1749–1761.

[3] T. Z. Q. Chen, S. Kitaev, B. Y. Sun. Word-representability of triangulations of grid-covered cylinder graphs. *Discr. Appl. Math.* **213(C)** (2016) 60–70.

[4] A. Collins, S. Kitaev, V. Lozin. New results on word-representable graphs. *Discr. Appl. Math.* **216** (2017) 136–141.

[5] A. Gao, S. Kitaev, P. Zhang. On 132-representable graphs. arXiv:1602.08965 (2016)

[6] M. Glen. Colourability and word-representability of near-triangulations. arXiv:1605.01688 (2016)

[7] M. Glen. Software available at `personal.cis.strath.ac.uk/sergey.kitaev/word-representable-graphs.html`

[8] M. Glen, S. Kitaev. Word-representability of triangulations of rectangular polyomino with a single domino tile. *J. Combin. Math. Combin. Comput.*, to appear.

[9] M. Halldórsson, S. Kitaev, A. Pyatkin. Graphs capturing alternations in words. *Lecture Notes Comp. Sci.* **6224** (2010) 436–437.

[10] M. Halldórsson, S. Kitaev, A. Pyatkin. Alternation graphs. *Lecture Notes Comp. Sci.* **6986** (2011) 191–202.

[11] M. Halldórsson, S. Kitaev, A. Pyatkin. Semi-transitive orientations and word-representable graphs. *Discr. Appl. Math.* **201** (2016) 164–171.

[12] M. Jones, S. Kitaev, A. Pyatkin, J. Remmel. Representing graphs via pattern avoiding words. *Electron. J. Combin.* **22(2)** (2015), #P2.53, 20 pp.

[13] J. Kim, M. Kim. Graph orientations on word-representable graphs. In preparation.

[14] S. Kitaev. Patterns in permutations and words, *Springer*, 2011.

[15] S. Kitaev. On graphs with representation number 3. *J. Autom., Lang. and Combin.* **18** (2013) 2, 97–112.

[16] S. Kitaev. Existence of u-representation of graphs, *J. Graph Theory*, to appear.

[17] S. Kitaev, V. Lozin. Words and Graphs, *Springer*, 2015.

[18] S. Kitaev, A. Pyatkin. On representable graphs. *J. Autom., Lang. and Combin.* **13** (2008) 1, 45–54.

[19] S. Kitaev, P. Salimov, C. Severs, H. Úlfarsson. On the representability of line graphs. *Lecture Notes Comp. Sci.* (2011)

[20] S. Kitaev, S. Seif. Word problem of the Perkins semigroup via directed acyclic graphs. *Order* **25** (2008) 3, 177–194.

[21] Y. Mandelshtam. On graphs representable by pattern-avoiding words. arXiv:1608.07614. (2016)

[22] O. Pretzel. On graphs that can be oriented as diagrams of ordered sets. *Order* **2** (1985) 1, 25–40.

[23] C. Thomassen. A short list color proof of Grötzsch's theorem. *J. Combin. Theory, Ser. B* **88(1)** (2003) 189 – 192.