Scientific Research

Word-Representability of Line Graphs

Sergey Kitaev[1,3*], Pavel Salimov[1,2†], Christopher Severs[1‡], Henning Ulfarsson[1§]
[1]*School of Computer Science, Reykjavik University, Reykjavik, Iceland*
[2]*Department of Theoretical Cybernetics, Sobolev Institute of Mathematics,*
Novosibirsk, Russia
[3]*Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK*
E-mail: *sergey.kitaev@gmail.com*

Abstract

A graph $G = (V, E)$ is representable 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) is in E for each x not equal to y. The motivation to study representable graphs came from algebra, but this subject is interesting from graph theoretical, computer science, and combinatorics on words points of view. In this paper, we prove that for n greater than 3, the line graph of an n-wheel is non-representable. This not only provides a new construction of non-representable graphs, but also answers an open question on representability of the line graph of the 5-wheel, the minimal non-representable graph. Moreover, we show that for n greater than 4, the line graph of the complete graph is also non-representable. We then use these facts to prove that given a graph G which is not a cycle, a path or a claw graph, the graph obtained by taking the line graph of G k-times is guaranteed to be non-representable for k greater than 3.

Keywords: Line Graph, Representability by Words, Wheel, Complete Graph

1. Introduction

A graph $G = (V, E)$ is representable 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 $x \neq y$. Such a W is called a *word-representant* of G. Note that in this paper we use the term graph to mean a finite, simple graph, even though the definition of representable is applicable to more general graphs.

It was shown by Kitaev and Pyatkin, in [1], that if a graph is representable by W, then one can assume that W is *uniform*, that is, it contains the same number of copies of each letter. If the number of copies of each letter in W is k, we say that W is k-*uniform*. For example, the graph to the left in **Figure 1** can be represented by the 2-uniform word 12312434 (in this word every pair of letters alternate, except 1 and 4, and 2 and 4), while the graph to the right, the Petersen graph,

can be represented by the 3-uniform word 02761859638243017296574908 3451 (the Petersen graph cannot be represented by a 2-uniform word as shown in [2])

The notion of a representable graph comes from algebra, where it was used by Kitaev and Seif to study the growth of the free spectrum of the well known *Perkins semigroup* [3]. There are also connections between representable graphs and robotic scheduling as described by Graham and Zang in [4]. Moreover, representable graphs are a generalization of *circle graphs*, which was shown by Halldórsson, Kitaev and Pyatkin in [5], and thus they are interesting from a graph theoretical point of view. Finally, representable graphs are interesting from a combinatorics on words point of view as they deal with the study of alternations in words.

Not all graphs are representable. Examples of minimal (with respect to the number of nodes) non-representable graphs given by Kitaev and Pyatkin in [1] are presented in **Figure 2**.

It was remarked in [5] that very little is known about the effect of the line graph operation on the representability of a graph. We attempt to shed some light on this subject by showing that the line graph of the smallest
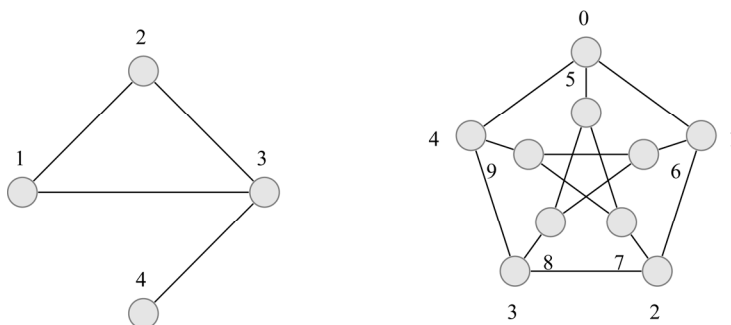
Figure 1. A graph representable by a 2-uniform word and the Petersen graph.
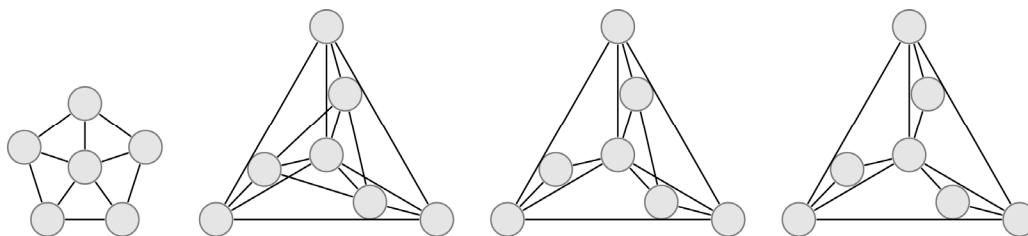

Figure 2. Minimal non-representable graphs.

known non-representable graph, the wheel on five vertices, W_5, is in fact non-representable. In fact we prove a stronger result, which is that $L(W_n)$ (where $L(G)$ denotes the line graph of G) is non-representable for $n \geq 4$. From the non-representability of $L(W_4)$ we are led to a more general theorem regarding line graphs. Our main result is that $L^k(G)$, where G is not a cycle, a path or the claw graph, is guaranteed to be non-representable for $k \geq 4$.

Although almost all graphs are non-representable (as discussed in [1]) and even though a criteria in terms of *semi-transitive orientations* is given in [5] for a graph to be representable, essentially only two explicit constructions of non-representable graphs are known. Apart from the so-called $co\text{-}(T_2)$ graph whose non-representability is proved in [2] in connection with solving an open problem in [1], the known constructions of non-representable graphs can be described as follows. Note that the property of being representable is hereditary, *i.e.*, it is inherited by all induced subgraphs, thus adding additional nodes to a non-representable graph and connecting them in an arbitrary way to the original nodes will also result in a non-representable graph.

- Adding an all-adjacent node to a *non-comparability graph* results in a non-representable graph (all of the graphs in **Figure 2** are obtained in this way). This construction is discussed in [1].
- Let H be a 4-chromatic graph with *girth* (the length of the shortest cycle) at least 10 (such graphs exist by a theorem of Erdös). For every path of length 3 in H add a new edge connecting the ends of the path. The

resulting graph will be non-representable as shown in [5]. This construction gives an example of triangle-free non-representable graphs whose existence was asked for in [1].

Our results showing that $L(W_n)$, $n \geq 4$, and $L(K_n)$, $n \geq 5$, are non-representable give two new constructions of non-representable graphs.

Our main result about repeatedly taking the line graph, shown in Section 5, also gives a new method for constructing non-representable graphs when starting with an arbitrary graph (excluding cycles, paths and the claw graph of course). Since we can start with an arbitrary graph this should also allow one to construct non-representable graphs with desired properties by careful selection of the original graph.

Although we have answered some questions about the line graph operation, there are still open questions related to the representability of the line graph, and in Sect. 6 we list some of these problems.

2. Preliminaries on Words and Basic Observations

2.1. Introduction to Words

We denote the set of finite words on an alphabet Σ by Σ^* and the empty word by ε.

A *morphism* φ is a mapping $\Sigma^* \to \Sigma^*$ that satisfies the property $\varphi(uv) = \varphi(u)\varphi(v)$ for all words u, v. Clearly, the morphism is completely defined by its action on the letters of the alphabet. The *erasing* of a set $\Sigma \setminus S$

of symbols is a morphism $\varepsilon_S : \Sigma^* \to \Sigma^*$ such that $\varepsilon_S(a) = a$ if $a \in S$ and $\varepsilon_S(a) = \varepsilon$ otherwise.

A word u *occurs* in a word $v = v_0 v_1 \cdots v_n$ at the position m and is called a *subword* of v if $u = v_m v_{m+1} \cdots v_{m+k}$ for some m, k. A subword that occurs at position 0 in some word is called a *prefix* of that word. A word is m-*uniform* if each symbol occurs in it exactly m times. We say that a word is uniform if it is m-uniform for some m.

Symbols a, b *alternate* in a word u if both of them occur in u and after erasing all other letters in u we get a subword of $abab\cdots$.

The *alternating graph* G of a word u is a graph on the symbols occurring in u such that G has an edge (a,b) if and only if a, b alternate in u. A graph G is *representable* if it is the alternating graph of some word u. We call u a *representant* of G in this case.

A key property of representable graphs was shown by Kitaev and Pyatkin in [1]:

Theorem 1 *Each representable graph has a uniform representant.*

Assuming uniformity makes dealing with the representant of a graph a much nicer task and plays a crucial role in some of our proofs.

2.2. Basic Observations

A *cyclic shift* of a word $u = u_0 u_1 \cdots u_n$ is the word $Cu = u_1 u_2 \cdots u_n u_0$.

Proposition 2 *Uniform words* $u = u_0 u_1 \cdots u_n$ *and* Cu *have the same alternating graph.*

Proof. Alternating relations of letters not equal to u_0 are not affected by the cyclic shift. Thus we need only to prove that u_0 has the same alternating relations with other symbols in Cu as it had in u.

Suppose u_0, u_i alternate in u. Due to u being uniform, it must be that $\varepsilon_{\{u_0, u_i\}}(u) = (u_0 u_i)^m$, where m is the uniform number of u. In this case, $\varepsilon_{\{u_0, u_i\}}(Cu) = u_i(u_0 u_i)^{m-1} u_0$ and hence the symbols u_0, u_i alternate in Cu.

Suppose u_0, u_i do not alternate u. Since u is uniform, $u_i u_i$ is a subword of $\varepsilon_{\{u_0, u_i\}}(u)$. Also, we know that $u_i u_i$ cannot be the prefix of u, so it must occur in $\varepsilon_{\{u_0, u_i\}}(Cu)$ too. Hence, u_0, u_i do not alternate in Cu.

Taking into account this fact, we may consider representants as cyclic or infinite words in order not to treat differently the end of the word while considering a local part of it.

Let us denote a clique on n vertices by K_n. One can easily prove the following proposition.

Proposition 3 *An m-uniform word that is a representant of K_n is a word of the form v^m where v is 1-uniform word containing n letters.*

Let us consider another simple case, the cycle C_n on n vertices.

Lemma 4 *The word $0, 1, 2, \cdots, n$ is not a subword of any uniform representant of C_{n+1} with vertices labeled in consecutive order, where $n \geq 3$.*

Proof. Suppose, u is a uniform representant of C_{n+1} and $v = 0, 1, 2, \ldots, n$ is a subword of u. Due to Proposition 2 we may assume that v is a prefix of u. Define a_i to be the position of the i-th instance of a in u for $a \in \{0, 1, \cdots, n\}$. Now for all adjacent vertices $a < b$ we have $a_i < b_i < a_{i+1}$ for each $i \geq 1$.

Vertices 0, 2 are not adjacent in C_{n+1} and so do not alternate in u. It follows that there is a $k \geq 1$ such that $0_k > 2_k$ or $2_k > 0_{k+1}$.

Suppose $2_k < 0_k$. Since $1, 2$ and $0, 1$ are adjacent, we have $1_i < 2_i$ and $0_i < 1_i$ for each i. Then we have a contradiction $0_k < 1_k < 2_k < 0_k$.

Suppose $0_{k+1} < 2_k$. Since all pairs $j, j+1$ and the pair $n, 0$ are adjacent, we have inequalities $j_i < (j+1)_i$ for each $j < n$, $i \geq 0$, and $n_i < 0_{i+1}$ for each $i \geq 0$. Thus we get a contradiction
$$2_k < 3_k < \cdots < n_k < 0_{k+1} < 2_k.$$

Here we introduce some notation. Let u be a representant of some graph G that contains a set of vertices $S = S_0 \cup S_1 \cup \{a\}$ such that $a \notin S_0 \cup S_1$ and $S_0 \cap S_1 = \varnothing$. We use the notation $\forall(aS_0 S_1 a)$ for the statement "Between every two consecutive occurrences of a in $C^n u$, for every n, each symbol of $S_0 \cup S_1$ occurs once and each symbol of S_0 occurs before any symbol of S_1" and the notation $\exists(aS_0 S_1 a)$ for the statement "There are two consecutive occurrences of a in $C^n u$, for at least one n, such that each symbol of $S_0 \cup S_1$ occurs between them and each symbol of S_0 occurs before any symbol of S_1" Note that $\forall(aS_0 S_1 a)$ implies $\exists(aS_0 S_1 a)$ and is contrary for $\exists(aS_1 S_0 a)$. The quantifiers in these statements operate on pairs of consecutive occurrences of a in all cyclic shifts of the given representant. This notation may be generalized to an arbitrary number of sets S_i with the same interpretation.

The following proposition illustrates the use of this notation.

Proposition 5 *Let a word u be a representant of some graph G containing vertices a, b, c, where a, b and a, c are adjacent. Then we have*

1) b, c *being not adjacent implies that both of the statements $\exists(abca)$, $\exists(acba)$ are true for u,*

2) b, c *being adjacent implies that exactly one of the statements $\forall(abca)$, $\forall(acba)$ is true for u.*

Proof. (Case 1) Since a,b and a,c alternate, at least one of $\exists(abca)$, $\exists(acba)$ is true. If only one of them is true for u, then b,c alternate in it, which is a contradiction with b,c being not adjacent.

(Case 2) the statement follows immediately from Proposition 3.

3. Line Graphs of Wheels

The *wheel graph*, denoted by W_n, is a graph we obtain from a cycle C_n by adding one external vertex adjacent to every other vertex.

A line graph $L(G)$ of a graph G is a graph on the set of edges of G such that in $L(G)$ there is an edge (a,b) if and only if edges a,b are adjacent in G.

Theorem 6 *The line graph* $L(W_{n+1})$ *is not representable for each* $n \geq 3$.

Proof. Let us describe $L(W_{n+1})$ first. Denote edges of the big (external) cycle of the wheel W_n by $e_0, e_1, \cdots, e_n$ in consecutive order and internal edges that connect the inside vertex to the big cycle by $i_0, i_1, \cdots, i_n$ so that an edge i_j is adjacent to e_j and e_{j+1} for $0 \leq j < n$ and i_n is adjacent to e_n, e_0.

In the line graph $L(W_{n+1})$ the vertices $e_0, e_1, \cdots, e_n$ form a cycle where they occur consecutively and the vertices $i_0, i_1, \cdots, i_n$ form a clique. In addition, vertices i_j are adjacent to e_j, e_{j+1} and i_n is adjacent to e_n, e_0.

Suppose that $L(W_{n+1})$ is the alternating graph of some word that, due to Theorem 1, can be chosen to be uniform. Now we deduce a contradiction with Lemma 4.

Let E be the alphabet $\{e_j : 0 \leq j \leq n\}$, I be the alphabet $\{i_j : 0 \leq j \leq n\}$ and a word u on the alphabet $E \cup I$ be the uniform representant of $L(W_{n+1})$. Due to Proposition 2, we may assume $u_0 = i_0$.

As we know from Proposition 3, the word $\varepsilon_I(u)$ is of the form v^m, where v is 1-uniform and $v_0 = i_0$. Let us prove that v is exactly $i_0, i_1, \cdots, i_n$ or $i_0, i_n, i_{n-1}, \cdots, i_1$.

Suppose there are some $\ell, k \in \{2, \cdots, n\}$ such that $\varepsilon_{\{i_0, i_1, i_\ell, i_k\}}(v) = i_0 i_\ell i_1 i_k$. Note, that $\ell \neq k$ due to v being 1-uniform. The supposition implies that the statement $\forall(i_0 i_\ell i_1 i_k i_0)$ is true for u. The vertex e_1 is neither adjacent to i_ℓ nor to i_k. By Proposition 5 this implies $\exists(i_0 e_1 i_\ell i_0)$ and $\exists(i_0 i_k e_1 i_0)$ are true for u. Taking into account the previous "for all" statement, we conclude that both of $\exists(i_0 e_1 i_1 i_0)$ and $\exists(i_0 i_1 e_1 i_0)$ are true for u, which contradicts Proposition 5 applied to i_0, i_1, e_1. So, there are only two possible cases, *i.e.*, $v = i_0, i_1, v_2, \cdots v_n$ and $v = i_0 v_1 \ldots i_1$.

Using the same reasoning on a triple i_j, i_{j+1}, e_{j+1}, by induction on $j \geq 1$, we get $v = i_0, i_1, \cdots, i_n$ for the first case and $v = i_0 i_n, i_{n-1}, \cdots, i_1$ for the second.

It is sufficient to prove the theorem only for the first case, since reversing a word preserves the alternating relation.

By Proposition 5 exactly one of the statements $\forall(i_0 e_0 e_1 i_0)$, $\forall(i_0 e_1 e_0 i_0)$ is true for u. Let us prove that it is the statement $\forall(i_0 e_0 e_1 i_0)$.

Applying Proposition 5 to the clique $\{i_0, i_1, e_1\}$ we have that exactly one of $\forall(i_0 i_1 e_1 i_0)$, $\forall(i_0 e_1 i_1 i_0)$ is true. Applying Proposition 5 to i_0, i_2, e_1 we have that both of $\exists(i_0 e_1 i_2 i_0)$ and $\exists(i_0 i_2 e_1 i_0)$ are true. The statement $\forall(i_0 e_1 i_1 i_0)$ contradicts $\exists(i_0 i_2 e_1 i_0)$ since we have $\forall(i_0 i_1 i_2 i_0)$. Hence $\forall(i_0 i_1 e_1 i_0)$ is true.

Now applying Proposition 5 to i_0, e_0 and i_1 we have $\exists(i_0 e_0 i_1 i_0)$. Taking into account $\forall(i_0 i_1 e_1 i_0)$ and Proposition 5 applied to the clique $\{i_0, e_0, e_1\}$ we conclude that $\forall(i_0 e_0 e_1 i_0)$ is true. In other words, between two consecutive i_0 in u there is e_0 that occurs before e_1.

Using the same reasoning, one can prove that the statement $\forall(i_n e_n e_0 i_n)$ and the statements $\forall(i_j e_j e_{j+1} i_j)$ for each $j < n$ are true for u. Let us denote this set of statements by (*).
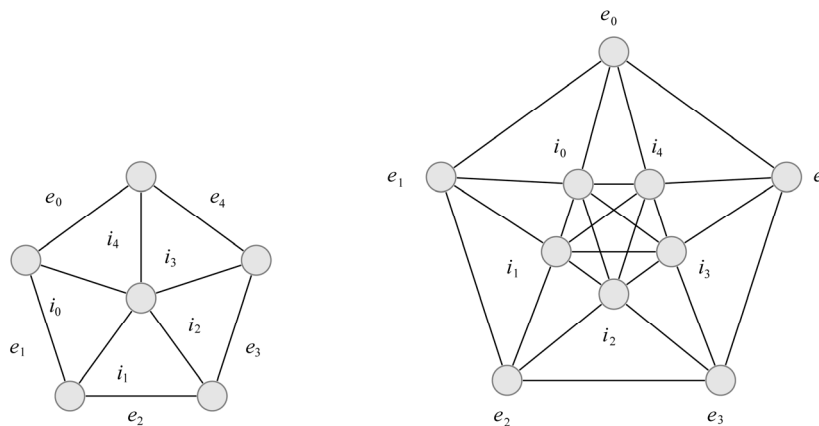

Figure 3. The wheel graph W_5 **and its line graph.**

The vertex e_0 is not adjacent to the vertex i_{n-1} but both of them are adjacent to i_0, hence, by Proposition 5, somewhere in $\varepsilon_{\{e_0, i_{n-1}, i_0\}}(u)$ the word $i_{n-1}e_0i_0$ occurs. Taking into account what we have already proved for v, this means that we found the structure

$$i_0 - i_1 - \cdots - i_{n-1} - e_0 - i_n - i_0 - i_1 - \cdots - e_n \quad \text{in} \quad u,$$ where symbols of I do not occur in gaps denoted by "$-$".

Now inductively applying the statements (*), we conclude that in u there is a structure

$$i_{n-1} - e_0 - e_1 - \cdots - e_n - i_{n-1}$$ where no symbol i_{n-1} occurs in the gaps. Suppose the symbol e_0 occurs somewhere in the gaps between e_0 and e_n. Since e_0 and e_n are adjacent, that would mean that between two e_0 another e_n also occurs and this contradicts the fact that e_n and i_{n-1} are adjacent. One may prove that no symbol of E occurs in the gaps between e_0 and e_n in the structure we found, by using induction and arriving at a contradiction similar to the one above. In other words, $e_0e_1\cdots e_n$ occurs in the word $\varepsilon_E(u)$ representing the cycle. This results in a contradiction with Lemma 4 which concludes the proof.

4. Line Graphs of Cliques

Theorem 7 *The line graph $L(K_n)$ is not representable for each $n \geq 5$.*

Proof. It is sufficient to prove the theorem for the case $n = 5$ since, as one can prove, any $L(K_{n\geq 5})$ contains an induced $L(K_5)$.

Let u be a representant of $L(K_5)$ with its vertices labeled as shown in **Figure 4**. Vertices $0, 1, a, b$ make a clique in $L(K_5)$. By applying Propositions 3 and 5 to this clique we see that exactly one of the following statements is true: $\forall(a\{0,1\}ba)$, $\forall(ab\{0,1\}a)$, $\forall(axb\overline{x}a)$, where $x \in \{0,1\}$ and $\overline{x}$ is the negation of x.

(Case 1) Suppose $\forall(axb\overline{x}a)$ is true. The vertex 3 is adjacent to a, b, but not to 0, 1. Keeping in mind that a is also adjacent to 0 and 1, then applying Proposition 5 we have that $\exists(a3\{0,1\}a)$ and $\exists(a\{0,1\}3a)$ are true. But between x, $\overline{x}$ there is b, so we have a contradiction $\exists(a3ba)$, $\exists(ab3a)$ with Proposition 5.

(Case 2a) Suppose $\forall(ab01a)$ is true. The vertex e is adjacent with a, 0, but not with b, 1. Applying Proposition 5 we have $\exists(aeba)$ and $\exists(a1ea)$. Taking into account the case condition, this implies $\exists(ae0a)$ and $\exists(a0ea)$ which is a contradiction.

(Case 2b) Suppose $\forall(ab10a)$ is true. The vertex 2 is adjacent with a, 1, but not with b, 0. Applying Proposition 5 we have $\exists(a2ba)$ and $\exists(a02a)$. Again, taking into account the case condition this implies $\exists(a21a)$ and $\exists(a12a)$, which gives a contradiction.

(Case 3a) If $\forall(a01ba)$ is true, a contradiction follows analogously to Case 2b.

(Case 3b) If $\forall(a10ba)$ is true, a contradiction follows analogously to Case 2a.

5. Iterating the Line Graph Construction

It was shown by van Rooji and Wilf [6] that iterating the line graph operator on most graphs results in a sequence of graphs which grow without bound. The exceptions are cycles, which stay as cycles of the same length, the claw graph $K_{1,3}$, which becomes a triangle after one iteration and then stays that way, and paths, which shrink to the empty graph. This unbounded growth results in graphs that are non-representable after a small number of iterations of the line graph operator since they contain the line graph of a large enough clique. A slight modification of this idea is used to prove our main result.

Theorem 8 *If a connected graph G is not a path, a cycle, or the claw graph $K_{1,3}$, then $L^n(G)$ is not representable for $n \geq 4$.*

Proof. Note that if H appears as a subgraph of G (not necessarily induced), then $L^n(H)$ is isomorphic to
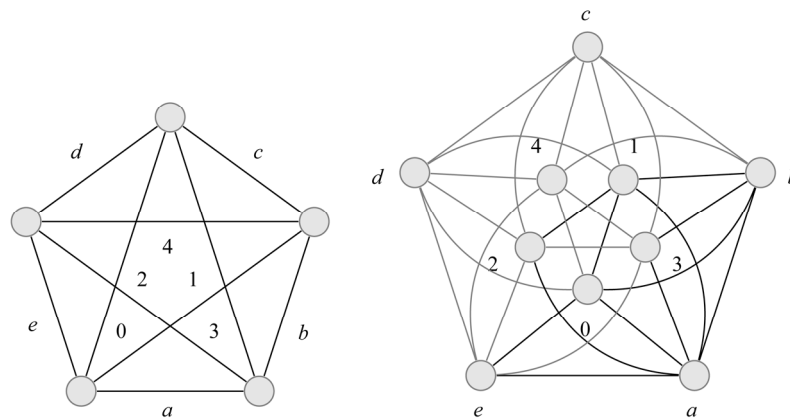

Figure 4. The clique K_5 and its line graph, where edges mentioned in the proof of Theorem 4 are drawn thicker.
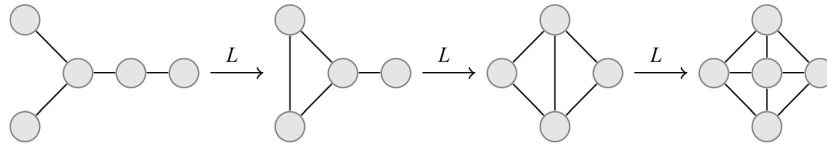
Figure 5. Iterating the line graph construction.

an induced subgraph of $L^n(G)$ for all $n \geq 1$.

We first consider the sequence of graphs in **Figure 5**. All but the leftmost graph are obtained by applying the line graph operator to the previous graph. The last graph in the sequence is W_4, and by Theorem 6, $L(W_4)$ is non-representable.

Now, let $G = (V, E)$ be a graph that is not a star and that satisfies the conditions of the theorem. G contains as a subgraph an isomorphic copy of either the leftmost graph of **Figure 5** or the second graph from the left. Thus $L^3(G)$, or respectively $L^4(G)$, is not representable, since it contains an induced line graph of the wheel W_4.

If G is a star $S_{k \geq 4}$ then $L(G)$ is the clique K_k and there is an isomorphic copy of the second from the left graph of **Figure 5** in G, and $L^4(G)$ is not representable again.

Note that there is an isomorphic copy of the second graph of **Figure 5** inside the third one. Therefore the same reasoning can be used for $L^{4+k}(G)$ for each $k \geq 1$, which concludes the proof.

6. Some Open Problems

We have the following open questions.

- Is the line graph of a non-representable graph always non-representable?

Our Theorem 8 shows that for any graph G, that is not a path, a cycle, or the claw $K_{1,3}$, the graph $L^n(G)$ is non-representable for all $n \geq 4$. It might be possible to find a graph G such that G is non-representable while $L(G)$ is.

- How many graphs on n vertices stay non-representable after at most i iterations, $i = 0, 1, 2, 3, 4$?

For a graph G define $\xi(G)$ as the smallest integer such that $L^k(G)$ is non-representable for all $k \geq \xi(G)$.

Theorem 8 shows that $\xi(G)$ is at most 4, for a graph that is not a path, a cycle, nor the claw $K_{1,3}$, while paths, cycles and the claw have $\xi(G) = +\infty$.

- Is there a finite classification of prohibited subgraphs in a graph G determining whether $L(G)$ is representable?

There is a classification of prohibited induced subgraphs which determine whether a graph G is the line graph of some other graph H. It would be nice to have such a classification, if one exists, to determine if $L(G)$ is representable.

7. References

[1] S. Kitaev and A. Pyatkin, "On Representable Graphs," *Journal of Automata, Languages and Combinatorics*, Vol. 13, No. 1, 2008, pp.45-54.

[2] M. M. Halldórsson, S. Kitaev and A. Pyatkin, "Graphs Capturing Alternations in Words," In *Proceedings of the 14th international conference on Developments in language theory*, London, August 17-20, 2010, pp. 436-437.

[3] S. Kitaev and S. Seif, "Word Problem of the Perkins Semigroup via Directed Acyclic Graphs," *Order*, Vol. 25, No. 3, 2008, pp. 177-194. doi:10.1007/s11083-008-9083-7

[4] R. Graham and N. Zang, "Enumerating Split-Pair Arrangements," *Journal of Combinatorial Theory*, *Series A*, Vol. 115, No. 2, 2008, pp. 293-303. doi:10.1016/j.jcta.2007.06.003

[5] M. M. Halldorsson, S. Kitaev and A. Pyatkin, "On Representable Graphs, Semi-Transitive Orientations, and the Representation Numbers," Submitted on 1 October 2008. http://arxiv.org/abs/0810.0310v1

[6] A. C. M. van Rooji and H. S. Wilf, "The Interchange Graph of a Finite Graph," *Acta Mathematica Academiae Scientiarum Hungaricae*, Vol. 16, No. 3-4, 1965, pp.263-269. doi:10.1007/BF01904834