Discussiones Mathematicae
Graph Theory 42 (2022) 1263–1280
https://doi.org/10.7151/dmgt.2344

REPRESENTING SPLIT GRAPHS BY WORDS

Herman Z.Q. Chen

School of Statistics and Data Science
Nankai University, P.R. China
e-mail: zqchern@163.com

Sergey Kitaev

Department of Mathematics and Statistics
University of Strathclyde
26 Richmond Street, Glasgow, G1 1XH, United Kingdom
e-mail: sergey.kitaev@cis.strath.ac.uk

and

Akira Saito

Department of Information Science
Nihon University
Sakurajosui 3-25-40 Setagaya-Ku Tokyo 156–8550, Japan
e-mail: asaito@chs.nihon-u.ac.jp.

Abstract

There is a long line of research in the literature dedicated to word-representable graphs, which generalize several important classes of graphs. However, not much is known about word-representability of split graphs, another important class of graphs.

In this paper, we show that threshold graphs, a subclass of split graphs, are word-representable. Further, we prove a number of general theorems on word-representable split graphs, and use them to characterize computationally such graphs with cliques of size 5 in terms of nine forbidden subgraphs, thus extending the known characterization for word-representable split graphs with cliques of size 4. Moreover, we use split graphs, and also provide an alternative solution, to show that gluing two word-representable graphs in any clique of size at least 2 may, or may not, result in a word-representable graph. The two surprisingly simple solutions provided by us answer a question that was open for about ten years.

Keywords: split graph, word-representability, semi-transitive orientation.

2010 Mathematics Subject Classification: 05C62.

1. Introduction

A graph $G = (V, E)$ is *word-representable* 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 $xy \in E$. Here, by *alternation* of x and y in w we mean that after removing *all* letters *but* the copies of x and y we either obtain a word $xyxy\cdots$, or a word $yxyx\cdots$. For example, the cycle graph C_5 labeled by 1–5 in clock-wise direction can be represented by the word 1521324354. It is easy to see that the class of word-representable graphs is *hereditary.* That is, removing a vertex in a word-representable graph results in a word-representable graph.

To date, many papers have been written on the subject [6], and the core of the book [8] is devoted to the theory of word-representable graphs. It should also be mentioned that the software produced by Marc Glen [3] is often of great help in dealing with such graphs. Word-representable graphs are important as they generalize several fundamental classes of graphs such as *circle graphs*, 3-*colorable graphs* and *comparability graphs* [8].

An orientation of a graph is *semi-transitive* if it is acyclic, and for any directed path $u_1 \to u_2 \to \cdots \to u_k$ either there is no edge between u_1 and u_k, or there is an edge $u_i \to u_j$ for all $1 \leq i < j \leq k$. A key result in the area is the following theorem.

Theorem 1 [5]**.** *A graph is word-representable if and only if it admits a semi-transitive orientation.*

In this paper, we will need the following simple lemma.

Lemma 2 [7]**.** *Let K_m be a clique in a graph G. Then any acyclic orientation of G induces a transitive orientation on K_m with a single source* (*a vertex with no incoming edges*) *and a single sink* (*a vertex with no outgoing edges*)*.*

Even though much is known about word-representable graphs, there is only one paper, namely [7], dedicated to the study of the word-representability of *split graphs* (considered, e.g. in [1, 2, 4, 9]), that is, graphs in which the vertices can be partitioned into a clique and an independent set. Section 2 overviews results in [7] that are most relevant to this paper and can be summarised as follows.

- Split graphs with cliques of size at most 3 are word-representable.
- Split graphs in which the clique is of size 4 are characterized by avoiding the four graphs in Figure 1 as induced subgraphs.
- Necessary and sufficient conditions for an orientation of a split graph to be semi-transitive are given.

The major results in this paper can be summarized as follows.

- The subclass of split graphs known as *threshold graphs* is shown to be word-representable in Theorem 12. Threshold graphs were first introduced by Chvátal and Hammer in [1]. A chapter on these graphs appears in [4], and the book [9] is devoted to them.
- Split graphs in which the clique is of size m and the clique's vertices are of degree at most m are word-representable (see Theorem 13).
- An upper bound on the number of vertices in the independent set of any given degree in a word-representable split graph is given (see Theorems 15 and 16).
- The upper bound is used to characterize computationally split graphs having the clique of size 5 in terms of rednine forbidden subgraphs — those in Figures 1 and 3 (see Section 5).
- Word-representability of split-graphs is used in Section 6 to show that gluing two word-representable graphs in a clique of size at least 2 may result in a non-word-representable graph, which answers a long standing, though unpublished until [6], open question. We also give an alternative solution to the problem, which is based on a generalization of a known result (see Section 6.2).

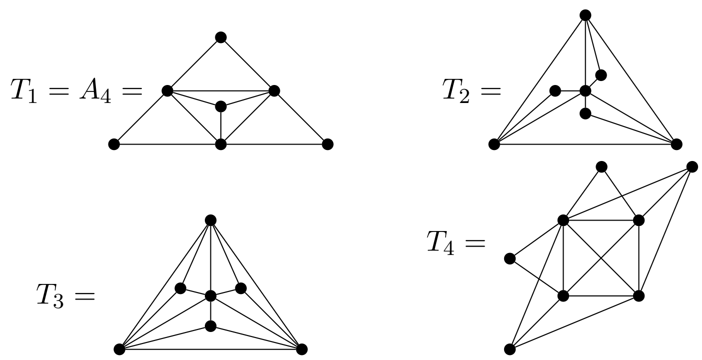

Figure 1. The minimal non-word-representable split graphs T_1, T_2, T_3, T_4.

2. SPLIT GRAPHS AND WORD-REPRESENTATION

Let $S_n = (E_{n-m}, K_m)$ be a split graph on n vertices, where the vertices of S_n are partitioned into a maximal clique K_m and an independent set E_{n-m} (the vertices in E_{n-m} are of degree at most $m-1$).

In this section, we overview results in [7] most relevant to us.

Lemma 3 [7]. *Let $S_n = (E_{n-m}, K_m)$ be a split graph, and a split graph S_{n+1} is obtained from S_n by either adding a vertex of degree* 0 *or* 1*, or by "copying" a* *vertex, that is, by adding a vertex whose neighbourhood is identical to the neighbourhood of a vertex in S_n (if copying a vertex in K_m, then the copy is connected to the original vertex). Then S_n is word-representable if and only if S_{n+1} is word-representable.*

Definition. For $\ell \geq 3$, the graph $K_\ell^\triangle$ is obtained from the complete graph K_ℓ with the vertices labeled by $1, \ldots, \ell$, by adding a vertex i' of degree 2 connected to vertices i and $i+1$ for each $i \in \{1, \ldots, \ell-1\}$. Also, a vertex ℓ' connected to the vertices 1 and ℓ is added.

Theorem 4 [7]**.** *$K_\ell^\triangle$ is word-representable.*

Definition. For $\ell \geq 4$, let A_ℓ be the graph obtained from $K_{\ell-1}^\triangle$ by adding a vertex ℓ connected to the vertices $1, \ldots, \ell-1$ and no other vertices. Note that $A_4 = T_1$ in Figure 1.

Theorem 5 [7]**.** *A_ℓ is a minimal non-word-representable graph.*

Theorem 6 [7]**.** *Let $S_n = (E_{n-4}, K_4)$ be a split graph. Then S_n is word-representable if and only if S_n does not contain the graphs T_1, T_2, T_3 and T_4 in Figure 1 as induced subgraphs.*

Let $S_n = (E_{n-m}, K_m)$ be a word-representable split graph. Then, by Theorem 1, S_n admits a semi-transitive orientation. Further, by Lemma 2 we know that any such orientation induces a transitive orientation on K_m with the longest directed path $\vec{P}$. Theorems 7 and 8 below describe the structure of semi-transitive orientations in an arbitrary word-representable split graph.

Theorem 7 [7]**.** *Any semi-transitive orientation of $S_n = (E_{n-m}, K_m)$ subdivides the set of all vertices in E_{n-m} into three, possibly empty, groups corresponding to each of the following types, where $\vec{P} = p_1 \to \cdots \to p_m$ is the longest directed path in K_m.*

- *A vertex in E_{n-m} is of* type A *if it is a source and is connected to all vertices in $\{p_i, p_{i+1}, \ldots, p_j\}$ for some $1 \leq i \leq j \leq m$.*
- *A vertex in E_{n-m} is of* type B *if it is a sink and is connected to all vertices in $\{p_i, p_{i+1}, \ldots, p_j\}$ for some $1 \leq i \leq j \leq m$.*
- *A vertex $v \in E_{n-m}$ is of* type C *if there is an edge $x \to v$ for each $x \in I_v = \{p_1, p_2, \ldots, p_i\}$ and there is an edge $v \to y$ for each $y \in O_v = \{p_j, p_{j+1}, \ldots, p_m\}$ for some $1 \leq i < j \leq m$.*

There are additional restrictions, given by the next theorem, on relative positions of the neighbours of vertices of types A, B and C.

Theorem 8 [7]**.** *Let $S_n = (E_{n-m}, K_m)$ be oriented semi-transitively with $\vec{P} = p_1 \to \cdots \to p_m$. For a vertex $x \in E_{n-m}$ of type C, there is no vertex $y \in E_{n-m}$ of type A or B, which is connected to both $p_{|I_x|}$ and $p_{m-|O_x|+1}$. Also, there is no vertex $y \in E_{n-m}$ of type C such that either I_y, or O_y contains both $p_{|I_x|}$ and $p_{m-|O_x|+1}$.*

One can now classify semi-transitive orientations on split graphs.

Theorem 9 [7]**.** *An orientation of a split graph $S_n = (E_{n-m}, K_m)$ is semi-transitive if and only if*

- *K_m is oriented transitively,*
- *each vertex in E_{n-m} is of one of the three types in Theorem* 7,
- *the restrictions in Theorem* 8 *are satisfied.*

The following corollary of Theorem 9 generalizes Theorem 4.

Corollary 10 [7]**.** *Let the split graph K_ℓ^k be obtained from the complete graph K_ℓ, whose vertices are drawn on a circle, by adding ℓ vertices so that*

- *each such vertex is connected to k consecutive* (*on the circle*) *vertices in K_ℓ;*
- *neighbourhoods of all these vertices are distinct; and*
- *$\ell \geq 2k - 1$.*

Then K_ℓ^k is word-representable.

The following theorem allows us to treat vertices of types A or B in the same way and to refer to them as vertices of type A&B.

Theorem 11 [7]**.** *Let $S_n = (E_{n-m}, K_m)$ be semi-transitively oriented. Then, any vertex in E_{n-m} of type A can be replaced by a vertex of type B, and vice versa, keeping orientation semi-transitive.*

3. Threshold Graphs and Split Graphs with Restricted Vertex Degree in the Clique

A *threshold graph* is a graph that can be constructed from the one-vertex graph by repeated applications of the following two operations.

(1) Addition of a single isolated vertex to the graph.

(2) Addition of a single dominating vertex to the graph, i.e., a single vertex that is connected to all other vertices.

It is not difficult to see that any threshold graph is a split graph.

Theorem 12. *Any threshold graph S_n is word-representable.*

Proof. Label the vertices in the order they were added to S_n: $1, 2, \ldots, n$. Note that no matter which operation is applied, the vertices 1 and 2 will have the same neighbourhood modulo the vertices 1 and 2 possibly being connected to each other. Thus, by Lemma 3, removing vertex 1 does not affect word-representability of the graph. But then, the vertices 2 and 3 will have the same neighbourhood modulo them possibly being connected to each other. Thus, by Lemma 3, removing vertex 2 does not affect word-representability of the graph. Continuing in the same way, we see that S_n is word-representable if and only if the one-vertex graph (labeled by n) is word-representable, which is trivially the case. ■

Theorem 13. *Let $S_n = (E_{n-m}, K_m)$ be a split graph such that each vertex v in K_m is of degree at most m, i.e., the degree of v is $m-1$ or m. Then S_n is word-representable.*

Proof. Orient K_m in an arbitrary transitive way, which will result, by Lemma 2, in a longest directed path $\vec{P} = p_1 \to \cdots \to p_m$. Because each vertex in K_m can be connected to at most one vertex in E_{n-m}, we can clearly permute the vertices in $\vec{P}$ (resulting in a different transitive orientation of K_m) so that the neighbourhood of each vertex in E_{n-m} consists of a number of consecutive vertices in $\vec{P}$, and these neighbourhoods do not overlap. Making each vertex in E_{n-m} either of type A, or of type B, we can apply Theorem 9 to see that S_n is semi-transitively oriented, and thus, by Theorem 1, S_n is word-representable. ■

Theorem 14. *Let $S_n = (E_{n-m}, K_m)$ be a split graph, where the neighbourhoods of all vertices in E_{n-m} are distinct. If K_m has a vertex v connected to at least $d+1$ vertices of degree $d \leq m-2$ in E_{n-m}, then S_n is not word-representable.*

Proof. Supposed S_n is word-representable, so that S_n can be oriented semi-transitively by Theorem 1. Let $v_1, v_2, \ldots, v_{d+1} \in E_{n-m}$ be vertices of degree d connected to v. By Theorem 9, the distinct neighbourhoods of v_i, $1 \leq i \leq d+1$, form consecutive cyclic intervals of d vertices on the directed path $\vec{P}$, each of which contains v. Contradiction with the fact that v can be covered by at most d distinct intervals of d vertices. ■

4. Properties of Degrees in the Independent Sets in Word-Representable Split Graphs

An immediate corollary of Lemma 3 is that in our studies of word-representable split graphs $S_n = (E_{n-m}, K_m)$ we can assume that

- no two vertices in S_n have the same set of neighbours modulo vertices being connected to each other, so that
- at most one vertex in K_m is not connected to any vertex in E_{n-m}, and each vertex in E_{n-m} is of degree at least 2.

However, when studying minimal non-word-representable subgraphs of a split graph, we can make other assumptions as well, which allow a reduction of the space of possible solutions, e.g. when proceeding with a computer-aided search. The following two theorems are very useful.

Theorem 15. *Let $S_n = (E_{n-m}, K_m)$ be a word-representable graph, $m \geq 3$, and $2 \leq d \leq \frac{m+1}{2}$. Then, E_{n-m} contains at most m vertices of degree d whose neighbourhoods are distinct. This bound is achievable.*

Proof. By Theorem 1, S_n admits a semi-transitive orientation, in which the neighbourhoods of the vertices in E_{n-m}, by Theorem 9, are consecutive on the directed path $\vec{P}$ when read cyclicly. There are m distinct consecutive (cyclic) intervals of length d, which gives the upper bound. Finally, since $d \leq \frac{m+1}{2}$, the restrictions in Theorem 8 are satisfied, which makes the bound achievable (letting $\ell = m$ and $k = d$ in Corollary 10, we obtain the graph achieving the bound). ∎

Theorem 16. *Let $S_n = (E_{n-m}, K_m)$ be a word-representable graph, $m \geq 4$, and $\frac{m+1}{2} < d \leq m-1$. Then, E_{n-m} contains at most $m-d+1$ vertices of degree d whose neighbourhoods are distinct. This bound is achievable.*

Proof. By Theorem 1, S_n admits a semi-transitive orientation. Note that $m - d + 1$ is the number of distinct non-cyclic consecutive intervals of vertices on the path $\vec{P}$. Any number of these intervals can be the neighbourhoods of type A&B vertices by Theorem 9, so there exists the split graph S_n with the maximum number of type A&B vertices showing that the bound is achievable.

Next we prove that the bound can never be exceeded. To do this, we use the schematic way to represent consecutive cyclic intervals of vertices on $\vec{P}$ given in Figure 2. In that figure, the vertices in K_m are placed on a circle in clockwise direction in the order they appear in the directed path $\vec{P}$, and the chord ab represents the (cyclic) interval of vertices of length d that starts at a and ends at b. If such an interval corresponds to the neighbourhood of a vertex v in E_{n-m}, then v is of type A&B if a is before b in $\vec{P}$, and v is of type C if b is before a in $\vec{P}$.

Our first observation is that no matter what the semi-transitive orientation of S is, no two chords corresponding to the neighbourhoods of vertices in E_{n-m} can share an endpoint. Indeed, suppose ab and bc are chords as in the leftmost picture in Figure 2. But then, because $d > \frac{m+1}{2}$, at least one of the cords ab and bc corresponds to the neighbourhood of a vertex in E_{n-m} of type C. Suppose ab corresponds to a vertex of type C (the second case is analogous). But then, the interval given by bc covers both of a and b, which contradicts Theorem 8.

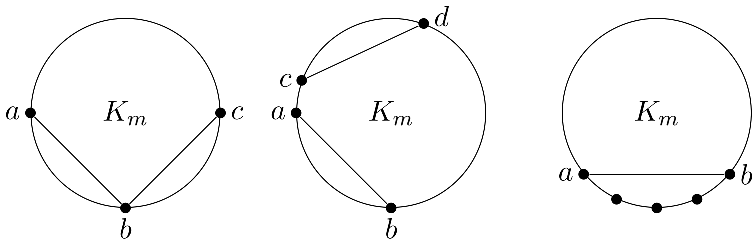

Figure 2. A schematic representation of consecutive cyclic intervals of vertices on $\vec{P}$ using chords to support the proof of Theorem 16.

Our second observation is that no matter what the semi-transitive orientation of S is, any two chords corresponding to the neighbourhoods of vertices in E_{n-m} must intersect each other, that is, the situation presented in the second picture in Figure 2 is not possible. Indeed, if ab and dc do not intersect each other, then at least one of them corresponds to a vertex in E_{n-m} of type C because $d > \frac{m+1}{2}$. But then, we obtain exactly the same contradiction with Theorem 8 as in the first observation.

Finally, suppose ab represents the neighbourhood of a vertex in E_{n-m} as in the rightmost picture in Figure 8. The chords representing any other neighbourhoods must have exactly one of their endpoints among the indicated $m-d$ vertices in that picture by the second observation. However, by the first observation, each of the $m-d$ vertices can be connected to at most one chord, which results in the maximum possible total amount of chords, and thus vertices in E_{n-m} of degree $d > \frac{m+1}{2}$, be $m-d+1$, as desired. ■

5. Characterizing Word-Representable Split Graphs with Cliques of Size 5

Applying Theorems 15 and 16 we see that in a word-representable graph $S_n = (E_{n-5}, K_5)$ we can have at most two vertices of degree 4, at most five vertices of degree 3, and at most five vertices of degree 2 (recall that vertices of degree 1 never affect word-representability).

Clearly, the minimal non-word-representable graphs in Figure 1 must be avoided when considering K_5. Computational experiments for $S_n = (E_{n-5}, K_5)$, which were possible due to the assumptions discussed above, reveal five more minimal non-word-representable graphs presented in Figure 3. One of these graphs is A_4 (see Definition 2) whose minimality and non-word-representability is given by Theorem 5. We conclude the section by proving that the graphs T_6–T_9 in Figure 3 are minimal non-word-representable graphs. Even though our proofs follow the same structure, each of them requires ad hoc arguments.

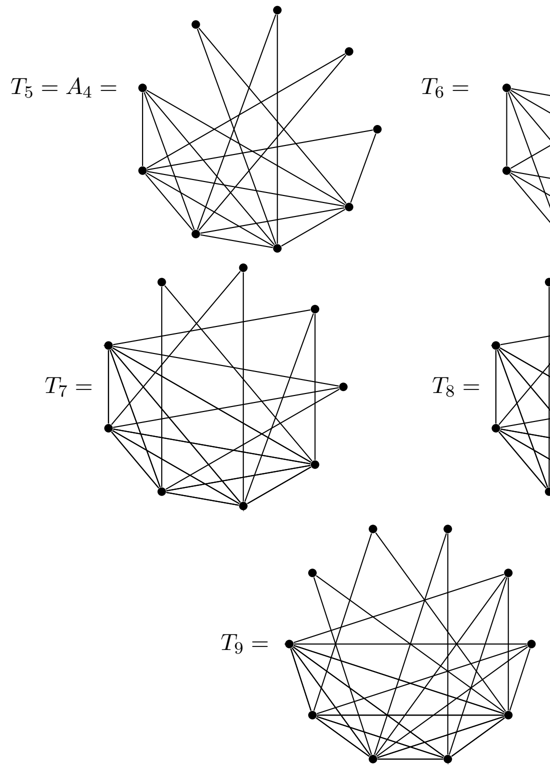

Figure 3. The minimal non-word-representable split graphs T_5–T_9.

Theorem 17. *The graph T_6 in Figure 3 is a minimal non-word-representable graph.*

Proof. We begin with proving non-word-representability of T_6. Suppose T_6 is word-representable, and thus, by Theorem 1, it can be oriented semi-transitively. Pick any such semi-transitive orientation of T_6. Then, by Theorem 9, the neighbourhood of a vertex of degree 2 must be two vertices staying next to each other, possibly cyclicly (if they are the source and the sink), on the path $\vec{P}$, as shown in Figure 4 (where the five vertices of $\vec{P}$ are placed on a circle; note that in our argument it is not important where the source and sink are). But then, the neighbourhood of the vertex d of degree 3 is forced to be non-consecutive vertices on $\vec{P}$. Contradiction with Theorem 9.

For proving the minimality of T_6, we consider removing each of the vertices in T_6 (one at a time) and, if necessary, describe a permutation of vertices of K_5, which results in all neighbourhoods of vertices in E_{n-5} be consecutive intervals on $\vec{P}$, or on whatever remains from $\vec{P}$ (which is still transitively oriented). Then, Theorem 9 can be used to obtain a semi-transitive orientation of the resulting graph proving its word-representability by Theorem 1.

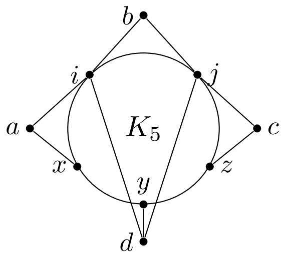

Figure 4. Proving non-word-representability of T_6.

- The vertices a and c are clearly symmetric, so we can consider removing a and skip considering removing c. In the case of a removed, swap x and y to obtain the desired result.
- If d is removed, all intervals become consecutive.
- If b is removed, place y between i and j to obtain the result.
- The vertices x and z are clearly symmetric, so we can consider removing x and skip considering removing y. The vertex a becomes of degree 1 and can be also removed by Lemma 3. All intervals become consecutive.
- If the vertex y is removed, then the vertices d and b have the same neighbourhoods, and one of them can be removed by Lemma 3. All intervals become consecutive.
- The vertices i and j are clearly symmetric, so we can consider removing i and skip considering removing j. If i is removed, a and b become of degree 1 and can be removed by Lemma 3. Swap x and y.

Our proof is completed. ■

Theorem 18. *The graph T_7 in Figure* 3 *is a minimal non-word-representable graph.*

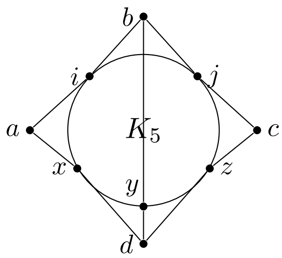

Figure 5. Proving non-word-representability of T_7.

Proof. We begin with proving non-word-representability of T_7. Suppose T_7 is word-representable, and thus, by Theorem 1, it can be oriented semi-transitively. Pick any such semi-transitive orientation of T_7. Then, by Theorem 9, the neighbourhoods of vertices of degree 2 must be consecutive, so since they are also disjoint, without loss of generality the degree 2 vertices are positioned as in Figure 5 (where the five vertices of $\vec{P}$ are placed on a circle; note that in our argument it is not important where the source and sink are). But then, since the vertices of degree 3 have symmetric properties (their neighbourhoods contain one vertex from each of vertices a and c neighbourhoods and vertex y) we see that there is no way for both neighbourhoods of b and d to be consecutive on $\vec{P}$. Contradiction with Theorem 9.

For proving the minimality of T_7, we consider removing each of the vertices in T_7 (one at a time) and, if necessary, describe a permutation of vertices of K_5, which results in all neighbourhoods of vertices in E_{n-5} be consecutive intervals on $\vec{P}$, or on whatever remains from $\vec{P}$ (which is still transitively oriented). Then, Theorem 9 can be used to obtain a semi-transitive orientation of the resulting graph proving its word-representability by Theorem 1.

- The vertices a and c are clearly symmetric, so we can consider removing a and skip considering removing c. In the case of a removed, swap x and y to obtain the desired result.
- If b is removed, all intervals become consecutive. If d is removed, then place y between i and j to obtain the desired result.
- The vertices i and j are clearly symmetric, so we can consider removing i and skip considering removing j. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- The vertices x and z are clearly symmetric, so we can consider removing x and skip considering removing z. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- If y is removed, then the graph is isomorphic to $K_4^{\triangle}$ and it is word-representable by Theorem 4.

Our proof is completed. ∎

Theorem 19. *The graph T_8 in Figure 3 is a minimal non-word-representable graph.*

Proof. We begin with proving non-word-representability of T_8. Suppose T_8 is word-representable, and thus, by Theorem 1, it can be oriented semi-transitively. Pick any such semi-transitive orientation of T_8. Then, by Theorem 9, the neighbourhoods of all vertices in the independent set must be consecutive intervals, and the only way to arrange this is shown in Figure 6 (where the five vertices of $\vec{P}$ are placed on a circle and the orientation of the longest path is assumed to be in clockwise direction). But then we obtain a contradiction with Theorem 8. Indeed, if b is of type C then d must be of type A&B, but x and z are in the neighbourhood of d. On the other hand, if d is of type C then b must be of type A&B, but x and z are in the neighbourhood of b. Thus, T_8 is not word-representable.

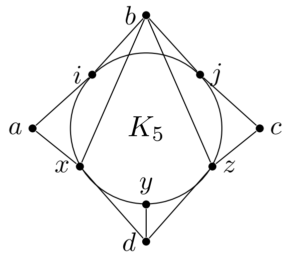

Figure 6. Proving non-word-representability of T_8.

For proving the minimality of T_8, we consider removing each of the vertices in T_8 (one at a time) and, if necessary, describe a permutation of vertices of K_5, which results in all neighbourhoods of vertices in E_{n-5} be consecutive intervals on $\vec{P}$, or on whatever remains from $\vec{P}$ (which is still transitively oriented). Then, Theorem 9 can be used to obtain a semi-transitive orientation of the resulting graph proving its word-representability by Theorem 1.

- The vertices a and c are clearly symmetric, so we can consider removing a and skip considering removing c. If a is removed, swap x and y and note that making x in the new position the source, both b and d become of type C, so there is no conflict with Theorem 8 (the neighbourhoods in question are still consecutive).
- If b is removed, or if d is removed, then clearly there is no conflict with Theorem 8, and the neighbourhoods in question are still consecutive.
- The vertices i and j are clearly symmetric, so we can consider removing i and skip considering removing j. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- The vertices x and z are clearly symmetric, so we can consider removing x and skip considering removing z. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- If y is removed, then the obtained graph is a subgraph of $K_5^{\triangle}$ and it is word-representable by Theorem 4.

Our proof is completed. ■

Theorem 20. *The graph T_9 in Figure 3 is a minimal non-word-representable graph.*

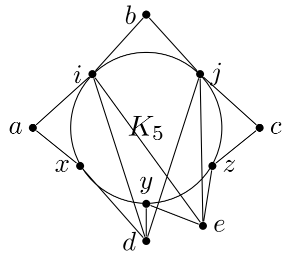

Figure 7. Proving non-word-representability of T_9.

Proof. We begin with proving non-word-representability of T_9. Suppose T_9 is word-representable, and thus, by Theorem 1, it can be oriented semi-transitively. Pick any such semi-transitive orientation of T_9. Then, by Theorem 9, the neighbourhoods of all vertices in the independent set must be consecutive intervals, and the only way to arrange this is shown in Figure 7 (where the five vertices of $\vec{P}$ are placed on a circle and the orientation of the longest path is assumed to be in clockwise direction). But then we obtain a contradiction with Theorem 8 given by vertices d and e. Indeed,

- if y is the source, or j is the source, or z is the source, then d is of type A&B and e is of type C; the problem is then with d being connected to y and i.
- if x is the source, or i is the source, then e is of type A&B and d is of type C; the problem is then with e being connected to j and y.

For proving the minimality of T_9, we consider removing each of the vertices in T_9 (one at a time) and, if necessary, describe a permutation of vertices of K_5, which results in all neighbourhoods of vertices in E_{n-5} be consecutive intervals on $\vec{P}$, or on whatever remains from $\vec{P}$ (which is still transitively oriented). Then, Theorem 9 can be used to obtain a semi-transitive orientation of the resulting graph proving its word-representability by Theorem 1.

- The vertices a and c are clearly symmetric, so we can consider removing a and skip considering removing c. In the case of a removed, swap x and y and note that making x in the new position the source, both d and e become of type A&B, so there is no conflict with Theorem 8 (all neighbourhoods in question are still consecutive).
- If d (respectively, e) is removed, we can make x (respectivley, z) the source and there will be no conflict with Theorem 8.
- If b is removed, then swapping x and i, as well as z and j, and making z the

source, we obtain both d and e being of type A&B, so there is no conflict with Theorem 8.

- The vertices i and j are clearly symmetric, so we can consider removing i and skip considering removing j. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- The vertices x and z are clearly symmetric, so we can consider removing x and skip considering removing z. The vertex a becomes of degree 1 and can be also removed by Lemma 3. The obtained graph is word-representable by Theorem 6.
- If y is removed, then the obtained graph is word-representable by Theorem 6.

Our proof is completed. ■

6. Word-Representability of Graphs Obtained by Gluing in a Clique

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, if at least one of two graphs is non-word-representable, then gluing the graphs in a clique will result in a non-word-representable graph. Moreover, it is known that gluing two word-representable graphs in a vertex (a clique of size 1) always results in a word-representable graph (e.g. see [6, Section 7.3] or [8, Section 5.4.3]). Further, it is not difficult to come up with examples when gluing two word-representable graphs in an arbitrary clique results in a word-representable graph; for a trivial such example, take two copies of a complete graph K_n, gluing which gives K_n, and K_n can be represented by any permutation of length n. However, there are examples of word-representable graphs gluing which in an edge (a clique of size 2), or a triangle (a clique of size 3), results in a non-word-representable graph. The respective examples can be found in [8, Section 5.4.3], and they are presented in Figures 8 and 9, respectively. Thus, the rightmost graphs in these pictures are non-word-representable, while the other graphs are word-representable.

The question on whether gluing two word-representable graphs in a clique of size 4, or more, may result in a non-word-representable graph was open, though unpublished until [6], for about ten years. In Subsection 6.1 we use split graphs to show that gluing two word-representable graphs in a clique of size 4, or more, may result in a non-word-representable graph. A significance of our solution to the problem is in showing that gluing two cliques may be sensitive to which vertices are glued to which vertices, as the word-representability of the resulting graph may depend on it. In either case, in Subsection 6.2, we give another, surprisingly simple solution to the problem, which is based on a generalization of the construction in Figure 9.

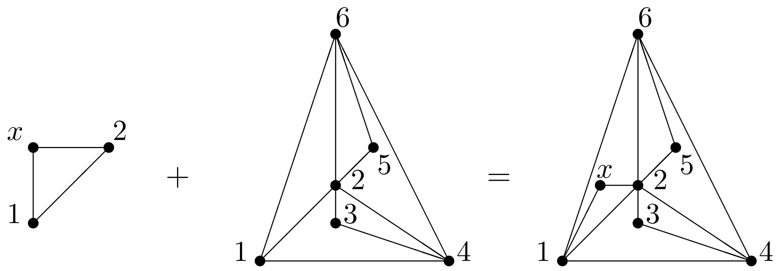

Figure 8. Gluing two word-representable graphs in an edge.

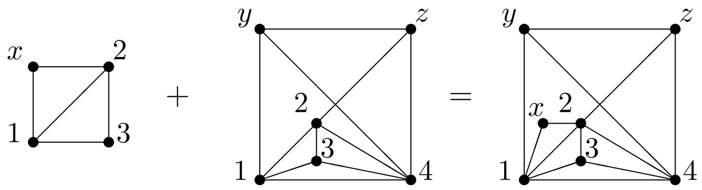

Figure 9. Gluing two word-representable graphs in a triangle.

6.1. Solving the problem via split graphs

Recall the definition of $K_\ell^{\triangle}$ in Section 2 (Definition 2) and the fact that $K_\ell^{\triangle}$ is word-representable by Theorem 4. Further, for $2 \leq i \leq \ell$, let K_ℓ^i be the graph obtained from the complete graph K_ℓ labeled by $1, 2, \ldots, \ell$, by adding a new vertex x of degree 2 connected to the vertices 1 and i. Clearly, any K_ℓ^i is isomorphic to K_ℓ^2, which is an induced subgraph of $K_\ell^{\triangle}$, and thus is word-representable.

Recall the definition of A_ℓ in Section 2 (Definition 2) and the fact that A_ℓ is not word-representable by Theorem 5.

We observe that, for $\ell \geq 4$, gluing two word-representable graphs $K_\ell^{\triangle}$ and K_ℓ^i, where $2 < i < \ell$, in the ℓ-clique so that a vertex j is glued with the vertex j for $1 \leq j \leq \ell$, results in a non-word-representable graph G_i. Indeed, G_i contains the non-word-representable A_i induced by the vertices $1, 2, \ldots, (i+1), 1', 2', \ldots, (i-1)', x$.

Note that even though K_ℓ^2 (respectively, K_ℓ^ℓ) is isomorphic to K_ℓ^i for $2 < i < \ell$, gluing the ℓ-cliques in $K_\ell^{\triangle}$ and K_ℓ^2 (respectively, K_ℓ^ℓ) as above results in a word-representable graph G_1 (respectively, G_ℓ). Indeed, both G_1 and G_ℓ are the graph $K_\ell^\triangle$ with the additional vertex x having the same neighbourhood as another vertex in $K_\ell^\triangle$ of degree 2. It is a direct corollary of Lemma 3 that word-representability of $K_\ell^\triangle$ implies word-representability of G_1. Thus, when gluing two word-representable graphs in a clique, the word-representability of the resulting graph may depend on how exactly we glue.

6.2. Generalizing the known construction

Here we present an alternative solution to the problem of gluing two graphs by generalizing the construction in Figure 9.

Let $n \geq 2$ and K'_n be the graph obtained from the complete graph K_n on the vertex set $\{1, 2, \ldots, n\}$ by adding a vertex x connected to the vertices 1 and 2. For example, the leftmost graphs in Figures 8, 9, 10 and 11 are K'_2, K'_3, K'_4 and K'_5, respectively. It is straightforward to check that the word $x12x34\cdots n$ represents K'_n for any $n \geq 2$.

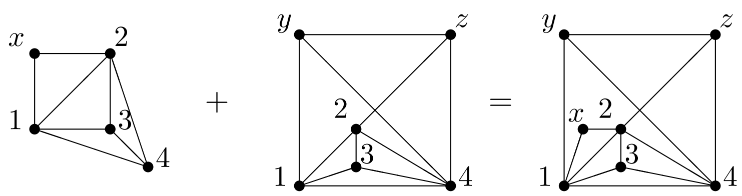

Figure 10. Gluing two word-representable graphs in K_4.

Let the middle graph in Figure 9 be denoted by M_4, and for $n \geq 5$, M_n is obtained by enlarging the clique formed by the vertices $1, 2, 3, 4$ in M_4. That is, M_n is obtained from M_{n-1} by adding the vertex n connected to all the vertices in $\{1, 2, \ldots, n-1\}$ but not the vertices y and z. For example, M_5 is the middle graph in Figure 11. It is straightforward to check that the word $y1z4y2z3567\cdots n$ represents M_n for any $n \geq 4$.

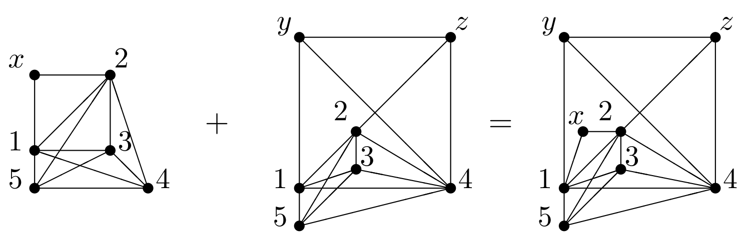

Figure 11. Gluing two word-representable graphs in K_5.

Finally, for $n \geq 4$, let B_n be obtained from M_n by adding a vertex x connected just to the vertices 1 and 2. For example, B_4 is the rightmost graph in Figures 9 and 10, and B_5 is the rightmost graph in Figure 11. Note that using the hereditary nature of word-representable graphs, B_n is not word-representable for any $n \geq 4$ since B_4 is not word-representable [6, 8].

Thus, for $n \geq 4$, gluing word-representable graphs K'_n and M_n in the clique formed by the vertices $1, 2, \ldots, n$ gives the non-word-representable graph B_n, as desired. See Figures 10 and 11 for the cases of $n = 4$ and $n = 5$, respectively.

7. Concluding Remarks

This paper extends our knowledge [7] on word-representable split graphs, and the general theorems we prove, Theorems 15 and 16, allow computational characterization of word-representable split graphs with cliques of size 5 in terms of nine forbidden subgraphs. Taking into account that tackling the general case seems to be not feasible for the moment, a natural next step is in using our general theorems in (computational) characterization of word-representable split graphs with cliques of size 6, which we leave as an open research direction.

Acknowledgments

The first author was supported by the National Natural Science Foundation of China (Grant Numbers 11901319) and the Fundamental Research Funds for the Central Universities (Grant Number 63191349).

References

[1] V. Chvátal and P.L. Hammer, *Aggregation of inequalities in integer programming*, in: Ann. Discrete Math. **1** (1977), P.L. Hammer, E.L. Johnson, B.H. Korte and G.L. Nemhauser (Ed(s)), (Studies in Integer Programming (Proc. Worksh. Bonn, 1975) 145–162.
https://doi.org/10.1016/S0167-5060(08)70731-3

[2] S. Foldes and P.L. Hammer, *Split graphs*, in: Proc. Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing, Congr. Numer. **XIX** (1977), (Louisiana State Univ., Baton Rouge, 1977) 311–315.

[3] M. Glen, *Software*.
http://personal.strath.ac.uk/sergey.kitaev/word-representable-graphs.html

[4] M. Golumbic, Algorithmic Graph Theory and Perfect Graphs (Elsevier Inc., 1980).
https://doi.org/10.1016/B978-0-12-289260-8.50010-8

[5] M.M. Halldórsson, S. Kitaev and A. Pyatkin, *Semi-transitive orientations and word-representable graphs*, Discrete Appl. Math. **201** (2016) 164–171.
https://doi.org/10.1016/j.dam.2015.07.033

[6] S. Kitaev, *A comprehensive introduction to the theory of word-representable graphs*, in: Developments in Language Theory, DLT 2017, E. Charlier, J. Leroy and M. Rigo (Ed(s)), Lecture Notes in Comput. Sci. **10396** (2017) 36–67. https://doi.org/10.1007/978-3-319-62809-7_2

[7] S. Kitaev, Y. Long, J. Ma and H. Wu, *Word-representability of split graphs*. arXiv:1709.09725

[8] S. Kitaev and V. Lozin, Words and Graphs (Springer, 2015). https://doi.org/10.1007/978-3-319-25859-1

[9] N.V.R. Mahadev and U.N. Peled, Threshold Graphs and Related Topics (Elsevier, 1995).

Received 23 November 2019
Revised 18 June 2020
Accepted 18 June 2020