On Word-Representable Split Graphs

Jon Rosario

March 12, 2024

Abstract

This report introduces split graphs and word-representable graphs, two interesting classes of graphs with unique properties. Split graphs are easily identifiable by a partition of the graph's vertices into a clique and and independent set. On the other hand, word-representable graphs are not as easily identifiable. The report first explores the ideas that led to the modern understanding of split graphs and word-representable graphs separately. Then, it explores the relationship between these two classes of graphs and presents the most recent results, including the characterization of word-representable split graphs in terms of forbidden subgraphs. The theorems presented in this report contribute to the understanding of the structure of split graphs and word-representable graphs, but do not entirely characterize the entire class of split graphs.

1 Introduction

The theory of split graphs was originally developed in a conference paper by Foldes and Hammer [FH77] in 1977. A set of vertices is a *clique* if all pairs of vertices share an edge, while a set of vertices is an *independent set* if no pair of vertices share an edge. A graph G is *split* if and only if its vertices can be partitioned into a clique and an independent set [FH85]. An example is shown in Figure 1. Since then, they have been studied extensively [Cla90, FH77, Gol04a, Gol04b, FH85].

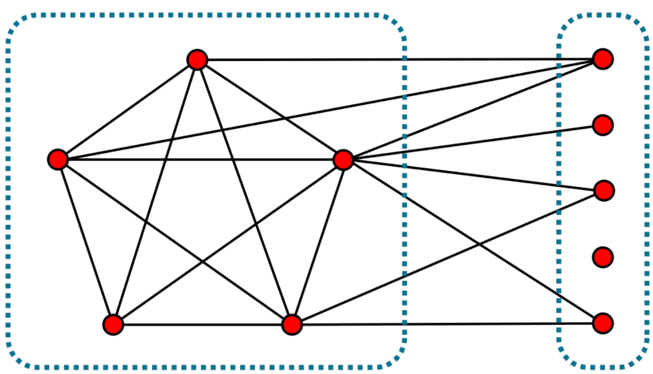

Figure 1: An example split graph. Figure from Wikipedia [Wik23].

A graph G is *word-representable* if one can form a symbolic word w by assigning symbols to each vertex $v \in V$, and two symbols u, v *alternate* in w if and only if the corresponding vertices are adjacent in G. See the left-hand graph in Figure 5 for an example of a word-representable graph. In general, the class of word-representable graphs is generalizes many families of graphs. Much research effort has been put into characterizing exactly which graph families are word-representable and not word-representable [HKP16, KL15, KP08, KS08].

The theory of split graphs and word-representable graphs are each individually interesting. However, very little is known about the intersection between split graphs and word-representable graphs. This report will summarize the known connections between these two objects. Firstly in Section 2, we will present the properties of split graphs that have found the most use in this theory. These mostly pertain to the fact that the split graphs can be characterized in terms of forbidden induced subgraphs.

Then in Section 3, we will formalize the notion of word-representable graphs. Word-representable graphs have a strong connection to what are called semi-transitive orientations of graphs, but will not be discussed here. For more information, see the Halldórsson-Kitaev-Pyatkin [HKP16] paper on semi-transitive orientations. In a recent paper by Kitaev, Long, Ma, and Wu, the authors proved the following result [KLMW17]:

Theorem 1.1. *Let G be a split graph in which vertices in the independent set always have degree at most 2. Then G is word-representable if and only if G contains certain induced subgraphs.*

After the sections on split graphs and word-represent graphs, we will then present the main results in Section 4. Finally, we will discuss open questions in this topic in Section 5.

2 Split Graphs

Definition 2.1. *Formally, a simple graph G is* ***split*** *if its vertices can be partitioned $V(G) = C \cup I$ where the vertices C form a clique and the vertices I form an independent set.*

The reason for the name "split" graphs is hopefully now clear: the definition states that the vertices can, in a sense, be split. There are two other ways in which the split graphs can be characterized. A simple graph G is *chordal* if for all cycles of length greater than or equal to 4, there is a chord, or an edge connecting two otherwise non-adjacent vertices of the cycle. If one defines $\bar{G} = (V, E')$ to be the graph on the same vertices as $G = (V, E)$ with edge $e \in E'$ if and only if $e \notin E$. All three are shown to be equivalent in Theorem 1 of Foldes' and Hammer's paper [FH85].

Theorem 2.1. *For any graph G, the following three conditions are equivalent:*

- *Both G and $\bar{G}$ are chordal;*
- *$V(G)$ can be partitioned into a clique C and independent set I;*
- *G does not contain an induced subgraph isomorphic to $2K_2$, C_4, or C_5. See Figure 3.*

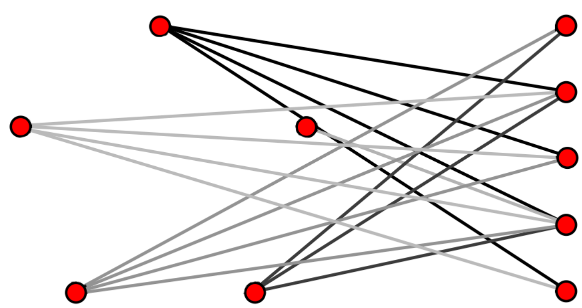

Figure 2: The edge compliment of the graph in Figure 1. One can verify that both graphs are chordal.

Additionally, the following lemmas on split graphs have found use in many contexts:

Theorem 2.2. *A graph G is split if and only if its edge complement $\bar{G}$ is split. [Gol04a]*

Define $\alpha(G)$ to be the size of the maximum independent set of G. Similarly, define $\omega(G)$ to be the size of the maximum clique of G.

Theorem 2.3. *Let G be a split graph with partition $V(G) = C \cup I$. Exactly one of the following conditions hold:*

- *$|I| = \alpha(G)$ and $|C| = \omega(G)$;*
- *$|I| = \alpha(G)$ and $|C| = \omega(G) - 1$;*
- *or $|I| = \alpha(G) - 1$ and $|C| = \omega(G)$*

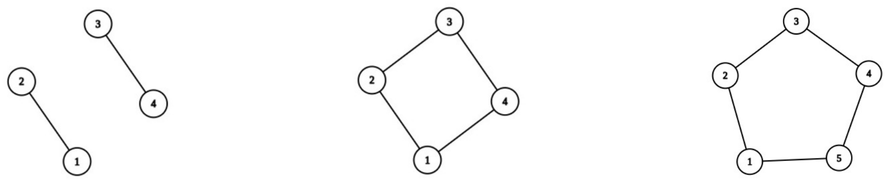

Figure 3: The graphs $2K_2$, C_4, and C_5, respectively.

We now move onto comparability graphs. For any strict partially ordered set with an ordering operation $(S, <)$, define the corresponding **comparability graph** to have vertices corresponding to elements of S, and whose edges are pairs of elements (u, v) if $u < v$. Then we have the following theorem characterizing comparability graphs within split graphs:

Theorem 2.4. *If G is a split graph, then G is a comparability graph if and only if G contains no induced subgraph isomorphic to H_1, H_2, or H_3 [FH77]. See Figure 4.*

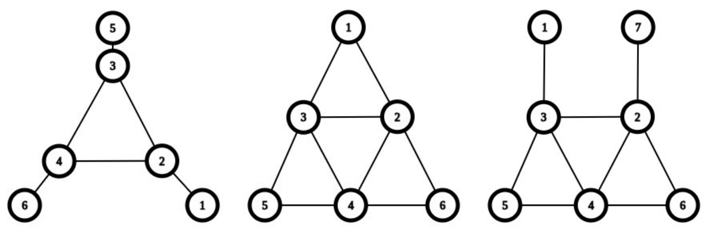

Figure 4: The graphs H_1, H_2, and H_3, respectively.

In the theory of split graphs, many other results utilize this notion of forbidden subgraphs. In this spirit of this theory, we will see in Section 5 that the main results presented in this report use this same idea to characterize word-representability in split graphs.

3 Word-Representable Graphs

Now we move onto word-representable graphs.

Definition 3.1. *Let w be a word over an alphabet V, and let x and y be two different letters in w. We say that x and y **alternate** in w if the deletion of all letter except the copies of x and y results in a word of type either $xyxy...$ or $yxyx...$ (with either odd or even length).*

Example 3.1. *Let $w_1 = xyzyxzx$. Then (x, y) do not alternate in w_1, but (y, z) and (x, z) both alternate.*

Let $w_2 = xzyyxz$. Then (x, y) are the only alternating letters in w_2. Note that if a letter appears twice, we can force the letter to not alternate with any other letters.

We now introduce word-representable graphs.

Definition 3.2. *A graph $G = (V, E)$ is **word-representable** if there exists a word w over an alphabet V such that (x, y) alternate in $w \iff (x, y) \in E$. In this case, we say w represents G [KP08, HKP16]. Note that w may not uniquely represent G, as shown in Figure 5.*

Word-representable graphs are closely related to *semi-transitive orientations*, a topic which is explained at depth in [HKP16]. In this paper, the authors managed to prove the following by exploring word-representability through the lens of *semi-transitive orientations* of graphs:

Theorem 3.1. *Any 3-colorable graph G is word-representable.*

Many graphs, but not all, are word-representable. See Figure 5 for an example. We can also define a parameter that relates, in some sense, to the difficulty of representing a graph G.

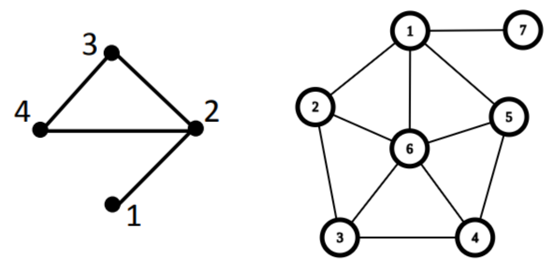

Figure 5: The left-hand graph M is represented by the word $w_M = 1213423$. It is also represented by the word $w'_M = 412134$. The right-hand graph N is an example of a minimal non-word-representable graph. Figure taken from Words and Graphs [KL15].

Definition 3.3. *A graph G is called k-word representable if there is some word w representing G where each letter appears exactly k times in w.*

Theorem 3.2. *A graph G is word-representable if and only if it is k-word representable for some k.*

Proving the backward direction of this theorem is nearly trivial. In the other direction, it is slightly harder to show, but nonetheless true. See [KP18] for a proof. We now take a detour to show the flavor of proofs involving word-representability:

Theorem 3.3. *The Petersen graph, shown in Figure 6 is not 2-word representable [KP18].*

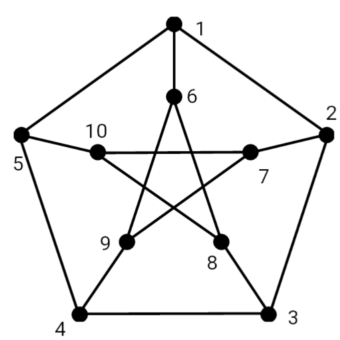

Figure 6: The Petersen graph. Figure taken from [Wen10].

The Petersen graph is frequently used as a small counterexample to statements that might be true for all graphs. Here is an example of where the Petersen graph fulfills this role it frequently takes on. We now prove the theorem:

Proof. We will prove this via contradiction. Refer to Figure 6, where we have labeled the vertices with numbers 1 to 10. The alphabet V will be the the numbers 1 to 10. Assume, for sake of contradiction, that the Petersen graph is 2-word representable. Then there exists some word w that represents the Petersen graph. There are a few observations will make use of to make our argument simpler.

Firstly, the Petersen graph is 3-regular, meaning that all vertices have degree 3. Furthermore the Petersen graph is edge-transitive, meaning that for any edges e_1 and e_2, there is an automorphism of the graph mapping e_1 to e_2.

These two properties allow us to assume, without loss of generality, that the word w starts with the symbol 1. Furthermore, we can assume that the vertex corresponding to 1 appears on the outer ring of the graph.

Now we can now assume that the word has form $w = 12561w_1 6w_2 5w_3 2w_4$, where w_i cannot contain the numbers 1, 2, 5, or 6. This is because we have assumed w is 2-word representable, and each of the aforementioned symbols already appear exactly twice.

We can also assume that there are exactly 3 symbols between the two copies of the symbol 1 in w. There are two cases to consider. If y appears once between the two copies of 1, then the other y must appear after the second copy, and then 1 alternates with y, meaning the corresponding vertex 1 has 4 edges, but we know the Petersen graph is 3-regular. Likewise, if y appears twice between two copies of 1, then again by the 3-regularity of the Petersen graph, we cannot have the vertex y alternating with a different three symbols than 1 does.

Finally we can present the contradiction. Note that vertex 8 must alternate with 6 but not with 5. Thus, we now know that 8 must appear in both w_1 and w_2. Similarly, in order to alternate with 2 but not 5, 3 must appear in w_3 and w_4. However, these two statements together imply that 8833 is a subword in w, meaning 8 and 3 cannot alternate in w. However, 3 and 8 have an edge in the Petersen graph, so this is a contradiction. Thus we have shown that the Petersen graph cannot be 2-word representable. □

It was conjectured in by Kitaev and Pyatkin [KP08] that the Petersen graph is non-word-representable at all. However, this is not the case. In 2010, Konovalov and Linton showed by computer that there are two 3-representations for the Petersen graph:

- 138729607493541283076850194562, and
- 134058679027341283506819726495.

Theorem 3.4. *The Petersen graph, shown in Figure 6, is word-representable.*

Proof. By the given 3-word-representations above and by Theorem 3.2, we have that the Petersen graph is word-representable. □

4 Main Results

The Petersen graph happens to have properties that make it keen for proving results on word-representability. The most useful of the properties is that the Petersen graph is 3-regular. Split graphs with maximum clique $\omega(G)$ greater than 2 are not k regular for any k, so it seems inappropriate to use the techniques displayed in the proof of Theorem 3.3.

Fortunately, there is another approach that one can try to characterize the split graphs in terms of word-representability. Additionally, this approach is much more natural in the context of split graphs.

Theorem 2.1 showed us that the split graphs can be characterized in terms of forbidden subgraphs. We introduce a theorem on graphs that relates subgraphs and word-representability.

Theorem 4.1. *If a graph G contains a non-word-representable graph N as an induced subgraph, then G is also non-word-representable.*

Intuitively, this should make sense. In order to represent such a graph G with a forbidden induced subgraph H, we would need show that every two vertices sharing an edge in H alternate in some word w. However, this would clearly produce a word w' that can represent H by just removing all symbols in w that don't refer to vertices in H—forming a contradiction.

At this point, we can state the main results that will be discussed in this report.

Theorem 4.2. *Let G be a split graph with maximum clique size $\omega(G) \leq 3$. Then G is word-representable.*

Proof. Any such G is 3-colorable. Begin by coloring the clique of G with each of the colors. Then color each vertex of the independent set with a color that it is not adjacent to. Since the maximum clique size $\omega(G)$ is less than or equal to 3, this must be possible. Thus by Theorem 3.1, we are done. □

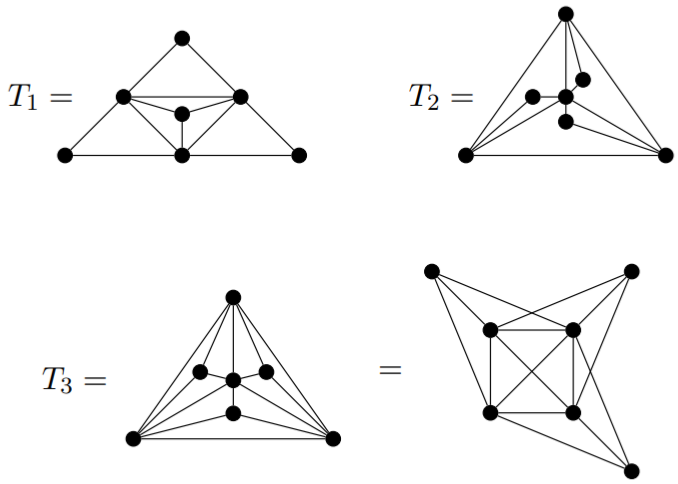

Figure 7: The graphs T_1, T_2, and T_3. Figures from [KLMW17].

Let the graphs T_i for $i = 1, 2, 3$ be defined as in Figure 7.

The graphs A_ℓ for $\ell \geq 4$ is formed by taking the complete graph $K_{\ell-1}$ and adding a vertex i' to form a triangle between vertices i and $i + 1 \mod \ell - 1$. Then add a vertex ℓ that is adjacent to each of the vertices $1, 2, \cdots, \ell - 1$. See Figure 8 for examples.

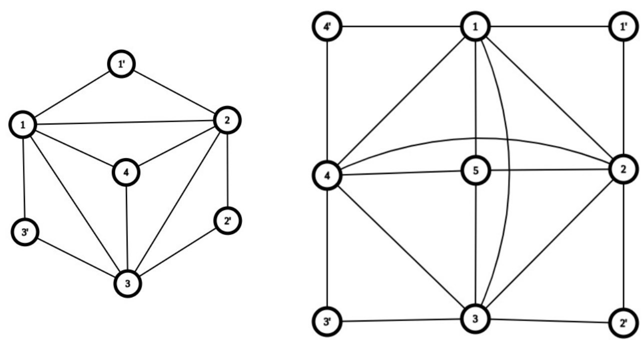

Figure 8: The graphs A_4 and A_5 on the left-hand and right-hand sides, respectively. The graphs A_l are explained in Section 4.

The following results by Kitaev, Long, Ma, and Wu characterize in terms of forbidden subgraphs certain word-representable split graphs [KLMW17]:

Theorem 4.3. *Let G be a split graph in which vertices in the independent set always have degree at most 2. Then G is word-representable if and only if G does not contain both T_2 and A_l as induced subgraphs.*

Theorem 4.4. *Let G be a split graph in which vertices in the independent set always have degree at most 2, and the size of the maximum clique $\omega(G)$ is exactly 4. Then G is word-representable if and only if G does not contain the graphs T_1, T_2, and T_3 as induced subgraphs.*

These two results are not easy to prove. The authors of the aforementioned paper manage to prove the first theorem by relying on properties of the graph A_l. This graph is minimally non-word-representable, meaning that every induced subgraph of A_l is word-representable, except for the trivial subgraph A_l. The following theorem is proved though intense casework involving several considerations over the possibilities of the graph, up to isomorphism. This proof also relies on demonstrating that several graphs emit semi-transitive orientations, a topic that is very closely related to word-representability.

5 Discussion

The theorems mentioned in this report (Theorem 4.3 and Theorem 4.4) are the most recent published results related to this topic. Unfortunately, this means that there are no known categorical results for arbitrary split graphs. The results shown here apply under very specific circumstances: when the maximum degree of the independent set is limited, when the size of the maximum clique of limited, or both.

One can consider the natural extensions of Theorem 4.3 and Theorem 4.4:

Open Question 5.1. *Let G be a split split in which vertices in the independent set always have degree at most d. When is G word-representable?*

It is unknown whether the original classification presented in Theorem 4.3 still holds for values of d greater than 2. Since this must be checked for split graphs of arbitrary size, verification via computer is not an option. However, it would certainly be a remarkable result if the original characterization of forbidden subgraphs holds for all $k \geq 2$.

Alternatively, we can consider extending the result of Theorem 4.4:

Open Question 5.2. *Let G be a split graph in which vertices in the independent set always have degree at most d, and the size of the maximum clique $\omega(G)$ is exactly k. Which subgraphs characterize when G is word-representable?*

This direction unfortunately seems much harder to extend than the former due to the fact that there are now two parameters to consider.

It is hard to imagine either of these questions being resolved anytime soon. It seems much more likely that results related to semi-transitive orientations may lead to success, perhaps due to their algebraic nature. For the key results relating word-representable graphs and semi-transitive orientations, see [HKP16].

References

[Cla90] R.J. Clarke. Covering a set by subsets. *Discrete Mathematics*, 81(2):147–152, 1990.

[FH77] Stephan Foldes and Peter L. Hammer. Split graphs. In *Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing (Louisiana State Univ., Baton Rouge, La., 1977)*, Congressus Numerantium, No. XIX, pages 311–315. Utilitas Math., Winnipeg, Man., 1977.

[FH85] Stephan Foldes and Peter L. Hammer. Split graphs of Dilworth number 2. *Discrete Mathematics*, 55(2):123–127, 1985.

[Gol04a] Martin Charles Golumbic. Chapter 6 - split graphs. In Martin Charles Golumbic, editor, *Algorithmic Graph Theory and Perfect Graphs*, volume 57 of *Annals of Discrete Mathematics*, pages 149–156. Elsevier, 2004.

[Gol04b] Martin Charles Golumbic. Chapter 9 - superperfect graphs. In Martin Charles Golumbic, editor, *Algorithmic Graph Theory and Perfect Graphs*, volume 57 of *Annals of Discrete Mathematics*, pages 203–218. Elsevier, 2004.

[HKP16] Magnús M. Halldórsson, Sergey Kitaev, and Artem Pyatkin. Semi-transitive orientations and word-representable graphs. *Discrete Applied Mathematics*, 201:164–171, 2016.

[KL15] Sergey Kitaev and Vadim Lozin. *Words and Graphs*. Monographs in Theoretical Computer Science. Springer-Verlag, 2015.

[KLMW17] Sergey Kitaev, Yangjing Long, Jun Ma, and Hehui Wu. Word-representability of split graphs. *Journal of Combinatorics*, 12, 09 2017.

[KP08] Sergey Kitaev and Artem Pyatkin. On representable graphs. *J. Autom. Lang. Comb.*, 13:45–54, 2008.

[KP18] Sergey Kitaev and Artem Pyatkin. Word-representable graphs: a survey. *Journal of Applied and Industrial Mathematics*, 12:278–296, 04 2018.

[KS08] Sergey Kitaev and Steve Seif. Word problem of the perkins semigroup via directed acyclic graphs. *Order*, 25:177–194, 08 2008.

[Wen10] Paul Wenger. Three existence problems in extremal graph theory. 08 2010.

[Wik23] Wikipedia. Split graph — Wikipedia, the free encyclopedia, 2023. [Online; accessed April-2023].