Word-representability of split graphs generated by morphisms

Kittitat Iamthong

May 3, 2021

Abstract

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 $xy \in E$. A split graph is a graph in which the vertices can be partitioned into a clique and an independent set. There is a long line of research on word-representable graphs in the literature, and recently, word-representability of split graphs has attracted interest.

In this paper, we first give a characterization of word-representable split graphs in terms of permutations of columns of the adjacency matrices. Then, we focus on the study of word-representability of split graphs obtained by iterations of a morphism, the notion coming from combinatorics on words. We prove a number of general theorems and provide a complete classification in the case of morphisms defined by $2 \times 2$ matrices.

1 Introduction

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 $xy \in E$. In this definition, letters $x$ and $y$ alternate in $w$ if after removing all letters but $x$ and $y$, we would either get a word of the form $xyxy\cdots$, or of the form $yxyx\cdots$, of even or odd length. Also, by definition, $w$ must contain each letter in $V$ at least once. For example, the cycle graph on four vertices labeled 1, 2, 3, 4 in clockwise direction is word-representable because it can be represented by the word 14213243.
There is a long line of research on word-representable graphs in the literature that is summarized in [5]. Word-representable graphs are important as they generalize several well-known and well-studied classes of graphs such as 3-colorable graphs, comparability graphs and circle graphs [4]. Note 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. The wheel graph $W_5$ is the smallest non-word-representable graph. One of the key results in the area is the following theorem, where an orientation of a graph is semi-transitive if it is acyclic, and for any directed path $v_0 \rightarrow v_1 \rightarrow \cdots \rightarrow v_k$ either there is no edge between $v_0$ and $v_k$, or $v_i \rightarrow v_j$ is an edge for all $0 \leq i < j \leq k$.

**Theorem 1** ([4]). A graph is word-representable if and only if it admits a semi-transitive orientation.

The following simple lemma will also be of use to us in this paper.

**Lemma 2** ([7]). Let $K_m$ be a clique in a graph $G$. Then any acyclic orientation of $G$ induces a transitive orientation on $K_m$ (where the presence of edges $u \rightarrow v$ and $v \rightarrow z$ implies the presence of the edge $u \rightarrow z$). In particular, any semi-transitive orientation of $G$ induces a transitive orientation on $K_m$. In either case, the orientation induced on $K_m$ contains a single source and a single sink.

### 1.1 Split graphs.

A split graph is a graph in which the vertices can be partitioned into a clique and an independent set [3]. The paper [7] initiated a systematic study of word-representatability of split graphs, which was extended in a follow up paper [1]. In particular, characterizations of split graphs in terms of forbidden induced subgraphs were obtained in [7] and [1] for cliques of sizes 4 and 5, respectively. Also, a characterization of semi-transitive orientations of split graphs was obtained in [7] (see below), and split graphs were used to solve a long standing problem in the theory of word-representation in [1]. We note though that currently a complete characterization of split graphs (e.g. in terms of forbidden subgraphs) seems to be a non-feasible problem, so a natural research direction is in understanding (non-)word-representable subclasses of split graphs.

Let $S_n = (E_{n-m}, K_m)$ be a word-representable split graph, where $K_m$ is the maximal clique, and $E_{n-m}$ is the independent set. 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 $\vec{P}$. Theorem [5] below characterizes semi-transitive orientations of split graphs.
Figure 1: Three types of vertices in $E_{n-m}$ in a semi-transitive orientation of $(E_{n-m}, K_m)$. The vertical oriented paths are a schematic way to show (parts of) $\vec{P}$

**Theorem 3** ($[7]$). Any semi-transitive orientation of a split graph $S_n = (E_{n-m}, K_m)$ subdivides the set of all vertices in $E_{n-m}$ into three, possibly empty, groups corresponding to each of the following types (also shown schematically in Figure 1), where $\vec{P} = p_1 \rightarrow \cdots \rightarrow p_m$ 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 $x \rightarrow v$ for each $x \in I_v = \{p_1, p_2, \ldots, p_i\}$ and there is an edge $v \rightarrow y$ 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 4** ($[7]$). Let $S_n = (E_{n-m}, K_m)$ be oriented semi-transitively with $\vec{P} = p_1 \rightarrow \cdots \rightarrow p_m$. 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 5** ($[7]$). An orientation of a split graph $S_n = (E_{n-m}, K_m)$ is semi-transitive if and only if

(i) $K_m$ is oriented transitively;

(ii) each vertex in $E_{n-m}$ is of one of the three types in Theorem 3;

(iii) the restrictions in Theorem 4 are satisfied.
1.2 Split graphs generated by morphisms

Let $A$ and $B$ be alphabets (possibly $A = B$). A map $\varphi : A^* \to B^*$ is called a morphism, if we have $\varphi(uv) = \varphi(u)\varphi(v)$ for any $u, v \in A^*$. A morphism $\varphi$ can be defined by defining $\varphi(a)$ for each $a \in A$. A particular property of a morphism $\varphi$ is that $\varphi(\varepsilon) = \varepsilon$, where $\varepsilon$ is the empty word. Morphisms are a central object in the area of combinatorics on words [8], and there is a natural extension of the notion to two, or more, dimensions. Indeed, one can begin with a matrix $M$ whose entries are elements of $A$, and then obtain $\varphi(M)$ by substituting each element in $M$ by matrices having the same dimensions and given by some substitution rules.

Relevance of (2-dimensional) morphisms to split graphs is coming through the ideas communicated in [2], where patterns in adjacency matrices are considered to study word-representability of graphs, and the notion of an infinite word-representable graph is introduced. In this paper, we study word-representability of families of split graphs defined by iteration of morphisms, and in particular, we give a complete classification in the case of $2 \times 2$ matrices. However, a key theorem we prove that characterizes word-representability of split graphs in terms of permutations of columns in the adjacency matrix is applicable to any split graph (see Theorem 15).

1.3 Organization of the paper

After introducing a number of preliminary results and examples in Section 2 that includes our key result (Theorem 15), in Section 3 we present a number of general results on split graphs generated by morphisms. In Section 4 we provide a complete classification of word-representable split graphs defined by iteration of morphisms using two $2 \times 2$ matrices; our results in this section are summarized in Tables 1 and 2. Finally, in Section 5 we state a number of open research questions.

2 Preliminaries

Definition 6. Let $M$ be a binary $m \times n$ matrix. Define $S(M)$ to be the matrix

\[
\begin{bmatrix}
L_n & M^T \\
M & O_m
\end{bmatrix}
\]

where $O_m$ is the $m \times m$ zero matrix and $L_n$ is the $n \times n$ matrix such that all diagonal entries are 0's and all other entries are 1's.
Table 1: The index of word-representability of infinite split graphs $G(A, B)$ for $2 \times 2$ matrices $A$ and $B$. 
| Case | A | B | IWR(A, B) | Ref. | Case | A | B | IWR(A, B) | Ref. |
|------|---|---|-----------|------|------|---|---|-----------|------|
| 60   | 0 | 1 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 82 | 0 | 1 | 0 | 0 | ∞ | Prop. 24 |
| 61   | 1 | 0 | 1 | 0 | 0 | 7 | Rem. 34 | 83 | 1 | 0 | 0 | 4 | Rem. 34 |
| 62   | 0 | 1 | 0 | 0 | 0 | 3 | Rem. 34 | 84 | 0 | 0 | 0 | ∞ | Rem. 34 |
| 63   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 85 | 0 | 0 | 1 | 4 | Rem. 34 |
| 64   | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 86 | 0 | 0 | 0 | ∞ | Rem. 34 |
| 65   | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 87 | 0 | 0 | 4 | Rem. 34 |
| 66   | 1 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 88 | 1 | 0 | 0 | ∞ | Prop. 24 |
| 67   | 1 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 89 | 0 | 0 | 1 | 1 | 3 | Rem. 19 |
| 68   | 0 | 1 | 0 | 0 | 0 | 3 | Rem. 34 | 90 | 0 | 1 | 0 | 1 | 3 | Rem. 19 |
| 69   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 91 | 0 | 1 | 0 | 4 | Rem. 34 |
| 70   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 92 | 0 | 0 | 1 | 1 | 4 | Rem. 34 |
| 71   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 93 | 0 | 0 | 1 | 1 | 4 | Rem. 34 |
| 72   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 94 | 0 | 0 | 1 | 1 | 4 | Rem. 34 |
| 73   | 0 | 0 | 1 | 0 | 0 | 3 | Rem. 34 | 95 | 0 | 0 | 1 | 1 | 4 | Rem. 34 |
| 74   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 96 | 0 | 1 | 1 | ∞ | case 94 |
| 75   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 97 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 76   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 98 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 77   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 99 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 78   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 100 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 79   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 101 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 80   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 102 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 81   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 103 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 82   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 104 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 83   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 105 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 84   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 106 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 85   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 107 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 86   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 108 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 87   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 109 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 88   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 110 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 89   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 111 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 90   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 112 | 0 | 1 | 1 | ∞ | Prop. 24 |
| 91   | 0 | 0 | 0 | 0 | 0 | 0 | 3 | Rem. 34 | 113 | 0 | 1 | 1 | ∞ | Prop. 24 |

Table 2: The remaining cases of the index of word-representability of infinite split graphs $G(A, B)$ for $2 \times 2$ matrices $A$ and $B$. 


It is easy to see that for any binary $m \times n$ matrix $M$, $S(M)$ is the adjacency matrix of a split graph with the maximal clique of order $n$ or $n + 1$. We denote the split graph by $G(M)$. Clearly, $M$ gives edges between the clique and independent set in $G(M)$, and the order of the maximal clique depends on the existence of a $11 \cdots 1$ row in $M$.

**Example 7.** If $M =$

$$
\begin{bmatrix}
1 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
$$

then

$$
S(M) =
\begin{bmatrix}
0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

is the adjacency matrix of the graph shown in Figure 2.

**Remark 8.** If $M$ is a zero matrix, then $G(M)$ is a disjoint union of a clique and isolated vertices and is word-representable, because the clique is semi-transitively (in fact, transitively) orientable and there are no other edges in $G(M)$.

The following lemma is Lemma 8 in [7].

**Lemma 9** ([7]). Let $S_n = (E_{n-m}, K_m)$ be a split graph with the maximum clique $K_m$, and a split graph $S_{n+1}$ is obtained from $S_n$ by either adding a vertex of degree 0 (to $E_{n-m}$), or adding a vertex of degree 1 (to $E_{n-m}$), or by “coping” a vertex (either in $E_{n-m}$ or in $K_m$), that is, by adding a vertex whose neighbourhood is identical to the neighbourhood of a vertex in $S_n$. Then $S_n$ is word-representable if and only if $S_{n+1}$ is word-representable.
To analyze word-representability of a given split graph, by Lemma 9 we can delete vertices of degree 0 and vertices of degree 1, as well as delete all but one vertex having the same neighborhood.

**Proposition 10.** Let $M$ be an $m \times n$ matrix. If every row, or every column of $M$ is of the form $00 \cdots 0$ or $11 \cdots 1$, then $G(M)$ is word-representable.

**Proof.** If every row of $M$ consists of all 0’s or all 1’s then in $G(M)$, each vertex in the independent set is an isolated vertex, or is connected to every vertex in the clique. By Lemma 9, word-representability of $G(M)$ is equivalent to word-representability of either the clique $K_{n+1}$, or the disjoint union of a clique ($K_n$ or $K_{n+1}$) and an isolated vertex, which are clearly semi-transitively orientable, and thus, by Theorem 1 word-representable.

On the other hand, if every column of $M$ consists of all 0’s or all 1’s then the neighborhood of each vertex in the independent set is the same and, by Lemma 9, word-representability of $G(M)$ is equivalent to word-representability of the clique $K_n$ with a vertex $x$ connected to some, maybe none or all of clique’s vertices. If w.l.o.g. $x$ is connected to vertices $1, 2, \ldots, p$, $0 \leq p \leq n$, in $K_n$ formed by the vertices $1, 2, \ldots, n$, then the word $x12\cdots px(p+1)(p+2)\cdots n$ represents the graph. □

It is obvious that if $M^*$ is a matrix obtained by a row or column permutation of a matrix $M$, then $G(M^*)$ is a split graph obtained by relabelling the vertices of the graph $G(M)$. Hence we get the following lemma.

**Lemma 11.** Let $M$ be an $m \times n$ binary matrix and $M^*$ is a matrix obtained from a sequence of row and/or column permutations of $M$. Then, $G(M)$ is word-representable if and only if $G(M^*)$ is word-representable.

**Lemma 12.** Let $M := [m_{ij}]_{m \times n}$ be an $m \times n$ binary matrix such that $m_{k1} = m_{k2} = \cdots = m_{kn} = 1$ for some $k \in \{1, 2, \ldots, m\}$. If

$$N = \begin{bmatrix}
  m_{11} & m_{12} & \cdots & m_{1n} & 0 \\
  m_{21} & m_{22} & \cdots & m_{2n} & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  m_{(k-1)1} & m_{(k-1)2} & \cdots & m_{(k-1)n} & 0 \\
  m_{(k+1)1} & m_{(k+1)2} & \cdots & m_{(k+1)n} & 0 \\
  \vdots & \vdots & \ddots & \vdots & \vdots \\
  m_{m1} & m_{m2} & \cdots & m_{mn} & 0
\end{bmatrix},$$

is an $(m - 1) \times (n + 1)$ binary matrix, then $G(M)$ is isomorphic to $G(N)$. 

8
Figure 3: Non-word-representable split graphs $T_1, T_2, T_3, T_4$.

Proof. Let $M^*$ be the matrix obtained from $M$ by making the $k^{th}$ row be the first row. That is,

$$M^* = \begin{bmatrix}
1 & 1 & \cdots & 1 \\
m_{11} & m_{12} & \cdots & m_{1n} \\
m_{21} & m_{22} & \cdots & m_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
m_{(k-1)1} & m_{(k-1)2} & \cdots & m_{(k-1)n} \\
m_{(k+1)1} & m_{(k+1)2} & \cdots & m_{(k+1)n} \\
\vdots & \vdots & \ddots & \vdots \\
m_{m1} & m_{m2} & \cdots & m_{mn}
\end{bmatrix}.$$  

Since $M^*$ is obtained by reordering rows of $M$, $G(M^*)$ is obtained by relabeling the vertices of $G(M)$, and thus $G(M^*)$ is isomorphic to $G(M)$. Note that $S(M^*) = S(N)$, and so $G(M^*)$ and $G(N)$ are the same graph. Hence $G(M)$ is isomorphic to $G(N)$.

It is known [5] that there are no non-word-representable graphs of order less than 6 and the only non-word-representable graph on 6 vertices is the wheel graph $W_5$, which is not a split graph. Thus, $G(M)$ is word-representable if $M$ is an $m \times n$ matrix and $m + n \leq 6$. In [7], it is shown that any split graph $S$ with maximum clique $K_4$ is word-representable if and only if $S$ does not contain the graphs $T_1, T_2, T_3$ and $T_4$ shown in Figure 3 as induced subgraphs. As a corollary to this result, we have the following theorem.

**Theorem 13.** Let $A$ be an $m \times 4$ binary matrix without all $1$’s rows. Then, $G(A)$ is word-representable if and only if the rows and columns of $A$ cannot be permuted to be a matrix containing

$$\begin{bmatrix}1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix}1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

9
or \[
\begin{bmatrix}
1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0
\end{bmatrix}
\]
as a submatrix.

Let \(x^n\) denote \(xx\cdots x\) where \(x\) is repeated \(n\) times. The next theorem gives a sufficient condition for word-representability of a given graph \(G(M)\).

**Theorem 14.** Let \(M\) be an \(m \times n\) binary matrix. If there is a sequence of column permutations of \(M\) giving the matrix such that each row of \(M\) is of the form \(0^r1^s0^t\) for some non-negative integers \(r, s, t\), then \(G(M)\) is word-representable.

**Proof.** Assume that \(M^*\) is the matrix obtained from a sequence of column permutations of \(M\) and each row of \(M^*\) is of the form \(0^r1^s0^t\) where \(r, s, t\) are non-negative integers. Let the \(i\)th row/column of the adjacency matrix \(S(M^*)\) correspond to vertex \(i\) in \(G(M^*)\). So the clique \(C\) in \(G(M^*)\) contains vertices \(1, 2, \ldots, n\) and the independent set \(I\) in \(G(M^*)\) contains vertices \(n + 1, n + 2, \ldots, n + m\). Assign the orientation of edges in \(G(M^*)\) as \(i \to j\) if and only if \(i < j\). We have that \(1 \to 2 \to \cdots \to n\) is the longest path in the transitively oriented \(C\), and the edges between \(C\) and \(I\) are oriented from \(C\) to \(I\). Thus, each edge in the independent set is of type B, and we are done by Theorems 1 and 5.

We have the following important generalization of Theorem 14.

**Theorem 15.** Let \(M\) be an \(m \times n\) binary matrix without all 1’s rows. The split graph \(G(M)\) is word-representable if and only if \(M\) satisfies the following conditions:

(i) there is a sequence of column permutations of \(M\) giving a matrix \(M^*\) where every row is of the form \(0^r1^s0^t\) or \(1^r0^s1^t\) for some nonnegative integers \(r, s, t\), and

(ii) for any row of \(M^*\) of the form \(1^a0^b1^c\) for some positive integer \(a, b, c\), there is no other row not of the form \(1^a0^b1^c\) in which the entries in positions \(a\) and \(a + b + 1\) are 1’s.

**Proof.** “\(\Leftarrow\)” Assign the orientation of edges in \(G(M^*)\) as \(i \to j\) if \(i < j\) except if \(j > n\) (i.e. \(j\) is a vertex in the independent set) and the row in \(M^*\) corresponding to \(j\) is of the form \(1^r0^s1^t\), in which case we still orient \(i \to j\) for \(1 \leq i \leq r\) but \(j \to i\) for \(r + s + 1 \leq i \leq n\). The vertices in the independent set will then be of types B and C, and taking into account condition (ii), Theorems 1 and 6 can be applied to see that \(G(M^*)\) is word-representable, and thus \(G(M)\) is word-representable by Lemma 1.
“⇒.” By Theorem 11, \( G(M) \) admits a semi-transitive orientation. By Theorem 5, under this orientation the clique is oriented transitively, and we can rename the vertices of the clique, if necessary so that the longest path would be formed by \( 1 \rightarrow 2 \rightarrow \cdots \rightarrow n \). Note that renaming vertices in the clique corresponds to permuting columns in \( M \) giving \( M^* \). But then, conditions (ii) and (iii) in Theorem 5 give conditions (i) and (ii) in this theorem.

**Remark 16.** If \( M \) has an all 1’s row, we can see that Theorems 13 and 15 cannot be applied. However, we can use Theorem 12 to change \( M \) into an \( (m-1) \times (n+1) \) matrix \( N \) which does not contain all 1’s row. So we can apply the theorems to matrix \( N \) instead of \( M \) because \( G(N) \) is isomorphic to \( G(M) \). This observation also applies to Corollary 23 below.

We can see that Theorem 15 allows us to answer the question on word-representability of \( G(M) \) by looking at permutations of columns in \( M \). Let \( M = [m_{ij}]_{m \times n} \) be an \( m \times n \) matrix and \( \rho = \rho_1 \rho_2 \cdots \rho_n \) is a permutation of \( \{1, 2, \ldots, n\} \) written in one-line notation. We say that \( M^* = [m_{\rho_1 m_{\rho_2} \cdots m_{\rho_n}}]_{m \times n} \) is the matrix obtained from reordering columns of \( M \) in the order given by \( \rho \). The key approach given by Theorem 15 is finding a permutation \( \rho \) that turns each row of \( M^* \) into the form \( 0^r 1^s 0^t \) or \( 1^r 0^s 1^t \) (so, all 1’s in \( M^* \) are cyclically consecutive). Interestingly, to prove word-representability results in this paper, only rows of the form \( 0^r 1^s 0^t \) are used, so that condition (ii) in Theorem 15 is not applicable.

**Example 17.** In the matrix \( M = \)

\[
\begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

we can ignore rows 2, 4, 6 and 8 because all entries in these rows are zero. Then we need to find a permutation \( \rho = \rho_1 \rho_2 \cdots \rho_8 \) making

- columns 1, 4, 6 and 7 be (cyclically) consecutive in rows 1;
- columns 1, 3, 5 and 7 be (cyclically) consecutive in row 3 and 7;
- columns 2, 3, 6 and 7 be (cyclically) consecutive in row 5.
It can be implied from the first and the second bullet points that 1 and 7 must be consecutive in \( \rho \) and then, w.l.o.g., 4 and 6 are next to the left of these numbers and 3 and 5 are next to the right of them (cyclically). Hence \( \rho \) contains
\[
\{4, 6\}, \{1, 7\}, \{3, 5\}
\]
where numbers in \( \{ \} \) are consecutive in \( \rho \) but are in some unknown to us order. But then, we get a contradiction with the second bullet point. Hence, there is no such \( \rho \) and \( G(M) \) is non-word-representable by Theorem 15.

3 General results on split graphs generated by morphisms

In this section, we discuss rather general results on split graphs generated by morphisms, thus preparing ourselves for a classification of the case of \( 2 \times 2 \) matrices coming in the next section.

**Definition 18.** Let \( A, B \) be \( m \times n \) binary matrices. The matrix \( M^k(A, B) \) is said to be the \( k \)th-iteration of the 2-dimensional morphism applied to the \( 1 \times 1 \) matrix \([0]\) which maps \([0] \to A\) and \([1] \to B\). Moreover, we write \( S^k(A, B) \) for the matrix \( S(M^k(A, B)) \) and \( G^k(A, B) \) for the graph with the adjacency matrix \( S^k(A, B) \).

**Example 19.** Let \( A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \) and \( B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \). Then we have
\[
M^0(A, B) = [0], \quad M^1(A, B) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad M^2(A, B) = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.
\]

Then, \( S^2(A, B) = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \) and \( G^2(A, B) \) is shown in Figure 4.

**Remark 20.** If \( A \) is a zero matrix, then \( M^k(A, B) \) is always a zero matrix for any \( m \times n \) matrix \( B \) and positive integer \( k \). So, by Remark 8, \( G^k(A, B) \) is word-representable.
3.1 The case of \( A = B \)

In this case, both \([0]\) and \([1]\) are mapped to the same matrix \( A \), so if \( A \) is an \( m \times n \) binary matrix, then

\[
S^k(A, A) = \begin{bmatrix} L_{n^k} & A_k^T \\ A_k & O_{m^k} \end{bmatrix}
\]

where

\[
A_k = \begin{bmatrix} A & A & \cdots & A \\ A & A & \cdots & A \\ \vdots & \vdots & \ddots & \vdots \\ A & A & \cdots & A \end{bmatrix}.
\]

Clearly, \( A_k \) is an \( n^k \times m^k \) matrix and \( S^k(A, A) = S(A_k) \), so \( G^k(A, A) \) is isomorphic to \( G(A_k) \).

**Theorem 21.** Let \( A \) be an \( m \times n \) binary matrix. For \( k \geq 1 \), \( G^k(A, A) \) is word-representable if and only if \( G(A) \) is word-representable.

**Proof.** Firstly, we label a vertex of \( G^k(A, A) \) by \( i \) if it is represented by the \( i^{th} \) column/row in \( S^k(A, A) \). Note that rows \( i, m+i, 2m+i, \ldots, (m^k-1)m+i \) in \( A_k \) are identical for any \( i \in \{1, 2, \ldots, m\} \), and columns \( j, n+j, 2n+j, \ldots, (n^k-1)n+j \) in \( A_k \) are also identical for any \( j \in \{1, 2, \ldots, n\} \). So, for any \( i \in \{1, 2, \ldots, m\} \), the vertices of \( G^k(A, A) \) in \( R_i := \{i+n^k, m+i+n^k, 2m+i+n^k, \ldots, (m^k-1)m+i+n^k\} \) have the same neighborhoods. Similarly, any two vertices of \( G^k(A, A) \) in \( C_j := \{j, n+j, 2n+j, \ldots, (n^k-1)n+j\} \) are connected to the same vertices in the independent set for any \( j \in \{1, 2, \ldots, n\} \). Thus, by Lemma 9, \( G^k(A, A) \) is word-representable if and only if the graph \( G \) obtained by deleting all vertices but the smallest one in \( R_i \) and \( C_j \) for all \( i, j \in \{1, 2, \ldots, n\} \) is word-representable. But \( G \) is exactly \( G(A) \), which complete the proof. \( \square \)

**Corollary 22.** If \( A \) is an \( m \times n \) binary matrix such that \( m + n \leq 6 \), then \( G^k(A, A) \) is word-representable for any \( k \geq 0 \).
Proof. Since the smallest non-word-representable split graph is of order 7, all split graphs of orders less than 7 are word-representable. Hence $G^k(A, A)$ is word-representable for any $m \times n$ matrix $A$ where $m + n \leq 6$.

Moreover, together with Theorem [13], we have the following result.

**Corollary 23.** Let $A$ be an $m \times 4$ binary matrix with no all 1’s row. For any integer $k$, the graph $G^k(A, A)$ is word-representable if and only if the rows and columns of $A$ cannot be permuted to be the matrix containing
\[
\begin{bmatrix}
1 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 \\
1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1
\end{bmatrix}
\]
as a submatrix.

### 3.2 The case of $A \neq B$

In what follows, $A$ and $B$ can be distinct.

**Proposition 24.** If every row, or every column, in $m \times n$ matrices $A$ and $B$ is either $0^n$ or $1^n$, then $G^k(A, B)$ is word-representable for any $k \geq 0$.

Proof. If every row (resp., column) in $A$ and $B$ is either $0^n$ or $1^n$, then every row (resp., column) in $M^k(A, B)$ is either $0^n$ or $1^n$, so by Proposition [10] $G^k(A, B)$ is word-representable.

**Theorem 25.** Let $A$ and $B$ be $m \times n$ binary matrices. Suppose that $A^*$ and $B^*$ are the matrices obtained from reordering columns of $A$ and $B$, respectively, in order given by a permutation $\sigma = \sigma_1 \sigma_2 \cdots \sigma_n$. Then $G^k(A, B)$ is word-representable if and only if $G^k(A^*, B^*)$ is word-representable for any $k \geq 0$.

Proof. The case $k = 0$ is trivial, so assume that $k \geq 1$. We claim that $M^k(A^*, B^*)$ is obtained from a permutation of columns in $M^k(A, B)$. We will prove the claim by induction on $k$. Note that $M^1(A, B) = A$ and $M^1(A^*, B^*) = A^*$. So $M^1(A^*, B^*)$ is the matrix obtained from reordering columns of $M^1(A, B)$. Suppose that $l$ is a positive integer and $M^l(A^*, B^*)$ is the matrix obtained from reordering columns of $M^l(A, B)$ in order given by a permutation $\tau = \tau_1 \tau_2 \cdots \tau_l$. Let $M^l(A, B) = [C_1, C_2, \ldots, C_n]$ where $C_i$ is the $i^{th}$ column of $M^l(A, B)$. Then $M^l(A^*, B^*) = [C_{\tau_1}, C_{\tau_2}, \ldots, C_{\tau_l}]$. For the next iteration of morphism, each column $C_i$ of $M^l(A, B)$ is mapped to $n$ columns $C_{i,1}, C_{i,2}, \ldots, C_{i,n}$, and each column $C_{\tau_i}$ of $M^l(A^*, B^*)$ is mapped to $n$ columns $C_{\tau_i, \tau_1}, C_{\tau_i, \tau_2}, \ldots, C_{\tau_i, \tau_n}$. So we have

$$M^{l+1}(A, B) = [C_{1,1}, C_{1,2}, \ldots, C_{1,n}, C_{n^*,1}, C_{n^*,2}, \ldots, C_{n^*,n}]$$
and
\[ M_{l+1}(A^*, B^*) = \begin{bmatrix}
C_{\tau_1, \sigma_1} & C_{\tau_1, \sigma_2} & \cdots & C_{\tau_1, \sigma_n} & \cdots & C_{\tau_n, \sigma_1} \\
C_{\tau_n, \sigma_2} & \cdots & C_{\tau_n, \sigma_n}
\end{bmatrix}. \]

A group of columns \( C_{i,1}, C_{i,2}, \ldots, C_{i,n} \) is called block \( B_i \). Firstly, we can see that reordering the blocks \( B_1, B_2, \ldots, B_n \) of \( M_{l+1}(A, B) \) in order given by \( \tau \), and then reordering columns in every block \( B_i \) in order given by \( \sigma \), yields the matrix \( M_{l+1}(A^*, B^*) \). Thus, \( M_{l+1}(A^*, B^*) \) is obtained by a column permutation of \( M_{l+1}(A, B) \) and our claim is true. Hence, by Lemma 11, \( G^k(A, B) \) is word-representable if and only if \( G^k(A^*, B^*) \) is word-representable for any positive integer \( k \).

Next theorem is a natural extension of Theorem 25 to the case of row permutations, and it can be proved in a similar way to the proof of Theorem 25, so we omit the proof.

**Theorem 26.** Let \( A \) and \( B \) be \( m \times n \) binary matrices. Suppose that \( A^* \) and \( B^* \) are the matrices obtained from reordering rows of \( A \) and \( B \), respectively, in order given by the same permutation. Then \( G^k(A, B) \) is word-representable if and only if \( G^k(A^*, B^*) \) is word-representable for any \( k \geq 0 \).

So, we can reorder rows and columns of given matrices \( A \) and \( B \) while preserving the word-representability of \( G^k(A, B) \). If \( A \) contains at least one 0, we can reorder rows and columns of \( A \) to make the leftmost bottom entry be a 0 (the matrix \( B \) will be changed by the same permutation of rows and columns as those applied to \( A \)). Thus, in what follows, if \( A \) is not all-one matrix, w.l.o.g. we can assume that the leftmost bottom entry of \( A \) is always 0. Then, the \( m \times n \) leftmost bottom submatrix of \( M^2(A, B) \) is \( A \) since \( M^1(A, B) = A \). Moreover, the \( m^{k-1} \times n^{k-1} \) leftmost bottom submatrix of \( M^k(A, B) \) is \( M^{k-1}(A, B) \). Thus, the limit \( \lim_{k \to \infty} M^k(A, B) \), called a fixed point of the morphism, is well-defined. So, we have that \( G^k(A, B) \) is an induced subgraph of \( G^k(A, B) \) if \( i \leq k \) and the notion of the infinite split graph \( G(A, B) \) is well-defined in the case when \( A \) has a 0 as the leftmost bottom entry. So we are interested in the smallest integer \( l \) (possibly non-existing) that \( G^l(A, B) \) is non-word-representable for given \( A \) and \( B \) (then \( G^l(A, B) \) is non-word-representable for \( i \geq l \)).

**Definition 27.** Suppose that a matrix \( A \) has a 0 as the leftmost bottom entry. The index of word-representability \( \text{IWR}(A, B) \) of an infinite split graph \( G(A, B) \) is the smallest integer \( l \) such that \( G^l(A, B) \) is non-word-representable. If such \( l \) does not exist, that is, if \( G^l(A, B) \) is word-representable for all \( l \), then \( l := \infty \). If the leftmost bottom entry of \( A \) is 1 (so that the \( \lim_{k \to \infty} M^k(A, B) \) may not be well-defined as the sequence of graphs \( G^k(A, B) \), for \( k \geq 0 \), may not be a chain of
induced subgraphs) then $\text{IWR}(A, B)$ is still defined in the same way even though $G(A, B)$ may not be defined.

Note that since $G^0(A, B)$ is a graph with one vertex, we have $\text{IWR}(A, B) \geq 1$. Even though Definition 27 is very similar to the respective definition of the index of word-representability of an infinite Toeplitz graph in [2] (where the index in our context would be the maximum $l$ such that $G^l(A, B)$ is word-representable), it is more flexible as it makes sense in the situation when the leftmost bottom entry of $A$ is 1.

**Theorem 28.** Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ binary matrices, and

$$C = \begin{bmatrix} a_{p_1q_1} & a_{p_1q_2} & \cdots & a_{p_1q_t} \\ a_{p_2q_1} & a_{p_2q_2} & \cdots & a_{p_2q_t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p_sq_1} & a_{p_sq_2} & \cdots & a_{p_sq_t} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} b_{p_1q_1} & b_{p_1q_2} & \cdots & b_{p_1q_t} \\ b_{p_2q_1} & b_{p_2q_2} & \cdots & b_{p_2q_t} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p_sq_1} & b_{p_sq_2} & \cdots & b_{p_sq_t} \end{bmatrix}$$

be $s \times t$ submatrices of $A$ and $B$, respectively, where $1 \leq p_1 < p_2 < \cdots < p_s \leq m$ and $1 \leq q_1 < q_2 < \cdots < q_t \leq n$. For any positive integer $k$, if $G^k(C, D)$ is non-word-representable, then $G^k(A, B)$ is non-word-representable.

**Proof.** First, we will prove by induction that $M^k(C, D)$ is a submatrix of $M^k(A, B)$ for any positive integer $k$. It is obvious that $M^1(C, D) = C$ is a submatrix of $M^1(A, B) = A$. Let $l$ be a positive integer such that $M^l(C, D)$ is a submatrix of $M^l(A, B)$ on the columns $c_1, c_2, \ldots, c_l$ and rows $r_1, r_2, \ldots, r_s$. For the next iteration of morphism, $M^{l+1}(A, B)$ is formed by replacing each entry of $M^l(A, B)$ with either $A$ or $B$. So the columns $(c_i - 1)n + p_j$ for $1 \leq i \leq t, 1 \leq j \leq s$, and rows $(r_i - 1)m + q_j$ for $1 \leq i \leq s, 1 \leq j \leq t$, form the matrix $M^{l+1}(C, D)$. Hence $M^k(C, D)$ is a submatrix of $M^k(A, B)$ for any $k \geq 1$. Therefore, $G^k(A, B)$ contains $G^k(C, D)$ as an induced subgraph for $k \geq 1$. As the property of word-representability is hereditary, we have that non-word-representability of $G^k(C, D)$ implies non-word-representability of $G^k(A, B)$.

Theorem 28 gives a useful tool to study non-word-representability of $G^k(A, B)$ for larger $A$ and $B$. Indeed, a starting point to justify suspected non-word-representability of $G^k(A, B)$ can be analysis of smaller submatrices of $A$ and $B$. This is one of our motivation points to conduct a systematic study of IWR$(A, B)$ for $2 \times 2$ matrices, to be done in the next section, as they are smallest submatrices that can be used to show non-word-representability of $G^k(A, B)$ for some $A$, $B$ and $k$. 

16
4 Classification of word-representable split graphs defined by iteration of morphisms using two $2 \times 2$ matrices

A summary of our classification of word-representability of $G^k(A, B)$ for $2 \times 2$ matrices $A$ and $B$ can be found in Tables 1 and 2 where the index of word-representability $\text{IWR}(A, B)$ is given along with a reference, or a comment to the respective result.

4.1 The case of $A$ is not all-one matrix

For any $2 \times 2$ matrices $A$ and $B$, the graph $G^1(A, B)$ is a split graph of order 4 which is always word-representable. Then, $\text{IWR}(A, B) \geq 2$. However, $2 \times 2$ matrices $A$ and $B$ such that $G^2(A, B)$ is non-word-representable can be found.

**Proposition 29.** For $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, $\text{IWR}(A, B) = 2$.

**Proof.** We have $M^2(A, B) = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Reordering columns of $M^2(A, B)$ in order given by the permutation 2314 yields the matrix

$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. So, by Theorem 13, $G^2(A, B)$ is non-word-representable ($G^2(A, B)$ contains $T_1$ as an induced subgraph).

**Remark 30.** Permuting rows or/and columns in $A$ and $B$ similarly to Proposition 29 we see that $\text{IWR}(A, B) = 2$ for $A$ and $B$ in Cases 31, 45, 81 and 101 in Tables 1 and 2 (in Case 81 $G^2(A, B)$ contains $T_3$, and in the other cases $G^2(A, B)$ contains $T_1$).

**Proposition 31.** For $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\text{IWR}(A, B) = 3$. 

17
Proof. We have

\[
M^2(A, B) = \begin{bmatrix}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\quad \text{and} \quad
M^3(A, B) = \begin{bmatrix}
1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
\]

Reordering columns of \(M^2(A, B)\) in order given by the permutation 4231 yields the matrix

\[
\begin{bmatrix}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

By Theorem 14, we have \(G^2(A, B)\) is word-representable. However, we have shown in Example 17 that \(G^3(A, B)\) is non-word-representable. So \(\text{IWR}(A, B) = 3\).

Remark 32. Proposition 31 gives Case 4 in Tables 1. In each of Cases 5, 10, 11, 12, 14, 15, 17, 18, 22, 23, 24, 25, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 42, 43, 44, 46, 47, 49, 58, 69, 70, 71, 73, 74, 75, 76, 77, 78, 80, 89, 99, 102, 104, 106, 108, 109, 111 and 112 in Tables 1 and 2, \(M^2(A, B)\) has a permutation satisfying the conditions of Theorem 25 and \(M^3(A, B)\) does not (similarly to Proposition 31). So \(\text{IWR}(A, B) = 3\) for \(A\) and \(B\) in these cases. Moreover, by Theorem 25, column and row permutations of \(A\) and \(B\) give the same \(\text{IWR}\). Consequently, we also have \(\text{IWR}(A, B) = 3\) for \(A\) and \(B\) in Cases 19, 22, 24, 25, 28, 29, 32, 33, 36, 37, 39, 40, 42, 46, 47, 49, 58, 69, 70, 71, 75, 76, 79, 80 and 90.

Proposition 33. For \(A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\) and \(B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\), \(\text{IWR}(A, B) = 4\).

Proof. We have

\[
M^3(A, B) = \begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
\]

Reordering columns of \(M^3(A, B)\) in order given by the permutation...
51732648 yields the matrix
\[
\begin{bmatrix}
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

By Theorem 14, we have \(G^3(A, B)\) is word-representable. For
\[
M^4(A, B) =
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

suppose that a reordering of columns \(\rho = \rho_1\rho_2\cdots\rho_{16}\) exists showing word-representability of \(G^4(A, B)\) by Theorem 14. Then,

- from row 3, columns 1, 3, 5, 7, 9, 13 and 15 must be (cyclically) consecutive;
- from row 5, columns 3, 5, 7, 11 and 15 must be (cyclically) consecutive;
- from row 7, columns 1, 3, 5, 9, 11, 13 and 15 must be (cyclically) consecutive.

But then, from the first and the third bullet points, columns 1, 3, 5, 9, 13 and 15 must be consecutive and then column 7 or 11 is next to the left of them and the other one is next to the right of them. This contradicts to the second bullet point. So there is no such \(\rho\) and \(G^4(A, B)\) is non-word-representable. Therefore, \(\text{IWR}(A, B) = 4\).

**Remark 34.** Proposition 33 gives Case 6 in Table 1. In each of Cases 53, 63, 83, 87, 93, 98 and 105 in Tables 1 and 2, \(M^3(A, B)\) has a permutations satisfying condition in Theorem 22 but \(M^4(A, B)\) does not (similarly to Proposition 33). So \(\text{IWR}(A, B) = 4\) for \(A\) and \(B\) in these
columns and row permutations of $A$ and $B$ give the same IWR. Consequently, we also have $\text{IWR}(A, B) = 3$ for $A$ and $B$ in Cases 21, 35, 54, 64, 85, 92 and 95.

By Remarks 30, 32 and 34, we can see that in many cases the index of word-representability is 2, 3 or 4. Next, we will introduce certain definitions and theorems to present the cases where the index of word-representability is infinity.

Let $M$ be an $m \times n$ binary matrix. For convenience, we will represent rows of $M$ by binary strings of length $n$. For example, we will represent three rows of $\begin{bmatrix}1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1\end{bmatrix}$ by 1101, 0100 and 0001.

**Definition 35.** Let $A$ and $B$ be $m \times n$ binary matrices. Define $R_k(A, B)$ to be the set of binary strings representing rows of $M_k(A, B)$. So every element of $R_k(A, B)$ is a binary string of length $n^k$. Each element of $R_k(A, B)$ is called a row pattern of $M_k(A, B)$.

**Definition 36.** Let $A = \begin{bmatrix}a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix}e & f \\ g & h \end{bmatrix}$ be $2 \times 2$ binary matrices and $B^n$ be the set of binary strings of length $n$. We define functions $u_{A,B} : \{0, 1\} \rightarrow B^2$ and $l_{A,B} : \{0, 1\} \rightarrow B^2$ by

$u_{A,B}(0) = ab$, $l_{A,B}(0) = cd$, $u_{A,B}(1) = ef$ and $l_{A,B}(1) = gh$.

Moreover, if $v = v_1v_2 \cdots v_k \in B^k$, $k \geq 2$, we extend the definition of the functions $u_{A,B}$ and $l_{A,B}$ to the case of $B^k \rightarrow B^{2k}$ by

$u_{A,B}(v) = u_{A,B}(v_1)u_{A,B}(v_2) \cdots u_{A,B}(v_k)$

and

$l_{A,B}(v) = l_{A,B}(v_1)l_{A,B}(v_2) \cdots l_{A,B}(v_k)$.

When $A$ and $B$ are clear from the context, we can omit the subscript and write $u$ and $l$ instead of $u_{A,B}$ and $l_{A,B}$, respectively.

**Example 37.** Let $A = \begin{bmatrix}1 & 1 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix}1 & 0 \\ 1 & 0 \end{bmatrix}$. Then we have

$M^3(A, B) = \begin{bmatrix}1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

and $R^4(A, B) = \{10111011, 10001000, 10101010, 11111111, 00000000\}$. In fact, we can find $R^4(A, B)$ by using the functions $u_{(A, B)}$ and $l_{A,B}$.
As we start with $M^0(A, B) = \{0\}$, we have $R^0(A, B) = \{0\}$. Then we apply the functions $u_{A,B}$ and $l_{A,B}$ to all elements in $R^0(A, B)$ to get all elements in $R^1(A, B)$:

$$u_{A,B}(0) = 11 \text{ and } l_{A,B}(0) = 00.$$  

So, $R^1(A, B) = \{11, 00\}$. Now,

$$u_{A,B}(11) = 1010 \text{ and } l_{A,B}(11) = 1010$$  

$$u_{A,B}(00) = 1111 \text{ and } l_{A,B}(00) = 0000$$

so $R^2(A, B) = \{1010, 1111, 0000\}$. Repeating the procedure one more time yields

$$u_{A,B}(1010) = 10111011 \text{ and } l_{A,B}(1010) = 10001000$$  

$$u_{A,B}(1111) = 10101010 \text{ and } l_{A,B}(1111) = 10101010$$  

$$u_{A,B}(0000) = 11111111 \text{ and } l_{A,B}(0000) = 00000000,$$

and so $R^3(A, B) = \{10111011, 10001000, 10101010, 11111111, 00000000\}$ which is the same as $R^4(A, B)$.

So, all elements in $R^k(A, B)$ are obtained by applying $u_{A,B}$ and $l_{A,B}$ to every element in $R^{k-1}(A, B)$. The next theorem generalizes this observation, and it can be proved easily by induction.

**Theorem 38.** Let $A$ and $B$ be $2 \times 2$ binary matrices. Then

$$R^k(A, B) = \{f_k(\cdots f_2(f_1(0))\cdots)|f_i \in \{u_{A,B}, l_{A,B}\} \text{ for any } k \geq 1.$$  

**Definition 39.** Let $v = v_1v_2\cdots v_k \in B^k$. Then, $\Gamma(v) := \{m \in \{1, 2, \ldots, k\}|v_m = 1\}$.

In order to study row patterns, we introduce a relation $\leq$ on $B^n$. Let $x = x_1x_2\cdots x_k$ and $y = y_1y_2\cdots y_k$ be in $B^k$. We say that $x \leq y$ if and only if $x_i = 1$ implies $y_i = 1$ for every $i \in \{1, 2, \ldots, k\}$. In other words, $x \leq y$ if and only if $\Gamma(x) \subseteq \Gamma(y)$. It is easy to see that $\leq$ is reflexive, antisymmetric and transitive, and thus $\leq$ is a partial order.

**Theorem 40.** Let $A$ and $B$ be $m \times n$ binary matrices. For any $k > 1$, if $(R^k(A, B), \leq)$ is a total order, then $G^k(A, B)$ is word-representable.

**Proof.** Let $R^k(A, B) = \{x_1, x_2, \ldots, x_l\}$ where $l \geq 1$ and $x_1, x_2, \ldots, x_l$ are binary strings of length $n^k$. Since $(R^k(A, B), \leq)$ is a total order, w.l.o.g., we assume that $x_1 \leq x_2 \leq \cdots \leq x_l$. That is $\Gamma(x_1) \subseteq \Gamma(x_2) \subseteq \cdots \subseteq \Gamma(x_l)$.
\[
\cdots \subseteq \Gamma(x_i). \text{ Let }
\]
\[
D_1 := \Gamma(x_1), \quad D_2 := \Gamma(x_2) \setminus \Gamma(x_1), \quad D_3 := \Gamma(x_3) \setminus \Gamma(x_2), \\
\vdots \\
D_l := \Gamma(x_l) \setminus \Gamma(x_{l-1}) \quad \text{and} \quad D_{l+1} := \{1, 2, \ldots, n^k\} \setminus \Gamma(x_l).
\]

If \(D_j = \{i_{j,1}, i_{j,2}, \ldots, i_{j,|D_j|}\}\) for \(i_{j,1} < i_{j,2} < \cdots < i_{j,|D_j|}\), then
\[
\rho = i_{1,1}i_{1,2} \cdots i_{1,|D_1|}i_{2,1}i_{2,2} \cdots i_{2,|D_2|} \cdots i_{l,1}i_{l,2} \cdots i_{l,|D_l|}
\]
is a \(n^k\)-permutation. Let \(M^*\) be the matrix obtained by reordering columns of \(M^k(A, B)\) according to the order given by \(\rho\). Then we want to prove that every row of \(M^*(A, B)\) is of the form \(1^s0^t\) for some \(s, t \geq 0\). Let \(y = y_1y_2 \cdots y_{n^k}\) be a row pattern of \(M^k(A, B)\) and \(y^* = y_1' \cdots y_{n^k}'\) be the row pattern after reordering \(y\). Since \(y \in R^k(A, B)\), then \(y = x_q\) for some \(q \in \{1, 2, \ldots, l\}\). So \(y_i = 1\) for all \(i \in \Gamma(x_q)\) and \(y_i = 0\) for all \(i \notin \Gamma(x_q)\). That is, \(\Gamma(y)\) is
\[
\{i_{1,1}, i_{1,2}, \ldots, i_{1,|D_1|}, i_{2,1}, i_{2,2}, \ldots, i_{2,|D_2|}, \ldots, i_{q,1}, i_{q,2}, \ldots, i_{l,|D_l|}\}.
\]
Hence \(y^* = 1^s0^t\) where \(s = |D_1| + |D_2| + \cdots + |D_q|\) and \(t = n^k - |D_1| - |D_2| - \cdots - |D_q|\). Therefore every row of \(M^*(A, B)\) is of the form \(1^s0^t\).

By Theorem 41 we have that \(G^k(A, B)\) is word-representable. \(\Box\)

**Theorem 41.** Let \(A\) and \(B\) be \(2 \times 2\) binary matrices. If \(x \leq y\) implies \(\{u_{A,B}(x), l_{A,B}(x), u_{A,B}(y), l_{A,B}(y)\}\) is comparable under \(\leq\) for binary strings \(x\) and \(y\) of the same length, then \(G^k(A, B)\) is word-representable for any \(k \geq 0\).

**Proof.** We will prove by induction that \(R^k(A, B)\) is comparable under \(\leq\). Because \(0 \leq 0\), we have \(\{u_{A,B}(0), l_{A,B}(0)\} = R^1(A, B)\) is comparable under \(\leq\). Suppose that \(R^l(A, B)\) is comparable under \(\leq\) for some \(l \geq 1\). Let \(v, w \in R^{l+1}(A, B)\). Then
\[
v = u_{A,B}(x) \quad \text{or} \quad v = l_{A,B}(x) \quad \text{for some} \quad x \in R^l(A, B)
\]
and
\[
w = u_{A,B}(y) \quad \text{or} \quad w = l_{A,B}(y) \quad \text{for some} \quad y \in R^l(A, B).
\]
By induction hypothesis, w.l.o.g., we can assume that \(x \leq y\). So, \(\{u_{A,B}(x), l_{A,B}(x), u_{A,B}(y), l_{A,B}(y)\}\) is comparable and \(v\) and \(w\) belong to this set. Thus we have that \(v\) and \(w\) are comparable. Hence \(R^k(A, B)\) is comparable under \(\leq\) for any \(k \geq 0\). Hence, by Theorem 41, \(G^k(A, B)\) is word-representable for any \(k \geq 0\). \(\Box\)
Theorem 42. Let $A$ and $B$ be $2 \times 2$ binary matrices. If $\{u_{A,B}(100), l_{A,B}(100), u_{A,B}(101), l_{A,B}(101)\}$ is comparable under $\le$, then $G^k(A, B)$ is word-representable for any $k \ge 0$.

Proof. For convenience, we write $u$ and $l$ instead of $u_{A,B}$ and $l_{A,B}$, respectively. Suppose that $\{u(100), l(100), u(101), l(101)\}$ is comparable under $\le$. Let $x = x_1x_2 \cdots x_m$ and $y = y_1y_2 \cdots y_m$ be binary strings of length $m$ such that $x \le y$. Also, let

$$R := \{i|x_i = 0, y_i = 0\},$$
$$S := \{i|x_i = 1, y_i = 1\}$$
and
$$T := \{i|x_i = 0, y_i = 1\}.$$

Note that $R, S, T$ partition the set $\{1, 2, \ldots, m\}$. Since $u(100)$ and $l(100)$ are comparable, we have two cases to consider.

If $u(100) \le l(100)$, then $u(1) \le l(1)$ and $u(0) \le l(0)$. So we have $u(101) \le l(101)$ and it is impossible that $l(101) \le u(100)$ and $l(100) \le u(101)$. So there are four possible cases here, which are

$$u(100) \le u(101) \le l(100) \le l(101),$$
$$u(100) \le u(101) \le l(101) \le l(100),$$
$$u(101) \le u(100) \le l(101) \le l(100)$$
and
$$u(101) \le u(100) \le l(100) \le l(101).$$

Similarly, in the case of $l(100) \le u(100)$, we have $l(1) \le u(1)$ and $l(0) \le u(0)$. So $l(101) \le u(101)$ and $u(100) \not\le l(100)$ and $u(101) \not\le l(100)$. So we have four more cases, which are

$$l(100) \le u(100) \le l(101) \le u(101),$$
$$l(100) \le l(101) \le u(101) \le u(100),$$
$$l(101) \le l(100) \le u(101) \le u(100)$$
and
$$l(101) \le l(100) \le u(100) \le u(101).$$

Next, we will consider comparability of the set $\{u(x), l(x), u(y), l(y)\}$ in each case.

- $u(100) \le u(101) \le l(100) \le l(101)$. So we have $u(0) \le u(1) \le l(0) \le l(1)$. Note that $u(x_i) \le u(y_i)$ and $l(x_i) \le l(y_i)$ for any $i \in T$. Then $u(x) \le u(y)$ and $l(x) \le l(y)$. Since $u(y_i) \le l(x_i)$ where $i$ belongs to $R, S$ or $T$, so $u(y) \le l(x)$. Hence, $u(x) \le u(y) \le l(x) \le l(y)$.

- $u(100) \le u(101) \le l(101) \le l(100)$. So we have $u(0) \le u(1) \le l(1) \le l(0)$. Note that $u(x_i) \le u(y_i)$ and $l(y_i) \le l(x_i)$ for any $i \in T$. Then $u(x) \le u(y)$ and $l(x) \le l(y)$. Since $u(y_i) \le l(y_i)$ where $i$ belongs to $R, S$ or $T$, so $u(y) \le l(y)$. Hence, $u(x) \le u(y) \le l(y) \le l(x)$.  

23
Remark 43. Theorem 42 can be applied to check word-representability of $G^k(A, B)$. We can see that \{u_{A,B}(100), l_{A,B}(100), u_{A,B}(101), l_{A,B}(101)\} is comparable under $\leq$ in Cases 2, 7, 8, 13, 51, 56, 61, 84, 94, 100, 103, 107 and 113 in Tables 2 and 3. Then IWR(A, B) = $\infty$ for A and B in these cases. By Theorem 25, column and row permutations of A and B in these cases preserve the index of word-representability. So we also have IWR(A, B) = $\infty$ for A and B in Cases 18, 23, 27, 30, 34, 43, 44, 48, 52, 59, 62, 86 and 96.
Proposition 44. Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $G^k(A, B)$ is word-representable for any $k \geq 0$.

Proof. The case of $k = 0$ is trivial. For $k \geq 1$, we let $S_0^k := \emptyset$,

$$S_j^k := \{s \in \{1, 2, \ldots, 2^k\} | s \equiv 2^j \pmod{2^j} \} \text{ for } j \in \{1, 2, \ldots, k\}$$

and

$$T^k := \{x_1x_2 \cdots x_{2^k} | x_i = \begin{cases} 0 & \text{if } i \notin S_j^k, \\ 1 & \text{if } i \in S_j^k \end{cases} \text{ for some } j \in \{0, 1, \ldots, 2^k\} \}.$$ 

We claim that $R^k(A, B) = T^k$ for any $k \geq 1$ and prove it by induction on $k$. It is obvious in the case of $k = 1$ because $R^1(A, B) = \{00, 10\}$ while $S_0 = \emptyset$ and $S_1 = \{1\}$. Suppose $l$ is a positive integer such that $R^l(A, B) = T^l$. Let $y \in R^{l+1}(A, B)$, then $y = u_{A,B}(z)$ or $y = l_{A,B}(z)$ for some $z \in R^l(A, B)$. Since $R^l(A, B) = T^l$, we have

$$z = z_1z_2 \cdots z_{2^l} \text{ where } z_i = \begin{cases} 0 & \text{if } i \notin S_j^l, \\ 1 & \text{if } i \in S_j^l \end{cases}.$$ 

If $y = u_{A,B}(z)$, then $y = 1010 \cdots 10$. That is, $y_i = \begin{cases} 0 & \text{if } i \notin S_{i+1}^{l+1}, \\ 1 & \text{if } i \in S_{i+1}^{l+1} \end{cases}$, and so $y \in T^{l+1}$. If $y = l_{A,B}(z)$, the number of 1’s in $y$ and $z$ is identical because 0 is mapped to 00 and 1 is mapped to 01. We can see that $y_{2i} = 1$ if and only if $x_i = 1$. Hence,

$$y_i = \begin{cases} 0 & \text{if } i \notin S_{i+1}^{l+1}, \\ 1 & \text{if } i \in S_{i+1}^{l+1} \end{cases}.$$

So $R^{l+1}(A, B) \subseteq T^{l+1}$. Conversely, let $v = v_1v_2 \cdots v_{2^{l+1}}$, where

$$v_i = \begin{cases} 0 & \text{if } i \notin S_i^{l+1}, \\ 1 & \text{if } i \in S_i^{l+1} \end{cases}$$

for some $t \in \{0, 1, \ldots, l + 1\}$. If $t = 0$, note that $v = 000 \cdots 0 = l_{A,B}(000 \cdots 0)$ and $000 \cdots 0 \in R^l(A, B)$, and so $v \in R^{l+1}(A, B)$. If $t = 1$, we have $v = 1010 \cdots 10 = u_{A,B}(\bar{v})$ for any $\bar{v} \in R^l(A, B)$. That is $v \in R^{l+1}(A, B)$. Suppose $t \in \{2, 3, \ldots, l + 1\}$ and $w = w_1w_2 \cdots w_{2^t}$ be an element of $T^t$ such that

$$w_i = \begin{cases} 0 & \text{if } i \notin S_i^{l-1}, \\ 1 & \text{if } i \in S_i^{l-1} \end{cases}.$$
By induction hypothesis, $w \in R^i(A, B)$. Note that $v = l_{A, B}(w)$ and then $v \in R^{i+1}(A, B)$. So we have $T^{i+1} \subseteq R^{i+1}(A, B)$. Hence we have already proved the claim.

Note that \{1, 2, \ldots, 2^k - 1\} is a disjoint union of $S_1, S_2, \ldots, S_k$. If $S_j^k = \{i_{j,1}, i_{j,2}, \ldots, i_{j,|S_j^k|}\}$ for $i_{j,1} < i_{j,2} < \cdots < i_{j,|S_j^k|}$, then we can let \( \rho \) be the permutation

\[
i_{1,1} i_{1,2} \cdots i_1,|S_j^k| i_{2,1} i_{2,2} \cdots i_2,|S_j^k| \cdots i_{k,1} i_{k,2} \cdots i_{k,|S_j^k|} 2^k.
\]

Since $R^i(A, B) = T^k$, then all 1’s in each row pattern in $R^k(A, B)$ is in columns $i_{j,1}, i_{j,2}, \ldots, i_{j,|S_j^k|}$ for some $1 < j < k$. Therefore we can see that every row of the matrix obtained by reordering columns of $M^k(A, B)$ according to the order given by \( \rho \) is of the form $0^a1^b0^c$ for some non-negative integers $a, b$ and $c$. By Theorem 14, $G^k(A, B)$ is word-representable for ant $k \geq 1$.

\[\Box\]

**Remark 45.** Proposition 44 gives IWR($A, B$) = \( \infty \) for $A$ and $B$ in Case 9 in Tables 3. By Theorem 25, column and row permutations of $A$ and $B$ give the same IWR. Consequently, we also have IWR($A, B$) = \( \infty \) for $A$ and $B$ in Cases 26 and 41.

**Proposition 46.** Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $G^k(A, B)$ is word-representable for any $k \geq 0$.

**Proof.** The case of $k = 0$ is trivial. For any binary string $x$, we define $\bar{x}$ to be the binary string obtained by changing digits of $x$ from 0 to 1 and from 1 to 0. We claim that, for $k \geq 1$, $R^k(A, B) = \{x, \bar{x}\}$ for some $x \in B^k$ and prove it by induction on $k$. As $R^1(A, B) = \{10, 01\}$ and $01 = \bar{10}$, the case of $k = 1$ is done. Suppose that $k > 1$ is a positive integer such that $R^k(A, B) = \{x, \bar{x}\}$ for some $x \in B^k$.

Then $R^{k+1}(A, B) = \{u_{A, B}(x), l_{A, B}(x), u_{A, B}(\bar{x}), l_{A, B}(\bar{x})\}$. Note that $u_{A, B}(0) = l_{A, B}(1)$ and $u_{A, B}(1) = l_{A, B}(0)$. Then $u_{A, B}(x) = l_{A, B}(\bar{x})$ and $l_{A, B}(x) = u_{A, B}(\bar{x})$. So we have $R^{k+1}(A, B) = \{u_{A, B}(x), l_{A, B}(x)\}$. Since $u_{A, B}(0) = \bar{l}_{A, B}(0)$ and $u_{A, B}(1) = \bar{l}_{A, B}(1)$, we have $u_{A, B}(x) = l_{A, B}(x)$. That is, $R^k(A, B) = \{u_{A, B}(x), u_{A, B}(\bar{x})\}$. Hence we have proved the claim by induction.

So, for any $k \geq 1$, $R^k(A, B) = \{x, \bar{x}\}$ for some binary string $x$. Suppose that there are $s$ 1’s in $x$. Let $\rho$ be a permutation such that reordering $[x]$ according to the order giving by $\rho$ make all 1’s in $x$ be together in the first $s$ columns. Then reordering $[\bar{x}]$ according to the order giving by $\rho$ makes all 1’s in $x$ be together in the last $2^k - s$ columns. Hence each row of the matrix obtained by reordering $M^k(A, B)$ according to the order giving by $\rho$ is $1^s0^{2^k-s} \text{ or } 0^s1^{2^k-s}$. By Theorem 14, $G^k(A, B)$ is word-representable for any $k \geq 0$. \[\square\]
4.2 The case when $A$ is the all-one matrix

There are only 16 cases when $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. If $B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, we have $M^0(A, B) = [0]$ and $M^k(A, B)$ is an all-one matrix for any $k \geq 1$ which is word-representable by Theorem 14. The rest of this paper, except Theorem 47, deals with the cases when $B$ has at least one 0.

**Theorem 47.** Let $A = \begin{bmatrix} 1 \end{bmatrix}_{m \times n}$ and $B$ be an $m \times n$ binary matrix. Then, $M^k(A, B)$ is a submatrix of $M^{k+2}(A, B)$, and $G^k(A, B)$ is an induced subgraph of $G^{k+2}(A, B)$, for any $k \geq 0$.

**Proof.** Since the case of $B$ being an all-one matrix is trivial, we assume that $B$ is not an all-one matrix. Let $B = \{b_{ij}\}_{m \times n}$ and $b_{rs} = 0$ for some $1 \leq r \leq m$ and $1 \leq s \leq n$. We will prove by induction on $k$ that $M^k(A, B)$ is contained in $M^{k+2}(A, B)$ as a submatrix by rows $(r-1)m^k + 1, (r-1)m^k + 2, \ldots, rm^k$ and columns $(s-1)n^k + 1, (s-1)n^k + 2, \ldots, sn^k$ for any $k \geq 0$.

Note that $M^0(A, B) = [0]$, $M^1(A, B) = [1]_{m \times n}$ and

$$M^2(A, B) = \begin{bmatrix} B & B & \cdots & B \\ B & B & \cdots & B \\ \vdots & \vdots & \ddots & \vdots \\ B & B & \cdots & B \end{bmatrix}$$

Let $M^2(A, B) = [m_{ij}]$, then $m_{rs} = 0$. So $M^0(A, B)$ is a submatrix of $M^2(A, B)$ by row $(r-1)m^0 + 1$ and column $(s-1)n^0 + 1$.

Let $l \geq 1$ be an integer such that $M^l(A, B)$ is a submatrix of $M^{l+2}(A, B)$ by rows $(r-1)m^k + 1, (r-1)m^k + 2, \ldots, rm^k$ and columns $(s-1)n^k + 1, (s-1)n^k + 2, \ldots, sn^k$. For the next iteration of morphism applied to $M^{l+2}(A, B)$, it is easy to see that $M^l(A, B)$ in the rows $(r-1)m^k + 1, (r-1)m^k + 2, \ldots, rm^k$ and the columns $(s-1)n^k + 1, (s-1)n^k + 2, \ldots, sn^k$ $M^{l+2}(A, B)$ is mapped to $M^{l+1}(A, B)$ in the rows $(r-1)m^{k+1} + 1, (r-1)m^{k+1} + 2, \ldots, rm^{k+1}$ and columns $(s-1)n^{k+1} + 1, (s-1)n^{k+1} + 2, \ldots, sn^{k+1}$ of $M^{l+3}(A, B)$. Hence, $M^k(A, B)$ is a submatrix of $M^{k+2}(A, B)$ for any $k \geq 0$. Consequently, $G^k(A, B)$ is an induced subgraph of $G^{k+2}(A, B)$ for any $k \geq 0$.

We know from Theorem 17 that $G^0(A, B) \preceq G^2(A, B) \preceq G^4(A, B) \preceq \cdots$ and $G^1(A, B) \preceq G^3(A, B) \preceq G^5(A, B) \preceq \cdots$ where $G \preceq H$ means $G$ is an induced subgraph of $H$. So we are interested in investigating the smallest integer $l$ such that $G^l(A, B)$ is non-word-representable in both cases of $l$ being even and $l$ being odd. The cases 114, 119, 120, 123, 124 and 129 in Table 2 are given by using Proposition 24. Theorem 12 can be applied to Cases 125, 126, 127 and 128, and the index of word-representability in these cases is infinity. The following propositions discuss the remaining cases.
Proposition 48. For \( A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \) and \( B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \), \( \text{IWR}(A, B) = 3 \).

Moreover, \( G^k(A, B) \) is not word-representable for \( k \geq 3 \).

Proof. Note that \( M^2(A, B) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \) and
\[
M^3(A, B) = \begin{bmatrix}
1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\
1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\
1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\
1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\
\end{bmatrix}.
\]

Reordering columns of \( M^2(A, B) \) in order given by 1324 yields the matrix
\[
\begin{bmatrix}
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 \\
\end{bmatrix}.
\]
By Theorem 14, \( G^2(A, B) \) is word-representable.

Let \( M^* \) be a matrix obtained from reordering columns of \( M^3(A, B) \) and every rows of \( M^* \) are of the form \( 0^*1^a0^* \) or \( 0^*1^a0^* \). Since every rows of \( M^3(A, B) \) has 2 zero entries, there is a row of \( M^* \) of the form \( 1^{a}0^{216-a} \) where \( 1 \leq a \leq 5 \). Note that there exist a row of \( M^* \) having 1’s in the columns \( a \) and \( a + 2 + 1 \). So \( M^* \) cannot satisfy the second condition of Theorem 15. Hence \( G^3(A, B) \) is non-word-representable, and so \( \text{IWR}(A, B) = 3 \).

Now we consider the word-representability of \( G^4(A, B) \). We have
\[
R^4(A, B) = \{1011101010111010, 0111010101111010, 1110101011101101, 1101010111010111, 1010111101011101, 0101110101111011, 1010111010111110, 0101110101111011 \}.
\]
In order to apply Theorem 15, we assume the existence of an order \( \rho = \rho_1 \rho_2 \cdots \rho_{16} \) with the following properties:

- from \( 1011101010111010 \in R^4(A, B) \), columns 2, 6, 8, 10, 14 and 16 must be cyclically consecutive;
- from \( 1101010111010111 \in R^4(A, B) \), columns 4, 6, 8, 12, 14 and 16 must be cyclically consecutive;
- from \( 1010111101011101 \in R^4(A, B) \), columns 2, 4, 6, 10, 12 and 14 must be cyclically consecutive.

28
It follows from the first and the second bullet points that 6, 8, 14, and 16 must be consecutive and then, w.l.o.g., 2 and 10 is next to the left of them and then 4 and 12 is next to the right them. That means that 2 and 4 cannot be cyclically consecutive, which contradicts the second bullet point. So there is no such \( \rho \) and \( G^4(A, B) \) is non-word-representable. As \( G^2(A, B) \) is non-word-representable, and the class of word-representable graphs is hereditary, by Theorem \( 47 \) \( G^{2k+1}(A, B) \) is non-word-representable for any \( k \geq 1 \). Similarly, \( G^{2k}(A, B) \) is non-word-representable for any \( k \geq 2 \) because \( G^4(A, B) \) is non-word-representable. Therefore, \( G^k(A, B) \) is not word-representable for \( k \geq 3 \).

Remark 49. Proposition 48 gives Case 121 in Tables 4 and we obtain IWR\((A, B) = 3\) in Case 122 by a column and row permutation of \( A \) and \( B \).

Proposition 50. For \( A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \) and \( B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \), IWR\((A, B) = 5\). Moreover, \( G^k(A, B) \) is not word-representable for \( k \geq 5 \).

Proof. It is easy to see that \( G^0(A, B), G^1(A, B) \) and \( G^2(A, B) \) are word-representable. Since reordering columns of \( M^3(A, B) \) in order given by the permutation 26153748 yields a matrix satisfying the condition in Theorem 14 \( G^3(A, B) \) is word-representable. Further, reordering columns of \( M^4(A, B) \) in order given by the permutation

\[
2(10)4(12)3(11)195(13)7(15)6(14)(16)
\]

also yields a matrix satisfying the condition in Theorem 14 so \( G^4(A, B) \) is word-representable.

Next, we consider \( G^5(A, B) \). We use \( u \) and \( l \) instead of \( u_{A,B} \) and \( l_{A,B} \), respectively. In order to apply Theorem 15 we assume the existence of an order \( \rho = \rho_1\rho_2\cdots\rho_{32} \) proving word-representability of \( G^5(A, B) \).

- \( l(l(l(u(0)))) = 101110111111110111101101111011 \in R^5(A, B) \)
  so columns 2, 6, 10, 14, 18, 22, 26 and 30 must be cyclically consecutive in \( \rho \).
- \( l(u(l(l(0)))) = 1010101011111111111011011111111 \in R^5(A, B) \)
  so columns 2, 4, 6, 8, 18, 20, 22 and 24 must be cyclically consecutive in \( \rho \).
- \( l(u(l(l(0)))) = 1111011011111111111101011111111 \in R^5(A, B) \)
  so columns 6, 8, 22 and 24 must be cyclically consecutive in \( \rho \).
- \( u(l(l(l(u(0)))) = 11110000111111111111000111111111 \in R^5(A, B) \)
  so columns 5, 6, 7, 8, 21, 22, 23 and 24 must be cyclically consecutive in \( \rho \).

29
It follows from the first and the second bullet points that 2, 6, 18, and 22 must be consecutive and then, w.l.o.g., 10, 14, 26 and 30 are next to the left of these numbers and 4, 8, 20 and 24 are next to the right them. Hence $\rho$ contains

$$\{10, 14, 26, 30\}, \{2, 6, 18, 22\}, \{4, 8, 20, 24\}$$

where numbers in $\{\}$ are consecutive in $\rho$ but are in some unknown to us order. From the third bullet point, $\rho$ contains

$$\{10, 14, 26, 30\}, \{2, 18\}, \{6, 22\}, \{8, 24\}, \{4, 20\}.$$

We obtain a contradiction with the fourth bullet point. So there is no such $\rho$ and $G^5(A, B)$ is non-word-representable.

Next, we consider word-representability of $G^6(A, B)$. In order to apply Theorem [15] we assume that there exists a permutation $\tau = \tau_1 \tau_2 \cdots \tau_{64}$ proving word-representability of $G^6(A, B)$.

- $l(u(u(u(u(0))))) \in R^6(A, B)$ is $1010101010101010101010101010101010$. So, all odd columns must be cyclically consecutive in $\tau$.

- $u(u(u(u(l(l(0)))))) \in R^6(A, B)$ is $11111111111111110000000000000000000000000000000$. So, columns in $\{1, 2, \ldots, 16, 33, 34, \ldots, 48\}$ must be cyclically consecutive in $\tau$.

- $u(l(l(u(u(u(0))))) \in R^6(A, B)$ is $0011000000110000001100000011000000110000001100000011000000$. So, columns in $\{3, 4, 11, 12, 19, 20, 27, 28, 35, 36, 43, 44, 51, 52, 59, 60\}$ must be cyclically consecutive in $\tau$.

Since all odd columns must be cyclically consecutive in $\tau$, there is a unique $s \in \{1, 2, \ldots, 64\}$ such that $\tau_s$ is odd and $\tau_{s+1}$ is even, where for $s = 64$, $s + 1 := 1$. From the second bullet point, we have

$$\{\tau_s, \tau_{s-1}, \ldots, \tau_{s-15}\} = \{1, 3, 5, \ldots, 15, 33, 35, 37, \ldots, 47\}$$

and

$$\{\tau_{s+1}, \tau_{s+2}, \ldots, \tau_{s+16}\} = \{2, 4, 6, \ldots, 16, 34, 36, 38, \ldots, 48\}$$

where for $i \geq 0$, $\tau_{-i} := \tau_{64-i}$ and for $s + i \geq 65$, $s + i := s + i - 64$. On the other hand, from the third bullet point, we have

$$\{\tau_s, \tau_{s-1}, \ldots, \tau_{s-7}\} = \{3, 11, 19, 27, 35, 43, 51, 59\}$$

and

$$\{\tau_{s+1}, \tau_{s+2}, \ldots, \tau_{s+8}\} = \{4, 12, 20, 28, 36, 44, 52, 60\}$$

where the indices less than 1 and larger than 64 are treated as above.

We obtain a contradiction because $19 \not\in \{\tau_s, \tau_{s-1}, \ldots, \tau_{s-15}\}$ but $19 \in \{\tau_s, \tau_{s-1}, \ldots, \tau_{s-7}\}$. Hence, there is no such $\tau$ and $G^6(A, B)$ is non-word-representable. Since the graphs $G^5(A, B)$ and $G^6(A, B)$ are not
word-representable and the class of word-representable graphs is hereditary, by Theorem \[47\] \(G^k(A, B)\) is non-word-representable for any \(k \geq 5\).

**Remark 51.** Proposition \[50\] is Case 117 in Table \[1\]. By Theorem \[25\], column and row permutations of \(A\) and \(B\) give the same IWR. Consequently, we also have IWR\((A, B) = 5\) for \(A\) and \(B\) in Cases 115, 116 and 117.

### 5 Concluding remarks

This paper is a major contribution to the study of word-representability of split graphs. Two key achievements in the paper are as follows:

- Theorem \[15\] can be used to study word-representability of split graphs via adjacency matrices, which is a novel approach.
- Necessary conditions for word-representability of split graphs obtained by iteration of morphism can be checked in polynomial time. Indeed, for a given such graph defined by matrices \(A\) and \(B\), we can go through all of \(2 \times 2\) submatrices in \(A\), and the respective submatrices in \(B\), and then use our classification results in Tables \[1\] and \[2\] to detect non-word-representability.

As for open directions of research, it would be useful to provide a classification, similar to that we provided for \(2 \times 2\) matrices in Tables \[1\] and \[2\] for larger matrices, for example, for \(2 \times 3\) matrices, or \(3 \times 3\) matrices, etc. This would enlarge our knowledge of word-representable split graphs obtained by iteration of morphisms, and hopefully, will eventually lead to a complete classification of such graphs.

For another research question, we noted in Tables \[1\] and \[2\] that if \(G^t(A, B)\) is word-representable, then \(IWR(A, B) = \infty\). In other words, the largest finite IWR in the case of \(2 \times 2\) matrices is 4. Is there a reason for that? Does there exist a positive integer \(t\) (a constant, or a function of \(n\) and \(m\)) making the following statement true “If \(A, B\) are \(n \times m\) matrices and \(G^t(A, B)\) is word-representable, then \(IWR(A, B) = \infty\)”?

Finally, recall that if the leftmost bottom entry of \(A\) is 1 then \(\lim_{k \to \infty} M^k(A, B)\) may not be well-defined as the sequence of graphs \(G^k(A, B)\), for \(k \geq 0\), may not be a chain of induced subgraphs. In all such cases, for \(2 \times 2\) matrices, we still have that non-word-representability of \(G^k(A, B)\) implies non-word-representability of \(G^{k+1}(A, B)\). Is it always the case for \(m \times n\) matrices \(A\) and \(B\)? If not, then how do we characterize the situations when it is the case?
References

[1] H. Z.Q. Chen, S. Kitaev, A. Saito. Representing split graphs by word, *Discussiones Mathematicae Graph Theory*, to appear 2020.

[2] G.-S. Cheon, M. Kim, J. Kim, S. Kitaev. Word-representable of Toeplitz graphs. *Discr. Appl. Math.* **270** (2019) 96–105.

[3] S. Foldes, P. L. Hammer. Split graphs. In *Proceedings of the Eighth Southeastern Conference on Combinatorics, Graph Theory and Computing* (Louisiana State Univ., Baton Rouge, 1977), pages 311–315. Congressus Numerantium, No. XIX.

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

[5] S. Kitaev. A Comprehensive Introduction to the Theory of Word-Representable Graphs, In: *Developments in Language Theory: 21st International Conference, DLT, Liege, Aug 7–11, 2017. Lecture Notes in Comp. Sci.* **10396** (2017) 36–67.

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

[7] S. Kitaev, Y. Long, J. Ma, H. Wu. Word-representability of split graphs, [arXiv:1709.09725](https://arxiv.org/abs/1709.09725).

[8] M. Lothaire. *Combinatorics on words*, Encyclopedia of Mathematics and its Applications, 17, Addison-Wesley Publishing Co., Reading, Mass, 1983.