



From Words to Graphs, and Back

Vadim Lozin^(✉)

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
V.Lozin@warwick.ac.uk

Abstract. In 1918, Heinz Prüfer discovered a fascinating relationship between labelled trees with n vertices and words of length $n - 2$ over the alphabet $\{1, 2, \dots, n\}$. Since the discovery of the Prüfer code for trees, the interplay between words and graphs has repeatedly been explored and exploited in both directions. In the present paper, we review some of the many results in this area and discuss a number of open problems related to this topic.

1 Introduction

In the beginning was the word. Graphs have appeared much later. Since then, these two notions frequently interact and cooperate. Graphs help reveal structure in words and (not necessarily formal) languages (see e.g. [34]). Words are used to represent graphs, which became an important issue with the advent of the computer era. We discuss this issue in Sect. 3 of the paper. Section 2 is devoted to the interplay between words and graphs around the notion of well-quasi-ordering. In the rest of the present section, we introduce basic terminology and notation used in the paper.

A binary relation is a *quasi-order* (also known as *pre-order*) if it is reflexive and transitive. A set of pairwise comparable elements in a quasi-ordered set is called a *chain* and a set of pairwise incomparable elements is called an *antichain*. A quasi-ordered set is *well-quasi-ordered* if it contains neither infinite strictly decreasing chains, nor infinite antichains.

Given a finite set B (an alphabet), we denote by B^* the set of all words over B . For a word $\alpha \in B^*$, $|\alpha|$ stands for the length of α and α_j for the j -th letter of α . A *factor* of a word α is a contiguous subword of α . The factor containment relation is a quasi-order, but not a well-quasi-order, since it contains infinite antichains, for instance, $\{101, 1001, 10001, \dots\}$. A language is *factorial* if it is closed under taking factors. It is well-known (and not difficult to see) that a factorial language L can be uniquely characterized by a set of minimal forbidden words, also known as the *antidictionary* of L , i.e. the set of minimal (with respect to the factor containment relation) words that do not belong to L .

All graphs in this paper are undirected, without loops and multiple edges. A graph G is an *induced subgraph* of H if G can be obtained from H by vertex deletions. The induced subgraph relation is a quasi-order, but not a well-quasi-order, since the cycles C_k , $k \geq 3$, constitute an infinite antichain. A class X

of graphs, also known as a graph property, is *hereditary* if it is closed under taking induced subgraphs. Clearly, if X is hereditary, then it can be uniquely characterized by a set M of minimal forbidden induced subgraphs, in which case we say that graphs in X are M -free. The *speed* of X is the number of n -vertex labelled graphs in X , studied as a function of n .

2 Words, Graphs and Well-quasi-ordering

Well-quasi-ordering (WQO) is a highly desirable property and frequently discovered concept in mathematics and theoretical computer science [19,26]. A simple but powerful tool for proving well-quasi-orderability is the celebrated Higman's Lemma, which can be stated as follows. Let M be a set with a quasi-order \leq . We extend \leq from M to M^* as follows: $a_1 \dots a_m \leq b_1 \dots b_n$ if and only if there is an order-preserving injection $f : \{a_1, \dots, a_m\} \rightarrow \{b_1, \dots, b_n\}$ with $a_i \leq f(a_i)$ for each $i = 1, \dots, m$. Higman's Lemma states the following.

Lemma 1. ([21]) *If (M, \leq) is a WQO, then (M^*, \leq) is a WQO.*

Kruskal [25] extended this result to the set of finite trees partially ordered under homeomorphic embedding. In other words, Kruskal's tree theorem restricted to paths becomes Higman's lemma. Later, Robertson and Seymour [31] generalized Kruskal's tree theorem to the set of all graphs partially ordered under the minor relation. However, the induced subgraph relation is not a well-quasi-order. Other examples of important relations that are not well-quasi-orders are pattern containment relation on permutations [35], embeddability relation on tournaments [16], minor ordering of matroids [22], factor containment relation on words [27]. On the other hand, each of these relations may become a well-quasi-order under some additional restrictions. Below we present some examples and discuss a number of open problems related to this topic.

2.1 An Introductory Example

A word can be interpreted as a graph in various ways. Consider, for instance, a binary word $\alpha = \alpha_1 \dots \alpha_n$, and let us associate with this word a graph G_α with vertices v_1, \dots, v_{n+1} such that for each $i = 2, \dots, n+1$, vertex v_i is adjacent to the vertices v_1, \dots, v_{i-1} if $\alpha_i = 1$ and v_i is not adjacent to the vertices v_1, \dots, v_{i-1} if $\alpha_i = 0$. In other words, G_α can be constructed from a single vertex recursively applying one of the following two operations: adding a dominating vertex (i.e. a vertex adjacent to every other vertex in the graph) or adding an isolated vertex. The graphs that can be constructed by means of these two operations are known in the literature as *threshold graphs*.

Obviously, not every graph is threshold. For instance, none of the following graphs is threshold, since none of them contains a dominating or isolated vertex: the path on 4 vertices P_4 , the cycle on 4 vertices C_4 and the complement of C_4 , denoted \overline{C}_4 . Moreover, no graph containing P_4, C_4 or \overline{C}_4 as an induced subgraph is threshold, i.e. threshold graphs are $(P_4, C_4, \overline{C}_4)$ -free. The inverse inclusion also

is true, which leads to the following conclusion: a graph is threshold if and only if it is $(P_4, C_4, \overline{C}_4)$ -free.

Let us note that the original definition of threshold graphs differs from both characterizations presented in the two previous paragraphs. The notion of threshold graphs was introduced in [17] and was inspired by the notion of *threshold Boolean function*. Since its original introduction, the notion of threshold graphs gave rise to a vast literature on the topic, including the book [28].

The relationship between binary words and threshold graphs described earlier is a bijection and it provides an easy way for counting *unlabelled* threshold graphs (observe that counting unlabelled graphs is generally a more difficult task than counting labelled graphs). This relationship also shows, with the help of Higman's Lemma, that the class of threshold graphs is well-quasi-ordered under the induced subgraph relation. The same conclusion can be derived from two other seemingly unrelated results, which we discuss in the next section.

2.2 Geometric Grid Classes of Permutations and Letter Graphs

Any collection of n points on the plane, with no two on a common vertical or horizontal line, uniquely defines a permutation π of n elements. This can be done, for instance, by labeling the points from 1 to n from bottom to top and then recording the labels reading left to right (see Fig. 1 for an illustration). By deleting any point, we obtain a permutation π' of $n - 1$ elements, in which case we say that π contains π' as a *pattern*.

The notion of a pattern defines a partial order on the set of permutations known as the *pattern containment* relation. A *pattern class of permutations*, or simply a *permutation class*, is any set of permutations which is downward closed under the pattern containment relation.

The pattern containment relation is not a well-quasi-order, since it contains infinite antichains [35]. However, under certain restrictions, this relation may become a well-quasi-order. To give an example, let us introduce the notion of monotone grid classes of permutations.

Let M be an $s \times t$ matrix with entries in $\{0, \pm 1\}$. An M -gridding of a permutation π represented by a collection of points on the plane is a partition of the plane into $s \times t$ cells by means of vertical and horizontal lines so that the cell in column i and row j of the partition is empty if $M(i, j) = 0$, contains the elements of π in an increasing order if $M(i, j) = +1$, and contains the elements of π in a decreasing order if $M(i, j) = -1$. Figure 1 represents two M -griddings of the permutation 351624 with $M = \begin{pmatrix} +1 & -1 \\ 0 & +1 \end{pmatrix}$ and $M = \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix}$. The grid class of M consists of all permutations which admit an M -gridding and is known as a *monotone grid class*.

The restriction to monotone grid classes is a strong restriction, but it is not strong enough to guarantee well-quasi-ordering. To describe more restrictions, we define the *cell graph* of M as follows: the vertices of this graph are the non-zero entries of M , in which two vertices are adjacent if and only if the corresponding

entries share a row or a column and there are no non-zero entries between them in this row or column.

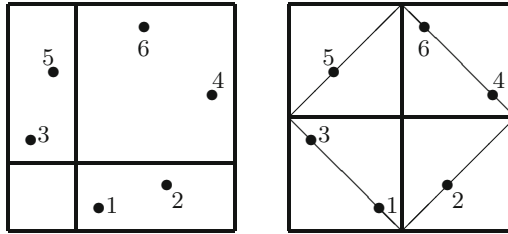


Fig. 1. Two griddings of the permutation 351624.

Theorem 1. ([38]) *For a $0/\pm 1$ matrix M , the grid class of M is well-quasi-ordered under the pattern containment relation if and only if the cell graph of M is a forest, i.e. a graph without cycles.*

Cycles in cell graphs can give rise to infinite antichains of permutations. However, if we require the elements of each cell in a monotone grid class to belong to a diagonal of the respective cell (see the gridding on the left of Fig. 1), then infinite antichains “magically” disappear and the class becomes well-quasi-ordered. This is known as *geometric gridding*, a notion introduced in [2]. The authors of [2] characterized geometric grid classes of permutations in various ways, of which we quote the following two results.

Theorem 2. ([2]) *Every geometrically griddable class of permutations is well-quasi-ordered and is in bijection with a regular language.*

Now we move from words to graphs and define the notion of letter graphs introduced in [29]. Let Σ be a finite alphabet and let $\mathcal{P} \subseteq \Sigma^2$ be a set of ordered pairs of symbols from Σ , known as a decoder. With each word $w = w_1w_2 \cdots w_n$ over Σ we associate a graph $G(\mathcal{P}, w)$, called the *letter graph* of w , by defining the vertex set of this graph to be $\{1, 2, \dots, n\}$ with i being adjacent to $j > i$ if and only if the ordered pair (w_i, w_j) belongs to \mathcal{P} .

It is not difficult to see that every graph G is a letter graph in a sufficiently large alphabet with an appropriate decoder \mathcal{P} . The minimum ℓ such that G is a letter graph in an alphabet of ℓ letters is the *lettericity* of G . A graph is a k -letter graph if its lettericity is at most k .

With the help of Higman’s Lemma it is not difficult to conclude that for each fixed value of k , the set of all k -letter graphs is well-quasi-ordered by the induced subgraph relation, which was formally proved in [29].

The notion of letter graphs was introduced 11 years earlier than the notion of geometric grid classes of permutations, and nothing in the definitions of these

two notions suggests any connection between them. However, there is intriguing relationship between these notions revealed recently in [3]. To describe this relationship, we define the permutation graph G_π of a permutation π on the set $\{1, 2, \dots, n\}$ to be the graph with vertex set $\{1, 2, \dots, n\}$ in which two vertices i and j are adjacent if and only if $(i - j)(\pi(i) - \pi(j)) < 0$.

Theorem 3. ([3]) *Let X be a class of permutations and \mathcal{G}_X the corresponding class of permutation graphs. If X is a geometrically griddable class, then \mathcal{G}_X is a class of k -letter graphs for a finite value of k .*

This theorem suggests the idea that geometrically griddable classes of permutations and letter graphs are two languages describing the same concept in the universe of permutations and permutation graphs, respectively. However, the inverse of Theorem 3 remains an open problem, which we state below as a conjecture.

Conjecture 1. Let X be a class of permutations and \mathcal{G}_X the corresponding class of permutation graphs. If \mathcal{G}_X is a class of k -letter graphs for a finite value of k , then X is a geometrically griddable class.

To support this conjecture, we return to the notion of threshold graphs introduced in Sect. 2.1 and observe that every threshold graph is a 2-letter graph. Indeed, consider the alphabet $\Sigma = \{0, 1\}$ and the decoder $\mathcal{P} = \{(1, 1), (1, 0)\}$, and let $w = w_1 w_2 \cdots w_n$ be any binary word. In the graph $G(\mathcal{P}, w)$, if $w_i = 1$, then i is adjacent to every vertex $j > i$ (since both pairs $(1, 1)$ and $(1, 0)$ belong to \mathcal{P}), and if $w_i = 0$ then i is not adjacent to any vertex $j > i$ (since neither $(0, 0)$ nor $(0, 1)$ belong to \mathcal{P}). Therefore, $G(\mathcal{P}, w)$ is a threshold graph. Notice that this representation is similar but not identical to the correspondence between graphs and words described in Sect. 2.1.

On the other hand, the *geometric* grid class of $M = \begin{pmatrix} -1 & +1 \\ +1 & -1 \end{pmatrix}$, also known as the X-class, consists of permutations that avoid the following four permutations as patterns: 2143, 3412, 2413, 3142 (see e.g. [18]). The permutation graphs of the first two of these permutations are, respectively, \overline{C}_4 and C_4 , while the last two permutations both represent a P_4 . Since a graph is threshold if and only if it is $(P_4, C_4, \overline{C}_4)$ -free, we conclude that the permutation graphs corresponding to the X-class are precisely the threshold graphs.

2.3 Deciding WQO

Deciding whether a permutation class is well-quasi-ordered is a difficult question. Decidability of this question for classes defined by finitely many forbidden permutations was stated as an open problem in [15]. Similar questions have been studied for the induced subgraph relation on graphs [24], the embeddability relation on tournaments [16], the minor ordering of matroids [22]. However, the decidability of this problem has been shown only for one or two forbidden elements (graphs, permutations, tournaments, matroids).

A breakthrough result in this area was recently obtained in [9], where decidability was proved for factorial languages. The solution is based on the analysis of the structure of an automaton describing the input language. The authors of [9] also discuss an alternative approach, which suggests a possible way to approach the same problem for graphs and permutations. This approach is based on the notion of a *periodic* infinite antichain. Speaking informally, an infinite antichain of words is called periodic of period p if each element of this set becomes a factor of some infinite periodic word of period p after dropping some prefix and suffix. For instance, the set $\{101, 1001, 10001, \dots\}$ is a periodic infinite antichain of period 1. The following theorem was proved in [9].

Theorem 4. ([9]) *Let $D = \{\alpha^1, \alpha^2, \dots, \alpha^k\}$ be a finite set of pairwise incompatible words and X be the factorial language with the antidictionary D . Then X is well-quasi-ordered by the factor containment relation if and only if it contains no periodic infinite antichains of period at most $|\alpha^1| + |\alpha^2| + \dots + |\alpha^k| + 1$.*

To apply the idea of periodic infinite antichains to graphs, we modify the notion of letter graphs by distinguishing between consecutive and nonconsecutive vertices corresponding to a word $w = w_1 w_2 \dots w_n$. For nonconsecutive vertices $i < j$ the definition remains the same: i and j are adjacent if and only if $(w_i, w_j) \in \mathcal{P}$. For consecutive vertices, we change the definition to the opposite: i and $i + 1$ are adjacent if and only if $(w_i, w_{i+1}) \notin \mathcal{P}$. Let us denote the graph obtained in this way from a word w by $G^*(\mathcal{P}, w)$. For instance, if a is a letter of Σ and $(a, a) \notin \mathcal{P}$, then the word $aaaaa$ defines a path on 5 vertices. With some restrictions, the induced subgraph relation on graphs defined in this way corresponds to the factor containment relation on words, i.e. $G^*(\mathcal{P}, w)$ is an induced subgraph of $G^*(\mathcal{P}, w')$ if and only if w is a factor of w' .

The graph $G^*(\mathcal{P}, w)$ constructed from a periodic word w is called a periodic graph. The period of w is called the period of $G^*(\mathcal{P}, w)$. To construct periodic antichains, we break the periodicity on both ends of the graph (word) by inserting an appropriate prefix and suffix. The following conjecture was proposed in [9] and was inspired by Theorem 4.

Conjecture 2. There is a function $f : \mathbb{N} \rightarrow \mathbb{N}$ such that the class X of graphs defined by a finite collection F of forbidden induced subgraphs is well-quasi-ordered by the induced subgraph relation if and only if X contains no periodic infinite antichains of period at most $f(t(F))$, where $t(F)$ stands for the total number of vertices of graphs in F .

To support this conjecture, let us mention the following decidability problem, which was recently solved in [8]: given a finite collection F of graphs, decide whether the speed of the class of F -free graphs is above or below the Bell number. The solution is based on a characterization of minimal classes with speeds above the Bell number by means of almost periodic words.

Definition 1. *A word w is almost periodic if for any factor f of w there is a constant k_f such that any factor of w of size at least k_f contains f as a factor.*

A jump to the Bell number for hereditary graph properties was identified in [12]. This paper distinguishes classes of graphs of two types: the classes where a certain graph parameter, called in [8] the *distinguishing number*, is finite and the classes where this parameter is infinite. For the case where the distinguishing number is infinite, the paper [12] provides a complete description of minimal classes above the Bell number, of which there are precisely 13. In the case where this parameter is finite, the family of minimal classes is infinite and all of them have been characterized in [8] via the notion of almost periodic words as follows.

Let A be a finite alphabet and \mathcal{P} a symmetric decoder, i.e. a decoder containing with each pair (a_i, a_j) the pair (a_j, a_i) . Also, let w be a word over A and $G^*(\mathcal{P}, w)$ the letter graph of w distinguishing between consecutive and non-consecutive letters, as defined earlier. Finally, let $X^*(\mathcal{P}, w)$ be the class of graphs containing all induced subgraphs of $G^*(\mathcal{P}, w)$.

Theorem 5. ([8]) *Let X be a hereditary class of graphs with a finite distinguishing number. Then X is a minimal class of speed above the Bell number if and only if there exists an infinite almost periodic word w over a finite alphabet and a symmetric decoder \mathcal{P} such that $X = X^*(\mathcal{P}, w)$.*

In [8], it was shown that for hereditary classes defined by a *finite* collection F of minimal forbidden induced subgraphs, the word “almost” can be omitted from this theorem. Moreover, the period of w in this case is bounded by a function of $t(F)$, where $t(F)$, as before, is the total number of vertices of graphs in F . This leads to a procedure deciding the Bell number for hereditary graph properties, as was shown in [8].

Interestingly, the same procedure decides well-quasi-ordering by induced subgraphs for classes with a finite distinguishing number, as was recently shown in [10]. In other words, this result verifies Conjecture 2 for classes with a finite distinguishing number.

We conclude this section by observing that Theorem 5 brings us back from graphs to words, and it also makes a bridge to the next section, where the speed of a hereditary graph property is an important issue.

3 Representing Graphs by Words

Representing graphs by words in a finite alphabet, or graph coding, is important in computer science for representing graphs in computer memory [20, 23, 37]. Without loss of generality we will assume that our alphabet is binary.

For a class X of graphs, we denote by X_n the set of graphs in X with the vertex set $\{1, 2, \dots, n\}$. *Coding* of graphs in the class X is a family of bijective mappings $\Phi = \{\phi_n : n = 1, 2, 3, \dots\}$, where $\phi_n : X_n \rightarrow \{0, 1\}^*$. A coding Φ is called *asymptotically optimal* if¹

$$\lim_{n \rightarrow \infty} \frac{\max_{G \in X_n} |\phi_n(G)|}{\log |X_n|} = 1.$$

¹ All logarithms are of base 2.

Every labelled graph G with n vertices can be represented by a binary word of length $\binom{n}{2}$, one bit per each pair of vertices, with 1 standing for an edge and 0 for a non-edge. Such a word can be obtained by reading the elements of the adjacency matrix above the main diagonal. The word obtained by reading these elements row by row, starting with the first row, is called the *canonical code* of G and is denoted $\phi_n^c(G)$.

If no a priori information about the graph is available, then $\binom{n}{2}$ is the minimum number of bits needed to represent the graph. However, if we know that our graph possesses some special properties, then this knowledge may lead to a shorter representation. For instance,

- if we know that our graph is bipartite, then we do not need to describe the adjacency of vertices that belong to the same part in its bipartition. Therefore, we need at most $n^2/4$ bits to describe the graph, the worst case being a bipartite graph with $n/2$ vertices in each of its parts.
- if we know that our graph is not an arbitrary bipartite graph but *chordal* bipartite, then we can further shorten the code and describe any graph in this class with at most $O(n \log^2 n)$ bits [36].
- a further restriction to trees (a proper subclass of chordal bipartite graphs) enables us to further shorten the code to $(n-2) \log n$ bits, which is the length of binary representation of the Prüfer code for trees [30].

How much can the canonical representation be shortened for graphs with a property X ? For hereditary properties this question can be answered through the notion of entropy.

3.1 Entropy of Hereditary Properties

In order to represent graphs in a class X , we need at least $|X_n|$ different binary words. Therefore, in the worst case the length of a binary code of an n -vertex graph in X cannot be shorter than $\log |X_n|$. Thus, the ratio $\log |X_n| / \binom{n}{2}$ can be viewed as the coefficient of compressibility for representing n -vertex graphs in X . Its limit value, for $n \rightarrow \infty$, was called by Alekseev in [4] the *entropy* of X . Moreover, in the same paper Alekseev showed that for every hereditary property X the entropy necessarily exists and in [5] he proved that its value takes the following form:

$$\lim_{n \rightarrow \infty} \frac{\log |X_n|}{\binom{n}{2}} = 1 - \frac{1}{k(X)}, \quad (1)$$

where $k(X)$ is a natural number, called the *index* of X . To define this notion let us denote by $\mathcal{E}_{i,j}$ the class of graphs whose vertices can be partitioned into at most i independent sets and j cliques. In particular, $\mathcal{E}_{2,0}$ is the class of bipartite graphs and $\mathcal{E}_{1,1}$ is the class of split graphs. Then $k(X)$ is the largest k such that X contains $\mathcal{E}_{i,j}$ with $i + j = k$. Independently, this result was obtained by Bollobás and Thomason [13, 14] and is known nowadays as the Alekseev-Bollobás-Thomason Theorem (see e.g. [7]).

3.2 Coding of Graphs in Classes of High Speed

In [4], Alekseev proposed a universal algorithm which gives an asymptotically optimal coding for graphs in every hereditary class X of index $k > 1$, i.e. of non-zero entropy. Below we present an adapted version of this algorithm.

Let $n > 1$ be a natural number and let p be a prime number between $\lfloor n/\sqrt{\log n} + 1 \rfloor$ and $2\lfloor n/\sqrt{\log n} \rfloor$. Such a number always exists by the Bertrand-Chebyshev theorem (see e.g. [1]). Define $k = \lfloor n/p \rfloor$. Then

$$p \leq 2n/\sqrt{\log n}, \quad k \leq \sqrt{\log n}, \quad n - kp < p. \quad (2)$$

Let G be an arbitrary graph with n vertices. Denote by D_n the set of all pairs of vertices of G . We split D_n into two disjoint subsets R_1 and R_2 as follows: R_1 consists of the pairs (a, b) such that $a \leq kp$, $b \leq kp$ and $\lfloor (a-1)/p \rfloor \neq \lfloor (b-1)/p \rfloor$, and R_2 consists of all the remaining pairs. Let us denote by $\mu^{(1)}$ the binary word consisting of the elements of the canonical code corresponding to the pairs of R_2 . This word will be included in the code of G we construct.

Now let us take care of the pairs in R_1 . For all $x, y \in \{0, 1, \dots, p-1\}$, we define

$$Q_{x,y} = \{pi + 1 + \text{res}_p(xi + y) \mid i = 0, 1, \dots, k-1\},$$

where $\text{res}_p(z)$ is the remainder on dividing z by p . Let us show that every pair of R_1 appears in exactly one set $Q_{x,y}$. Indeed, if $(a, b) \in Q_{x,y}$ ($a < b$), then

$$xi_1 + y \equiv a \pmod{p}, \quad xi_2 + y \equiv b \pmod{p},$$

where $i_1 = \lfloor (a-1)/p \rfloor$, $i_2 = \lfloor (b-1)/p \rfloor$. Since $i_1 \neq i_2$ (by definition of R_1), there exists a unique solution of the following system

$$\begin{aligned} x(i_1 - i_2) &\equiv a - b \pmod{p} \\ y(i_1 - i_2) &\equiv ai_2 - bi_1 \pmod{p}. \end{aligned} \quad (3)$$

Therefore, by coding the graphs $G_{x,y}$ induced by $Q_{x,y}$ and combining their codes with the word $\mu^{(1)}$ (that describes the pairs in R_2) we obtain a complete description of G .

To describe the graphs $G_{x,y}$ induced by $Q_{x,y}$ we first relabel their vertices according to

$$z \rightarrow \lfloor (z-1)/p \rfloor + 1.$$

In this way, we obtain p^2 graphs $G'_{x,y}$, each on the vertex set $\{1, 2, \dots, k\}$. Some of these graphs may coincide. Let m ($m \leq p^2$) denote the number of pairwise different graphs in this set and H_0, H_1, \dots, H_{m-1} an (arbitrarily) ordered list of m pairwise different graphs in this set. In other words, for each graph $G'_{x,y}$ there is a unique number i such that $G'_{x,y} = H_i$. We denote the binary representation of this number i by $\omega(x, y)$ and the length of this representation by ℓ , i.e. $\ell = \lceil \log m \rceil$. Also, denote

$$\mu^{(2)} = \phi_k^c(H_0)\phi_k^c(H_1)\dots\phi_k^c(H_{m-1}),$$

$$\mu^{(3)} = \omega(0,0)\omega(0,1)\dots\omega(0,p-1)\omega(1,0)\dots\omega(p-1,p-1).$$

The word $\mu^{(2)}$ describes all graphs H_i and the word $\mu^{(3)}$ indicates for each pair x, y the interval in the word $\mu^{(2)}$ containing the information about $G'_{x,y}$. Therefore, the words $\mu^{(2)}$ and $\mu^{(3)}$ completely describe all graphs $G_{x,y}$. In order to separate the word $\mu^{(2)}\mu^{(3)}$ into $\mu^{(2)}$ and $\mu^{(3)}$, it suffices to know the number ℓ , because $|\mu^{(2)}| = \ell p^2$ and the number p is uniquely defined by n . Since $m \leq 2^{\binom{k}{2}}$, the number ℓ can be described by at most

$$\lceil \log \ell \rceil = \lceil \log \lceil \log m \rceil \rceil \leq \lceil \log \binom{k}{2} \rceil \leq \lceil \log k^2 \rceil \leq \lceil \log \log n \rceil$$

binary bits. Let $\mu^{(0)}$ be the binary representation of the number ℓ of length $\lceil \log \log n \rceil$, and let

$$\phi_n^*(G) = \mu^{(0)}\mu^{(1)}\mu^{(2)}\mu^{(3)}, \quad \Phi^* = \{\phi_n^* \mid n = 2, 3, \dots\}.$$

Theorem 6. [4] Φ^* is an asymptotically optimal coding for any hereditary class X with $k(X) > 1$.

3.3 Representing Graphs in Hereditary Classes of Low Speed

The universal algorithm presented in the previous section is, unfortunately, not optimal for classes X of index $k(X) = 1$, also known as *unitary classes*, since equation (1) does not provide the asymptotic behavior of $\log |X_n|$ in this case. This is unfortunate, because the family of unitary classes contains a variety of properties of theoretical or practical importance, such as line graphs, interval graphs, permutation graphs, threshold graphs, forests, planar graphs and, even more generally, all proper minor-closed graph classes, all classes of graphs of bounded vertex degree, of bounded tree- and clique-width, etc.

A systematic study of hereditary properties of low speed was initiated by Scheinerman and Zito in [32]. In particular, they distinguished the first four lower layers in the family of unitary classes: constant (classes X with $|X_n| = \Theta(1)$), polynomial ($|X_n| = n^{\Theta(1)}$), exponential ($|X_n| = 2^{\Theta(n)}$) and factorial ($|X_n| = n^{\Theta(n)}$). Independently, similar results have been obtained by Alekseev in [6]. Moreover, Alekseev described the set of minimal classes in all the four lower layers and the asymptotic structure of properties in the first three of them. A more detailed description of the polynomial and exponential layers was obtained by Balogh, Bollobás and Weinreich in [11]. However, the factorial layer remains largely unexplored and the asymptotic structure is known only for properties at the bottom of this layer, below the Bell numbers [11, 12]. On the other hand, the factorial properties constitute the core of the unitary family, as all the interesting classes mentioned above (and many others) are factorial.

We conclude the paper with an important conjecture, which deals with representing graphs in the factorial classes.

Definition 2. A representation of an n -vertex graph G is said to be implicit if it assigns to each vertex of G a binary code of length $O(\log n)$ so that the adjacency of two vertices is a function of their codes.

This notion was introduced in [23], where the authors identified a variety of graph classes admitting an implicit representation. Clearly, not every class admits an implicit representation, since a bound on the total length of the code implies a bound on the number of graphs admitting such a representation. More precisely, only classes containing $2^{O(n \log n)}$ graphs with n vertices can admit an implicit representation. This restriction, however, is not sufficient to represent graphs implicitly. A simple counter-example can be found in [37]. This example deals with a non-hereditary graph property, which leaves the question of implicit representation for *hereditary* classes of speed $2^{O(n \log n)}$ open. These are precisely the classes with at most factorial speed of growth. It is known that every hereditary class with a sub-factorial speed admits an implicit representation [33]. For hereditary classes with factorial speeds the question of implicit representation is generally open and is known as the implicit graph representation conjecture (see e.g. [37]).

Conjecture 3. Any hereditary class of speed $n^{\Theta(n)}$ admits an implicit representation.

References

1. Aigner, M., Ziegler, G.M.: Proofs from THE BOOK, 4th edn. Springer, Berlin (2010). <https://doi.org/10.1007/978-3-642-00856-6>
2. Albert, M.H., Atkinson, M.D., Bouvel, M., Ruskuc, N., Vatter, V.: Geometric grid classes of permutations. *Trans. Am. Math. Soc.* **365**, 5859–5881 (2013)
3. Alecu, B., Lozin, V., Zamaraev, V., de Werra, D.: Letter graphs and geometric grid classes of permutations: characterization and recognition. In: Brankovic, L., Ryan, J., Smyth, W.F. (eds.) IWOCOA 2017. LNCS, vol. 10765, pp. 195–205. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78825-8_16
4. Alekseev, V.E.: Coding of graphs and hereditary classes. *Probl. Cybernet.* **39**, 151–164 (1982). (in Russian)
5. Alekseev, V.E.: Range of values of entropy of hereditary classes of graphs. *Diskret. Mat.* **4**(2), 148–157 (1992). (in Russian; translation in *Discrete Mathematics and Applications*, 3 (1993), no. 2, 191–199)
6. Alekseev, V.E.: On lower layers of a lattice of hereditary classes of graphs. *Diskretn. Anal. Issled. Oper. Ser.* **4**(1), 3–12 (1997). (Russian)
7. Alon, N., Balogh, J., Bollobás, B., Morris, R.: The structure of almost all graphs in a hereditary property. *J. Comb. Theory B* **79**, 131–156 (2011)
8. Atminas, A., Collins, A., Foniok, J., Lozin, V.: Deciding the Bell number for hereditary graph properties. *SIAM J. Discrete Math.* **30**, 1015–1031 (2016)
9. Atminas, A., Lozin, V., Moshkov, M.: WQO is decidable for factorial languages. *Inf. Comput.* **256**, 321–333 (2017)
10. Atminas, A., Brignall, R.: Well-quasi-ordering and finite distinguishing number. <https://arxiv.org/abs/1512.05993>
11. Balogh, J., Bollobás, B., Weinreich, D.: The speed of hereditary properties of graphs. *J. Comb. Theory Ser. B* **79**(2), 131–156 (2000)
12. Balogh, J., Bollobás, B., Weinreich, D.: A jump to the Bell number for hereditary graph properties. *J. Comb. Theory B* **95**, 29–48 (2005)
13. Bollobás, B., Thomason, A.: Projections of bodies and hereditary properties of hypergraphs. *Bull. London Math. Soc.* **27**, 417–424 (1995)

14. Bollobás, B., Thomason, A.: Hereditary and monotone properties of graphs. In: Graham, R.L., Nešetřil, J. (eds.) *The Mathematics of Paul Erdős, II*. AC, vol. 14, pp. 70–78. Springer, Berlin (1997). https://doi.org/10.1007/978-3-642-60406-5_7
15. Brignall, R., Ruškuc, N., Vatter, V.: Simple permutations: decidability and unavoidable substructures. *Theor. Comput. Sci.* **391**, 150–163 (2008)
16. Cherlin, G.L., Latka, B.J.: Minimal antichains in well-founded quasi-orders with an application to tournaments. *J. Comb. Theory B* **80**, 258–276 (2000)
17. Chvátal, V., Hammer, P.L.: Aggregation of inequalities in integer programming. *Ann. Discrete Math.* **1**, 145–162 (1977)
18. Elizalde, S.: The X-class and almost-increasing permutations. *Ann. Comb.* **15**, 51–68 (2011)
19. Finkel, A., Schnoebelen, P.: Well-structured transition systems everywhere!. *Theor. Comput. Sci.* **256**, 63–92 (2001)
20. Galperin, H., Wigderson, A.: Succinct representations of graphs. *Inf. Control* **56**, 183–198 (1983)
21. Higman, G.: Ordering by divisibility of abstract algebras. *Proc. London Math. Soc.* **2**, 326–336 (1952)
22. Hine, N., Oxley, J.: When excluding one matroid prevents infinite antichains. *Adv. Appl. Math.* **45**, 74–76 (2010)
23. Kannan, S., Naor, M., Rudich, S.: Implicit representation of graphs. *SIAM J. Discrete Math.* **5**, 596–603 (1992)
24. Korpelainen, N., Lozin, V.V.: Two forbidden induced subgraphs and well-quasi-ordering. *Discrete Math.* **311**, 1813–1822 (2011)
25. Kruskal, J.B.: Well-quasi-ordering, the tree theorem, and Vazsonyi’s conjecture. *Trans. Am. Math. Soc.* **95**, 210–225 (1960)
26. Kruskal, J.B.: The theory of well-quasi-ordering: a frequently discovered concept. *J. Comb. Theory A* **13**, 297–305 (1972)
27. de Luca, A., Varricchio, S.: Well quasi-orders and regular languages. *Acta Informatica* **31**, 539–557 (1994)
28. Mahadev, N.V.R., Peled, U.N.: *Threshold graphs and related topics*. Annals of Discrete Mathematics, vol. 56. North-Holland, Amsterdam (1995)
29. Petkovšek, M.: Letter graphs and well-quasi-order by induced subgraphs. *Discrete Math.* **244**, 375–388 (2002)
30. Prüfer, H.: Neuer Beweis eines Satzes über Permutationen. *Arch. Math. Phys.* **27**, 742–744 (1918)
31. Robertson, N., Seymour, P.D.: Graph minors. XX. Wagner’s conjecture. *J. Comb. Theory Ser. B* **92**, 325–357 (2004)
32. Scheinerman, E.R., Zito, J.: On the size of hereditary classes of graphs. *J. Comb. Theory Ser. B* **61**(1), 16–39 (1994)
33. Scheinerman, E.R.: Local representations using very short labels. *Discrete Math.* **203**, 287–290 (1999)
34. Sonderegger, M.: Applications of graph theory to an English rhyming corpus. *Comput. Speech Lang.* **25**, 655–678 (2011)
35. Spielman, D.A., Bóna, M.: An infinite antichain of permutations. *Electron. J. Comb.* **7**, 2 (2000)
36. Spinrad, J.P.: Nonredundant 1’s in Γ -free matrices. *SIAM J. Discrete Math.* **8**, 251–257 (1995)
37. Spinrad, J.P.: *Efficient Graph Representations*. Fields Institute Monographs, vol. 19, xiii+342 pp. American Mathematical Society, Providence (2003)
38. Vatter, V., Waton, S.: On partial well-order for monotone grid classes of permutations. *Order* **28**, 193–199 (2011)