On Colouring $(2P_2, H)$-Free and (P_5, H)-Free Graphs*

Konrad K. Dabrowski and Daniël Paulusma

Department of Computer Science, Durham University,
Science Laboratories, South Road, Durham DH1 3LE, United Kingdom
{konrad.dabrowski,daniel.paulusma}@durham.ac.uk

Abstract. The Colouring problem asks whether the vertices of a graph can be coloured with at most k colours for a given integer k in such a way that no two adjacent vertices receive the same colour. A graph is (H_1, H_2)-free if it has no induced subgraph isomorphic to H_1 or H_2. A connected graph H_1 is almost classified if Colouring on (H_1, H_2)-free graphs is known to be polynomial-time solvable or NP-complete for all but finitely many connected graphs H_2. We show that every connected graph H_1 apart from the claw $K_{1,3}$ and the 5-vertex path P_5 is almost classified. We also prove a number of new hardness results for Colouring on $(2P_2, H)$-free graphs. This enables us to list all graphs H for which the complexity of Colouring is open on $(2P_2, H)$-free graphs and all graphs H for which the complexity of Colouring is open on (P_5, H)-free graphs. In fact we show that these two lists coincide. Moreover, we show that the complexities of Colouring for $(2P_2, H)$-free graphs and for (P_5, H)-free graphs are the same for all known cases.

1 Introduction

Graph colouring is an extensively studied concept in both Computer Science and Mathematics due to its many application areas. A *k-colouring* of a graph $G = (V, E)$ is a mapping $c : V \to \{1, \ldots, k\}$ such that $c(u) \neq c(v)$ whenever $uv \in E$. The Colouring problem that of deciding whether a given graph G has a k-colouring for a given integer k. If k is fixed, then we write k-Colouring instead. It is well known that even 3-Colouring is NP-complete [22].

Due to the computational hardness of Colouring, it is natural to restrict the input to special graph classes. A class is hereditary if it is closed under vertex deletion. Hereditary graph classes form a large collection of well-known graph classes for which the Colouring problem has been extensively studied. A classical result in the area is due to Grötschel, Lovász, and Schrijver [14], who showed that Colouring is polynomial-time solvable for perfect graphs.

Graphs with no induced subgraph isomorphic to a graph in a set $\mathcal{H}$ are said to be *$\mathcal{H}$-free*. It is readily seen that a graph class $\mathcal{G}$ is hereditary if and only if it there exists a set $\mathcal{H}$ such that every graph in $\mathcal{G}$ is $\mathcal{H}$-free. If the graphs of $\mathcal{H}$ are required to be minimal under taking induced subgraphs, then $\mathcal{H}$ is unique. For example, the set $\mathcal{H}$ of minimal forbidden induced subgraphs for the class of perfect graphs consists of all odd holes and odd antiholes [6].

Král', Kratochvíl, Tuza, and Woeginger [21] classified the complexity of Colouring for the case where $\mathcal{H}$ consists of a single graph H. They proved that Colouring on H-free graphs is polynomial-time solvable if H is an induced subgraph of P_4 or $P_1 + P_3$ and NP-complete otherwise.[1]

Král' et al. [21] also initiated a complexity study of Colouring for graph classes defined by two forbidden induced subgraphs H_1 and H_2. Such graph classes are said to be *bigenic*. For bigenic graph classes, no dichotomy is known or even conjectured, despite many results [1,2,4,5,7,8,10,15,16,18,20,21,23,26,27,28,31]. For instance, if we forbid two graphs H_1 and H_2 with $|V(H_1)| \leq 4$ and $|V(H_2)| \leq 4$, then there are three open cases left, namely when $(H_1, H_2) \in \{(K_{1,3}, 4P_1), (K_{1,3}, 2P_1 + P_2), (C_4, 4P_1)\}$ (see [23] and Fig. 1). If H_1 and H_2 are connected with $|V(H_1)| \leq 5$ and $|V(H_2)| \leq 5$, then there are four open cases left, namely when $H_1 = P_5$ and $H_2 \in \{\overline{C_3 + 2P_1}, \overline{C_3 + P_2}, \overline{P_1 + 2P_2}\}$ (see [20] and Fig. 2) and when $H_1 = K_{1,3}$ and

* This paper was supported by EPSRC (EP/K025090/1) and the Leverhulme Trust (RPG-2016-258).

[1] We refer to Section 2 for notation used throughout Section 1.

$H_2 = \overline{C_4 + P_1}$ (see [28] and Fig. 2). To give another example, Blanché et al. [1] determined the complexity of Colouring for $(H, \overline{H})$-free graphs for every graph H except when $H = P_3 + sP_1$ for $s \geq 3$ or $H = P_4 + sP_1$ for $s \geq 2$.

Fig. 1. The graphs from the three pairs $(H_1, H_2) \in \{(K_{1,3}, 4P_1), (K_{1,3}, 2P_1 + P_2), (C_4, 4P_1)\}$ of graphs on at most four vertices, for which the complexity of Colouring on (H_1, H_2)-free graphs is still open.

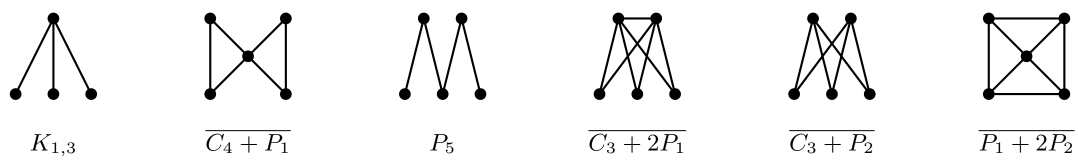


Fig. 2. The graphs from the four pairs $(H_1, H_2) \in \{(K_{1,3}, \overline{C_4 + P_1}), (P_5, \overline{C_3 + 2P_1}), (P_5, \overline{C_3 + P_2}), (P_5, \overline{P_1 + 2P_2})\}$ of connected graphs on at most five vertices, for which the complexity of Colouring on (H_1, H_2)-free graphs is still open.

The related problems Precolouring Extension and List Colouring have also been studied for bigenic graph classes. For the first problem, we are given a graph G, an integer k and a k-colouring c' defined on an induced subgraph of G. The question is whether G has a k-colouring c extending c'. For the second problem, each vertex u of the input graph G has a list $L(u)$ of colours. Here the question is whether G has a colouring c that respects L, that is, with $c(u) \in L(u)$ for all $u \in V(G)$. For the Precolouring Extension problem no classification is known and we refer to the survey [12] for an overview on what is known. In contrast to the incomplete classifications for Colouring and Precolouring Extension, Golovach and Paulusma [13] showed a dichotomy for the complexity of List Colouring on bigenic graph classes.

Our Approach. To get a handle on the computational complexity classification of Colouring for bigenic graph classes, we continue the line of research in [2,16,20,26,27,28] by considering pairs (H_1, H_2), where H_1 and H_2 are both connected. We introduce the following notion. We say that a connected graph H_1 is *almost classified* if Colouring on (H_1, H_2)-free graphs is known to be either polynomial-time solvable or NP-complete for all but finitely many connected graphs H_2. This leads to the following research question:

Which connected graphs are almost classified?

Our Results. In Section 3 we show, by combining known results from the literature, that every connected graph H_1 apart from the claw $K_{1,3}$ and the 5-vertex path P_5 is almost classified. In fact we show that the number of pairs (H_1, H_2) of connected graphs for which the complexity of Colouring is unknown is finite if neither H_1 nor H_2 is isomorphic to $K_{1,3}$ or P_5. In Section 4 we prove a number of new hardness results for Colouring restricted to $(2P_2, H_2)$-free graphs (which form a subclass of (P_5, H_2)-free graphs). We do the latter by adapting the NP-hardness construction from [11] for List Colouring restricted to complete bipartite graphs. In Section 5, we first summarize our knowledge on the complexity of Colouring restricted to $(2P_2, H)$-free graphs and (P_5, H)-free graphs. Afterwards, we list all graphs H for which the complexity of

COLOURING on $(2P_2, H)$-free graphs is still open, and all graphs H for which the complexity of COLOURING on (P_5, H)-free graphs is still open. As it turns out, these two lists coincide. Moreover, the complexities of COLOURING for $(2P_2, H)$-free graphs and for (P_5, H)-free graphs turn out to be the same for all cases that are known.

2 Preliminaries

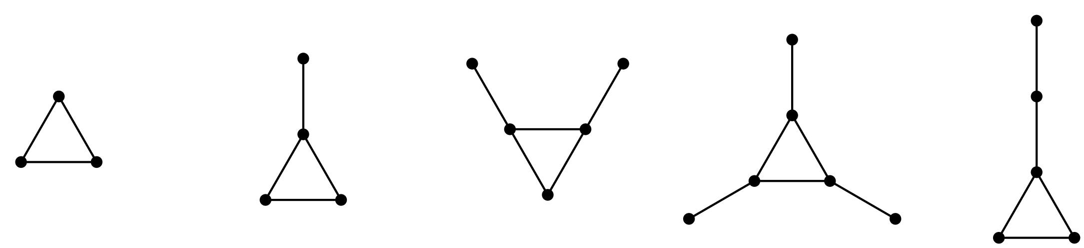
We consider only finite, undirected graphs without multiple edges or self-loops. Let $G = (V, E)$ be a graph. The *complement* $\overline{G}$ of G is the graph with vertex set $V(G)$ and an edge between two distinct vertices if and only if these two vertices are not adjacent in G. For a subset $S \subseteq V$, we let $G[S]$ denote the subgraph of G *induced* by S, which has vertex set S and edge set $\{uv \mid u, v \in S, uv \in E\}$.

Let $\{H_1, \ldots, H_p\}$ be a set of graphs. A graph G is $(H_1, \ldots, H_p)$*-free* if G has no induced subgraph isomorphic to a graph in $\{H_1, \ldots, H_p\}$. If $p = 1$, we may write H_1-free instead of (H_1)-free. The *disjoint union* $G + H$ of two vertex-disjoint graphs G and H is the graph $(V(G) \cup V(H), E(G) \cup E(H))$. The disjoint union of r copies of a graph G is denoted by rG. A *linear forest* is the disjoint union of one or more paths.

The graphs C_r, K_r and P_r denote the cycle, complete graph and path on r vertices, respectively. The graph K_3 is also known as the *triangle*. The graph $K_{r,s}$ denotes the complete bipartite graph with partition classes of size r and s, respectively. The graph $K_{1,3}$ is also called the *claw*.

The graph $S_{h,i,j}$, for $1 \leq h \leq i \leq j$, denotes the *subdivided claw*, that is, the tree that has only one vertex x of degree 3 and exactly three leaves, which are at distance h, i and j from x, respectively. Observe that $S_{1,1,1} = K_{1,3}$. The graph $S_{1,1,2}$ is also known as the *fork* or the *chair*.

The graph $T_{h,i,j}$ with $0 \leq h \leq i \leq j$ denotes the graph with vertices $a_0, \ldots, a_h$, $b_0, \ldots, b_i$ and $c_0, \ldots, c_j$ and edges a_0b_0, b_0c_0, c_0a_0, a_pa_{p+1} for $p \in \{0, \ldots, h-1\}$, b_pb_{p+1} for $p \in \{0, \ldots, i-1\}$ and c_pc_{p+1} for $p \in \{0, \ldots, j-1\}$. Note that $T_{0,0,0} = C_3$. The graph $T_{0,0,1} = \overline{P_1 + P_3}$ is known as the *paw*, the graph $T_{0,1,1}$ as the *bull*, the graph $T_{1,1,1}$ as the *net*, and the graph $T_{0,0,2}$ is known as the *hammer*; see also Fig. 3. Also note that $T_{h,i,j}$ is the line graph of $S_{h+1,i+1,j+1}$.



$T_{0,0,0} = C_3$ $T_{0,0,1} =$ paw $T_{0,1,1} =$ bull $T_{1,1,1} =$ net $T_{0,0,2} =$ hammer

Fig. 3. Examples of $T_{h,i,j}$ graphs.

Let $\mathcal{T}$ be the class of graphs for which every component is isomorphic to a graph $T_{h,i,j}$ for some $1 \leq h \leq i \leq j$ or a path P_r for some $r \geq 1$. The following result, which is due to Schindl and which we use in Section 5, shows that the $T_{h,i,j}$ graphs play an important role for our study.

Theorem 1 ([31]). *For $p \geq 1$, let $H_1, \ldots, H_p$ be graphs whose complement is not in $\mathcal{T}$. Then* COLOURING *is* NP*-complete for $(H_1, \ldots, H_p)$-free graphs.*

3 Almost Classified Graphs

In this section we prove the following result, from which it immediately follows that every connected graph apart from $K_{1,3}$ and P_5 is almost classified. In Section 5 we discuss why $K_{1,3}$ and P_5 are not almost classified.

Theorem 2. *There are only finitely many pairs (H_1, H_2) of connected graphs with $\{H_1, H_2\} \cap \{K_{1,3}, P_5\} = \emptyset$, such that the complexity of* Colouring *on (H_1, H_2)-free graphs is unknown.*

Proof. We first make a useful observation. Let H be a tree that is not isomorphic to $K_{1,3}$ or P_5 and that is not an induced subgraph of P_4. If H contains a vertex of degree at least 4 then it contains an induced $K_{1,4}$. If H has maximum degree 3, then since H is connected and not isomorphic to $K_{1,3}$, it must contain an induced $S_{1,1,2}$. If H has maximum degree at most 2, then it is a path, and since it is not isomorphic to P_5 and not an induced subgraph of P_4, it follows that H must be a path on at least six vertices. We conclude that if H is a tree that is not isomorphic to $K_{1,3}$ or P_5 and that is not an induced subgraph of P_4, then H contains $K_{1,4}$ or $S_{1,1,2}$ as an induced subgraph or H is a path on at least six vertices.

Now let (H_1, H_2) be a pair of connected graphs with $\{H_1, H_2\} \cap \{K_{1,3}, P_5\} = \emptyset$. If H_1 or H_2 is an induced subgraph of P_4, then Colouring is polynomial-time solvable for (H_1, H_2)-free graphs, as Colouring is polynomial-time solvable for P_4-free graphs (see, for example, [21]). Hence we may assume that this is not the case. If H_1 and H_2 both contain at least one cycle [9] or both contain an induced $K_{1,3}$ [17], then even 3-Colouring is NP-complete. Hence we may also assume that at least one of H_1, H_2 is a tree and that at least one of H_1, H_2 is a $K_{1,3}$-free graph. This leads, without loss of generality, to the following two cases.

Case 1. H_1 is a tree and $K_{1,3}$-free.
Then H_1 is a path. First suppose that H_1 has at least 22 vertices. It is known that 4-Colouring is NP-complete for (P_{22}, C_3)-free graphs [19] and that Colouring is NP-complete for (P_9, C_4)-free graphs [10] and for $(2P_2, C_r)$-free graphs for all $r \geq 5$ [21]. Hence we may assume that H_2 is a tree. By the observation at the start of the proof, this implies that H_2 contains an induced $K_{1,4}$, $S_{1,1,2}$ or P_6. Therefore H contains an induced $4P_1$ or $2P_1 + P_2$. Since H_1 is a path on at least 22 vertices, H_1 contains an induced $4P_1$. As Colouring is NP-complete for $(4P_1, 2P_1 + P_2)$-free graphs [21], Colouring is NP-complete for (H_1, H_2)-free graphs.

Now suppose that H_1 has at most 21 vertices. By the observation at the start of the proof, H_1 is a path on at least six vertices. It is known that 5-Colouring is NP-complete for P_6-free graphs [18]. As K_6 is not 5-colourable, this means that 5-Colouring is NP-complete for (P_6, K_6)-free graphs, as observed in [12]. Therefore we may assume that H_2 is K_6-free. Recall that Colouring is NP-complete for $(2P_1 + P_2, 4P_1)$-free graphs [21], which are contained in the class of $(P_6, 4P_1)$-free graphs. Therefore we may assume that H_2 is $4P_1$-free. Since H_2 is $(K_6, 4P_1)$-free, Ramsey's Theorem [29] implies that $|V(H_2)|$ is bounded by a constant. We conclude that both H_1 and H_2 have size bounded by a constant.

Case 2. H_1 is a tree and not $K_{1,3}$-free, and H_2 is $K_{1,3}$-free and not a tree.
Then H_1 contains a vertex of degree at least 3 and H_2 contains an induced cycle C_r for some $r \geq 3$. It is known that 3-Colouring is NP-complete for $(K_{1,5}, C_3)$-free graphs [25] and for $(K_{1,3}, C_r)$-free graphs whenever $r \geq 4$ [21]. We may therefore assume that H_1 is a tree of maximum degree at most 4 and that H_2 contains at least one induced C_3 but no induced cycles on more than three vertices. Recall that 4-Colouring is NP-complete for (P_{22}, C_3)-free graphs [19]. Hence we may assume that H_1 is a P_{22}-free tree. As H_1 has maximum degree at most 4, we find that H_1 has a bounded number of vertices.

By assumption, H_1 contains a vertex of degree at least 3. As Colouring is NP-complete for $(K_{1,3}, K_4)$-free graphs [21], we may assume that H_2 is K_4-free. By the observation at the start of the proof, H_1 must contain an induced $K_{1,4}$ or $S_{1,1,2}$. Recall that Colouring is NP-complete for the class of $(2P_1 + P_2, 4P_1)$-free graphs [21], which is contained in the class of $(K_{1,4}, S_{1,1,2}, 4P_1)$-free graphs. Hence we may assume that H_2 is $4P_1$-free. Since H_2 is $(K_4, 4P_1)$-free, Ramsey's Theorem [29] implies that $|V(H_2)|$ is bounded by a constant. Again, we conclude that in this case both H_1 and H_2 have size bounded by a constant. □

Corollary 1. *Every connected graph apart from $K_{1,3}$ and P_5 is almost classified.*

4 Hardness Results

In this section we prove that Colouring restricted to $(2P_2, H)$-free graphs is NP-complete for several graphs H. To prove our results we adapt a hardness construction from Golovach and Heggernes [11] for proving that List Colouring is NP-complete for complete bipartite graphs. As observed in [13], a minor modification of this construction yields that List Colouring is NP-complete for complete split graphs, which are the graphs obtained from complete bipartite graphs by changing one of the bipartition classes into a clique.

We first describe the construction of [11], which uses a reduction from the NP-complete [30] problem Not-All-Equal 3-Satisfiability with positive literals only. To define this problem, let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of logical variables, and let $\mathcal{C} = \{C_1, C_2, \ldots, C_m\}$ be a set of 3-literal clauses over X in which all literals are positive and every literal appears at most once in each clause. The question is whether X has a truth assignment such that each clause in $\mathcal{C}$ contains at least one true literal and at least one false literal. If so, we say that such a truth assignment is *satisfying*.

Let $(X, \mathcal{C})$ be an instance of Not-All-Equal 3-Satisfiability with positive literals only. We construct an instance (G_1, L) of List Colouring as follows. For each x_i we introduce a vertex, which we also denote by x_i and which we say is of *x-type*. We define $L(x_1) = \{1, 2\}$, $L(x_2) = \{3, 4\}, \ldots, L(x_n) = \{2n-1, 2n\}$. In this way, each x_i has one odd colour and one even colour in its list, and all lists $L(x_i)$ are pairwise disjoint. For each C_j we introduce two vertices, which we denote by C_j and C'_j and which we say are of *C-type*. If $C_j = \{x_g, x_h, x_i\}$ with $L(x_g) = \{a, a+1\}$, $L(x_h) = \{b, b+1\}$ and $L(x_i) = \{c, c+1\}$, then we set $L(C_j) = \{a, b, c\}$ and $L(C'_j) = \{a+1, b+1, c+1\}$. Hence each C_j has only odd colours in its list and each C'_j has only even colours in its list. To obtain the graph G_1 we add an edge between every vertex of x-type and every vertex of C-type. Note that G_1 is a complete bipartite graph with bipartition classes $\{x_1, \ldots, x_n\}$ and $\{C_1, \ldots, C_m\} \cup \{C'_1, \ldots, C'_m\}$.

We also construct an instance (G_2, L) where G_2 is obtained from G_1 by adding edges between every pair of vertices of x-type. Note that G_2 is a complete split graph.

The following lemma is straightforward. We refer to [11] for a proof for the case involving G_1. The case involving G_2 follows from this proof and the fact that the lists $L(x_i)$ are pairwise disjoint, as observed in [13].

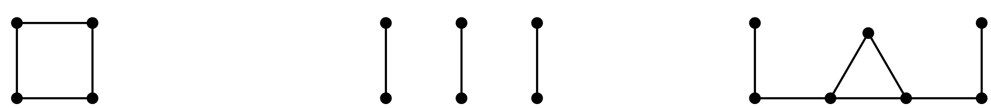
Lemma 1 ([11]). *$(\mathcal{C}, X)$ has a satisfying truth assignment if and only if G_1 has a colouring that respects L if and only if G_2 has a colouring that respects L.*

We now extend G_1 and G_2 into graphs G'_1 and G'_2, respectively, by adding a clique K consisting of $2n$ new vertices $k_1, \ldots, k_{2n}$ and by adding an edge between a vertex k_ℓ and a vertex u of the original graph if and only if $\ell \notin L(u)$. We say that the vertices $k_1, \ldots, k_{2n}$ are of *k-type*.

Lemma 2. *$(\mathcal{C}, X)$ has a satisfying truth assignment if and only if G'_1 has a $2n$-colouring if and only if G'_2 has a $2n$-colouring.*

Proof. Let $i \in \{1, 2\}$. By Lemma 1, we only need to show that G_i has a colouring that respects L if and only if G'_i has a $2n$-colouring. First suppose that G_i has a colouring c that respects L. We extend c to a colouring c' of G'_i by setting $c'(k_\ell) = \ell$ for $\ell \in \{1, \ldots, 2n\}$. Now suppose that G'_i has a $2n$-colouring c'. As the k-type vertices form a clique, we may assume without loss of generality that $c'(k_\ell) = \ell$ for $\ell \in \{1, \ldots, 2n\}$. Hence the restriction of c' to G_i yields a colouring c that respects L. □

In the next two lemmas we show forbidden induced subgraphs in G'_1 and G'_2, respectively. The complements of these forbidden graphs are shown in Figs. 4 and 5, respectively.



$\overline{2P_2} = C_4$ $3P_2$ $T_{0,2,2}$

Fig. 4. Graphs that are not induced subgraphs of the complement of G'_1.

Lemma 3. *The graph G'_1 is $(2P_2, \overline{3P_2}, \overline{T_{0,2,2}})$-free.*

Proof. We will prove that $\overline{G'_1}$ is $(C_4, 3P_2, T_{0,2,2})$-free. Observe that in $\overline{G'_1}$, the set of x-type vertices is a clique, the set of C-type vertices is a clique and the set of k-type vertices is an independent set. Furthermore, in $\overline{G'_1}$, no x-type vertex is adjacent to a C-type vertex.

$\mathbf{C_4}$-freeness. For contradiction, suppose that $\overline{G'_1}$ contains an induced subgraph H isomorphic to C_4; say the vertices of H are u_1, u_2, u_3, u_4 in that order. As the union of the set of x-type and C-type vertices induces a P_3-free graph in $\overline{G'_1}$, there must be at least two vertices of the C_4 that are neither x-type nor C-type. Since the k-type vertices form an independent set, we may assume without loss of generality that u_1 and u_3 are of k-type. It follows that u_2 and u_4 cannot be of k-type. As the set of x-type vertices and the set of C-type vertices each from a clique in $\overline{G'_1}$, but u_2 is non-adjacent to u_4, we may assume without loss of generality that u_2 is of x-type and u_4 is of C-type. Then u_4 is adjacent to the two k-type neighbours of an x-type vertex, which correspond to an even and odd colour. This is not possible as u_4, being a C-type vertex, is adjacent in $\overline{G'_1}$ to (exactly three) k-type vertices, which correspond either to even colours only or to odd colours only. We conclude that $\overline{G'_1}$ is C_4-free.

$\mathbf{3P_2}$-freeness. For contradiction, suppose that $\overline{G'_1}$ contains an induced subgraph H isomorphic to $3P_2$. As the C-type vertices and x-type vertices each form a clique in $\overline{G'_1}$, one edge of H must consist of two k-type vertices. This is not possible, as k-type vertices form an independent set in $\overline{G'_1}$. We conclude that $\overline{G'_1}$ is $3P_2$-free.

$\mathbf{T_{0,2,2}}$-freeness. For contradiction, suppose that $\overline{G'_1}$ contains an induced subgraph H isomorphic to $T_{0,2,2}$ with vertices $a_0, a_1, a_2, b_0, b_1, b_2, c_0$ and edges a_0b_0, b_0c_0, c_0a_0, a_0a_1, a_1a_2, b_0b_1, b_1b_2. As the k-type vertices form an independent set in $\overline{G'_1}$, at least one of a_1, a_2 and at least one of b_1, b_2 is of x-type or C-type. As the x-type vertices and the C-type vertices form cliques in $\overline{G'_1}$, we may assume without loss of generality that at least one of a_1, a_2 is of C-type and at least one of b_1, b_2 is of x-type. As the C-type vertices and the x-type vertices each form a clique in $\overline{G'_1}$, this means that c_0 must be of k-type, a_0 cannot be of x-type and b_0 cannot be of C-type. As k-type vertices form an independent set in $\overline{G'_1}$, a_0 and b_0 cannot be of k-type. Therefore a_0 is of C-type and b_0 is of x-type. This is a contradiction, as C-type vertices are non-adjacent to x-type vertices. We conclude that $\overline{G'_1}$ is $T_{0,2,2}$-free. □

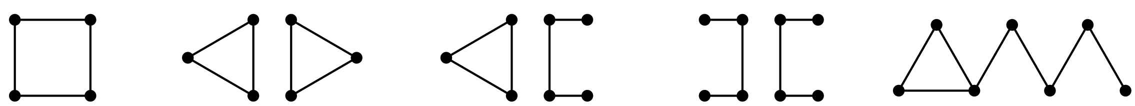


Fig. 5. Graphs that are not induced subgraphs of the complement of G'_2.

Lemma 4. *The graph G'_2 is $(2P_2, \overline{2C_3}, \overline{C_3 + P_4}, \overline{2P_4}, \overline{T_{0,0,4}})$-free.*

Proof. We will prove that $\overline{G'_2}$ is $(C_4, 2C_3, C_3 + P_4, 2P_4, T_{0,0,4})$-free. Observe that in $\overline{G'_2}$, the set of C-type vertices is a clique, the set of x-type vertices is an independent set and the set of k-type vertices is an independent set. Furthermore, in $\overline{G'_2}$, no x-type vertex is adjacent to a C-type vertex and every x-type vertex has degree 2. In fact, the union of the set of x-type vertices and the k-type vertices induces a disjoint union of P_3s in $\overline{G'_2}$.

$\mathbf{C_4}$-freeness. For contradiction, suppose that $\overline{G'_2}$ contains an induced subgraph H isomorphic to C_4, say the vertices of H are u_1, u_2, u_3, u_4 in that order. As the union of the set of x-type and C-type vertices induces a P_3-free graph in $\overline{G'_2}$, there must be at least two vertices of the C_4 that are neither x-type nor C-type. Since the k-type vertices form an independent set, we may assume without loss of generality that u_1 and u_3 are of k-type. It follows that u_2 and u_4 cannot be of k-type. As the set of vertices of C-type form a clique in $\overline{G_2'}$, at least one of u_2, u_4, say u_2, is of x-type. If u_4 is also of x-type, then u_2 and u_4 are x-type vertices with the same two colours in their list, namely those corresponding to u_1 and u_3. This is not possible. Thus u_4 must be of C-type. Then u_4 is adjacent to the two k-type neighbours of an x-type vertex, which correspond to an even and odd colour. This is not possible as u_4, being a C-type vertex, is adjacent in $\overline{G_2'}$ to (exactly three) k-type vertices, which correspond either to even colours only or to odd colours only. We conclude that $\overline{G_2'}$ is C_4-free.

$\mathbf{(2C_3, C_3+P_4, 2P_4)}$**-freeness.** For contradiction, suppose that $\overline{G_2'}$ contains an induced subgraph H isomorphic to $2C_3$, C_3+P_4 or $2P_4$. As the k-type and x-type vertices induce a disjoint union of P_3s in $\overline{G_2'}$, both components of H must contain a C-type vertex. This is not possible, as C-type vertices form a clique in $\overline{G_2'}$. We conclude that $\overline{G_2'}$ is $(2C_3, C_3+P_4, 2P_4)$-free.

$\mathbf{T_{0,0,4}}$**-freeness.** For contradiction, suppose that $\overline{G_2'}$ contains an induced subgraph H isomorphic to $T_{0,0,4}$ with vertices $a_0, a_1, a_2, a_3, a_4, b_0, c_0$ and edges a_0b_0, b_0c_0, c_0a_0, a_0a_1, a_1a_2, a_2a_3, a_3a_4.

First suppose, that neither b_0 nor c_0 is of C-type. Since the union of the set of x-type vertices and the k-type vertices induces a disjoint union of P_3s in $\overline{G_2'}$ it follows that a_0 is of C-type. Since b_0 and c_0 are not of C-type and no vertex of C-type has a neighbour of x-type in $\overline{G_2'}$, it follows that b_0 and c_0 must be of k-type. This is not possible, because the k-type vertices form an independent set in $\overline{G_2'}$.

Now suppose that at least one of b_0, c_0 is of C-type. Since the vertices of C-type induce a clique in $\overline{G_2'}$, it follows that no vertex in $A := \{a_1, a_2, a_3, a_4\}$ is of C-type. Since A induces a P_4 in $\overline{G_2'}$, but the union of the set of x-type vertices and the k-type vertices induces a disjoint union of P_3s in $\overline{G_2'}$, this is a contradiction. We conclude that $\overline{G_2'}$ is $T_{0,0,4}$-free. □

We are now ready to state the two main results of this section. It is readily seen that Colouring belongs to NP. Then the first theorem follows from Lemma 2 combined with Lemma 3, whereas the second one follows from Lemma 2 combined with Lemma 4. Note that $\overline{2C_3}$ is isomorphic to $K_{3,3}$.

Theorem 3. Colouring *is* NP-*complete for* $(2P_2, \overline{3P_2}, \overline{T_{0,2,2}})$-*free graphs.*

Theorem 4. Colouring *is* NP-*complete for* $(2P_2, \overline{2C_3}, \overline{C_3+P_4}, \overline{2P_4}, \overline{T_{0,0,4}})$-*free graphs.*

5 Conclusions

We showed that every connected graph is almost classified except for the claw and the P_5. Our notion of almost classified graphs originated from recent work [16,20,27,28] on Colouring for (H_1, H_2)-free graphs for connected graphs H_1 and H_2, in particular when $H_1 = P_5$. We decreased the number of open cases for the latter graph by showing new NP-hardness results for $(2P_2, H)$-free graphs. In the following theorem we summarize all known results for Colouring restricted to $(2P_2, H)$-free graphs and (P_5, H)-free graphs.

Theorem 5. *Let H be a graph on n vertices. Then the following two statements hold:*

(i) If $\overline{H}$ contains a graph in $\{C_3+P_4, 3P_2, 2P_4\}$ as an induced subgraph, or $\overline{H}$ is not an induced subgraph of $T_{1,1,3}+P_{2n-1}$, then Colouring *is* NP-*complete for $(2P_2, H)$-free graphs.*

(ii) If $\overline{H}$ is an induced subgraph of a graph in $\{2P_1+P_3, P_1+P_4, P_2+P_3, P_5, T_{0,0,1}+P_1, T_{0,1,1}, T_{0,0,2}\}$ or of sP_1+P_2 for some integer $s \geq 0$, then Colouring *is polynomial-time solvable for (P_5, H)-free graphs.*

Proof. If $\overline{H}$ contains a graph in $\{2C_3, C_3+P_4, 3P_2, 2P_4, T_{0,2,2}, T_{0,0,4}\}$ as an induced subgraph, then Colouring is NP-complete for $(2P_2, H)$-free graphs due to Theorems 3 and 4. We may therefore assume that $\overline{H}$ is $(2C_3, C_3+P_4, 3P_2, 2P_4, T_{0,2,2}, T_{0,0,4})$-free. (Note that $T_{1,1,3}+P_{2n-1}$ is $(2C_3, T_{0,2,2}, T_{0,0,4})$-free.)

Recall that $\mathcal{T}$ is the class of graphs for which every component is isomorphic to a graph $T_{h,i,j}$ for some $1 \leq h \leq i \leq j$ or a path P_r for some $r \geq 1$. Note that $\overline{2P_2} = C_4 \notin \mathcal{T}$. Therefore, if $\overline{H} \notin \mathcal{T}$, then Colouring is NP-complete for $(2P_2, H)$-free graphs by Theorem 1. We may therefore assume that $\overline{H} \in \mathcal{T}$. Since $\overline{H}$ is $2C_3$-free, $\overline{H}$ can contain at most one component that is not a path. Since $\overline{H}$ is $(T_{0,2,2}, T_{0,0,4})$-free, if $\overline{H}$ does have a component that is not a path, then this component must be an induced subgraph of $T_{1,1,3}$. The union of components of $\overline{H}$ that are isomorphic to paths form an induced subgraph of P_{2n-1}. Therefore $\overline{H}$ is an induced subgraph of $T_{1,1,3} + P_{2n-1}$.

It is known that COLOURING is polynomial-time solvable for (P_5, H)-free graphs if $\overline{H}$ is an induced subgraph of $2P_1 + P_3$ [27], $P_1 + P_4$ (this follows from the fact that $(P_5, \overline{P_1 + P_4})$-free graphs have clique-width at most 5 [3]; see also [2] for a linear-time algorithm), $P_2 + P_3$ [28], P_5 [16], $T_{0,0,1} + P_1$ [20], $T_{0,1,1}$ [20], $T_{0,0,2}$ [20] or $sP_1 + P_2$ for some integer $s \geq 0$ [28]. □

Theorem 5 leads to the following open problem, which shows how the P_5 is not almost classified. Recall that $T_{0,0,0} = C_3$.

Open Problem 1 *Determine the complexity of* COLOURING *for* $(2P_2, H)$*-free graphs and for* (P_5, H)*-free graphs if*

- $\overline{H} = sP_1 + P_t + T_{h,i,j}$ *for* $0 \leq h \leq i \leq j \leq 1$, $s \geq 0$ *and* $2 \leq t \leq 3$
- $\overline{H} = sP_1 + T_{h,i,j}$ *for* $0 \leq h \leq i \leq 1 \leq j \leq 3$ *and* $s \geq 0$ *such that* $h + i + j + s \geq 3$
- $\overline{H} = sP_1 + T_{0,0,0}$ *for* $s \geq 2$
- $\overline{H} = sP_1 + P_t$ *for* $s \geq 0$ *and* $3 \leq t \leq 7$ *such that* $s + t \geq 6$
- $\overline{H} = sP_1 + P_t + P_u$ *for* $s \geq 0$, $2 \leq t \leq 3$ *and* $3 \leq u \leq 4$ *such that* $s + t + u \geq 6$
- $\overline{H} = sP_1 + 2P_2$ *for* $s \geq 1$.

Open Problem 1 shows the following.

- The open cases for COLOURING restricted to $(2P_2, H)$-free graphs and (P_5, H)-free graphs coincide.
- The graph H in each of the open cases is connected.
- The number of minimal open cases is 10, namely when $\overline{H} \in \{C_3 + 2P_1, C_3 + P_2, P_1 + 2P_2\}$ (see also Section 1) and when $\overline{H} \in \{3P_1 + P_3, 2P_1 + P_4, 2P_3, P_6, T_{0,1,1} + P_1, T_{0,1,2}, T_{1,1,1}\}$.

As every graph H listed in Open Problem 1 appears as an induced subgraph in both the graph G'_1 and the graph G'_2 defined in Section 4, we need new arguments to solve the open cases in Problem 1.

The complexity of COLOURING for $(K_{1,3}, H)$-free graphs is less clear. As mentioned in Section 1, the cases where $H \in \{4P_1, 2P_1 + P_2, \overline{C_4 + P_1}\}$ are still open. Moreover, $K_{1,3}$ is not almost classified, as the case $H = P_t$ is open for all $t \geq 6$ (polynomial-time solvability for $t = 5$ was shown in [26]). Note that $|E(\overline{H})|$ may be arbitrarily large, while Open Problem 1 shows that $|E(\overline{H})| \leq 8$ in all open cases for the P_5. Since we have no new results for the case $H_1 = K_{1,3}$, we refer to [24] for further details or to the summary of COLOURING restricted to (H_1, H_2)-free graphs in [12].

References

1. A. Blanché, K. K. Dabrowski, M. Johnson, and D. Paulusma. Hereditary graph classes: When the complexities of Colouring and Clique Cover coincide. *CoRR*, abs/1607.06757, 2016.
2. H. L. Bodlaender, A. Brandstädt, D. Kratsch, M. Rao, and J. Spinrad. On algorithms for (P_5,gem)-free graphs. *Theoretical Computer Science*, 349(1):2–21, 2005.
3. A. Brandstädt, H.-O. Le, and R. Mosca. Chordal co-gem-free and (P_5,gem)-free graphs have bounded clique-width. *Discrete Applied Mathematics*, 145(2):232–241, 2005.
4. H. Broersma, P. A. Golovach, D. Paulusma, and J. Song. Determining the chromatic number of triangle-free $2P_3$-free graphs in polynomial time. *Theoretical Computer Science*, 423:1–10, 2012.
5. K. Cameron and C. T. Hoàng. Solving the clique cover problem on (bull,C_4)-free graphs. *CoRR*, abs/1704.00316, 2017.
6. M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas. The strong perfect graph theorem. *Annals of Mathematics*, 164(1):51–229, 2006.
7. K. K. Dabrowski, F. Dross, and D. Paulusma. Colouring diamond-free graphs. *Journal of Computer and System Sciences*, 89:410–431, 2017.

8. K. K. Dabrowski, P. A. Golovach, and D. Paulusma. Colouring of graphs with Ramsey-type forbidden subgraphs. *Theoretical Computer Science*, 522:34–43, 2014.
9. T. Emden-Weinert, S. Hougardy, and B. Kreuter. Uniquely colourable graphs and the hardness of colouring graphs of large girth. *Combinatorics, Probability and Computing*, 7(04):375–386, 1998.
10. S. Gaspers, S. Huang, and D. Paulusma. Colouring square-free graphs without long induced paths. *Proc. STACS 2018, LIPIcs*, 96:36:1–36:15, 2018.
11. P. A. Golovach and P. Heggernes. Choosability of P_5-free graphs. *Proc. MFCS 2009, LNCS*, 5734:382–391, 2009.
12. P. A. Golovach, M. Johnson, D. Paulusma, and J. Song. A survey on the computational complexity of colouring graphs with forbidden subgraphs. *Journal of Graph Theory*, 84(4):331–363, 2017.
13. P. A. Golovach and D. Paulusma. List coloring in the absence of two subgraphs. *Discrete Applied Mathematics*, 166:123–130, 2014.
14. M. Grötschel, L. Lovász, and A. Schrijver. Polynomial algorithms for perfect graphs. *Annals of Discrete Mathematics*, 21:325–356, 1984.
15. P. Hell and S. Huang. Complexity of coloring graphs without paths and cycles. *Discrete Applied Mathematics*, 216, Part 1:211–232, 2017.
16. C. T. Hoàng and D. A. Lazzarato. Polynomial-time algorithms for minimum weighted colorings of $(P_5, \overline{P_5})$-free graphs and similar graph classes. *Discrete Applied Mathematics*, 186:106–111, 2015.
17. I. Holyer. The NP-Completeness of edge-coloring. *SIAM Journal on Computing*, 10(4):718–720, 1981.
18. S. Huang. Improved complexity results on k-coloring P_t-free graphs. *European Journal of Combinatorics*, 51:336–346, 2016.
19. S. Huang, M. Johnson, and D. Paulusma. Narrowing the complexity gap for colouring (C_s, P_t)-free graphs. *The Computer Journal*, 58(11):3074–3088, 2015.
20. T. Karthick, F. Maffray, and L. Pastor. Polynomial cases for the vertex coloring problem. *CoRR*, abs/1709.07712, 2017.
21. D. Král', J. Kratochvíl, Zs. Tuza, and G. J. Woeginger. Complexity of coloring graphs without forbidden induced subgraphs. *Proc. WG 2001, LNCS*, 2204:254–262, 2001.
22. L. Lovász. Coverings and coloring of hypergraphs. *Congressus Numerantium*, VIII:3–12, 1973.
23. V. V. Lozin and D. S. Malyshev. Vertex coloring of graphs with few obstructions. *Discrete Applied Mathematics*, 216, Part 1:273–280, 2017.
24. V. V. Lozin and C. Purcell. Coloring vertices of claw-free graphs in three colors. *Journal of Combinatorial Optimization*, 28(2):462–479, 2014.
25. F. Maffray and M. Preissmann. On the NP-completeness of the k-colorability problem for triangle-free graphs. *Discrete Mathematics*, 162(1–3):313–317, 1996.
26. D. S. Malyshev. The coloring problem for classes with two small obstructions. *Optimization Letters*, 8(8):2261–2270, 2014.
27. D. S. Malyshev. Two cases of polynomial-time solvability for the coloring problem. *Journal of Combinatorial Optimization*, 31(2):833–845, 2016.
28. D. S. Malyshev and O. O. Lobanova. Two complexity results for the vertex coloring problem. *Discrete Applied Mathematics*, 219:158–166, 2017.
29. F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286, 1930.
30. T. J. Schaefer. The complexity of satisfiability problems. *Proc. STOC 1978*, pages 216–226, 1978.
31. D. Schindl. Some new hereditary classes where graph coloring remains NP-hard. *Discrete Mathematics*, 295(1–3):197–202, 2005.