Solving Graph Coloring Problems with Abstraction and Symmetry*

Michael Codish[1], Michael Frank[1], Avraham Itzhakov[1], and Alice Miller[2]

[1] Department of Computer Science, Ben-Gurion University of the Negev, Israel
[2] School of Computing Science, University of Glasgow, Scotland

Abstract. This paper introduces a general methodology, based on abstraction and symmetry, that applies to solve hard graph edge-coloring problems and demonstrates its use to provide further evidence that the Ramsey number $R(4,3,3) = 30$. The number $R(4,3,3)$ is often presented as the unknown Ramsey number with the best chances of being found "soon". Yet, its precise value has remained unknown for more than 50 years. We illustrate our approach by showing that: (1) there are precisely 78,892 $(3,3,3;13)$ Ramsey colorings; and (2) if there exists a $(4,3,3;30)$ Ramsey coloring then it is (13,8,8) regular. Specifically each node has 13 edges in the first color, 8 in the second, and 8 in the third. We conjecture that these two results will help provide a proof that no $(4,3,3;30)$ Ramsey coloring exists implying that $R(4,3,3) = 30$.

1 Introduction

This paper introduces a general methodology that applies to solve graph edge-coloring problems and demonstrates its application in the search for Ramsey numbers. These are notoriously hard graph coloring problems that involve assigning k colors to the edges of a complete graph. In particular, $R(4,3,3)$ is the smallest number n such that any coloring of the edges of the complete graph K_n in three colors will either contain a K_4 sub-graph in the first color, a K_3 sub-graph in the second color, or a K_3 sub-graph in the third color. The precise value of this number has been sought for more than 50 years. Kalbfleisch [13] proved in 1966 that $R(4,3,3) \geq 30$, Piwakowski [17] proved in 1997 that $R(4,3,3) \leq 32$, and one year later Piwakowski and Radziszowski [18] proved that $R(4,3,3) \leq 31$. We demonstrate how our methodology applies to provide further evidence that $R(4,3,3) = 30$.

Solving hard search problems on graphs, and graph coloring problems in particular, relies heavily on breaking symmetries in the search space. When searching for a graph, the names of the vertices do not matter, and restricting the search modulo graph isomorphism is highly beneficial. When searching for a graph coloring, on top of graph isomorphism, solutions are typically closed under permutations of the colors: the names of the colors do not matter and the term often used is "weak isomorphism" [18] (the equivalence relation is weaker because both node names and edge colors do not matter). When

* Supported by the Israel Science Foundation, grant 182/13. Computational resources provided by an IBM Shared University Award (Israel). The first three authors acknowledge the support of the Frankel Center for Computer Science at Ben-Gurion University.

the problem is to compute the set of all solutions modulo (weak) isomorphism the task is even more challenging. Often one first attempts to compute all the solutions of the coloring problem, and to then apply one of the available graph isomorphism tools, such as `nauty` [14] to select representatives of their equivalence classes modulo (weak) isomorphism. However, typically the number of solutions is so large that this approach is doomed to fail even though the number of equivalence classes itself is much smaller. The problem is that tools such as `nauty` apply after, and not during, search. To this end, we first observe that the technique described in [5] for graph isomorphism applies also to weak isomorphism, facilitating symmetry breaks during the search for solutions to graph coloring problems. This form of symmetry breaking is an important component in our methodology but on its own cannot provide solutions to hard graph coloring problems.

When confronted with hard computational problems, a common strategy is to consider approximations which focus on "abstract" solutions which characterize properties of the actual "concrete" solutions. To this end, given a graph coloring problem with k colors on n nodes, we introduce the notion of an $n \times k$ *degree matrix* in which each of n rows describes the degrees of a corresponding node in the k colors. In case the graph coloring problem is too hard to solve directly, we seek, possibly an over approximation of, all of the degree matrices of its solutions. This enables an independent search for solutions "per degree matrix" facilitating so called "embarrassingly parallel" search.

After laying the ground for a methodology based on symmetry breaking and abstraction we apply it to the problem of computing the Ramsey number $R(4,3,3)$ which reduces to determining if there exists a $(4,3,3)$ coloring of the complete graph K_{30}. We first characterize the degrees of the nodes in each of the three colors in any such coloring, if one exists. To this end, we show that if there is such a graph coloring then, up to swapping the colors two and three, all of its vertices have degrees in the three colors corresponding to the following triples: $(13,8,8)$, $(14,8,7)$, $(15,7,7)$, $(15,8,6)$, $(16,7,6)$, $(16,8,5)$. Then, we demonstrate that any potential $(4,3,3;30)$ coloring with a node with degrees (d_1,d_2,d_3) in the corresponding colors must have three corresponding embedded graphs G_1,G_2,G_3 which are $(3,3,3;d_1)$, $(4,2,3;d_2)$, and $(4,3,2;d_3)$ colorings. For all of the cases except when the degrees are $(13,8,8)$ these sets of colorings are known and easy to compute. Based on this, we show using a SAT solver that there can be no nodes with degrees $(14,8,7)$, $(15,7,7)$, $(15,8,6)$, $(16,7,6)$ or $(16,8,5)$ in any $(4,3,3;30)$ coloring. Thus, we prove that any such coloring would have to be $(13,8,8)$ regular, meaning that all nodes are of degree 13 in the first color and of degree 8 in the second and third color.

In order to apply the same proof technique for the case where the graph is $(13,8,8)$ regular we need to first compute the set of all $(3,3,3;13)$ colorings, modulo weak isomorphism. This set of graphs does not appear in previously published work. So, we address the problem of computing the set of all $(3,3,3;13)$ Ramsey colorings, modulo weak isomorphism. This results in a set of 78,892 graphs. The set of $(3,3,3;13)$ Ramsey colorings has recently been independently computed by at least three other researchers: Richard Kramer, Ivan Livinsky, and Stanislaw Radziszowski [20].

Finally, we describe the ongoing computational effort to prove that there is no $(13,8,8)$ regular $(4,3,3;30)$ Ramsey coloring. Using the embedding approach, and given the 78,892 (3,3,3;13) colorings there are $78{,}892 \times 3 \times 3 = 710{,}028$ instances to consider. Over the period of 3 months we have verified using a SAT solver that an equivalent of $557{,}451$ of these are not satisfiable. When this ongoing effort completes we will know if the value of $R(4, 3, 3)$ is 30 or 31.

Throughout the paper we express graph coloring problems in terms of constraints via a "mathematical language". Our implementation uses the BEE, finite-domain constraint compiler [16], which solves constraints by encoding them to CNF and applying an underlying SAT solver. The solver can be applied to find a single (first) solution to a constraint, or to find all solutions for a constraint modulo a specified set of (integer and/or Boolean) variables. We have performed all computations using CryptoMiniSAT [22] as the underlying solver. In some of the experiments we also report on results using MiniSAT [8, 9], and Glucose [2, 3]. It is straightforward to configure BEE to work with any of these. All computations were performed on a cluster with a total of 228 Intel E8400 cores clocked at 2 GHz each, able to run a total of 456 parallel threads. Each of the cores in the cluster has computational power comparable to a core on a standard desktop computer. Each SAT instance is run on a single thread.

The notion of a "degree matrix" arises in the literature with several different meanings. Degree matrices with the same meaning as we use in in this paper are considered in [4]. Gent and Smith [11], building on the work of Puget [19], study symmetries in graph coloring problems and recognize the importance of breaking symmetries during search. Meseguer and Torras [15] present a framework for exploiting symmetries to heuristically guide a depth first search, and show promising results for $(3, 3, 3; n)$ Ramsey colorings with $14 \leq n \leq 17$. Al-Jaam [1] proposes a hybrid meta-heuristic algorithm for Ramsey coloring problems, combining tabu search and simulated annealing. While all of these approaches report promising results, to the best of our knowledge, none of them have been successfully applied to solve open instances or improve the known bounds on classical Ramsey numbers. Our approach focuses on symmetries due to weak-isomorphism for graph coloring and models symmetry breaking in terms of constraints introduced as part of the problem formulation. This idea, advocated by Crawford *et al.* [7], has previously been explored in [5] (for graph isomorphism), and in [19] (for graph coloring).

Graph coloring has many applications in computer science and mathematics, such as scheduling, register allocation and synchronization, path coloring and sensor networks. Specifically, many finite domain CSP problems have a natural representation as graph coloring problems. Our main contribution is a general methodology that applies to solve graph edge coloring problems. The application to potentially compute an unknown Ramsey number is attractive, but the importance here is in that it shows the utility of the methodology.

2 Preliminaries

An $(r_1, \dots, r_k; n)$ Ramsey coloring is an assignment of one of k colors to each edge in the complete graph K_n such that it does not contain a monochromatic complete sub-graph K_{r_i} in color i for $1 \leq i \leq k$. The set of all such colorings is denoted $\mathcal{R}(r_1, \dots, r_k; n)$. The Ramsey number $R(r_1, \dots, r_k)$ is the least $n > 0$ such that no

$(r_1,\ldots,r_k;n)$ coloring exists. In the multicolor case ($k>2$), the only known value of a nontrivial Ramsey number is $R(3,3,3)=17$. The value of $R(4,3,3)$ is known to be equal either to 30 or to 31. The numbers of $(3,3,3;n)$ colorings are known for $14\le n\le 16$ but prior to this paper the number of colorings for $n=13$ was unpublished. Recently, the set of all $(3,3,3;13)$ colorings has also been computed by other researchers [20], and the number 78,892 as reported also in this paper. More information on recent results concerning Ramsey numbers can be found in the electronic dynamic survey by Radziszowski [21].

In this paper, graphs are always simple, i.e. undirected and with no self loops. Colors are associated with graph edges. The set of neighbors of a node x is denoted $N(x)$ and the set of neighbors by edges colored c, by $N_c(x)$. For a natural number n denote $[n]=\{1,2,\ldots,n\}$. A graph coloring, in k colors, is a pair (G,κ) consisting of a simple graph $G=([n],E)$ and a mapping $\kappa\colon E\to[k]$. When κ is clear from the context we refer to G as the graph coloring. The sub-graph of G induced by the color $c\in[k]$ is the graph $G^c=([n],\left\{\, e\in E \,\middle|\, \kappa(e)=c \,\right\})$. The sub-graph of G on the c colored neighbors of a node x is the projection of the labeled edges in G to $N_c(x)\times N_c(x)$ and denoted G^c_x. We typically represent G as an $n\times n$ adjacency matrix, A, defined such that

$$A_{i,j}=\begin{cases}\kappa(i,j) & \text{if } (i,j)\in E\\ 0 & \text{otherwise}\end{cases}$$

If A is the adjacency matrix representing the graph G, then we denote the Boolean adjacency matrix corresponding to G^c as $A[c]$. We denote the i^{th} row of a matrix A by A_i. The color-c degree of a node x in G is denoted $deg_{G^c}(x)$ and is equal to the degree of x in the induced sub-graph G^c. When clear from the context we write $deg_c(x)$. Let $G=([n],E)$ and π be a permutation on $[n]$. Then $\pi(G)=(V,\left\{\,(\pi(x),\pi(y)) \,\middle|\, (x,y)\in E\,\right\})$. Permutations act on adjacency matrices in the natural way: If A is the adjacency matrix of a graph G, then $\pi(A)$ is the adjacency matrix of $\pi(G)$ obtained by simultaneously permuting with π both rows and columns of A.

$$\varphi^{n,k}_{adj}(A)=\bigwedge_{1\le q<r\le n}\big(1\le A_{q,r}\le k \;\wedge\; A_{q,r}=A_{r,q} \;\wedge\; A_{q,q}=0\big) \tag{1}$$

$$\varphi^{n,c}_{K_3}(A)=\bigwedge_{1\le q<r<s\le n}\neg\Big(A_{q,r}=A_{q,s}=A_{r,s}=c\Big) \tag{2}$$

$$\varphi^{n,c}_{K_4}(A)=\bigwedge_{1\le q<r<s<t\le n}\neg\big(A_{q,r}=A_{q,s}=A_{q,t}=A_{r,s}=A_{r,t}=A_{s,t}=c\big) \tag{3}$$

$$\varphi_{(3,3,3;n)}(A)=\varphi^{n,3}_{adj}(A)\wedge\bigwedge_{1\le c\le 3}\varphi^{n,c}_{K_3}(A) \tag{4}$$

$$\varphi_{(4,3,3;n)}(A)=\varphi^{n,3}_{adj}(A)\wedge\bigwedge_{1\le c\le 2}\varphi^{n,c}_{K_3}(A)\wedge\varphi^{n,3}_{K_4}(A) \tag{5}$$

Fig. 1. Graph labeling problems: Ramsey colorings $(3,3,3;n)$ and $(4,3,3;n)$

A graph coloring problem is a formula $\varphi(A)$ where A is an $n \times n$ adjacency matrix of integer variables together with a set (conjunction) of constraints φ on these variables. A solution is an assignment of integer values to the variables in A which satisfies φ and determine both the graph edges and their colors. We often refer to a solution as an integer adjacency matrix and denote the set of solutions as $sol(\varphi(A))$. Figure 1 illustrates the two graph coloring problems we focus on in this paper: $(3,3,3;n)$ and $(4,3,3;n)$ Ramsey colorings. In Constraint (1), $\varphi^{n,k}_{adj}(A)$, states that the graph A has n vertices, is k colored, and is simple (symmetric, and with no self loops). In Constraints (2) and (3), $\varphi^{n,c}_{K_3}(A)$ and $\varphi^{n,c}_{K_4}(A)$ state that the n vertex graph A has no embedded sub-graph K_3, and respectively K_4, in color c. In Constraints (4) and (5), the formulas state that a graph A is a $(3,3,3;n)$ and respectively a $(4,3,3;n)$ Ramsey coloring.

For graph coloring problems, solutions are typically closed under permutations of vertices and of colors. Restricting the search space for a solution modulo such permutations is crucial when trying to solve hard graph coloring problems. It is standard practice to formalize this in terms of graph (coloring) isomorphism.

Definition 1 ((weak) isomorphism of graph colorings). *Let (G,κ_1) and (H,κ_2) be k-color graph colorings with $G = ([n], E_1)$ and $H = ([n], E_2)$. We say that (G,κ_1) and (H,κ_2) are weakly isomorphic, denoted $(G,\kappa_1) \approx (H,\kappa_2)$ if there exist permutations $\pi\colon [n] \to [n]$ and $\sigma\colon [k] \to [k]$ such that $(u,v) \in E_1 \iff (\pi(u),\pi(v)) \in E_2$ and $\kappa_1(u,v) = \sigma(\kappa_2(\pi(u),\pi(v)))$. When σ is the identity permutation, (i.e. $\kappa_1(u,v) = \kappa_2(\pi(u),\pi(v))$) we say that (G,κ_1) and (H,κ_2) are isomorphic. We denote such a weak isomorphism thus: $(G,\kappa_1) \approx_{\pi,\sigma} (H,\kappa_2)$.*

The following lemma emphasizes the importance of weak graph isomorphism as it relates to Ramsey numbers. Many classic coloring problems exhibit the same property.

Lemma 1 ($\mathcal{R}(r_1, r_2, \ldots, r_k; n)$ is closed under $\approx$). *Let (G,κ_1) and (H,κ_2) be graph colorings in k colors such that $(G,\kappa_1) \approx_{\pi,\sigma} (H,\kappa_2)$. Then,*

$$(G,\kappa_1) \in \mathcal{R}(r_1, r_2, \ldots, r_k; n) \iff (H,\kappa_2) \in \mathcal{R}(\sigma(r_1), \sigma(r_2), \ldots, \sigma(r_k); n).$$

Proof. Assume that $(G,\kappa_1) \in \mathcal{R}(r_1, r_2, \ldots, r_k; n)$ and in contradiction that $(H,\kappa_2) \notin \mathcal{R}(\sigma(r_1), \sigma(r_2), \ldots, \sigma(r_k); n)$. Let R denote a monochromatic clique of size r_s in H and R^{-1} the inverse of R in G. From Definition 1, $(u,v) \in R \iff (\pi^{-1}(u), \pi^{-1}(v)) \in R^{-1}$ and $\kappa_2(u,v) = \sigma^{-1}(\kappa_1(u,v))$. Consequently R^{-1} is a monochromatic clique of size r_s in (G,κ_1) in contradiction to $(G,\kappa_1) \in \mathcal{R}(r_1, r_2, \ldots, r_k; n)$.

Codish *et al.* introduce in [5] an approach to break symmetries due to graph isomorphism (without colors) during the search for a solution to general graph problems. Their approach involves adding a symmetry breaking predicate $\mathsf{sb}^*_\ell(A)$, as advocated by Crawford *et al.* [7], on the variables of the adjacency matrix, A, when solving graph problems. In [6] the authors show that the symmetry breaking approach of [5] holds also for graph coloring problems where the adjacency matrix consists of integer variables (the proofs for the integer case are similar to those for the Boolean case).

Definition 2. ***[5].*** *Let A be an $n \times n$ adjacency matrix. Then, viewing the rows of A as strings, $\mathsf{sb}^*_\ell(A) = \bigwedge \left\{ A_i \preceq_{\{i,j\}} A_j \,\middle|\, i < j \right\}$ where $s \preceq_{\{i,j\}} s'$ is the lexicographic order on strings s and s' after simultaneously omitting the elements at positions i and j.*

Table 1 illustrates the impact of the symmetry breaking technique introduced by Codish *et al.* in [6] on the search for $(3,3,3;n)$ Ramsey colorings. The column headed by "$\#\backslash_{\approx}$" specifies the known number of colorings modulo weak isomorphism [21]. The columns headed by "#vars" and "#clauses" indicate, respectively, the number of variables and clauses in the corresponding CNF encodings of the coloring problems with and without the symmetry breaking constraint. The columns headed by "time" indicate the time (in seconds, on a single thread of the cluster) to find all colorings iterating with a SAT solver. The timeout assumed here is 24 hours. The column headed by "#" specifies the number of colorings found when solving with the symmetry break. These include colorings which are weakly isomorphic, but far fewer than the hundreds of thousands generated without the symmetry break (until the timeout).

n	$\#\backslash_{\approx}$	no sym break			with sym break			
		#vars	#clauses	time	#vars	#clauses	time	#
17	0	408	2584	3042.10	4038	20734	0.15	0
16	2	360	2160	∞	3328	17000	0.14	6
15	2	315	1785	∞	2707	13745	0.37	66
14	115	273	1456	∞	2169	10936	259.56	24635
13	?	234	1170	∞	1708	8540	∞	?

Table 1. The search for $(3,3,3;n)$ Ramsey colorings with and without the symmetry break defined in [6] (time in seconds with 24 hr. timeout).

$$
\begin{bmatrix}
0&1&1&1&1&1&2&2&2&2&2&3&3&3&3&3\\
1&0&2&2&3&3&1&1&2&2&3&1&1&2&3&3\\
1&2&0&3&2&3&1&2&1&3&2&2&3&1&1&3\\
1&2&3&0&3&2&2&1&3&1&2&3&2&1&3&1\\
1&3&2&3&0&2&2&3&1&2&1&1&3&3&2&1\\
1&3&3&2&2&0&3&2&2&1&1&3&1&3&1&2\\
2&1&1&2&2&3&0&3&3&1&1&2&3&2&3&1\\
2&1&2&1&3&2&3&0&1&3&1&3&2&2&1&3\\
2&2&1&3&1&2&3&1&0&1&3&2&1&3&2&3\\
2&2&3&1&2&1&1&3&1&0&3&1&2&3&3&2\\
2&3&2&2&1&1&1&1&3&3&0&3&3&1&2&2\\
3&1&2&3&1&3&2&3&2&1&3&0&2&1&1&2\\
3&1&3&2&3&1&3&2&1&2&3&2&0&1&2&1\\
3&2&1&1&3&3&2&2&3&3&1&1&1&0&2&2\\
3&3&1&3&2&1&3&1&2&3&2&1&2&2&0&1\\
3&3&3&1&1&2&1&3&3&2&2&2&1&2&1&0
\end{bmatrix}
\quad
\begin{bmatrix}
0&1&1&1&1&1&2&2&2&2&2&3&3&3&3&3\\
1&0&2&2&3&3&1&1&2&2&3&1&1&2&3&3\\
1&2&0&3&2&3&2&3&1&1&2&1&2&3&1&3\\
1&2&3&0&3&2&1&2&1&3&2&2&3&1&3&1\\
1&3&2&3&0&2&3&2&2&1&1&3&1&3&2&1\\
1&3&3&2&2&0&2&1&3&2&1&3&3&1&1&2\\
2&1&2&1&3&2&0&3&3&1&1&2&3&2&1&3\\
2&1&3&2&2&1&3&0&1&1&3&3&2&2&3&1\\
2&2&1&1&2&3&3&1&0&3&1&2&1&3&3&2\\
2&2&1&3&1&2&1&1&3&0&3&3&2&1&2&3\\
2&3&2&2&1&1&1&3&1&3&0&1&3&3&2&2\\
3&1&1&2&3&3&2&3&2&3&1&0&2&1&2&1\\
3&1&2&3&1&3&3&2&1&2&3&2&0&1&1&2\\
3&2&3&1&3&1&2&2&3&1&3&1&1&0&2&2\\
3&3&1&3&2&1&1&3&3&2&2&2&1&2&0&1\\
3&3&3&1&1&2&3&1&2&3&2&1&2&2&1&0
\end{bmatrix}
\quad
\begin{bmatrix}
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5\\
5&5&5
\end{bmatrix}
$$

Fig. 2. Two 16×16 non-isomorphic $(3,3,3;16)$ Ramsey colorings (left and middle) with their common 16×3 degree matrix (right).

Figure 2 depicts, on the left and in the middle, the two non-isomorphic colorings $(3,3,3;16)$ represented as adjacency graphs in the form found using the approach of Codish *et al.* [6]. Note the lexicographic order on the rows in both matrices. These graphs are isomorphic to the two colorings reported in 1968 by Kalbfleish and Stanton [12] where it is also proven that there are no others (modulo weak isomorphism). The 16×3 degree matrix (right) describes the degrees of each node in each color as defined below in Definition 3. The results reported in Table 1 also illustrate that the approach of Codish *et al.* is not sufficiently powerful to compute the number of $(3,3,3;13)$ colorings. Likewise, it does not facilitate the computation of $R(4,3,3)$.

In the following we make use of the following results from [18].

Theorem 1. $30 \leq R(4,3,3) \leq 31$ *and,* $R(4,3,3) = 31$ *if and only if there exists a* $(4,3,3;30)$ *coloring* κ *of* K_{30} *such that: (1) For every vertex* v *and* $i \in \{2,3\}$*,* $5 \leq deg_i(v) \leq 8$*, and* $13 \leq deg_1(v) \leq 16$*. (2) Every edge in the third color has at* *least one endpoint v with $deg_3(v) = 13$. (3) There are at least 25 vertices v for which $deg_1(v) = 13$, $deg_2(v) = deg_3(v) = 8$.*

The following is a direct consequence of Theorem 1.

Corollary 1. *If G is a $(4, 3, 3; 30)$ coloring, and assuming without loss of generality that the degree of color two is greater equal to the degree of color three, then every vertex in G has degrees in the corresponding colors corresponding to one of the triplets $(13, 8, 8)$, $(14, 8, 7)$, $(15, 7, 7)$, $(15, 8, 6)$, $(16, 7, 6)$, $(16, 8, 5)$.*

Consider a vertex v in a $(4, 3, 3; n)$ coloring and focus on the three subgraphs induced by the neighbors of v in each of the three colors. The following states that these must be corresponding Ramsey colorings.

Corollary 2. *Let G is a $(4, 3, 3; n)$ coloring and v be any vertex with degrees (d_1, d_2, d_3) in the corresponding colors. Then, $d_1 + d_2 + d_3 = n - 1$ and G_v^1, G_v^2, and G_v^3 are respectively $(3, 3, 3; d_1)$, $(4, 2, 3; d_2)$, and $(4, 3, 2; d_3)$ colorings.*

Note that by definition a $(4, 2, 3; n)$ coloring is a $(4, 3; n)$ coloring in colors 1 and 3 and likewise a $(4, 3, 2; n)$ coloring is a $(4, 3; n)$ coloring in colors 1 and 2. For $n \in \{14, 15, 16\}$, the set of all $(3, 3, 3; n)$ colorings modulo (weak) isomorphism are known and consist respectively of 2, 2 and 15 colorings. Similarly, for $n \in \{5, 6, 7, 8\}$ the set of all $(4, 3; n)$ colorings modulo (weak) isomorphism are known and consist respectively of 9, 15, 9, and 3 Ramsey colorings.

3 Searching for Ramsey Colorings with Embeddings

In this section we apply a general approach where, when seeking a $(r_1, \dots, r_k; n)$ Ramsey coloring one selects a "preferred" vertex, call it v_1, and based on its degrees in each of the k colors, embeds k subgraphs which are corresponding smaller colorings. Using this approach, we apply Corollaries 1 and 2 to establish that a $(4, 3, 3; 30)$ coloring, if one exists, must be $(13, 8, 8)$ regular. Specifically, all vertices have 13 neighbors by way of edges in the first color and 8 each, by way of edges in the second and third colors.

Theorem 2. *Any $(4, 3, 3; 30)$ coloring, if one exists, is $(13, 8, 8)$ regular.*

Proof. By computation as described in the rest of this section.

We seek a $(4, 3, 3; 30)$ coloring of K_{30}, represented as a 30×30 adjacency matrix A. We focus on the degrees, (d_1, d_2, d_3) in each of the three colors, of the vertex v_1, corresponding to the first row in A, as prescribed by Corollary 1. For each such degree triplet, except for the case $(13, 8, 8)$, we take each of the known corresponding colorings for the subgraphs $G_{v_1}^1$, $G_{v_1}^2$, and $G_{v_1}^3$ and embed them in A. We then apply a SAT solver, to complete the remaining cells in A to satisfy Constraint (5) of Figure 1. If the SAT solver fails, then no such completion exists.

To illustrate the approach, consider the case where v_1 has degrees $(14, 8, 7)$ in the three colors. Figure 3 details one of the embeddings corresponding to this case. The first row of A specifies the colors of the edges of the 29 neighbors of v_1. The symbol

“_” indicates an integer variable that takes a value between 1 and 3. The neighbors of v_1 in color 1 form a submatrix of A embedded in rows (and columns) 2–15 of the matrix in the Figure. By Corollary 2 these are a $(3,3,3;14)$ Ramsey coloring and there are 115 possible such colorings modulo weak isomorphism. The Figure details one of them. Similarly, there are 3 possible subgraphs for the neighbors of v_1 in color 2, (the 3 $(4,2,3;8)$ colorings). In Figure 3, rows (and columns) 16–23 detail one such coloring. Finally, there are 9 possible subgraphs for the neighbors of v_1 in color 3, (the 9 $(4,3,2;7)$ colorings). In Figure 3, rows (and columns) 24–30 detail one such coloring.

$$
\begin{bmatrix}
0&1&1&1&1&1&1&1&1&1&1&1&1&1&1&2&2&2&2&2&2&2&2&3&3&3&3&3&3&3\\
1&0&1&1&1&1&2&2&2&2&3&3&3&3&3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&1&0&2&2&3&1&1&3&3&1&2&2&3&3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&1&2&0&3&2&1&2&1&3&1&1&3&2&3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&1&2&3&0&2&2&1&3&1&1&3&1&3&2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&1&3&2&2&0&1&1&2&2&2&3&3&1&1&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&2&1&1&2&1&0&3&3&1&2&3&1&2&3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&2&1&2&1&1&3&0&1&3&2&1&3&3&2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&2&3&1&3&2&3&1&0&1&3&2&3&2&1&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&2&3&3&1&2&1&3&1&0&3&3&2&1&2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&3&1&1&1&2&2&2&3&3&0&2&2&1&1&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&3&2&1&3&3&3&1&2&3&2&0&1&1&2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&3&2&3&1&3&1&3&3&2&2&1&0&2&1&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&3&3&2&3&1&2&3&2&1&1&1&2&0&2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
1&3&3&3&2&1&3&2&1&2&1&2&1&2&0&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&0&2&2&2&1&1&1&1&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&2&0&1&1&2&2&1&1&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&2&1&0&1&2&1&2&1&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&2&1&1&0&1&1&2&2&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&2&2&1&0&1&1&2&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&2&1&1&1&0&2&2&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&1&2&2&1&2&0&1&\_&\_&\_&\_&\_&\_&\_\\
2&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&1&1&2&2&2&1&0&\_&\_&\_&\_&\_&\_&\_\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&0&2&2&1&1&1&1\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&2&0&1&2&1&1&1\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&2&1&0&1&2&1&1\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&2&1&0&1&2&1\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&1&2&1&0&1&2\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&1&1&2&1&0&2\\
3&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&\_&1&1&1&1&2&2&0
\end{bmatrix}
$$

Fig. 3. One embedding in the search for a $(4,3,3;30)$ coloring when v_1 has degrees $(14,8,7)$.

To summarize, Figure 3 is a partial instantiated adjacency matrix in which the first row determines the degrees of v_1, in the three colors, and where 3 corresponding subgraphs are embedded. The uninstantiated values in the matrix must be completed to obtain a solution that satisfies Constraint (5) of Figure 1. This can be determined using a SAT solver. For the specific example in Figure 3, the CNF generated using our tool set consists of 33,959 clauses, involves 5,318 Boolean variables, and is shown to be unsatisfiable in 52 seconds of computation time. For the case where v_1 has degrees $(14,8,7)$ in the three colors this is one of $115 \times 3 \times 9 = 3105$ instances that need to be checked.

Table 2 summarizes the experiment which proves Theorem 2. For each of the possible degrees of vertex 1 in a $(4,3,3;30)$ coloring as prescribed by Corollary 1, except $(13,8,8)$, and for each possible choice of colorings for the derived subgraphs $G^1_{v_1}$, $G^2_{v_1}$, and $G^3_{v_1}$, we apply a SAT solver to show that Constraint (5) of Figure 1 cannot be satisfied. The table details for each degree triple, the number of instances, their average size (number of clauses and Boolean variables), and the average and total times to show that the constraint is not satisfiable.

v_1 degrees	# instances	# clauses (avg.)	# vars (avg.)	unsat (avg)	unsat (total)
(16,8,5)	54 (2*3*9)	32432	5279	51 sec.	0.77 hrs.
(16,7,6)	270 (2*9*15)	32460	5233	420 sec.	31.50 hrs.
(15,8,6)	90 (2*3*15)	33607	5450	93 sec.	2.32 hrs.
(15,7,7)	162 (2*9*9)	33340	5326	1554 sec.	69.94 hrs.
(14,8,7)	3105 (115*3*9)	34069	5324	294 sec.	253.40 hrs.

Table 2. Proving that any $(4, 3, 3; 30)$ Ramsey coloring is $(13, 8, 8)$ regular (summary).

To gain confidence in our implementation, we illustrate its application to find a $(4, 3, 3; 29)$ coloring which is known to exist. This experiment involves some reverse engineering. In 1966 Kalbfleisch [13] reported the existence of a circulant $(3, 4, 4; 29)$ coloring. Encoding Constraint (5) with $n = 29$, together with a constraint that states that the adjacency matrix A is circulant, results in a CNF with 146,506 clauses and 8,394 variables. Using a SAT solver, we obtain a corresponding $(4, 3, 3; 29)$ coloring in less than two seconds of computation time. The solution is $(12, 8, 8)$ regular and permuting its first row to be of the form 01111111111112222222233333333 we extract from it three corresponding subgraphs: $G^1_{v_1}$, $G^2_{v_1}$ and $G^3_{v_1}$ which are respectively $(3, 3, 3; 12)$, $(4, 2, 3; 8)$ and $(4, 3, 2; 8)$ Ramsey colorings. An embedding of these three in a 29×29 adjacency matrix is depicted as Figure 4 (the boldface elements).

$$
\begin{bmatrix}
0&1&1&1&1&1&1&1&1&1&1&1&1&2&2&2&2&2&2&2&2&3&3&3&3&3&3&3&3\\
1&0&1&3&1&3&2&2&2&2&1&3&1&2&2&1&3&1&1&1&2&3&3&1&2&3&1&1&3\\
1&1&0&1&3&3&1&2&1&3&2&1&3&1&2&3&3&1&2&2&2&1&3&1&1&3&1&2&2\\
1&3&1&0&1&2&3&2&2&3&3&2&1&1&1&3&1&1&2&3&2&1&1&3&3&2&1&2&1\\
1&1&3&1&0&1&3&1&2&1&3&3&2&3&2&2&2&1&2&3&1&2&1&3&1&1&1&2&3\\
1&3&3&2&1&0&1&3&3&2&2&1&2&1&2&1&2&1&3&3&1&2&1&2&3&1&3&1&1\\
1&2&1&3&3&1&0&1&3&1&2&2&2&1&1&1&2&2&2&1&3&3&3&2&1&1&3&3&1\\
1&2&2&2&1&3&1&0&1&3&3&1&2&3&1&2&2&2&1&1&1&1&3&3&1&1&2&3&3\\
1&2&1&2&2&3&3&1&0&1&2&3&3&1&3&3&1&2&1&2&1&1&1&3&1&3&2&1&2\\
1&2&3&3&1&2&1&3&1&0&1&3&1&1&3&2&1&2&2&2&3&3&2&1&1&1&3&1&2\\
1&1&2&3&3&2&2&3&2&1&0&1&3&2&3&2&1&1&3&1&1&1&2&1&2&3&3&1&1\\
1&3&1&2&3&1&2&1&3&3&1&0&1&2&2&2&1&3&3&2&1&2&2&1&3&1&1&3&1\\
1&1&3&1&2&2&2&2&3&1&3&1&0&2&1&1&1&3&1&2&2&3&1&1&3&2&1&3&3\\
2&2&1&1&3&1&1&3&1&1&2&2&2&0&1&3&3&3&1&1&1&3&1&1&3&3&2&2&2\\
2&2&2&1&2&2&1&1&3&3&3&2&1&1&0&1&3&1&1&3&1&1&3&1&3&2&2&1&3\\
2&1&3&3&2&1&1&2&3&2&2&2&1&3&1&0&1&1&3&1&1&1&1&2&1&2&3&3&3\\
2&3&3&1&2&2&2&2&1&1&1&1&1&3&3&1&0&1&1&1&3&1&2&3&2&1&3&3&2\\
2&1&1&1&1&1&2&2&2&2&1&3&3&3&1&1&1&0&1&3&3&2&3&3&1&2&3&2&1\\
2&1&2&2&2&3&2&1&1&2&3&3&1&1&1&3&1&1&0&1&3&3&3&3&2&1&2&1&1\\
2&1&2&3&3&3&1&1&2&2&1&2&2&1&3&1&1&3&1&0&1&3&1&2&2&3&1&3&1\\
2&2&2&2&1&1&3&1&1&3&1&1&2&1&1&1&3&3&3&1&0&2&2&2&3&3&1&1&3\\
3&3&1&1&2&2&3&1&1&3&1&2&3&3&1&1&1&2&3&3&2&0&1&1&2&1&2&1&2\\
3&3&3&1&1&1&3&3&1&2&2&2&1&1&3&1&2&3&3&1&2&1&0&1&1&2&2&2&1\\
3&1&1&3&3&2&2&3&3&1&1&1&1&1&1&2&3&3&3&2&2&1&1&0&1&1&2&2&2\\
3&2&1&3&1&3&1&1&1&1&2&3&3&3&3&1&2&1&2&2&3&2&1&1&0&2&1&2&1\\
3&3&3&2&1&1&1&1&3&1&3&1&2&3&2&2&1&2&1&3&3&1&2&1&2&0&1&1&2\\
3&1&1&1&1&3&3&2&2&3&3&1&1&2&2&3&3&3&2&1&1&2&2&2&1&1&0&1&1\\
3&1&2&2&2&1&3&3&1&1&1&3&3&2&1&3&3&2&1&3&1&1&2&2&2&1&1&0&1\\
3&3&2&1&3&1&1&3&2&2&1&1&3&2&3&3&2&1&1&1&3&2&1&2&1&2&1&1&0
\end{bmatrix}
$$

Fig. 4. Embedding (in boldface) and Solution (in gray text) for a $(3, 3, 4; 29)$ Ramsey coloring

Applying a SAT solver to complete this embedding to a $(4, 3, 3; 29)$ coloring satisfying Constraint (5) involves a CNF with 30,944 clauses and 4,736 variables and requires under two hours of computation time. The obtained coloring is depicted as Figure 4.

Proving that $R(4,3,3) = 30$. To apply the embedding approach described in this section to prove that there is no $(4,3,3;30)$ Ramsey coloring which is $(13,8,8)$ regular would require considering all $(3,3,3;13)$ colorings modulo weak isomorphism. Then, showing unsatisfiability of the SAT instances derived from all of the corresponding embeddings would constitute a proof that $R(4,3,3) = 30$. We defer this discussion until after Section 7 where we describe how we compute the set of all 78,892 $(3,3,3;13)$ Ramsey colorings modulo weak isomorphism.

4 Degree Matrices for Graph Colorings

We introduce an abstraction on graph colorings defined in terms of *degree matrices* and an equivalence relation on degree matrices. Our motivation is to solve graph coloring problems by first focusing on an over approximation of their degree matrices. The equivalence relation on degree matrices enables us to break symmetries during search when solving graph coloring problems. Intuitively, degree matrices are to graph edge-colorings as degree sequences are to graphs.

Definition 3 (abstraction, degree matrix). *Let A be a graph coloring on n vertices with k colors. The* degree matrix *of A, denoted $\alpha(A)$ is an $n \times k$ matrix, M such that $M_{i,j} = deg_j(i)$ is the degree of vertex i in color j. For a set $\mathcal{A}$ of graph colorings we denote $\alpha(\mathcal{A}) = \left\{ \alpha(A) \,\middle|\, A \in \mathcal{A} \right\}$.*

A degree matrix, M, is said to *represent* the set of graphs weakly-isomorphic to a graph with degrees as in M. We say that two degree matrices are equivalent if they represent the same sets of graph colorings.

Definition 4 (concretization and equivalence). *Let M and N be $n \times k$ degree matrices. Then, $\gamma(M) = \left\{ A \,\middle|\, A \approx A',\ \alpha(A') = M \right\}$ is the set of graph colorings represented by M and we say that $M \equiv N \Leftrightarrow \gamma(M) = \gamma(N)$. For a set $\mathcal{M}$ of degree matrices we denote $\gamma(\mathcal{M}) = \cup \left\{ \gamma(M) \,\middle|\, M \in \mathcal{M} \right\}$.*

Due to properties of weak-isomorphism (vertices as well as colors can be reordered) we can exchange both rows and columns of a degree matrix without changing the set of graphs it represents. In our construction we assume that the rows and columns of a degree matrix are sorted lexicographically. Observe also that the columns of a degree matrix each form a graphic sequence (when sorted).

Definition 5 (lex sorted degree matrix). *For an $n \times k$ degree matrix M we denote by $lex(M)$ the smallest matrix with rows and columns in the lexicographic order (non-increasing) obtained by permuting rows and columns of M.*

The following implies that for degree matrices we can assume without loss of generality that rows and columns are lexicographically ordered.

Theorem 3. *If M, N are degree matrices then $M \equiv N$ if and only if there exists permutations $\pi\colon [n] \to [n]$ and $\sigma\colon [k] \to [k]$ such that, for $1 \le i \le n$, $1 \le j \le k$, $M_{i,j} = N_{\pi(i),\sigma(j)}$.*

Proof. Let M and N be degree matrices. Then,
$M \equiv N \overset{\text{Defn. 4}}{\Longleftrightarrow} \gamma(M) = \gamma(N) \overset{\text{Defn. 4}}{\Longleftrightarrow} \forall_{G \approx H}.G \in \gamma(M) \leftrightarrow H \in \gamma(N) \overset{\text{Defn. 1}}{\Longleftrightarrow}$
$\exists_{\pi,\sigma}.\alpha(G)_{i,j} = \alpha(H)_{\pi(i),\sigma(j)} \overset{\text{Defn. 3}}{\Longleftrightarrow} M_{i,j} = N_{\pi(i),\sigma(j)}$

Corollary 3. $M \equiv lex(M)$.

Proof. The result follows from Theorem 3 because M and $lex(M)$ are related by permutations of rows and columns.

Example 1. In Figure 2, the degree matrix (right) describes both graphs (left).

5 Solving Graph Coloring Problems with Degree Matrices

Let $\varphi(A)$ be a graph coloring problem in k colors on an $n \times n$ adjacency matrix, A. Assuming that $\mathcal{A} = sol(\varphi(A))$ is too hard to compute, either because the number of solutions is too large or because finding even a single solution is too hard, our strategy is to first compute an over-approximation $\mathcal{M}$ of degree matrices such that $\gamma(\mathcal{M}) \supseteq \mathcal{A}$ and to then use $\mathcal{M}$ to guide the computation of $\mathcal{A}$. We denote the set of solutions of the graph coloring problem, $\varphi(A)$, which have a given degree matrix, M, by $sol_M(\varphi(A))$ and we have

$$sol_M(\varphi(A)) = sol(\varphi(A) \wedge \alpha(A){=}M) \tag{6}$$

Note that $M \notin \alpha(sol(\varphi(A))) \Rightarrow sol_M(\varphi(A)) = \emptyset$. Hence, for $\mathcal{M} \supseteq \alpha(sol(\varphi(A)))$,

$$sol(\varphi(A)) = \bigcup_{M \in \mathcal{M}} sol_M(\varphi(A)) \tag{7}$$

Equation (7) implies that, using any over-approximation $\mathcal{M} \supseteq \alpha(sol(\varphi(A)))$, we can compute the solutions to a graph coloring problem by computing the independent sets $sol_M(\varphi(A))$ for each $M \in \mathcal{M}$. This facilitates the computation of $sol(\varphi(A))$ for three reasons: (1) The problem is now broken into a set of independent sub-problems for each $M \in \mathcal{M}$ which can be solved in parallel. (2) The computation of each individual $sol_M(\varphi(A))$ is now directed using M, and (3) Symmetry breaking is facilitated.

There are two symmetry breaks when solving $\varphi(A)$. First, we compute $\mathcal{M}$ to consist of canonical degree matrices, sorted lexicographically by rows and by columns. Second, we impose an additional symmetry breaking constraint $\mathsf{sb}^*_\ell(A, M)$ as explained below.

Consider a computation of all solutions of the constraint in the right side of Equation (6). Consider a permutation π of the rows and columns of A, such that $\alpha(\pi(A)) = \alpha(A) = M$. Then, both A and A' are solutions and they are weakly isomorphic. The following equation

$$sol_M(\varphi(A)) = sol(\varphi(A) \wedge (\alpha(A){=}M) \wedge \mathsf{sb}^*_\ell(A, M)) \tag{8}$$

refines Equation (6) introducing a symmetry breaking constraint similar to the (partitioned lexicographic) symmetry break predicate introduced by Codish *et al.* in [5] for Boolean adjacency matrices.

$$\mathsf{sb}^*_\ell(A, M) = \bigwedge_{i<j} \big(\big(M_i = M_j \Rightarrow A_i \preceq_{\{i,j\}} A_j\big) \big) \tag{9}$$

where $s \preceq_{\{i,j\}} s'$ denotes the lexicographic order on strings s and s' after simultaneously omitting the elements at positions i and j.

To justify that Equations (6) and (8) both compute $sol_M(\varphi(A))$, modulo weak isomorphism, we must show that whenever $\mathsf{sb}^*_\ell(A, M)$ excludes a solution then there is another weakly isomorphic solution that is not excluded. To this end, we introduce a definition and then a theorem.

Definition 6 (degree matrix preserving permutation). *Let A be an adjacency matrix with a lexicographically ordered degree matrix $\alpha(A) = M$. We say that permutation π is* degree matrix preserving *for M and A if $\alpha(\pi(A)) = M$.*

Theorem 4 (correctness of $\mathsf{sb}^*_\ell(A, M)$). *Let A be an adjacency matrix with a lexicographically ordered degree matrix $\alpha(A) = M$. Then, there exists a degree matrix preserving permutation π such that $\alpha(\pi(A)) = M$ and $\mathsf{sb}^*_\ell(\pi(A), M)$ holds.*

Proof. If the rows of M are distinct, then the theorem holds with π the identity permutation. Assume that some rows of M are equal. Denote by P the set of degree matrix preserving permutations for M and A. Assume the premise and that no $\pi \in P$ satisfies $\mathsf{sb}^*_\ell(\pi(A), M)$. Let $\pi \in P$ be such that $\pi(A) = \min \left\{ \pi'(A) \in P \,\middle|\, \pi' \in P \right\}$ (in the lexicographical order viewing matrices as strings). From the assumption, there exist $i < j$ such that $M_i = M_j$ and $\pi(A)_i \not\preceq_{\{i,j\}} \pi(A)_j$. Hence there exists a minimal index $k \notin \{i, j\}$ such that $\pi(A)_{i,k} > \pi(A)_{j,k}$. Let A' be the matrix obtained by permuting nodes i and j in $\pi(A)$. Since $M_i = M_j$ it follows that $\alpha(A') = M$. Thus there is a $\pi' \in P$ such that $\pi'(A) = A'$. If $k < i$: for $1 \leq l < k$ we have $\pi(A)_l = A'_l$. Thus k is the first row for which A' and $\pi(A)$ differ. Permuting nodes i and j changes row k by simply swapping elements $\pi(A)_{k,i}$ and $\pi(A)_{k,j}$. Since $\pi(A)_{k,i} > \pi(A)_{k,j}$, clearly $A'_k \prec \pi(A)_k$ hence $A' \prec \pi(A)$ which is a contradiction. Similarly if $k > i$ the same argument applies to show that i is the first row for which A' and $\pi(A)$ differ, thus obtaining the same contradiction for row i.

The following corollary clarifies that if a solution A is eliminated when introducing the symmetry break predicate $\mathsf{sb}^*_\ell(A, \alpha(A))$ to a graph coloring problem then there always remains an isomorphic solution A' which satisfies the predicate $\mathsf{sb}^*_\ell(A', \alpha(A'))$.

Corollary 4. *Let A be an adjacency matrix. Then there exists A' isomorphic to A such that $\alpha(A')$ is lex ordered and $\mathsf{sb}^*_\ell(A', \alpha(A'))$ holds.*

Proof. Let $M = \alpha(A)$. From Corollary 3 we know that $M \equiv lex(M)$, thus there exists A'' isomorphic to A such that $\alpha(A'') = lex(M)$. From Theorem 3 it follows that there exists a degree matrix preserving permutation π such that $\alpha(\pi(A'')) = lex(M)$ and $\mathsf{sb}^*_\ell(\pi(A''), \alpha(\pi(A'')))$ holds. If $A' = \pi(A'')$ then A' is isomorphic to A, $\alpha(A')$ is lex ordered and $\mathsf{sb}^*_\ell(A', \alpha(A'))$ holds.

6 Computing Degree Matrices for $R(3, 3, 3; 13)$

This section describes how we compute a set $\mathcal{M}$ of degree matrices that approximate those of the solutions of Constraint (4). We apply a strategy in which we mix SAT solving with brute-force enumeration as follows. The computation of the degree matrices is summarized in Table 3. In the first step, we compute bounds on the degrees of the nodes in any $R(3,3,3;13)$ coloring.

Lemma 2. *Let A be a $R(3,3,3;13)$ coloring then for every vertex x in A, and color $c \in \{1,2,3\}$, $2 \leq deg_c(x) \leq 5$.*

Proof. By solving Constraint (4) together with $\mathsf{sb}^*_\ell(A, M)$ seeking a graph with minimal degree less than 2 or maximal degree greater than 5. The CNF encoding is of size 13672 clauses with 2748 Boolean variables and takes under 15 seconds to solve and yields an UNSAT result which implies that such a graph does not exist.

In the second step, we enumerate the degree sequences with values within the bounds specified by Lemma 2. Recall that the degree sequence of an undirected graph is the non-increasing sequence of its vertex degrees. Not every non-increasing sequence of integers corresponds to a degree sequence. A sequence that corresponds to a degree sequence is said to be graphical. The number of degree sequences of graphs with 13 vertices is 836,315 (see Sequence number `A004251` of The On-Line Encyclopedia of Integer Sequences, published electronically at`http://oeis.org`). However, when the degrees are bound by Lemma 2 there are only 280.

Lemma 3. *There are 280 degree sequences with values between* 2 *and* 5.

Proof. By straightforward enumeration using the algorithm of Erdos and Gallai [10].

In the third step, we test each of the 280 degree sequences identified by Lemma 3 to determine how many of them might occur as the left column in a degree matrix.

Lemma 4. *Let A be a $R(3,3,3;13)$ coloring and let M be the canonical form of $\alpha(A)$. Then, (a) the left column of M is one of the 280 degree sequences identified in Lemma 3; and (b) there are only 80 degree sequences from the 280 which are the left column of $\alpha(A)$ for some coloring A in $R(3,3,3;13)$.*

Proof. By solving Constraint (4) with each degree sequence from Lemma 3 to test if it is satisfiable. This involves 280 instances with average CNF size: 10861 clauses and 2215 Boolean variables. The total solving time is 375.76 hours and the hardest instance required about 50 hours. These instances were solved in parallel on the cluster described in Section 1.

In the fourth step we extend the 80 degree sequences identified in Lemma 4 to obtain all possible degree matrices.

Lemma 5. *Given the 80 degree sequences identified in Lemma 4 as potential left columns of a degree matrix, there are 11,933 possible degree matrices.*

Proof. By straightforward enumeration. The rows and columns are lex sorted, must sum to 12, and the columns must be graphical (when sorted). We first compute all of the degree matrices and then select the smallest representatives under permutations of rows and columns. The computation requires a few seconds.

Step	Notes	ComputationTimes	CNF Size	
1	compute degree bounds (Lemma 2) (1 instance, unsat)	12.52 sec.	#Vars 2748	#Clauses 13672
2	enumerate 280 possible degree sequences (Lemma 3)	Prolog, fast (seconds)		

3	test degree sequences (Lemma 4)
	(280 instances: 200 unsat, 80 sat)
		16.32 hrs.
		hardest: 1.34 hrs
			#Vars
			1215 (avg)
				#Clauses
				7729(avg)

4	enumerate 11,933 degree matrices (Lemma 5)	Prolog, fast (seconds)		
5	test degree matrices (Lemma 6) (11,933 instances: 10,934 unsat, 999 sat)	126.55 hrs. hardest: 0.88 hrs.	#Vars 1520 (avg)	#Clauses 7632 (avg)

Table 3. Computing the degree matrices for $\mathcal{R}(3,3,3;13)$ step by step.

Step	Notes	Computation Times		
1	compute all Ramsey $(3,3,3;13)$ colorings per degree matrix (999 instances, 129,188 solutions)	MiniSAT total: 308.23 hr. hardest: 9.15 hr.	CryptoMiniSAT total: 136.31 hr. hardest: 4.3 hr.	Glucose total: 373.2 hr. hardest:17.67 hr.
2	reduce modulo $\approx$. (78,892 solutions)	`nauty`, fast (minutes)		

Table 4. Computing $\mathcal{R}(3,3,3;13)$ step by step.

In the fifth step, we test each of the 11,933 degree matrices identified by Lemma 4 to determine how many of them are the abstraction of some $R(3,3,3;13)$ coloring.

Lemma 6. *From the 11,933 degree matrices identified in Lemma 5, 999 are $\alpha(A)$ for a coloring A in $R(3,3,3;13)$.*

Proof. By solving Constraint (4) together with a given degree matrix to test if it is satisfiable. This involves 11,933 instances with average CNF size: 7632 clauses and 1520 Boolean variables. The total solving time is 126.55 hours and the hardest instance required 0.88 hours. These instances were solved in parallel on the cluster described in Section 1.

7 Computing $R(3,3,3;13)$ from Degree Matrices

We describe the computation of the set of all $(3,3,3;13)$ colorings starting from the 3805 degree matrices identified in Section 6. Table 4 summarizes the two step experiment reporting the computation on three different SAT solvers: MiniSAT [8, 9], CryptoMiniSAT [22], and Glucose [2, 3].

step 1: For each degree matrix we compute, using a SAT solver, all corresponding solutions of Equation (8), where $\varphi(A)$ is constraint (4) and M is one of the 999 degree matrices identified in (Lemma 6). These instances were solved in parallel on the cluster described in Section 1. This generates in total 129,188 $(3,3,3;13)$ Ramsey colorings. Table 4 details the total solving time for these instances and the solving times for the hardest instance for each SAT solver. The largest number of graphs generated by a single instance is 3720.

step 2: The 129,188 $(3,3,3;13)$ colorings from step 1 are reduced modulo weak-isomorphism using `nauty`[3] [14]. This process results in a set with 78,892 graphs.

[3] Note that `nauty` does not handle edge colored graphs and weak isomorphism directly. We applied an approach called k-layering described at `https://computationalcombinatorics.wordpress.com/2012/09/20/canonical-labelings-with-nauty`.

Does $\mathbf{R(4,3,3) = 30}$? In order to prove that there are no $(4,3,3;30)$ colorings with degrees $(13,8,8)$ using the embedding approach, we need to check that $78{,}892 \times 3 \times 3 = 710{,}028$ corresponding instances are unsatisfiable. However, the three $(2,3,4;8)$ colorings are described by the single matrix with constraints as portrayed in Figure 5 (which includes also a fourth solution that is not a $(2,3,4;8)$ coloring) and similar for the three $(3,2,4;8)$ colorings. So, in fact we have a total of only 78,892 embedding instances to consider. Over the past months, using the cluster described in Section 1, we have determined 78,872 instances (99.97%) to be unsatisfiable and found no satisfiable instance. The average size of an instance is 36,259 clauses with 5187 variables. The average solving time is 13.91 hours per instance. Table 5 specifies, in the second column, the total number of instances (from the 78,872 solved so far) that can be shown unsatisfiable within the specified number of hours. The third column indicates the increment in percentage (within 10 hours we solve 71.46%, within 20 hours we solve an additional 12.11%, etc). The last row in the table indicates that there are 5 instances (0.01%) which require between 1000 and 1600 hours of computation. Only 20 instance remain to be solved. Each of these is running on a single thread of the cluster. We expect these to finish within days, and conclude that $R(4,3,3) = 30$.

$$\begin{bmatrix} 0 & 2 & 2 & 2 & 3 & 3 & 3 & 3 \\ 2 & 0 & 3 & 3 & 2 & \mathtt{A} & 3 & 3 \\ 2 & 3 & 0 & 3 & \mathtt{A} & \mathtt{B} & 2 & 3 \\ 2 & 3 & 3 & 0 & 3 & \mathtt{B} & \mathtt{A} & 2 \\ 3 & 2 & \mathtt{A} & 3 & 0 & \mathtt{B} & \mathtt{C} & \mathtt{A} \\ 3 & \mathtt{A} & \mathtt{B} & \mathtt{B} & \mathtt{B} & 0 & \mathtt{A} & \mathtt{A} \\ 3 & 3 & 2 & \mathtt{A} & \mathtt{C} & \mathtt{A} & 0 & \mathtt{B} \\ 3 & 3 & 3 & 2 & \mathtt{A} & \mathtt{A} & \mathtt{B} & 0 \end{bmatrix}$$

Fig. 5. Approximating the three (2,3,4;8) colorings by a single matrix with the constraints $\mathtt{A}, \mathtt{B}, \mathtt{C} \in \{2,3\}$ and $\mathtt{A} \neq \mathtt{B}$.

8 Conclusion

We have applied SAT solving techniques to show that any $(4,3,3;30)$ Ramsey coloring must be $(13,8,8)$ regular. In order to apply the same technique to show that there is no $(13,8,8)$ regular coloring we would need to make use of the set of all $(3,3,3;13)$ colorings. We have computed this set modulo weak isomorphism. To this end we applied a technique involving abstraction and symmetry breaking to reduce the redundancies in the number of isomorphic solutions obtained when applying the SAT solver. Ongoing computation is proceeding to determine the precise value of $R(4,3,3)$.

time (hrs)	# instances	% instances (Δ)
10	56,363	71.46 %
20	65,914	12.11 %
100	77,263	14.39 %
500	78,791	1.94 %
1000	78,867	0.09 %
1600	78,872	0.01 %

Table 5. Time required per instance for proof that there are no $(4,3,3;30)$ colorings with degrees $(13,8,8)$: $78{,}872 \setminus 78{,}892$ instances.

Acknowledgments

We thank Stanislaw Radziszowski for his guidance and comments which helped to improve the presentation of this paper. In particular Stanislaw proposed to show that our technique is able to find the $(4,3,3;29)$ coloring depicted as Figure 4.

References

1. J. M. Al-Jaam. Experiments of intelligent algorithms on Ramsey graphs. *Int. Arab J. Inf. Technol.*, 4(2):161–167, 2007.
2. G. Audemard and L. Simon. *Glucose SAT Solver*. `http://www.labri.fr/perso/lsimon/glucose/`.
3. G. Audemard and L. Simon. Predicting learnt clauses quality in modern SAT solvers. In C. Boutilier, editor, *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, pages 399–404, 2009.
4. J. R. Carroll and G. Isaak. Degree matrices realized by edge-colored forests. Available from `www.lehigh.edu/~gi02/ecforest.pdf` (unpublished), 2009.
5. M. Codish, A. Miller, P. Prosser, and P. J. Stuckey. Breaking symmetries in graph representation. In F. Rossi, editor, *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*. IJCAI/AAAI, 2013.
6. M. Codish, A. Miller, P. Prosser, and P. J. Stuckey. Constraints for symmetry breaking in graph representation. Full version of [5] (in preparation)., 2014.
7. J. M. Crawford, M. L. Ginsberg, E. M. Luks, and A. Roy. Symmetry-breaking predicates for search problems. In L. C. Aiello, J. Doyle, and S. C. Shapiro, editors, *KR*, pages 148–159. Morgan Kaufmann, 1996.
8. N. Eén and N. Sörensson. *MiniSAT SAT Solver*. `http://minisat.se/Main.html`.
9. N. Eén and N. Sörensson. An extensible sat-solver. In E. Giunchiglia and A. Tacchella, editors, *Theory and Applications of Satisfiability Testing, 6th International Conference, SAT 2003. Santa Margherita Ligure, Italy, May 5-8, 2003 Selected Revised Papers*, volume 2919 of *Lecture Notes in Computer Science*, pages 502–518. Springer, 2003.
10. P. Erdös and T. Gallai. Graphs with prescribed degrees of vertices (in Hungarian). *Mat. Lapok*, pages 264–274, 1960. Available from `http://www.renyi.hu/~p_erdos/1961-05.pdf`.
11. I. P. Gent and B. M. Smith. Symmetry breaking in constraint programming. In W. Horn, editor, *ECAI 2000, Proceedings of the 14th European Conference on Artificial Intelligence, Berlin, Germany, August 20-25, 2000*, pages 599–603. IOS Press, 2000.
12. J. Kalbfleisch and R. Stanton. On the maximal triangle-free edge-chromatic graphs in three colors. *Journal of Combinatorial Theory*, 5:9–20, 1968.
13. J. G. Kalbfleisch. *Chromatic Graphs and Ramsey's Theorem*. PhD thesis, University of Waterloo, January 1966.
14. B. McKay. *nauty* user's guide (version 1.5). Technical Report TR-CS-90-02, Australian National University, Computer Science Department, 1990.
15. P. Meseguer and C. Torras. Exploiting symmetries within constraint satisfaction search. *Artif. Intell.*, 129(1-2):133–163, 2001.
16. A. Metodi, M. Codish, and P. J. Stuckey. Boolean equi-propagation for concise and efficient SAT encodings of combinatorial problems. *J. Artif. Intell. Res. (JAIR)*, 46:303–341, 2013.
17. K. Piwakowski. On Ramsey number r(4, 3, 3) and triangle-free edge-chromatic graphs in three colors. *Discrete Mathematics*, 164(1-3):243–249, 1997.
18. K. Piwakowski and S. P. Radziszowski. $30 \leq R(3,3,4) \leq 31$. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 27:135–141, 1998.
19. J. Puget. On the satisfiability of symmetrical constrained satisfaction problems. In H. J. Komorowski and Z. W. Ras, editors, *Methodologies for Intelligent Systems, 7th International Symposium, ISMIS '93, Trondheim, Norway, June 15-18, 1993, Proceedings*, volume 689 of *Lecture Notes in Computer Science*, pages 350–361. Springer, 1993.
20. S. P. Radziszowski. Personal communication. January, 2015.

21. S. P. Radziszowski. Small Ramsey numbers. *Electronic Journal of Combinatorics*, 1994. Revision #14: January, 2014.
22. M. Soos. CryptoMiniSAT, v2.5.1. `http://www.msoos.org/cryptominisat2`, 2010.