

# Edge coloring of graphs, uses, limitation, complexity

Sándor SZABÓ

University of Pécs

Institute of Mathematics and

Informatics, Ifjúság u. 6

7624 Pécs, HUNGARY

email: [sszabo7@hotmail.com](mailto:sszabo7@hotmail.com)

Bogdán ZAVÁLNIJ

University of Pécs

Institute of Mathematics and

Informatics, Ifjúság u. 6

7624 Pécs, HUNGARY

email: [bogdan@ttk.pte.hu](mailto:bogdan@ttk.pte.hu)

**Abstract.** The known fact that coloring of the nodes of a graph improves the performance of practical clique search algorithm is the main motivation of this paper. We will describe a number of ways in which an edge coloring scheme proposed in [8] can be used in clique search. The edge coloring provides an upper bound for the clique number. This estimate has a limitation. It will be shown that the gap between the clique number and the upper bound can be arbitrarily large. Finally, it will be shown that to determine the optimal number of colors in an edge coloring is NP-hard.

## 1 Introduction

Let  $G = (V, E)$  be a graph. Here  $V$  is the set of nodes of the graph  $G$  and  $E$  is the set of edges of the graph. In this paper we will be dealing exclusively with finite graph, that is, it will be assumed that the sets  $V$  and  $E$  are finite. We further narrow the class of graphs we consider throughout this paper by

---

**Computing Classification System 1998:** G.2.2

**Mathematics Subject Classification 2010:** 05C15

**Key words and phrases:** clique, maximum clique, maximal clique, clique search algorithms, vertex coloring, edge coloring.

excluding the graphs that have either loops or double edges. In other words we will be working with finite simple graph.

Let  $k$  be a fixed positive integer. A subgraph  $\Delta$  of  $G$  is called a  $k$ -clique in  $G$  if each two distinct nodes of  $\Delta$  are adjacent in  $G$  and  $\Delta$  has  $k$  nodes. The number of edges in  $\Delta$  is equal to  $k(k-1)/2$ . We refer to  $k$  as the size of the clique  $\Delta$ . Sometimes we call  $\Delta$  a clique of size  $k$  instead of a  $k$ -clique. A  $k$ -clique  $\Delta$  in  $G$  is called a maximal clique if  $\Delta$  is not a subgraph of any  $(k+1)$ -clique in  $G$ . A  $k$ -clique  $\Delta$  in  $G$  is called a maximum clique if  $G$  does not contain any  $(k+1)$ -clique. A graph may have several maximum cliques. All maximum cliques in  $G$  have a well defined common size. This well defined number is referred as the clique number of  $G$ , and it is denoted by  $\omega(G)$ .

The expression “clique search problem” refers to a number of problems related to finding cliques in a given graph. One may look for maximal cliques or maximum cliques. One might be interested in listing all maximal cliques or listing all maximum cliques. We maybe content with locating only one maximum clique. Or we maybe satisfied with just learning the clique size of  $G$  without exhibiting any maximum clique. We describe some relevant clique search problems more formally.

**Problem 1** *Given a finite simple graph  $G$  and given a positive integer  $k$ . Decide if  $G$  contains a  $k$ -clique.*

Problem 1 is a decision problem and it is well-known that it belongs to the NP-complete complexity class. (See [7].)

**Problem 2** *Given a finite simple graph  $G$  and a positive integer  $k$ . List all  $k$ -cliques in  $G$ .*

Problem 2 is not a decision problem. It is clear that Problem 2 cannot be computationally less demanding than Problem 1.

Determining the clique number  $\omega(G)$  of  $G$  is not a decision problem either. It is an optimization problem. But again it must be clear that finding  $\omega(G)$  is computationally at least as challenging as Problem 1.

Clique search problems have many practical applications and there is a considerable amount of research devoted to them. For details see for example [1], [3], [4], [5]. Many practical clique search algorithms utilize the coloring the nodes of a graph. We color the nodes of a given finite simple graph  $G$  with  $k$  colors satisfying the following conditions.

- (1) Each node receives exactly one of the colors.

- (2) Adjacent nodes never receive the same color.

This is the most commonly encountered coloring of the nodes of a graph. It is referred as a legal or a well or a proper coloring of the nodes of  $G$ . In connection with each finite simple graph  $G$  there is a number of colors  $k$  such that the nodes of  $G$  have a legal coloring with  $k$  colors and the nodes of  $G$  do not have any legal coloring with  $k - 1$  colors. This well defined number  $k$  is called the chromatic number of  $G$  and it is denoted by  $\chi(G)$ . Coloring of graphs is a vast subject on its own. In this paper we take a rather narrow view of coloring. We are interested in coloring only from one reason. Coloring provides upper estimates for  $\omega(G)$ . Namely,  $\omega(G) \leq \chi(G)$ .

One can devise further coloring schemes to get new upper bounds for  $\omega(G)$ . For example we may color the edges of a graph  $G$  with  $k$  colors in the following way.

- (1) Each edge receives exactly one color.
- (2) If  $x, y, z$  are distinct nodes of a 3-clique in  $G$ , then the edges  $\{x, y\}$ ,  $\{y, z\}$ ,  $\{x, z\}$  must receive three distinct colors.
- (3) If  $x, y, u, v$  are distinct nodes of a 4-clique in  $G$ , then the edges  $\{x, y\}$ ,  $\{x, u\}$ ,  $\{x, v\}$ ,  $\{y, u\}$ ,  $\{y, v\}$ ,  $\{u, v\}$  must receive six distinct colors.

We call this type of coloring of the edges of  $G$  a legal or well or proper edge coloring.

For a given finite simple graph  $G$  there is a number of colors  $k$  such that the edges of  $G$  has a legal coloring with  $k$  colors and the edges of  $G$  does not admit any legal coloring using  $k - 1$  colors. This minimum number of colors is called the edge chromatic number of  $G$  and it is denoted by  $\chi_{(e)}(G)$ .

The reader can observe that the inequality

$$\omega(G) (\omega(G) - 1) / 2 \leq \chi_{(e)}(G)$$

must hold. So coloring the edges of a graph  $G$  can be used to establish an upper bound for  $\omega(G)$ .

The edge coloring scheme described here was proposed in [8]. It was mentioned that a large scale numerical experiment indicates that typically edge coloring provides better bounds for the clique number than the node coloring. On the hand the edge coloring is computationally more expensive than the node coloring. In Tables 1, 2, and 3 we presented some of the numerical results. These results were not reported earlier. The graphs we used are related

to the construction of certain codes. The captions of the tables simply refer to the related codes.

In the tables the columns labeled by  $|V|$  and  $|E|$  contain the numbers of the nodes and the edges of the graphs, respectively. Using the simplest sequential node and edge coloring procedures give the number of colors listed in the columns labeled by the words “node” and “edge”, respectively. Finally, the column labeled by the word “estimate” lists the upper estimate of the clique size computed from the number of colors of the edges.

Let  $G = (V, E)$  be a finite simple graph. Using  $G$  we construct a new auxiliary graph  $\Gamma = (W, F)$ . We set  $W = E$  and the distinct edges  $\{u, v\}, \{x, y\}$  of  $G$  will be adjacent nodes of the graph  $\Gamma$  if each of the unordered pairs

$$\{u, x\}, \{u, y\}, \{v, x\}, \{v, y\}$$

is an edge of the graph  $G$ . The reader can verify that a legal coloring of the nodes of the auxiliary graph  $\Gamma$  corresponds to a legal coloring of the edges of the original graph  $G$ . We may refer to  $\Gamma$  as the derived graph of  $G$ .

## 2 Applications of edge coloring

In this section we discuss the relevance of edge coloring to clique search algorithms.

The edge coloring can be used as a preconditioning technique.

Suppose that the edge coloring of the graph  $G = (V, E)$  is given by the function  $f : E \rightarrow \{1, \dots, t\}$ . Here  $f(\{x, y\}) = c$  means that the edge  $\{x, y\}$  receives color  $c$ . The function  $f$  can be stored conveniently as a matrix  $M$ . The typical entry  $m(x, y)$  of  $M$  is defined by

$$m(x, y) = \begin{cases} f(\{x, y\}), & \text{if } \{x, y\} \in E, \\ 0, & \text{if } \{x, y\} \notin E. \end{cases}$$

The number of the colors used for coloring the neighbors of the node  $x$  is equal to the number of the distinct colors appearing in row  $x$  of the matrix  $M$ . We may call this number the color degree of the node  $x$  in the graph  $G$ .

If  $x$  is a node of a  $k$ -clique  $\Delta$  in  $G$ , then the color degree of the node  $x$  must be at least  $k - 1$ . Therefore, if the color degree of the node  $x$  is less than  $k - 1$ , then node  $x$  can be deleted from the graph  $G$  without losing the clique  $\Delta$ . In other words, when we are looking for a  $k$ -clique in the graph  $G$  we can delete safely each node whose color degree is less than or equal to  $k - 2$ .

There is further way to exploit the edge coloring for preconditioning. The greedy sequential coloring of the edges of the given graph  $G$  as a side result provides us with the adjacency matrix or the linked lists representation of the derived graph  $\Gamma$  of  $G$ . An inspection can help to delete nodes or edges of  $\Gamma$ .

The edge coloring can be used as a pruning rule. One simply can add an edge colored graph to the Carraghan-Pardalos [2] clique search algorithm.

The Carraghan-Pardalos algorithm at a particular stage of its execution maintains two sets of nodes. The first one is the set of nodes  $U$  which consists of the nodes of an  $r$ -clique  $\Delta$ . The second one is the set of nodes  $L$  which contains the nodes of  $G$  that have a chance to extend the clique  $\Delta$ . We restrict the graph  $G$  to the set  $L$ . Let  $H$  be the graph spanned by the set  $L$  in  $G$ . The edges of  $H$  are colored as  $H$  inherits a coloring from  $G$ .

One can count the number of the distinct colors appearing as edge colors in  $H$ . Let this number be  $s$ . Using  $s$  one can estimate the clique size of  $H$ . Namely, if  $\omega(H) = t$ , then  $t(t-1)/2 \leq s$  must hold. This means that  $t \leq (1 + \sqrt{1 + 8s})/2$  must hold.

Suppose we are looking for a  $k$ -clique in the given graph  $G$ . It is clear that if  $\omega(\Delta) + \omega(H) \leq k-1$ , then choosing nodes from the set  $L$  the clique  $\Delta$  cannot be extended to a  $k$ -clique. Thus, if

$$r + (1 + \sqrt{1 + 8s})/2 \leq k-1,$$

then at this node of the search tree one can terminate the search. In other words at this node we can prune the search tree. One can record the edge coloring of  $G$  using the matrix  $M$  described earlier. One can easily store two different edge colorings of  $G$  in the matrix  $M$ . Using two edge colorings enhances the efficiency of the pruning.

When the largest color class is very large compared with the others, then the edge coloring offers a new opportunity to estimate the clique size.

Let  $\Delta$  be a maximum clique in the given graph  $G$ . Suppose that the color classes of the edges in  $G$  are  $C_1, \dots, C_k$  and  $|C_1| \geq \dots \geq |C_k|$  holds. If we delete the edges appearing in  $C_1$ , then we get a new graph  $G'$  from  $G$ . If  $\Delta$  does not contain any edge from  $C_1$ , then  $\Delta$  is a maximum clique in  $G'$  too. In this case  $\omega(G) = \omega(G')$ . If  $\Delta$  does contain an edge from  $C_1$ , then it may happen that  $\omega(G) = \omega(G') + 1$ . In either case  $\omega(G) \leq \omega(G') + 1$ . Since  $G'$  has fewer edges than  $G$  computing or estimating  $\omega(G')$  can be simpler than computing or estimating  $\omega(G)$ . In this way we may collect some information about the clique size of  $G$ .

The edge coloring can be used as a branching rule to devise a parallel clique search algorithm.

Let us assume that we are interested in deciding if the given graph  $G$  contains a  $k$ -clique. Here  $k$  is a given positive integer. Suppose that the edges of  $G$  are legally colored with  $t$  colors. Let  $C_1, \dots, C_t$  be the colors classes of the edges such that  $|C_1| \geq \dots \geq |C_t|$ . The edges of  $k$ -clique  $\Delta$  in  $G$  can be colored with  $r = (k(k-1))/2$  colors and cannot be colored with fewer colors. If  $t < r$ , then clearly  $G$  cannot contain any  $k$ -clique. For the remaining part of the argument we assume that  $t \geq r$ . Let

$$e_1 = \{x_1, y_1\}, \dots, e_s = \{x_s, y_s\} \quad (1)$$

be all the edges in the color classes  $C_1, \dots, C_t$ . Let  $G_i$  be the subgraph of  $G$  spanned by the set of nodes  $N(x_i) \cap N(y_i)$  in  $G$  for each  $i$ ,  $1 \leq i \leq s$ . Here  $N(x)$  denotes the set of neighbors of the node  $x$  in the graph  $G$ . If  $G_i$  contains a  $(k-2)$ -clique for some  $i$ ,  $1 \leq i \leq s$ , then  $G$  contains a  $k$ -clique. In this case our problem is solved.

For the remaining part we may assume that  $G_i$  does not contain any  $k$ -clique for each  $i$ ,  $1 \leq i \leq s$ . It means that we can delete the edge  $e_i = \{x_i, y_i\}$  from  $G$  without removing any  $k$ -clique from  $G$ . (When we delete the edge  $e_i$  we do not delete any of the nodes  $x_i$  and  $y_i$ .) Deleting the edges (1) from  $G$  we end up with a graph  $G'$  whose edges are legally colored with  $r-1$  colors. Consequently, this graph  $G'$  cannot contain any  $k$ -clique.

The summary of our consideration is that locating a  $k$ -clique in  $G$  can be reduced to locating a  $(k-2)$ -clique in the graphs  $G_1, \dots, G_s$ . We replaced the original clique search problem by a large number of smaller clique search instances. These smaller problems can be attacked independently of each other and we can solve them using a number of processors simultaneously. In this sense the edge coloring can form the base of a parallel clique search algorithm.

When the number of the nodes is overly large, then we can divide the set of nodes of the graph into two disjoint sets. The clique sizes of the smaller subgraphs induced by these sets provide lower and upper bounds for the clique size of the original graph. Using edge coloring the upper bound can be improved.

Let  $G$  be a finite simple graph. We divide the set of nodes  $V$  into two disjoint subsets  $U$  and  $W$ . Let  $H, K$  be the subgraphs of  $G$  induced by the subsets  $U$  and  $W$ , respectively. We consider a bipartite subgraph  $L$  of  $G$  induced by the subsets  $U$  and  $W$ . Note that

$$\max\{\omega(H), \omega(K)\} \leq \omega(G) \leq \omega(H) + \omega(K)$$

holds. Setting  $h = \omega(H)$ ,  $k = \omega(K)$  the upper estimate is  $\omega(G) \leq h + k$ . Coloring the edges of the bipartite graph  $L$  provides a correction term  $r$  to modify the estimate to  $\omega(G) \leq h + k - r$ .

Suppose  $G$  contains a  $(h+k)$ -clique. In this case  $\omega(G) = h+k$  and the bipartite graph  $L$  must contain a  $(h,k)$ -biclique. Note that this biclique has  $hk$  edges and the colors of these edges are pair-wise distinct. In other words, if the edges of  $L$  can be colored legally using less than  $hk$  colors, then the equation  $\omega(G) = h+k$  cannot hold.

Suppose that the edges of  $L$  can be colored legally using  $s$  colors. If  $s$  is small compared to  $hk$ , then the upper estimate for  $\omega(G)$  can be lowered. For the sake of definiteness let us assume that  $h \geq k$ . Choose an integer  $r$  such that

$$(k-r)h \leq s < (k-r+1)h.$$

It follows that

$$k - (s/h) \leq r < k - (s/h) + 1.$$

This  $r$  is the correction term to improve the estimate for  $\omega(G)$ .

The edge coloring can be used to construct cuts in the linear programming reformulation of the maximum clique problem.

The maximum clique problem can be reformulated in terms of a 0-1 linear program. Let  $G = (V, E)$  be a finite simple graph with  $V = \{v_1, \dots, v_n\}$  and let  $\Delta$  be a clique in  $G$  such that  $U$  is the set of nodes of  $\Delta$ . To the clique  $\Delta$  we assign an  $n$ -dimensional vector  $\mathbf{x}^T = [x_1, \dots, x_n]$  such that

$$x_i = \begin{cases} 1, & \text{if } v_i \in U, \\ 0, & \text{if } v_i \notin U. \end{cases}$$

We may call  $\mathbf{x}$  the characteristic vector of the clique  $\Delta$ .

We consider the following 0-1 linear program  $P$ . Maximize the objective function  $x_1 + \dots + x_n$  subject to the constraints  $x_i + x_j \leq 1$ , where the unordered pair  $\{v_i, v_j\}$  is not an edge of  $G$ . Replacing the condition  $x_i \in \{0, 1\}$  by  $0 \leq x_i \leq 1$  for each  $i$ ,  $1 \leq i \leq n$  we get the real relaxation  $P'$  of the 0-1 linear program  $P$ . An optimum solution of  $P'$  provides an upper bound for the clique size  $\omega(G)$  of  $G$ .

Feeding the program  $P'$  into a linear program solver we typically get the optimum solution  $[1/2, \dots, 1/2]^T$  which leads to the estimate  $\omega(G) \leq n/2$ .

The polyhedron of the program  $P'$  may have many vertices with non-integer components. There are inequalities in the form  $a_1x_1 + \dots + a_nx_n \leq b$  that slices down non-integer solutions from the polyhedron but not slicing down any integer solutions. Such inequalities are called cuts and they can be added to the constraints of the program  $P'$  to improve the estimate for  $\omega(G)$ .

Suppose that the edges of  $G$  are legally colored using  $t$  colors and  $C_1, \dots, C_t$  are the color classes of the edges of  $G$  such that

$$e_1 = \{a_1, b_1\}, \dots, e_s = \{a_s, b_s\}$$

are the edges in the color class  $C_i$ . Let  $v_{i(1)}, \dots, v_{i(r)}$  be all the distinct nodes of  $G$  among  $a_1, b_1, \dots, a_s, b_s$ . Note that the set  $\{v_{i(1)}, \dots, v_{i(r)}\}$  cannot contain the nodes of any 3-clique in the graph  $G$ . Consequently, the inequality

$$x_{i(1)} + \dots + x_{i(r)} \leq 2$$

can be added, as a cut, to the linear program  $P'$ .

### 3 The gap phenomenon

In 1955 J. Mycielski [6] has proved the next result.

**Theorem 3** *For each positive integer  $k$  there is a graph  $G$  such that  $\chi(G) = k$  and  $G$  does not contain any 3-clique.*

Since  $G$  does not contain any triangle, the edges of  $G$  have a proper coloring with exactly 1 color. On the other hand, as  $\chi(G) = k$  the nodes of this graph do not have any legal coloring with  $k + 1$  colors. In other words, the upper estimates for the clique size  $\omega(G)$  provided by the coloring of edges of  $G$  coincides with  $\omega(G)$  and the upper bound of the clique size provided by the coloring of the nodes can be arbitrarily large. This makes the edge coloring looking very good in comparison with the node coloring.

In this section we will construct a family of graphs for which the gap between the clique number and the edge chromatic number can be arbitrarily large.

**Theorem 4** *Let us choose an integer  $k$  with  $k \geq 3$ . There is a graph  $L^{(k)}$  such that  $\omega(L^{(k)}) \leq 4$  and  $\chi_{(e)}(L^{(k)}) \geq k$ .*

**Proof.** Let  $M^{(k)}$  be the Mycielski graph with parameter  $k$ . Let  $u_1, \dots, u_n$  be the nodes of  $M^{(k)}$ . The graph  $L^{(k)}$  will have  $2n$  nodes  $x_1, \dots, x_n, y_1, \dots, y_n$ . The unordered pair  $\{x_i, y_i\}$  is an edge of  $L^{(k)}$  for each  $i$ ,  $1 \leq i \leq n$ . Further if the unordered pair  $\{u_i, u_j\}$  is an edge of  $M^{(k)}$ , then we add the edges

$$\{x_i, x_j\}, \{x_i, y_j\}, \{y_i, x_j\}, \{y_i, y_j\}$$

to  $L^{(k)}$ . Figure 1 shows the construction in the special case  $k = 3$ .

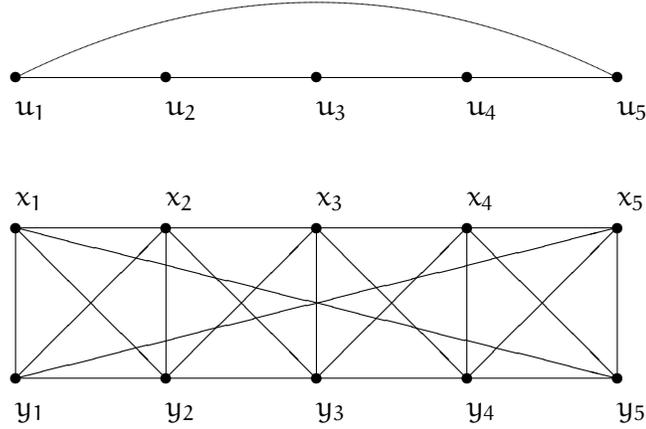


Figure 1: The construction in the proof of Theorem 4 when  $k = 3$ .

We claim that  $\omega(L^{(k)}) \leq 4$ .

In order to prove this claim let us assume the contrary that  $\omega(L^{(k)}) \geq 5$ . Let  $\Delta$  be a 5-clique in  $L^{(k)}$ . Note that a set  $\{x_i, y_i\}$  may contain at most 2 nodes of the clique  $\Delta$ . It follows that there must be at least 3 distinct values of  $i$  such that the set  $\{x_i, y_i\}$  contains at least one node of the clique  $\Delta$ . For the sake of definiteness let us suppose that each of the sets

$$\{x_\alpha, y_\alpha\}, \{x_\beta, y_\beta\}, \{x_\gamma, y_\gamma\}$$

contains a node of the clique  $\Delta$  and  $z_\alpha, z_\beta, z_\gamma$  are the nodes of the clique  $\Delta$  for which

$$z_\alpha \in \{x_\alpha, y_\alpha\}, \quad z_\beta \in \{x_\beta, y_\beta\}, \quad z_\gamma \in \{x_\gamma, y_\gamma\}.$$

Here  $\alpha, \beta, \gamma$  are pair-wise distinct elements of the set  $\{1, \dots, n\}$ .

The unordered pair  $\{z_\alpha, z_\beta\}$  can be an edge of the graph  $L^{(k)}$  only if the unordered pair  $\{u_\alpha, u_\beta\}$  is an edge of the graph  $M^{(k)}$ . From this observation it follows that the nodes  $u_\alpha, u_\beta, u_\gamma$  are the nodes of a 3-clique in  $M^{(k)}$ . But the Mycielski graph  $M^{(k)}$  does not contain any 3-clique since  $\omega(M^{(k)}) \leq 2$ .

This contradiction completes the proof of the claim.

Next we claim that  $\chi_{(e)}(L^{(k)}) \geq k$ .

In order to prove the claim let us assume on the contrary that  $\chi_{(e)}(L^{(k)}) \leq k-1$ . Let  $E$  be the set of edges of  $L^{(k)}$ . Further let  $f : E \rightarrow \{1, \dots, k-1\}$  be the

n	V	E	node	edge	estimate
3	27	189	6	10	5
4	64	1296	12	37	9
5	125	5500	20	113	15
6	216	17550	30	273	23
7	343	46305	42	565	34
8	512	106624	56	1063	46
9	729	221616	72	1807	60
10	1000	425250	90	2922	76
11	1331	765325	110	4477	95
12	1728	1306800	132	6602	115
13	2197	2135484	156	9390	137
14	2744	3362086	182	12998	161
15	3375	5126625	210	17600	188

Table 1: Monotonic matrices. The 2nd and 3rd columns contain the number of nodes and edges of the graphs. The estimates of the clique size are in the 4th and 6th columns.

map which defines a legal coloring of the edges of  $L^{(k)}$  using at most  $k-1$  colors. Guided by the map  $f$  we construct a coloring of the nodes of the graph  $M^{(k)}$ . Let us set  $U = \{u_1, \dots, u_n\}$  and let us consider the map  $h : U \rightarrow \{1, \dots, k-1\}$  defined by  $h(u_i) = f(\{x_i, y_i\})$  for each  $i$ ,  $1 \leq i \leq n$ .

At this point we should observe that the map  $h$  defines a legal coloring of the nodes of the graph  $M^{(k)}$ . The only thing which needs verification is that if  $u_i$  and  $u_j$  are distinct adjacent nodes of the graph  $M^{(k)}$ , then the inequality  $h(u_i) \neq h(u_j)$  must hold.

Since  $u_i$  and  $u_j$  are adjacent nodes in the graph  $M^{(k)}$ , the nodes  $x_i, y_i, x_j, y_j$  are the nodes of a 4-clique in the graph  $L^{(k)}$ . As the map  $f$  defines a legal coloring of the edges of the graph  $L^{(k)}$ , it follows that  $f(\{x_i, y_i\}) \neq f(\{x_j, y_j\})$ . Using

$$h(u_i) = f(\{x_i, y_i\}) \quad \text{and} \quad h(u_j) = f(\{x_j, y_j\})$$

we get  $h(u_i) \neq h(u_j)$ , as required.

Therefore the map  $h$  describes a legal coloring of the nodes of the graph  $M^{(k)}$ . In this coloring at most  $k-1$  colors occur. But this is not possible since  $\chi(M^{(k)}) \geq k$ .

This contradiction completes the proof of the claim.  $\square$

n	V	E	node	edge	estimate
3	8	9	2	1	2
4	16	57	4	6	4
5	32	305	8	17	6
6	64	1473	14	60	11
7	128	6657	26	221	21
8	256	28801	50	875	42
9	512	121089	101	3406	83
10	1024	499713	199	13081	162
11	2048	2037761	395	49268	314
12	4096	8247297	782	186246	610

Table 2: Deletion error detecting codes. The 2nd and 3rd columns contain the number of nodes and edges of the graphs. The estimates of the clique size are in the 4th and 6th columns.

n	V	E	node	edge	estimate
6	15	45	4	3	3
7	35	385	10	23	7
8	70	1855	20	107	15
9	126	6615	35	391	28
10	210	19425	56	1131	48
11	330	49665	84	2754	74
12	495	114345	120	5918	109
13	715	242385	165	11610	152
14	1001	480480	220	21172	206
15	1365	900900	286	36514	270
16	1820	1611610	364	60054	347
17	2380	2769130	455	95038	436
18	3060	4594590	560	145441	539

Table 3: Johnson codes. The 2nd and 3rd columns contain the number of nodes and edges of the graphs. The estimates of the clique size are in the 4th and 6th columns.

## 4 A complexity result

In this section we will be interested in the computational complexity of the following problem.

**Problem 5** *Given a finite simple graph  $G$  and given a positive integer  $k$ . Decide if the edges of  $G$  admit a legal coloring using  $k$  colors.*

We will show that for  $k \geq 6$  Problem 5 is NP hard. The intuitive meaning of this result is that finding the optimal number of colors in the edge coloring of a given graph is computationally hard. Thus in practical computations we should approximate the edge chromatic number of a graph instead of exactly computing it.

We will show that Problem 5 can be reduced to the following problem which is known to be NP-complete. (See [7].)

**Problem 6** *Given a finite simple graph  $G$  and given a positive integer  $k$ . Decide if the nodes of  $G$  have a legal coloring using  $k$  colors.*

The reduction of Problem 5 to Problem 6 can be accomplished using an algorithm which runs in polynomial time and uses polynomial size memory.

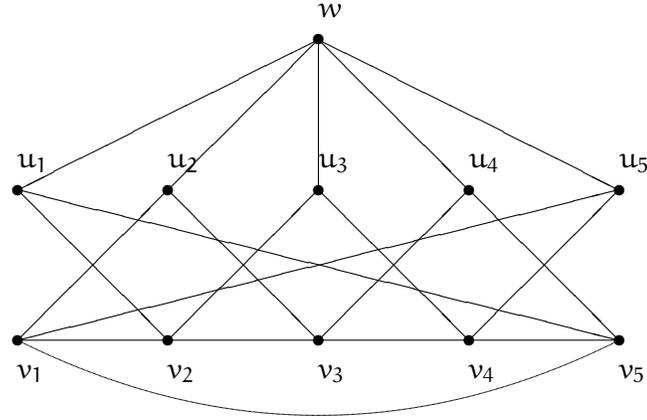
Let  $M^{(k)}$  be the Mycielski graph with parameter  $k$  and let  $e = \{x, y\}$  be an edge of  $M^{(k)}$ . We delete the edge  $e$  from  $M^{(k)}$  but we do not delete any of the end points of the edge  $e$ . We denote the resulting graph by  $N^{(k)}$ .

**Lemma 7** *For the graph  $N^{(k)}$  defined above the following holds.*

- (1) *The nodes of the graph  $N^{(k)}$  can be colored legally using  $k - 1$  colors.*
- (2) *In each legal coloring of the nodes of the graph  $N^{(k)}$  using  $k - 1$  colors the nodes  $x$  and  $y$  must receive the same color.*

**Proof.** The Mycielski graph  $M^{(3)}$  is a circle consisting of 5 nodes and 5 edges. After deleting an edge from  $M^{(3)}$  the nodes of the remaining graph can be colored legally with 2 colors. The end points of the deleted edge must receive the same colors since otherwise one puts back the deleted edge and the nodes of the graph  $M^{(3)}$  could be legally colored with 2 colors. Thus the special case  $k = 3$  is settled.

The Mycielski graph  $M^{(4)}$  has 11 nodes  $v_1, \dots, v_5, u_1, \dots, u_5, w$ . The set of nodes  $\{v_1, \dots, v_5\}$  induces a subgraph  $H$  in  $M^{(4)}$  such that  $H$  is isomorphic to  $M^{(3)}$ . Figures 2 and 3 illustrate this step of the proof.

Figure 2: The Mycielski graph  $M^{(4)}$ .

The node  $u_i$  is adjacent to the neighbors of  $v_i$  for each  $i$ ,  $1 \leq i \leq 5$ . The node  $w$  is adjacent to  $u_i$  for each  $i$ ,  $1 \leq i \leq 5$ . After deleting the edge  $\{v_1, v_5\}$  from  $H$  the nodes of the resulting graph  $H'$  can be legally colored with 3 colors. To exhibit a legal coloring of the nodes let us color the nodes  $u_1, \dots, u_5$  with colors 1. The nodes  $v_1, \dots, v_5$  can be colored with the colors 2, 3. Finally, we color the node  $w$  with color 3. We have ended up with a legal coloring of the nodes of the graph  $N^{(4)}$  using 3 colors. This settles the case  $k = 4$ .

The Mycielski graph  $M^{(5)}$  has 23 nodes  $v_1, \dots, v_{11}, u_1, \dots, u_{11}, w$ . The set of nodes  $\{v_1, \dots, v_{11}\}$  induces a subgraph  $H$  in  $M^{(5)}$  such that  $H$  is isomorphic to  $M^{(4)}$ . The node  $u_i$  is connected by an edge to each of the neighbors of the node  $v_i$  for each  $i$ ,  $1 \leq i \leq 11$ . Finally, we connect the node  $w$  to the node  $u_i$  by an edge for each  $i$ ,  $1 \leq i \leq 11$ . The graph  $H$  has an edge  $e$  such that after deleting  $e$  from  $H$  the nodes of the resulting graph  $H'$  can be colored legally with 3 colors.

We assign color 1 to node  $u_i$  for each  $i$ ,  $1 \leq i \leq 11$ . The nodes  $v_1, \dots, v_{11}$  can be colored legally with 3 colors. We will use the colors 2, 3, 4. Finally, we assign color 4 to node  $w$ . This provides a legal coloring of the nodes of the graph  $N^{(5)}$  using 4 colors. Therefore the case  $k = 5$  has been settled.

After working out the  $k = 5$  special case we are well prepared to settle the general case using an induction on  $k$ .

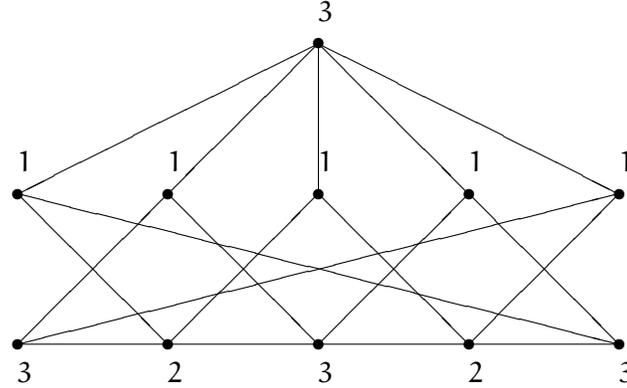


Figure 3: The graph  $N^{(4)}$  with colored nodes.

Set  $V = \{v_1, \dots, v_n\}$  and  $U = \{u_1, \dots, u_n\}$ . Let  $U \cup V \cup \{w\}$  be the set of nodes of the Mycielski graph  $M^{(k)}$ . The set of nodes  $V$  induces the subgraph  $H$  which is isomorphic to  $M^{(k-1)}$ . Note that  $N(w) = U$  and  $N(u_i) \cap V = N(v_i) \cap V$  for each  $i$ ,  $1 \leq i \leq n$ .

By the inductive assumption  $H$  has an edge  $e$  such that after deleting  $e$  from  $H$  the nodes of the resulting graph  $H'$  can be colored legally using  $k-2$  colors.

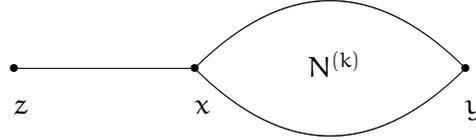
We color each node in  $U$  with color 1. The nodes in  $V$  can be colored legally using the colors  $2, \dots, k-1$ . The node  $w$  can be colored with color  $k-1$ . In this way we ended up with a legal coloring of the nodes of  $N^k$  using  $k-1$  colors.

Putting back the edge  $e$  gives back the graph  $M^{(k)}$ . The nodes of  $M^{(k)}$  can be colored legally with  $k$  colors but not with  $k-1$  colors. Consequently, the end points of the edge  $e$  must receive the same colors in any legal coloring of the nodes of  $N^k$  using  $k-1$  colors.

This completes the proof.  $\square$

Using the graph  $N^{(k)}$  we construct a new graph  $L^{(k)}$  by adding a new edge  $\{z, x\}$  to  $N^{(k)}$ . Figure 4 depicts the graph  $L^{(k)}$ . The newly constructed graph  $L^{(k)}$  clearly has the following properties.

- (1) The nodes of the graph  $L^{(k)}$  can be colored legally using  $k-1$  colors.
- (2) In each legal coloring of the nodes of the graph  $L^{(k)}$  using  $k-1$  colors the nodes  $z$  and  $y$  must receive distinct colors.

Figure 4: The graph  $L^{(k)}$ .

- (3) If the nodes  $z$  and  $y$  are colored with distinct colors, then this partial coloring of the nodes of  $L^{(k)}$  can be extended to a legal coloring of the nodes of  $L^{(k)}$ .

In order avoid notational difficulties first we will deal with the  $k$  coloring problem in the special case when  $k = 6$  and so we will use the graph  $L^{(7)}$  as an auxiliary graph. The reader can verify that the graph  $L^{(7)}$  has 95 nodes and 640 edges.

Let  $G = (V, E)$  be a finite simple graph and let  $u_1, \dots, u_n$  be the edges of  $G$ . From the given graph  $G$  we construct a new graph  $G'$ .

Let  $L_{i,j}$  be an isomorphic copy of the graph  $L^{(7)}$ . We assume that  $w_{i,j,1}, \dots, w_{i,j,95}$  are all the nodes of the graph  $L_{i,j}$  and the node  $w_{i,j,1}$  of  $L_{i,j}$  corresponds to the node  $z$  of the graph  $L^{(7)}$ . Further we assume that the node  $w_{i,j,95}$  of  $L_{i,j}$  corresponds to the node  $y$  of the graph  $L^{(7)}$ .

From the given graph  $G = (U, E)$  we construct a new graph  $G'$ . Let the unordered pair  $\{u_i, u_j\}$  be an edge of the graph  $G$ . We replace this edge of  $G$  by  $L_{i,j}$ . We identify the node  $u_i$  of  $G$  with the node  $w_{i,j,1}$  of the graph  $L_{i,j}$  and we identify the node  $u_j$  of  $G$  with the node  $w_{i,j,95}$  of the graph  $L_{i,j}$ . In the graph  $L^{(k)}$  the roles of the nodes  $y$  and  $z$  are not symmetric. In order to avoid ambiguity in the construction we assume that for the edge  $\{u_i, u_j\}$  the condition  $i < j$  holds.

From the graph  $G'$  we construct a new graph  $\Gamma$ . Let us suppose that  $w_1, \dots, w_m$  are all the nodes of  $G'$ . The graph  $G'$  has  $m$  nodes and the graph  $\Gamma$  will have  $2m$  nodes  $x_1, \dots, x_m, y_1, \dots, y_m$ . To the node  $w_i$  of  $G'$  we assign an edge  $\{x_i, y_i\}$  of the graph  $\Gamma$  for each  $i$ ,  $1 \leq i \leq m$ . Next if the unordered pair  $\{w_i, w_j\}$  is an edge of  $G'$ , then we add the edges

$$\{x_i, x_j\}, \{x_i, y_j\}, \{y_i, x_j\}, \{y_i, y_j\}$$

to  $\Gamma$ .

The pivotal result of our consideration in pursuing the  $k = 6$  particular case is summarized in the following assertions.

**Theorem 8** *For the graph  $\Gamma$  defined above the following holds.*

- (1) *If the edges of the graph  $\Gamma$  have a legal coloring using 6 colors, then the nodes of the graph  $G$  have a legal coloring using 6 colors.*
- (2) *If the nodes of the graph  $G$  admit a legal coloring with 6 colors, then the edges of the graph  $\Gamma$  admit a legal coloring with 6 colors.*

**Proof.** In order to prove claim (1) let us assume that the edges of the graph  $\Gamma = (W, F)$  have a legal coloring using 6 colors. Let us suppose that the map  $f : F \rightarrow \{1, \dots, 6\}$  describes this coloring. Using this coloring of the edges of the graph  $\Gamma$  we construct a coloring of the nodes of the graph  $G$ . To the node  $u_i$  of  $G$  we assign the colors of the edge  $\{x_i, y_i\}$  of  $\Gamma$ . In other words we define a map  $g : U \rightarrow \{1, \dots, 6\}$  by setting  $g(u_i)$  to be equal to  $f(\{x_i, y_i\})$ .

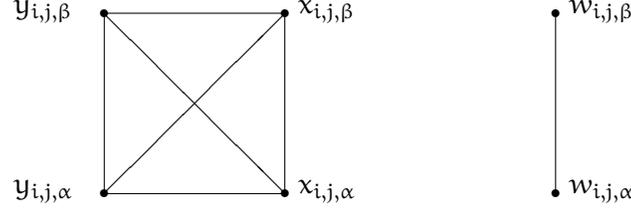
The map  $g$  describes a coloring of the nodes of the graph  $G$  using 6 colors. We claim that  $g$  describes a legal coloring of the nodes of the graph  $G$ . In order to verify the claim it is sufficient to show that if the unordered pair  $\{u_i, u_j\}$  is an edge of the graph  $G$ , then  $g(u_i) \neq g(u_j)$  must hold. The node  $u_i$  of  $G$  corresponds to the node  $w_{i,j,1}$  of the graph  $G'$  and the node  $u_j$  of  $G$  corresponds to the node  $w_{i,j,95}$  of the graph  $G'$ . The node  $w_{i,j,1}$  corresponds to the edge  $\{x_{i,j,1}, y_{i,j,1}\}$  of the graph  $\Gamma$ . Similarly, the node  $w_{i,j,95}$  corresponds to the edge  $\{x_{i,j,95}, y_{i,j,95}\}$  of the graph  $\Gamma$ .

We know that the edges  $\{x_{i,j,1}, y_{i,j,1}\}$  and  $\{x_{i,j,95}, y_{i,j,95}\}$  receive distinct colors, that is,  $f(\{x_{i,j,1}, y_{i,j,1}\}) \neq f(\{x_{i,j,95}, y_{i,j,95}\})$ . Consequently  $g(u_i) \neq g(u_j)$ , as required.

In order to prove assertion (2) let us assume that the nodes of  $G$  have a legal coloring with 6 colors. We assume that the map  $g : U \rightarrow \{1, \dots, 6\}$  describes this coloring. Suppose that the unordered pair  $\{u_i, u_j\}$  is an edge of the graph  $G$  and  $i < j$  holds.

When we constructed the graph  $G'$  from the graph  $G$  we replaced the edge  $\{u_i, u_j\}$  of  $G$  by  $L_{i,j}$  which is an isomorphic copy of the graph  $L^{(7)}$ . Let  $x_{i,j}$ ,  $y_{i,j}$ ,  $z_{i,j}$  be the nodes of the graph  $L_{i,j}$  that correspond to the nodes  $x$ ,  $y$ ,  $z$  of the graph  $L^{(7)}$  at the isomorphism.

The node  $z_{i,j}$  of the graph  $G'$  gets the color of the edge  $u_i$  of the graph  $G$ . The node  $y_{i,j}$  of the graph  $G'$  gets the color of the edge  $u_j$  of the graph  $G$ . The node  $x_{i,j}$  of the graph  $G'$  gets the color of the edge  $x$  of the graph  $G$ . We know that this partial coloring of the nodes of the graph  $L_{i,j}$  can be extended to the coloring of all nodes of the graph  $L_{i,j}$ .

Figure 5: The 4-clique  $\Delta$  in the graph  $L_{i,j}$ .

For the sake of a convenient notation we rename the nodes of the graph  $L_{i,j}$ . Let  $w_{i,j,1}, \dots, w_{i,j,95}$  be the nodes of the graph  $L_{i,j}$ . We will assume that

$$w_{i,j,1} = z_{i,j}, \quad w_{i,j,2} = x_{i,j}, \quad w_{i,j,95} = y_{i,j}.$$

When we constructed the graph  $\Gamma$  from the graph  $G'$  we have assigned an edge  $\{x_{i,j,\alpha}, y_{i,j,\alpha}\}$  of  $\Gamma$  to the node  $w_{i,j,\alpha}$  of  $G'$  for each  $i, j, \alpha, 1 \leq i < j \leq m, 0 \leq \alpha \leq 95$ . Now we assign the color of the node  $w_{i,j,\alpha}$  of the graph  $G'$  to the edge  $\{x_{i,j,\alpha}, y_{i,j,\alpha}\}$  of the graph  $\Gamma$ .

When the nodes  $w_{i,j,\alpha}$  and  $w_{i,j,\beta}$  were adjacent in the graph  $G'$ , then we added the edges

$$\{x_{i,j,\alpha}, x_{i,j,\beta}\}, \{x_{i,j,\alpha}, y_{i,j,\beta}\}, \{y_{i,j,\alpha}, x_{i,j,\beta}\}, \{y_{i,j,\alpha}, y_{i,j,\beta}\} \quad (2)$$

to the graph  $\Gamma$  during the construction of the graph  $\Gamma$  from the graph  $G'$ . The situation is shown by Figure 5. The edge  $\{x_{i,j,\alpha}, y_{i,j,\alpha}\}$  of the graph  $\Gamma$  receives the color of the node  $w_{i,j,\alpha}$  of the graph  $G'$  and the edge  $\{x_{i,j,\beta}, y_{i,j,\beta}\}$  of the graph  $\Gamma$  receives the color of the node  $w_{i,j,\beta}$  of the graph  $G'$ .

The nodes of  $G'$  are colored using 6 colors. We intend to color the edges of  $\Gamma$  using these 6 colors. Out of the 6 colors we intend to use for the coloring the edges of the graph  $\Gamma$  4 colors are still available to color the edges (2). In this way we get a legal coloring of the edges of the graph  $\Gamma$ .  $\square$

The main result of this section is the following theorem.

**Theorem 9** *Problem 5 is NP hard for each integer  $k \geq 6$ .*

**Proof.** Theorem 8 settles the special case  $k = 6$ . For the remaining part of the proof we assume that  $k \geq 7$ . We start with the graph  $N^{(k+1)}$  and follow a reasoning analogous to the proof of Theorem 8.

A few elementary estimates are still missing to complete the argument. Let  $G = (V, E)$ ,  $G' = (V', E')$ ,  $\Gamma = (W, F)$ . It is clear that there is a positive constant  $c_1$  such that  $|V'| \leq c_1|V|$ . One can choose  $c_1$  to be the number of nodes of the auxiliary graph  $L^{(k+1)}$ . There is a positive constant  $c_2$  for which  $|E'| \leq c_2|V|^2$ . Indeed,

$$|E'| \leq (1/2)|V'|^2 \leq (1/2)c_1^2|V|^2 = c_2|V|^2.$$

The computation

$$|W| \leq 2|V'| = 2c_1|V| = c_3|V|$$

shows that there is a positive constant  $c_3$  such that  $|W| \leq c_3|V|$ . Finally, there is a positive constant  $c_4$  with the property that  $|F| \leq c_4|V|^2$ . This can be seen from

$$|F| \leq 6|E'| \leq 6c_2|V|^2 = c_4|V|^2.$$

The essential point is that the quantities  $|E|$ ,  $|W|$ ,  $|F|$  can be over estimated by a polynomial in terms of  $|V|$ . As we can see the degree of this polynomial is two. The leading coefficient can be very large. As a matter of fact it is an exponential function of  $k$ . But for each fixed  $k$  the leading coefficient is a constant.  $\square$

## Acknowledgements

We would like to express our thanks for the anonymous referee.

## References

- [1] I. M. [Bomze](#), M. Budinich, P. M. [Pardalos](#), M. Pelillo, The Maximum Clique Problem, Handbook of Combinatorial Optimization Vol. 4, [Kluwer Academic Publisher](#), 1999. [⇒64](#)
- [2] R. Carraghan, P. M. [Pardalos](#), An exact algorithm for the maximum clique problem, *Operation Research Letters* **9** (1990), 375–382. [⇒67](#)
- [3] J. Hasselberg, P. M. [Pardalos](#), and G. Vairaktarakis, Test case generators and computational results for the maximum clique problem, *Journal of Global Optimization* **3** (1993), 463–482. [⇒64](#)

- 
- [4] D. [Kumlander](#), *Some Practical Algorithms to Solve the Maximal Clique problem* PhD. Thesis, Tallin University of Technology, 2005. [⇒64](#)
  - [5] C. Morgan, *A Combinatorial Search with Dancing Links*, PhD. Thesis, Univ. of Warwick, 1999–2000. [⇒64](#)
  - [6] J. [Mycielski](#), Sur le coloriage des graphes, *Colloq. Math.* **3** (1955), 161–162. [⇒70](#)
  - [7] C. H. [Papadimitriou](#), *Computational Complexity*, Addison-Wesley Publishing Company, Inc., 1994. [⇒64, 74](#)
  - [8] S. Szabó, Parallel algorithms for finding cliques in a graph, *Journal of Physics: Conference Series* **268** (2011) 012030 [⇒63, 65](#)

*Received: February 16, 2016 • Revised: April 17, 2016*