

# National Research University Higher School of Economics Laboratory of Algorithms and Technologies for Networks Analysis Nizhny Novgorod

# "A fast greedy sequential heuristic for the vertex coloring problem"

Workshop on "Clustering and Search techniques in large scale networks"

November 4 – 7, 2014 Nizhny Novgorod

> Larisa Komosko, Mikhail Batsyn

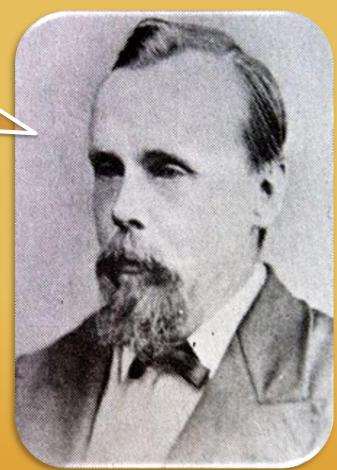
### History: four color problem 1852

no more than four colors are required to color the regions of the map so that no two adjacent regions have the same color



proof: 1976

- Kenneth Appel
- Wolfgang Haken

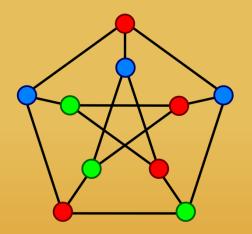


**Thomas Gutherie** 

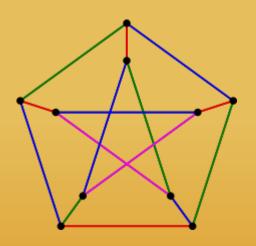
# Coloring methods



vertex coloring



edge coloring



#### Vertex coloring problem:

finding the minimum number of colors to paint the vertices of a graph in such a way that any two vertices joined by an edge always have different colors

#### NP-hard: 1979

Garey, Johnson

#### Applications:

- timetabling
- register allocation
- traffic lights phasing
- map coloring
- storage and transportation of goods
- computing upper bounds in branch-and-bound algorithms for the maximum clique problem

#### Mathematical model:

$$\mathbf{x}_{ih} = \begin{cases} 1, & \text{if color } h \text{ is assigned to vertex } i \\ 0, & \text{otherwise} \end{cases}$$

$$y_h = \begin{cases} 1, & \text{if color } h \text{ is used in the solution} \\ 0, & \text{otherwise} \end{cases}$$

#### Model:

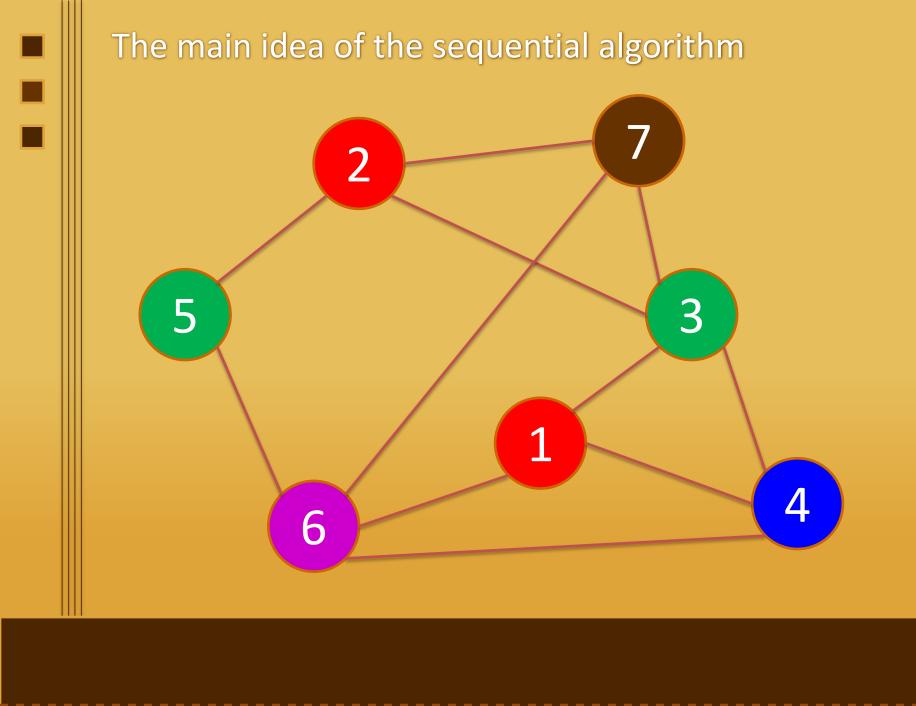
$$\sum_{h=1}^{n} y_h \rightarrow \min$$

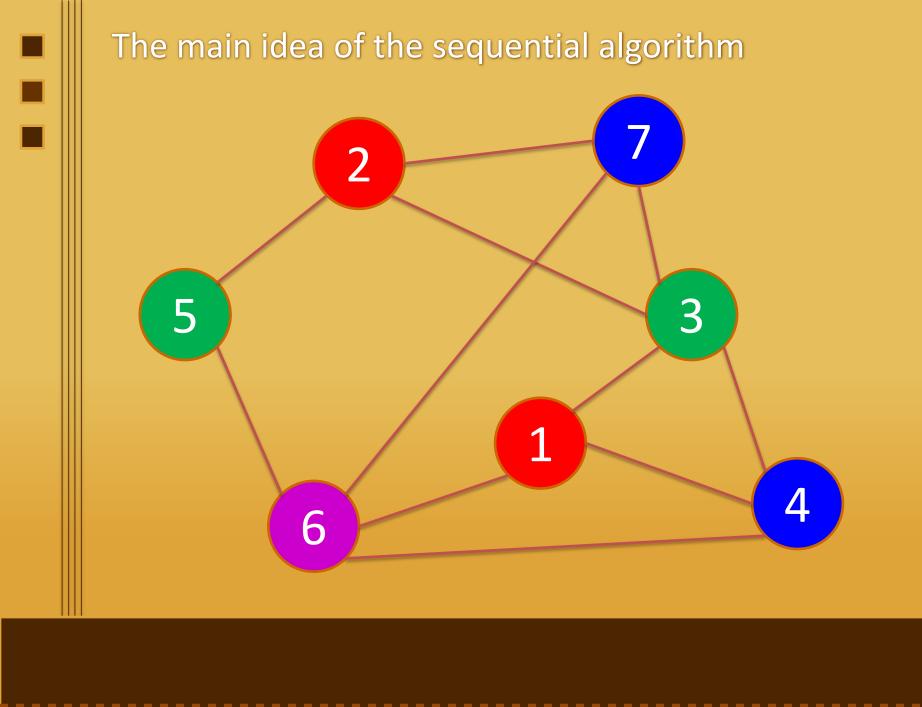
$$\sum_{h=1}^{n} x_{ih} = 1, i \in V$$

$$x_{ih} + x_{jh} \le y_h$$
,  $(i, j) \in E, h = 1, ..., n$ 

$$x_{ih} \in \{0, 1\}, i \in V, h = 1, ..., n$$

$$y_h \in \{0, 1\}, h = 1, ..., n$$





#### Classical coloring algorithms

- Greedy-Color
- Color-with-Interchange
- Random-Sequential-Color
- Largest-First-Color
- Smallest-Last-Color
- Random-Sequential-Interchange-Color

- Largest-First-Interchange-Color
- Smallest-Last-Interchange-Color
- Connected-Sequential-Color
- Saturation-Color
- Greedy Independent Sets-Color

#### Pseudo code

#### **Procedure** Coloring(A,C,n)

```
begin
maxColor = 0;
for vertex=0 to n
     for k=0 to n
            if c_{k,vertex} = 0 then
                  color = k;
                  break;
            end if
            if color > maxColor then
                  recolour(A,C,n,vertex,color);
                  maxColor = color;
            end if
     end for
C_{color} := C_{color} \lor A_{vertex};
end for
end
```

#### adjacency matrix

#### forbidden colors matrix

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 3 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 5 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 6 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 7 \end{bmatrix}$$

$$a_{ij} = \begin{cases} 1, & \text{if vertices } i \text{ and } j \text{ are} \\ & \text{connected with an edge} \\ 0, & \text{otherwise} \end{cases}$$

0, otherwise

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 5 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 6 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 7 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 3 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 \end{bmatrix}$$

# The main idea of the developed algorithm

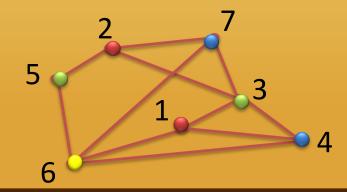
# The main idea of the developed algorithm

The main idea of the developed algorithm

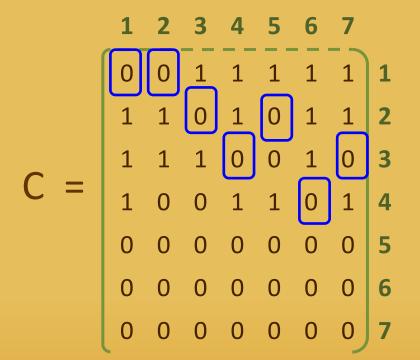
# The result

	1	2	3	4	5	6	7	
	0	0	1	1	1	1	1	1
	1	1	0	1	0	1	1	2
_	1	1	1	0	0	1	0	3
C =	1	0	0	1	1	0	1	4
	0	0	0	0	0	0	0	5
	0	0	0	0	0	0	0	6
	0	0	0	0	0	0	0	7

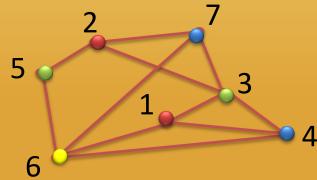
vertex	color
1	1
2	1
3	2
4	3
5	2
6	4
7	3



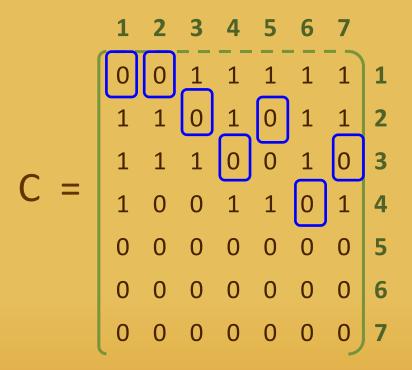
### The result



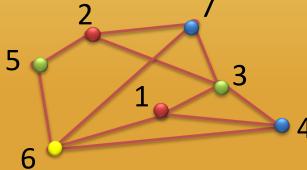
vertex	color
1	1
2	1
3	2
4	3
5	2
6	4
7	3



The result



vertex	color
1	1
2	1
3	2
4	3
5	2
6	4
7	3



COLOURS:  $1 - \{1,2\}$ ;  $2 - \{3,5\}$ ;  $3 - \{4,7\}$ ;  $4 - \{6\}$ 

Experimental results: speed

### time, ms (for 10000 runs)

G(V,E)	V	E	Bit-Greedy-Color	<b>Greedy-Color</b>	speedup
brock800_4	800	207643	3869	22043	6
c-fat500-10	500	46627	1835	12333	7
C1000.9	1000	450079	6563	37622	6
dsjc500.5.col.txt	500	62624	1260	10163	8
frb30-15-5	450	83231	871	7851	9
gen400_p0.9_55	400	71820	1353	8364	6
hamming10-4	1024	434176	5423	54761	10
johnson32-2-4	496	107880	815	4501	6
keller5	776	225990	2083	21203	10
MANN_a9	45	918	36	804	22
p_hat1000-3	1000	371746	6550	25553	4
san1000	1000	250500	3315	20426	6
sanr400_0.7	400	55869	1406	5335	4

### Advantages of the developed algorithm:

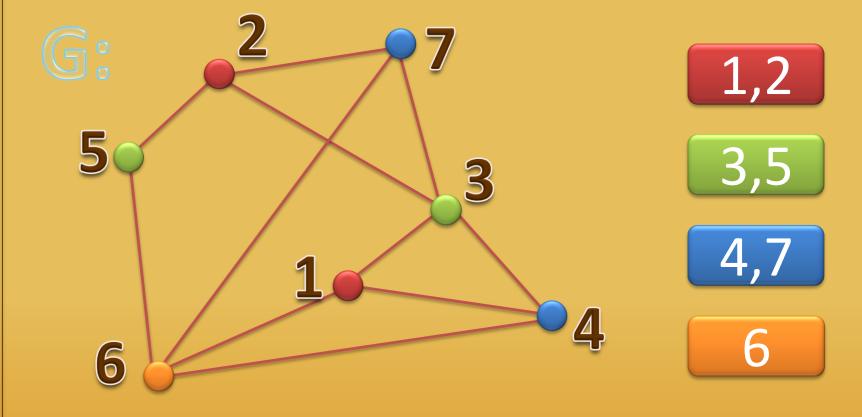
- the usage of the bit representations of the adjacency and forbidden colors matrices;
- bitwise disjunction operation on the rows of the adjacency and forbidden colors matrixes.

average acceleration- **6.8 times** 

### Pseudo code

**Procedure** Recoloring(A,C,n,vertex,color)

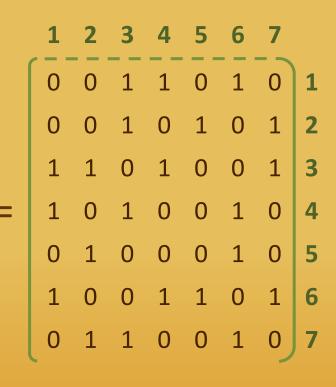
```
begin
for i=0 to color
     for c=0 to n
          neighbors = count_number_of_neighbors(A, vertex);
          if neighbors >= 2 then
               break;
          end if
          if neighbors =1 then
               for g=0 to color
                    value=find _neighbor_to_exchange_colors_with(c,C,g);
               end for
          end if
     end for
     C_{color} := C_{color} \vee A_{i};
end for
end
```



try to recolor 6-th vertex – it is only one in the color class



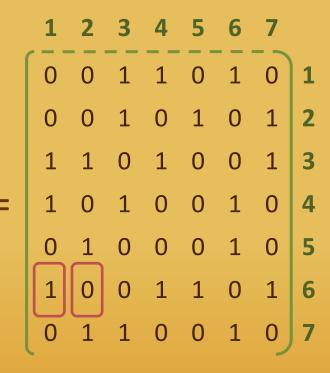
count the number of neighbors of the 6<sup>th</sup> vertex in each color class





### Step 1:

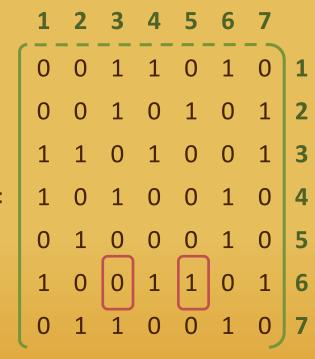
count the number of neighbors of the 6<sup>th</sup> vertex in each color class





### Step 1:

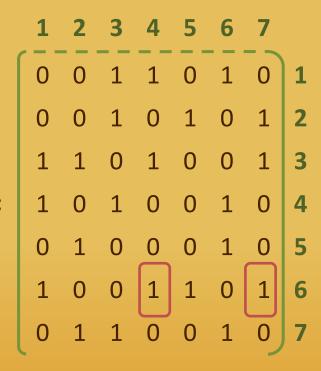
 count the number of neighbors of the 6<sup>th</sup> vertex in each color class





### Step 1:

 count the number of neighbors of the 6<sup>th</sup> vertex in each color class



### Step 2:

consider only color classes with only one neighbor **if 0** – the vertex would be in that color class

if more than 1 – the elements in the color class wouldn't be independent



### Step 2:

consider onlycolor classes withonly one neighbor



**if 0** – the vertex would be in that color class

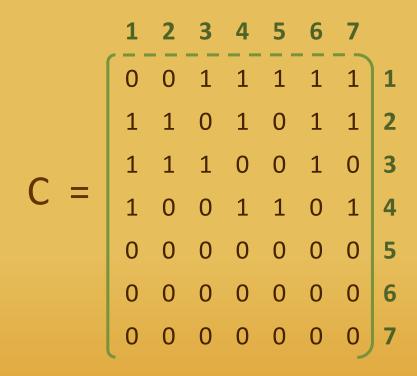
if more than 1 – the elements in the color class wouldn't be independent

### Step 3:

take the vertices
 from that classes
 which are the
 neighbors of the 6<sup>th</sup>

### Step 4:

for vertices 1 and 5 check if they can be recolored



### Step 4:

for vertices 1 and 5 check if they can be recolored

don't take into account the last color class containing the 6<sup>th</sup> vertex

	1	2	3	4	5	6	7	
	0	0	1	1	1	1	1	1
	1	1	0	1	0	1	1	2
	1	1	1	0	0	1	0	3
=.	1	0	0	1	1	0	1	4
	0	0	0	0	0	0	0	5
	Û	Û	Û	Û	Û	Û	Û	Ô
	0	0	0	0	0	0	0	7

### Step 4:

for vertices 1 and 5check if they can berecolored

$$C = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 3 \end{bmatrix}$$

5<sup>th</sup> vertex can be recolored

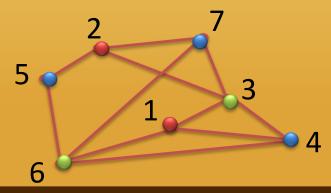
### Step 5:

- New color for the 5<sup>th</sup> vertex is № 3:
  - move the 5<sup>th</sup> vertex to the 3<sup>rd</sup> color class
  - assign the 6<sup>th</sup> vertex color Nº2 – former color of the 5<sup>th</sup> vertex









Experimental results: speed

### time, ms (for 1000 runs)

G(V,E)	V	E	GC + recolor	BGC + recolor	speedup
brock200_4	200	13089	1415	383	3,7
c-fat500-1	500	4459	4634	298	15,6
C500.9	500	112332	4023	3064	1,3
frb30-15-2	450	83151	4039	221	18,3
gen400_p0.9_65	400	71820	6388	1562	4,1
hamming6-4	64	704	464	10	46,4
johnson8-2-4	28	210	145	4	36,3
MANN_a9	45	918	336	12	28,0
dsjc500.1.col.txt	500	12458	3036	428	7,1
p_hat500-3	500	93800	7267	3047	2,4
san200_0.9_3	200	17910	1693	226	7,5
sanr400_0.5	400	39984	4211	841	5,0

average acceleration- 5 times

Experimental results: quality

Lxpermenta		4163. G	<u>aant</u>	<b>Y</b>			
G(V,E)	V	E	χ(G)	GC	BGC	GC + recolor	BGC + recolor
jean.col	80	254	10	10	10	10	10
queen13_13.col	169	6656	13	21	21	20	20
queen15_15.col	225	10360	15	25	25	23	23
myciel3.col	11	20	4	4	4	4	4
myciel7.col	191	2360	8	8	8	8	8
mulsol.i.3.col	184	3916	31	31	31	31	31
mulsol.i.5.col	185	3973	31	31	31	31	31
zeroin.i.2.col	211	3541	30	30	30	30	30
zeroin.i.3.col	206	3540	30	30	30	30	30
anna.col	138	493	11	12	12	11	11
david.col	87	406	11	12	12	12	12
fpsol2.i.3.col	425	8688	30	30	30	30	30
games120.col	120	638	9	10	10	9	9
homer.col	561	1629	13	15	15	13	13
huck.col	74	301	11	11	11	11	11

# THANK YOU for your attention!

vertices		500			1000			5000			10000	
density	usual	bool	bit	usual	bool	bit	usual	bool	bit	usual	bool	bit
0.01	1135,94	328,78	15,26	2377,93	1323,83	55,63	17821,60	32674,00	1259,33	16880,00	65455,50	6400,50
0.1	1030,44	332,57	17,24	2231,97	1331,53	62,43	16302,62	32767,60	1353,47	15250,00	65913,00	6660,00
0.2	956,34	333,76	18,80	2148,63	1332,70	68,07	15496,82	32942,40	1468,27			
0.3	709,38	327,72	20,45	1733,77	1343,90	74,47	15660,89	33554,22	1605,93			
0.4	574,52	329,06	22,21	1380,57	1353,17	80,63	14242,80	34632,00	1778,00			
0.5	471,52	330,09	24,09	1181,80	1323,37	88,00	13635,00	35568,00	1986,67			
0.6	399,44	332,60	26,21	1017,13	1333,30	96,60	13614,60	38034,00	2641,33			
0.7	342,56	334,84	28,95	870,33	1337,43	110,07	14111,00	40092,00	3150,67			
0.8	282,64	338,33	32,43	749,53	1349,33	128,70						
0.9	229,68	344,26	37,30	607,10	1369,50	159,33						
0.99	175,82	358,24	51,23	539,80	1447,27	245,83						

vertices	500	1000	5000	10000
	300			10000
density		spee	edup	
0.01	74,43	42,74	14,15	2,64
0.1	59,76	35,75	12,05	2,29
0.2	50,87	31,57	10,55	
0.3	34,69	23,28	9,75	
0.4	25,86	17,12	8,01	
0.5	19,57	13,43	6,86	
0.6	15,24	10,53	5,15	
0.7	11,83	7,91	4,48	
0.8	8,72	5,82		
0.9	6,16	3,81		
0.99	3,43	2,20		

- DSJ: Random graphs used in David Johnson's paper with Aragon, McGeoch, and Schevon
- DSJR are geometric graphs, with being complements of geometric graphs.
- **CUL:** Quasi-random coloring problem.
- ▶ REG: Problem based on register allocation for variables in real codes.
- LEI: Leighton graphs with guaranteed coloring size.
- ▶ <u>SCH:</u>Class scheduling graphs, with and without study halls.
- LAT: Latin square problem.
- ▶ SGB: Graphs from Donald Knuth's Stanford GraphBase. These can be divided into:
  - Book Graphs. Given a work of literature, a graph is created where each node represents a character. Two nodes are connected by an edge if the corresponding characters encounter each other in the book. Knuth creates the graphs for five classic works: Tolstoy's Anna Karenina (anna), Dicken's David Copperfield (david), Homer's Iliad (homer), Twain's Huckleberry Finn (huck), and Hugo's Les Mis\'erables (jean).
  - <u>Game Graphs.</u> A graph representing the games played in a college football season can be represented by a graph where the nodes represent each college team. Two teams are connected by an edge if they played each other during the season. Knuth gives the graph for the 1990 college football season.
  - Miles Graphs. These graphs are similar to geometric graphs in that nodes are placed in space with two nodes connected if they are close enough. These graphs, however, are not random. The nodes represent a set of United States cities and the distance between them is given by by road mileage from 1947. These graphs are also due to Kuth.
  - Oueen Graphs. Given an n by n chessboard, a queen graph is a graph on n^2 nodes, each corresponding to a square of the board. Two nodes are connected by an edge if the corresponding squares are in the same row, column, or diagonal. Unlike some of the other graphs, the coloring problem on this graph has a natural interpretation: Given such a chessboard, is it possible to place n sets of n queens on the board so that no two queens of the same set are in the same row, column, or diagonal? The answer is yes if and only if the graph has coloring number n. Martin Gardner states without proof that this is the case if and only if \$n\$ is not divisible by either 2 or 3. In all cases, the maximum clique in the graph is no more than n, and the coloring value is no less than n.

MYC: Graphs based on the Mycielski transformation. These graphs are difficult to solve because they are triangle free (clique number 2) but the coloring number increases in problem size.

# Coloring methods



vertex coloring



edge coloring

