

Coloring problems

Outline

Why vertex coloring problems?

Overview of the family

A short review of graphs

The vertex coloring problem

Assumptions

Two simple bounds

Counting the solutions, a first solution and the bounds

Remarks

Typical mistakes

In summary

Learning objectives

At the end of this chapter you will be able to:

- ▶ use the basic notions on graphs: neighbourhood, degree, induced graph, complement graph, clique and stable set;
- ▶ state the definition of the vertex coloring problem and the assumptions on the graph;
- ▶ count the candidate solutions of a small instance and recognize that they are partitions;
- ▶ find a coloring by hand and compute the clique bound and the stable set bound;
- ▶ explain why the clique bound is not always tight.

Prerequisites

- ▶ From this module, the chapter «[Combinatorial optimization problems](#)»: instances, feasible solutions, primal and dual bounds, the gap, and the Bell numbers that count the partitions of a set
- ▶ From the chapter «[Packing problems](#)»: the weight bound as an example of a dual bound
- ▶ No knowledge of graph theory is needed: the definitions used here are recalled in a short section

Why vertex coloring problems?

Introduction: Why vertex coloring problems?

- ▶ Coloring problems model **conflicts**: pairs of objects that cannot receive the same resource
- ▶ We start from a small story, the scheduling of exams, and see how it becomes a problem on a graph
- ▶ We then recall a few milestones of the history of the problem
- ▶ We close with a challenge on a famous graph, which the chapter answers

A story: scheduling the exams

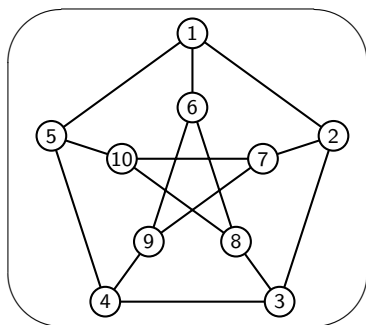
- ▶ Six exams must be scheduled in the exam session
- ▶ Two exams that share at least one student cannot take place on the same day
- ▶ What is the **minimum number of days**?

Exam session	Graph
the exams	the vertices
two exams with students in common	an edge between them
the days	the colors

History

- ▶ **Brooks (1941)** proved that a connected graph that is neither complete nor an odd cycle can be colored with at most Δ colors, where Δ is the largest degree
- ▶ **Karp (1972)** included the chromatic number among his 21 NP-complete problems
- ▶ **Garey, Johnson and Stockmeyer (1976)** proved that deciding whether 3 colors are enough remains NP-complete on planar graphs of maximum degree 4

A challenge



the Petersen graph

Before going on, try to answer two questions on the Petersen graph, the reference instance of the chapter:

1. How many colors are needed?
2. How can we prove that fewer colors are not enough, without listing all the partitions?

The clique bound is not enough here, and the chapter shows why.

Summary

Conflicts and colors:

- ▶ The objects in conflict are the vertices of a graph, the conflicts are its edges, the resources are the colors
- ▶ Scheduling the exams in the fewest days is a coloring problem

A classical hard problem:

- ▶ The chromatic number is one of Karp's 21 NP-complete problems; 3 colors are already hard on planar graphs

The challenge:

- ▶ How many colors does the Petersen graph need, and how do we prove it?

Overview of the family

Introduction: Overview of the family

- ▶ In this section we present the family of **coloring problems**: their data, their solutions and their objective
- ▶ We recall the few notions on graphs that we need: neighbourhood, degree, induced graph, complement graph, clique and stable set
- ▶ These notions are the common notation of the whole family, and the language of the rest of the chapter

Coloring problems

- ▶ The **coloring problems** share the same data: a simple undirected graph $G = (V, E)$, whose edges are the pairs of vertices in conflict
- ▶ A solution assigns a color to every vertex, so that the two endpoints of every edge receive different colors: the vertices with the same color form a stable set of G
- ▶ The objective is to minimize the number of colors used

The family in one sentence

A coloring problem partitions the vertices of a graph into the **fewest stable sets**.

The notions on graphs that follow are the common notation of the family.

The members of the family

- ▶ The table lists the members of the family, with the section that defines each of them

Problem	Acronym	Solution	Section
vertex coloring problem	VCP	a partition of V into stable sets	The vertex coloring problem

We now recall the notions on graphs
that we use in the chapter:

first graphs, neighbourhoods,
induced graphs and complement graphs;
then cliques and stable sets.

Definition 1: graph, neighbourhood, induced graph and complement graph (1/2)

- ▶ A **graph** $G = (V, E)$ is given by a finite set V of **vertices** and a set E of **edges**, each edge being a 2-element subset of V . We write $n = |V|$ and $m = |E|$; the vertices are labelled $1, 2, \dots, n$. All the graphs of this chapter are simple and undirected.
- ▶ Two distinct vertices u and v are **adjacent** if $\{u, v\} \in E$. The **neighbourhood** of a vertex v is the set $N(v) = \{u \in V : \{u, v\} \in E\}$, and its **degree** is $\deg(v) = |N(v)|$.

Definition: graph, neighbourhood, induced graph and complement graph (2/2)

- ▶ For a subset $U \subseteq V$ the **induced graph** $G[U] = (U, E[U])$ has the vertices of U and the edges of E with both endpoints in U .
- ▶ The **complement graph** $\overline{G} = (V, \overline{E})$ has the same vertices as G , and two distinct vertices are adjacent in $\overline{G}$ if and only if they are not adjacent in G .

Definition 2: clique and stable set

Let $G = (V, E)$ be a graph.

1. A **clique** is a subset $K \subseteq V$ of pairwise adjacent vertices: $\{u, v\} \in E$ for all $u, v \in K$ with $u \neq v$.
2. A **stable set** (also called **independent set**) is a subset $Q \subseteq V$ of pairwise non-adjacent vertices: $\{u, v\} \notin E$ for all $u, v \in Q$ with $u \neq v$.

The **clique number** $\omega(G)$ is the largest number of vertices of a clique, and the **stability number** $\alpha(G)$ is the largest number of vertices of a stable set.

Cliques, stable sets and the complement graph

- ▶ A stable set of G is a clique of the complement graph $\overline{G}$, and vice versa
- ▶ A vertex alone and the empty set are both a clique and a stable set

Cliques, stable sets and colors

A clique is a group of vertices that pairwise *conflict*;
a stable set is a group of vertices with no conflict.

In a coloring, the vertices of a clique
need **different colors**,
and the vertices of a stable set
can share one.

Summary (1/2)

Coloring problems:

- ▶ Data: a simple undirected graph $G = (V, E)$; the edges are the pairs of vertices in conflict
- ▶ Solution: a color for every vertex, different on the two endpoints of every edge
- ▶ Objective: minimize the number of colors used

Graphs:

- ▶ Neighbourhood $N(v)$ and degree $\deg(v) = |N(v)|$ of a vertex
- ▶ Induced graph $G[U]$ and complement graph $\overline{G}$

Summary (2/2)

Cliques and stable sets:

- ▶ A clique has pairwise adjacent vertices: the clique number is $\omega(G)$
- ▶ A stable set has pairwise non-adjacent vertices: the stability number is $\alpha(G)$
- ▶ The stable sets of G are the cliques of $\overline{G}$, and vice versa

Colors:

- ▶ A clique needs different colors; a stable set can share one
- ▶ A coloring partitions the vertices into the fewest stable sets

The vertex coloring problem

Introduction: The vertex coloring problem

- ▶ In this section we define the **vertex coloring problem** (VCP): its instances, its feasible solutions and its optimal value, the chromatic number $\chi(G)$
- ▶ We meet two reference instances: a triangle and the Petersen graph
- ▶ We state two assumptions on the graph and two simple dual bounds, the clique bound and the stable set bound
- ▶ With these tools we answer the challenge on the Petersen graph

Definition 3: Vertex Coloring Problem

Given:

1. *a simple graph $G = (V, E)$, where*

$$V = \{1, 2, \dots, n\}$$

is the set of n vertices, and

$$E \subseteq \{ \{u, v\} : u, v \in V, u \neq v \}$$

is the set of m edges,

*the **Vertex Coloring Problem** (VCP) asks to decide how to assign a color to each vertex, so that no two adjacent vertices share the same color, and the total number of colors used is minimized.*

A VCP instance

A **VCP instance** is specified by:

1. two integer numbers:
 - ▶ the number of vertices $n \in \mathbb{Z}_{\geq 1}$,
 - ▶ the number of edges $m \in \mathbb{Z}_{\geq 1}$,
2. two vectors of integer numbers of length m :
 - ▶ $\mathbf{u} \in \mathbb{Z}_{\geq 1}^m$, containing the first endpoint of the edges,
 - ▶ $\mathbf{v} \in \mathbb{Z}_{\geq 1}^m$, containing the second endpoint of the edges.

A VCP feasible solution (1/2)

A **VCP feasible solution** is a family of **color classes**, i.e., subsets of vertices each assigned to one color,

$$\mathcal{Q} \subseteq 2^V$$

such that:

(i) $\mathcal{Q}$ forms a partition of V into *nonempty* color classes, namely

$$\begin{aligned} Q &\neq \emptyset, \quad \text{for all } Q \in \mathcal{Q}, \\ Q \cap Q' &= \emptyset, \quad \text{for all } Q, Q' \in \mathcal{Q} \text{ with } Q \neq Q', \\ \bigcup_{Q \in \mathcal{Q}} Q &= V; \end{aligned}$$

A VCP feasible solution (2/2)

(ii) each color class in $\mathcal{Q}$ is a *stable set*, namely

$$\{u, v\} \notin E, \quad \text{for all } Q \in \mathcal{Q}, \text{ for all } u, v \in Q \text{ with } u \neq v.$$

Accordingly, the VCP can be equivalently stated as the problem of determining a partition of the vertex set into color classes

- ▶ such that each color class is a stable set (i.e., its vertices can receive the same color),
- ▶ and the number of color classes (colors) is minimized.

The VCP in formulas

Given a VCP instance, the VCP can be formally stated as:

$$\tilde{\mathcal{Q}} \in \arg \min_{\mathcal{Q} \subseteq 2^V} \left\{ |\mathcal{Q}| : \begin{array}{l} Q \neq \emptyset, \quad \forall Q \in \mathcal{Q}, \\ Q \cap Q' = \emptyset, \quad \forall Q, Q' \in \mathcal{Q} \text{ with } Q \neq Q', \\ \bigcup_{Q \in \mathcal{Q}} Q = V, \\ \{u, v\} \notin E, \quad \forall Q \in \mathcal{Q}, \forall u, v \in Q \text{ with } u \neq v \end{array} \right\}.$$

The chromatic number

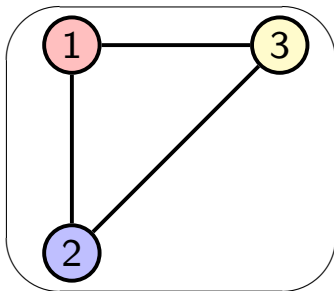
- ▶ The optimal value of the VCP is the **chromatic number** $\chi(G)$ of the graph: the minimum number of colors (stable sets) needed to color all the vertices
- ▶ The color classes of a solution are stable sets, and the number of classes is the number of colors used

Two reference instances

We now see two VCP instances:
a **triangle**,
and the **Petersen graph**, the reference instance of the
chapter.

Example: a triangle

We consider the following graph $G = (V, E)$ with 3 vertices ($n = 3$) and 3 edges ($m = 3$):

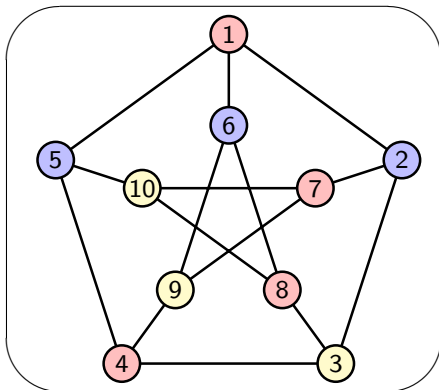


The vertex set V and the edge set E are:

$$V = \{ 1, 2, 3 \}, \quad E = \{ \{1, 2\}, \{1, 3\}, \{2, 3\} \}.$$

Example: the Petersen graph (1/2)

We consider now the following graph $G = (V, E)$ with 10 vertices ($n = 10$) and 15 edges ($m = 15$):



Example: the Petersen graph (2/2)

The vertex set V and the edge set E are:

$$V = \{ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \},$$

$$E = \{ \{1, 2\}, \{1, 5\}, \{1, 6\}, \{2, 3\}, \{2, 7\}, \\ \{3, 4\}, \{3, 8\}, \{4, 5\}, \{4, 9\}, \{5, 10\}, \\ \{6, 8\}, \{6, 9\}, \{7, 9\}, \{7, 10\}, \{8, 10\} \}.$$

This VCP instance is the **reference instance** of the chapter.

Assumptions

We make the following assumptions:

1. The graph has no isolated vertices. An isolated vertex can receive any color used by the rest of the graph; if the graph has no edges, the optimal value is one. Formally:

$$\deg(v) \geq 1, \quad \forall v \in V.$$

2. No vertex has its neighbourhood contained in the neighbourhood of another vertex. Formally:

$$N(u) \not\subseteq N(v), \quad \forall u, v \in V, u \neq v.$$

The second assumption is without loss of generality

The second assumption can be made without loss of generality, because:

1. if $N(u) \subseteq N(v)$, the vertex u can be removed without changing the chromatic number: every coloring of $G \setminus \{u\}$ extends to u with the color of v ;
2. the vertices u and v are not adjacent: otherwise $v \in N(u) \subseteq N(v)$, which is impossible in a simple graph, so u and v can share a color.

Bounds on the chromatic number

- ▶ A coloring found by hand gives a **primal bound**: $\chi(G)$ is at most its number of colors
- ▶ To prove that fewer colors are not enough we need **dual bounds**
- ▶ We now see two simple ones: the first uses a clique, the second a stable set

Observation 1: clique bound

For every graph $G = (V, E)$ we have

$$\omega(G) \leq \chi(G).$$

Proof

- ▶ All vertices in any clique are pairwise adjacent
- ▶ Hence they must all receive different colors in any VCP feasible solution
- ▶ Thus, at least $\omega(G)$ different colors are needed to color the vertices of any maximum clique, implying

$$\omega(G) \leq \chi(G).$$



Observation 2: stable set bound

For every graph $G = (V, E)$ we have

$$\left\lceil \frac{n}{\alpha(G)} \right\rceil \leq \chi(G).$$

Proof

- ▶ Consider a feasible solution with k color classes: every class is a stable set, so it has at most $\alpha(G)$ vertices
- ▶ The n vertices are covered by the classes, hence

$$n \leq k \alpha(G)$$

- ▶ The number k is an integer, so

$$k \geq \left\lceil \frac{n}{\alpha(G)} \right\rceil$$



Two dual bounds

- ▶ The **clique bound** is useful when the graph has a large clique
- ▶ The **stable set bound** is useful when its stable sets are small
- ▶ Both are **dual bounds**: a clique and a stable set are found without knowing $\chi(G)$

We now apply the definitions
to the two reference instances:

we count the candidate solutions,
find a first coloring by eye
and compute the two dual bounds.

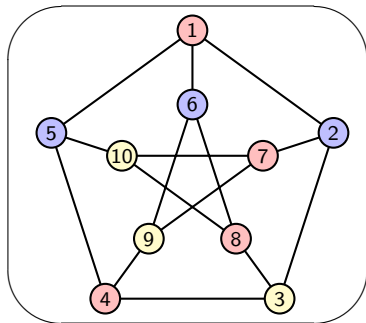
Candidate solutions

- ▶ The candidate solutions are the partitions of V , counted by the **Bell numbers** (see the chapter «Combinatorial optimization problems»)
- ▶ **Triangle**: we have $B(3) = 5$ partitions and only one is feasible: three singletons, since every pair of vertices is adjacent. Hence

$$\chi(G) = 3.$$

- ▶ **Petersen graph**: we have $B(10) = 115\,975$ partitions, and 7 430 of them are feasible, that is, have only stable sets as classes

A first solution, by eye



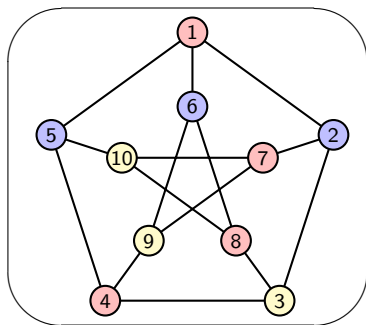
- ▶ The coloring shown in the Petersen graph example uses 3 colors:

$$\mathcal{Q} = \left\{ \begin{array}{l} \{1, 4, 7, 8\}, \\ \{2, 5, 6\}, \\ \{3, 9, 10\} \end{array} \right\}$$

- ▶ No edge joins two vertices of the same class, so it is feasible
- ▶ It is a **primal bound**:

$$\text{UB} = 3 \geq \chi(G)$$

Dual bounds on the Petersen graph (1/3)



- ▶ The graph has no triangle, so

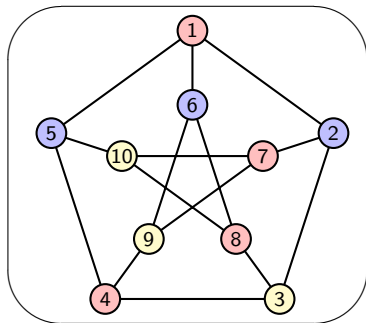
$$\omega(G) = 2$$

- ▶ The clique bound gives only

$$LB = 2$$

- ▶ With $UB = 3$, a **gap of 1** remains

Dual bounds on the Petersen graph (2/3)



- ▶ The vertices 1, 2, 3, 4, 5 form a cycle with 5 vertices, and so do 6, 7, 8, 9, 10 (the edges of the figure)
- ▶ A stable set contains at most 2 vertices of each cycle, so $\alpha(G) \leq 4$
- ▶ The set $\{1, 4, 7, 8\}$ shows that

$$\alpha(G) = 4$$

Dual bounds on the Petersen graph (3/3)

- ▶ The stable set bound gives

$$\left\lceil \frac{n}{\alpha(G)} \right\rceil = \left\lceil \frac{10}{4} \right\rceil = 3$$

- ▶ The two bounds coincide:

$$\text{LB} = 3 = \text{UB}$$

- ▶ The **gap is zero** and

$$\chi(G) = 3$$

Back to the challenge

- ▶ **How many colors are needed?** Three: the coloring by eye uses 3 colors, and no coloring uses fewer
- ▶ **How do we prove that two colors are not enough?** Not by listing the 115 975 partitions: the stable set bound $\lceil 10/4 \rceil = 3$ proves it
- ▶ **Why is the clique bound not enough?** The Petersen graph has no triangle, so the clique bound gives only 2

Summary (1/2)

The vertex coloring problem:

- ▶ A feasible solution is a partition of V into nonempty color classes, each a stable set
- ▶ The optimal value is the chromatic number $\chi(G)$

Assumptions:

- ▶ No isolated vertices: $\deg(v) \geq 1$ for all $v \in V$
- ▶ No neighbourhood contained in another one: $N(u) \not\subseteq N(v)$ for all $u \neq v$

Summary (2/2)

Bounds:

- ▶ A coloring found by eye is a primal bound: $UB \geq \chi(G)$
- ▶ Clique bound: $\omega(G) \leq \chi(G)$
- ▶ Stable set bound: $\lceil n/\alpha(G) \rceil \leq \chi(G)$

Reference instances:

- ▶ Triangle: $B(3) = 5$ partitions, one feasible, $\chi(G) = 3$
- ▶ Petersen graph: 7 430 feasible partitions out of 115 975;
 $\omega(G) = 2$, $\alpha(G) = 4$, $\chi(G) = 3$

Remarks

Introduction: Remarks

- ▶ In this section we collect three classical results on the chromatic number
- ▶ The first one is an upper bound that depends only on the largest degree
- ▶ The other two show that the problem is NP-complete, even on very special graphs

Three classical results

- ▶ **Brooks (1941)** proved that a connected graph that is neither complete nor an odd cycle can be colored with at most Δ colors, where Δ is the largest degree
- ▶ **Karp (1972)** included the chromatic number among his 21 NP-complete problems
- ▶ **Garey, Johnson and Stockmeyer (1976)** proved that deciding whether 3 colors are enough remains NP-complete on planar graphs of maximum degree 4

Summary

An upper bound from the degrees:

- ▶ Brooks: at most Δ colors for a connected graph that is neither complete nor an odd cycle

Complexity:

- ▶ The chromatic number is NP-complete (Karp)
- ▶ Deciding whether 3 colors are enough is NP-complete even on planar graphs of maximum degree 4

Typical mistakes

Introduction: Typical mistakes

- ▶ In this section we list the most common mistakes on coloring problems
- ▶ They concern cliques and stable sets, the bounds, the counting of the solutions and the assumptions
- ▶ Recognizing them helps to read and check a coloring correctly

Observation 3: typical mistakes (1/2)

- ▶ ***Confusing a clique with a stable set.*** In a clique all the pairs are edges and the vertices need different colors; in a stable set no pair is an edge and the vertices can share a color.
- ▶ ***Believing that the clique bound is tight.*** It is not: the Petersen graph has $\omega(G) = 2$ and $\chi(G) = 3$, and the same happens for a cycle with 5 vertices.
- ▶ ***Counting the permutations of the colors as different solutions.*** A solution is a partition of the vertices into stable sets: exchanging the names of two colors gives the same solution.

Observation: typical mistakes (2/2)

- ▶ *Using a class that is not a stable set.* A color class with two adjacent vertices makes the solution infeasible, whatever the number of colors.
- ▶ *Forgetting the assumptions.* An isolated vertex and a vertex with $N(u) \subseteq N(v)$ can be removed before solving, without changing the chromatic number.

Summary

Before trusting a coloring:

- ▶ Check that every color class is a stable set
- ▶ Count partitions, not permutations of the colors

Before trusting a bound:

- ▶ The clique bound may leave a gap, as on the Petersen graph and on a cycle with 5 vertices
- ▶ Apply the assumptions first: an isolated vertex and a vertex with $N(u) \subseteq N(v)$ can be removed

In summary

Introduction: In summary

- ▶ We collect the key ideas of the chapter in a few points
- ▶ We close with the dashboard of the reference instances: the best lower and upper bounds found so far

Summary (1/2)

- ▶ The VCP partitions the vertices of a graph into the minimum number of stable sets; the optimal value is the chromatic number $\chi(G)$
- ▶ A clique needs different colors and a stable set can share one; the cliques of G are the stable sets of the complement graph
- ▶ A coloring found by eye is a primal bound; the clique bound $\omega(G)$ and the stable set bound $\lceil n/\alpha(G) \rceil$ are dual bounds
- ▶ The bounds do not always meet: on the Petersen graph the clique bound leaves a gap, the stable set bound closes it

Summary (2/2)

Dashboard of the reference instances.

instance (min)	lower bound	upper bound	from
triangle	3 (clique bound)	3 (by eye)	this chapter
Petersen graph	3 (stable set bound)	3 (by eye)	this chapter