Graph Theory Notes*

Vadim Lozin

Institute of Mathematics
University of Warwick

1 Introduction

A graph $G = (V, E)$ consists of two sets V and E. The elements of V are called the vertices and the elements of E the edges of G. Each edge is a pair of vertices. For instance, the sets $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$ define a graph with 5 vertices and 4 edges.

Graphs have natural visual representations in which each vertex is represented by a point and each edge by a line connecting two points.

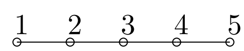

Figure 1: Graph $G = (V, E)$ with $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$

By altering the definition, we can obtain different types of graphs. For instance,

- by replacing the set E with a set of *ordered pairs* of vertices, we obtain a *directed graph* or *digraph*, also known as *oriented graph* or *orgraph*. Each edge of a directed graph has a specific orientation indicated in the diagram representation by an arrow (see Figure 2). Observe that in general two vertices i and j of an oriented graph can be connected by *two* edges directed opposite to each other, i.e. (i, j) and (j, i).
- by allowing E to contain both directed and undirected edges, we obtain a *mixed graph*.
- by allowing repeated elements in the set of edges, i.e. by replacing E with a multiset, we obtain a *multigraph*.
- by allowing edges to connect a vertex to itself (a loop), we obtain *pseudographs*.
- by allowing the edges to be arbitrary subsets of vertices, not necessarily of size two, we obtain *hypergraphs*.
- by allowing V and E to be infinite sets, we obtain *infinite graphs*.

*The notes are under constant construction. The text may and will be modified, updated, extended, improved, polished in all possible respects, i.e. in terms of its content, structure, style, etc. Suggestions for improvement are welcome.

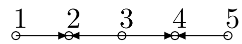

Figure 2: An oriented graph $G = (V, E)$ with $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1, 2), (3, 2), (3, 4), (4, 5)\}$

Definition 1. A simple graph is a finite undirected graph without loops and multiple edges.

All graphs in these notes are simple, unless stated otherwise.

2 Terminology, notation and introductory results

- The sets of vertices and edges of a graph G will be denoted $V(G)$ and $E(G)$, respectively.
- For notational convenience, instead of representing an edge by $\{a, b\}$ we shall denote it by ab.

Definition 2. [Adjacency, neighbourhood, vertex degree]

Let u, v be two vertices of a graph G.

- If $uv \in E(G)$, then u, v are said to be *adjacent*, in which case we also say that u is *connected* to v or u is a *neighbour* of v. If $uv \notin E(G)$, then u and v are *nonadjacent* (*not connected*, *non-neighbours*).
- The *neighbourhood* of a vertex $v \in V(G)$, denoted $N(v)$, is the set of vertices adjacent to v, i.e. $N(v) = \{u \in V(G) \mid vu \in E(G)\}$. The *closed neighbourhood* of v is denoted and defined as follows: $N[v] = N(v) \cup \{v\}$.
- If $e = uv$ is an edge of G, then e is *incident* to u and v. We also say that u and v are the *endpoints* of e.
- The *degree* of $v \in V(G)$, denoted $deg(v)$, is the number of edges incident to v. Alternatively, $deg(v) = |N(v)|$. If $deg(v) = 0$, then vertex v is called *isolated*. If $deg(v) = 1$, then vertex v and the only edge incident to v are called *pendant*. The maximum vertex degree and the minimum vertex degree in a graph G are denoted by $\Delta(G)$ and $\delta(G)$, respectively.

Definition 3. [Path, connectedness, distance, diameter]

- A *path* in a graph is a sequence of distinct vertices $v_1, v_2, \ldots, v_k$ such that $v_i v_{i+1}$ is an edge for each $i = 1, \ldots, k-1$. The *length* of a path P is the number of edges in P. A *chord* in a path is an edge connecting two non-consecutive vertices. A *chordless path* is a path without chords.
- A graph G is *connected* if every pair of distinct vertices is joined by a path. Otherwise it is *disconnected*.
- The *distance* between two vertices a and b, denoted $dist(a, b)$, is the length of a shortest path joining them.
- The *diameter* of a connected graph, denoted $diam(G)$, is $\max\limits_{a,b \in V(G)} dist(a, b)$.

Definition 4. [Notation for special graphs]

- K_n is the *complete graph* with n vertices, i.e. the graph with n vertices every two of which are adjacent.

- O_n is the *empty (edgeless) graph* with n vertices, i.e. the graph with n vertices no two of which are adjacent.

- P_n is a *chordless path* with n vertices, i.e. $V(P_n) = \{v_1, v_2, \ldots, v_n\}$ and $E(P_n) = \{v_1v_2, \ldots, v_{n-1}v_n\}$.

- C_n is a *chordless cycle* with n vertices, i.e. $V(C_n) = \{v_1, v_2, \ldots, v_n\}$ and $E(C_n) = \{v_1v_2, \ldots, v_{n-1}v_n, v_nv_1\}$.

- Q_n is a *hypercube*, i.e. the graph whose vertex set is the set of all binary strings $a_1a_2 \ldots a_n$ of length n (i.e. $a_i \in \{0, 1\}$ for each i), where two vertices $a_1a_2 \ldots a_n$ and $b_1b_2 \ldots b_n$ are adjacent if and only if there exists an i such that $a_i \neq b_i$ and $a_j = b_j$ for all $j \neq i$ (see Figure 3 for Q_3.)

- $G + H$ is the union of two disjoint graphs G and H, i.e. if $V(G) \cap V(H) = \emptyset$, then $V(G+H) = V(G) \cup V(H)$ and $E(G+H) = E(G) \cup E(H)$. In particular, nG denotes the disjoint union of n copies of G.

- $G \times H$ is the graph obtained from $G + H$ by adding all possible edges between G and H. We will call $G \times H$ the *join* of G and H.

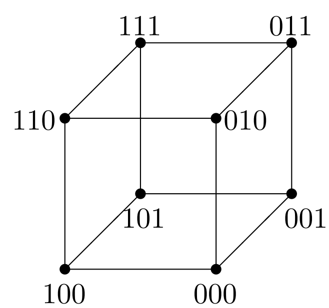

Figure 3: Q_3

Exercises

- Find the diameter of $K_n, P_n, C_n, Q_n,$ $P_n \times C_n$.

Definition 5. [Isomorphism] Two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are *isomorphic* if there is a bijection $f : V_1 \to V_2$ that preserves the adjacency, i.e. $uv \in E_1$ if and only if $f(u)f(v) \in E_2$.

Exercises

- Find all pairwise non-isomorphic graphs with 2,3,4,5 vertices.

Definition 6. [GRAPH COMPLEMENT] The *complement* of a graph $G = (V, E)$ is a graph with vertex set V and edge set E' such that $e \in E'$ if and only if $e \notin E$. The complement of a graph G is denoted $\overline{G}$ and sometimes is called co-G.

> EXERCISES
>
> - Find the complements of C_4, C_5, P_4, P_5.
> - Show that if $diam(G) \geq 3$, then $diam(\overline{G}) \leq 3$.

Definition 7. [SELF-COMPLEMENTARY GRAPHS] A graph G is *self-complementary* if G is isomorphic to its complement.

> EXERCISES
>
> - Find self-complementary graphs with 4,5,6 vertices.
> - Is it possible for a self-complementary graph with 100 vertices to have exactly one vertex of degree 50?

Definition 8. [GRAPH CONTAINMENT RELATIONS] Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, the graph G_1 is said to be

- a *subgraph* of $G_2 = (V_2, E_2)$ if $V_1 \subseteq V_2$ and $E_1 \subseteq E_2$, i.e. G_1 can be obtained from G_2 by deleting some vertices and some edges;
- a *spanning subgraph* of G_2 if $V_1 = V_2$ and $E_1 \subseteq E_2$, i.e. G_1 can be obtained from G_2 by deleting some edges but not vertices;
- an *induced subgraph* of G_2 if G_1 is a subgraph of G_2 and every edge of G_2 with both endpoints in V_1 is also an edge of G_1, i.e. G_1 can be obtained from G_2 by deleting some vertices but not edges.

> EXERCISES
>
> - Is any of the three relations applicable to the pair P_5 and C_5? P_4 and $\overline{P}_5$?
> - What are the subgraphs, induced subgraphs and spanning subgraphs of K_n?

Definition 9. [INDUCED SUBGRAPH RELATION]

Given a graph G and a subset $U \subseteq V(G)$, we denote by

- $G[U]$ the subgraph of G induced by U, i.e. the graph with vertex set U whose vertices are adjacent if and only if they are adjacent in G,
- $G - U$ the subgraph of G induced by $V(G) - U$, i.e. the graph obtained from G by deleting the vertices of U.

We say that G contains a graph H as an induced subgraph if H is isomorphic to an induced subgraph of G, in which case we also say that H is contained in G as an induced subgraph, or simply, H is an induced subgraph of G. We denote by

- $H < G$ the fact that H is an induced subgraph of G.

Definition 10. [Connected component, co-component]

- A maximal (with respect to inclusion) connected subgraph of G is called a *connected component* of G.
- A *co-component* in a graph is a connected component of its complement.

Exercises
• Is it true that the complement of a connected graph is necessarily disconnected?   • Prove that the complement of a disconnected graph is necessarily connected.   • Prove that a graph is connected if and only if for every partition of its vertex set into two non-empty sets A and B there is an edge $ab \in E(G)$ such that $a \in A$ and $b \in B$.   • Show that if a graph with n vertices has more than $\binom{n-1}{2}$ edges, then it is connected.

Definition 11. [Degree sequence, regular graph]

- If $V(G) = \{v_1, v_2, \ldots, v_n\}$, then the sequence $(deg(v_1), deg(v_2), \ldots, deg(v_n))$ is called the *degree sequence* of G.
- If $deg(v_1) = deg(v_2) = \ldots = deg(v_n)$, then G is a *regular* graph.
- If the degree of each vertex is d, then the graph is d-regular.
- 3-regular graphs are called *cubic*.

Exercises
• Which of the following graphs are regular: $K_n, P_n, C_n, 2K_2$?   • Find all pairwise non-isomorphic regular graphs of degree $n-2$.   • Find all pairwise non-isomorphic graphs with the degree sequence (0,1,2,3,4).   • Find all pairwise non-isomorphic graphs with the degree sequence (1,1,2,3,4).   • Find all pairwise non-isomorphic graphs with the degree sequence (2,2,3,3,4,4).

Lemma 12. *Let $G = (V, E)$ be a graph with m edges. Then $\sum\limits_{v \in V} deg(v) = 2m$.*

Proof. In counting the sum $\sum\limits_{v \in V} deg(v)$, we count each edge of the graph twice, because each edge is incident to exactly two vertices. $\square$

Corollary 13. *In every graph, the number of vertices of odd degree is even.*

Exercises

- Prove that if a graph has exactly two vertices of odd degrees, then they are connected by a path.

Definition 14. [Clique, independent set]

- In a graph, a set of pairwise adjacent vertices is called a *clique*. The size of a maximum clique in G is called the *clique number* of G and is denoted $\omega(G)$.
- A set of pairwise non-adjacent vertices is called an *independent set* (also known as *stable set*). The size of a maximum independent set in G is called the *independence number* (also know as *stability number*) of G and is denoted $\alpha(G)$.

Exercises

- Find the clique number and the independence number of $K_n, P_n, C_n, Q_n, P_n + C_n$, $P_n \times C_n$.

3 Classes of graphs

A *class of graphs* (also known as *graph property*) is a set of graphs closed under isomorphism. We have seen already various classes of graphs: complete graphs, cycles, paths, connected graphs. Many more classes can be defined by means of various graph parameters. For instance, for a natural number k, we can define the class of graphs of maximum vertex degree at most k, of diameter k, of clique number at least k, etc. In these notes, of special interest will be classes closed under taking induced subgraph. More formally:

Definition 15. A class X of graphs containing with each graph G all induced subgraphs of G is called hereditary. In other words, X is hereditary if and only if $G \in X$ implies $G - v \in X$ for any vertex v of G.

Exercises

- Which of the following sets of graphs are hereditary: complete graphs, paths, cycles, connected graphs, graphs of diameter k, graphs of vertex degree k, graphs of independence number at most k?

An important property of hereditary classes is that they admit so-called forbidden induced subgraph characterization. To make things more precise, let us denote by

- $Free(M)$ the set of graphs containing no graph from M as an induced subgraph.

We will say that graphs in M are *forbidden induced subgraphs* for the class $Free(M)$ and that graphs in $Free(M)$ are M-free.

Theorem 16. *A class of graphs X is hereditary if and only if there is a set M such that $X = Free(M)$.*

Proof. Assume $X = Free(M)$ for a set M. Consider a graph $G \in X$ and an induced subgraph H of G. Then H is M-free, since otherwise G contains a forbidden graph from M. Therefore, $H \in X$ and hence X is hereditary.

Conversely, if X is hereditary, then $X = Free(M)$ with M being the set of all graphs not in X. □

To illustrate the theorem, consider the set X of all complete graphs. Clearly, this set is hereditary and $X = Free(M)$ with M being the set of all non-complete graphs. On the other hand, it is not difficult to see that $X = Free(\overline{K}_2)$, since a graph G is complete if and only if $\overline{K}_2 \not\leq G$, i.e. G has no pair of non-adjacent vertices. This example motivates the following definition.

Definition 17. A graph G is a *minimal* forbidden induced subgraph for a hereditary class X if and only if $G \notin X$ but every proper induced subgraph of G belongs to X (or alternatively, the deletion of any vertex from G results in a graph that belongs to X).

Let us denote the set of all minimal forbidden induced subgraphs for a hereditary class X by $MFIS(X)$.

Theorem 18. *For any hereditary class X, we have $X = Free(MFIS(X))$. Moreover, $MFIS(X)$ is the unique minimal set with this property.*

Proof. To prove that $X = Free(MFIS(X))$, we show two inclusions: $X \subseteq Free(MFIS(X))$ and $Free(MFIS(X)) \subseteq X$. Assume first $G \in X$, then by definition all induced subgraphs of G belong to X and hence no graph from $MFIS(X)$ is an induced subgraph of G, since none of them belongs to X. As a result, $G \in Free(MFIS(X))$, which proves that $X \subseteq Free(MFIS(X))$.

Conversely, assume that $G \in Free(MFIS(X))$, and suppose by contradiction that $G \notin X$. Let H be a minimal induced subgraph of G which is not in X (possibly $H = G$). But then $H \in MFIS(X)$ contradicting the fact that $G \in Free(MFIS(X))$. This contradiction shows that $G \in X$ and hence proves that $Free(MFIS(X)) \subseteq X$.

To prove the minimality of the set $MFIS(X)$, we will show that for any set N such that $X = Free(N)$ we have $MFIS(X) \subseteq N$. Assume this is not true and let H be a graph in $MFIS(X) - N$. By the minimality of the graph H, any proper induced subgraph of H is in X, and hence is in $Free(N)$. Together with the fact that H does not belong to N, we conclude that $H \in Free(N) = X = Free(MFIS(X))$. But this contradicts the fact that $H \in MFIS(X)$. □

Exercises

- Show that the union of two hereditary classes is a hereditary class.
- Show that the intersection of two hereditary classes is a hereditary class.
- Show that if $X = Free(M)$ and $Y = Free(N)$, then $X \cap Y = Free(M \cup N)$.
- Show that if $X = Free(M)$, then $\overline{X} = Free(\overline{M})$, where for an arbitrary set A of graphs $\overline{A}$ denotes the set of complements of graphs in A.

Theorem 19. *$Free(M_1) \subseteq Free(M_2)$ if and only if for every graph $G \in M_2$ there is a graph $H \in M_1$ such that H is an induced subgraph of G.*

Proof. Assume $Free(M_1) \subseteq Free(M_2)$, and suppose to the contrary that a graph $G \in M_2$ contains no induced subgraphs from M_1. By definition, this means that $G \in Free(M_1) \subseteq Free(M_2)$. On the other hand, G belongs to the set of forbidden graphs for $Free(M_2)$, a contradiction.

Conversely, assume that every graph in M_2 contains an induced subgraph from M_1. By contradiction, let $G \in Free(M_1) - Free(M_2)$. Since G does not belong to $Free(M_2)$, by definition G contains an induced subgraph $H \in M_2$. Due to the assumption, H contains an induced subgraph $H' \in M_1$. Then obviously H' is an induced subgraph of G which contradicts the fact that $G \in Free(M_1)$. □

Exercises

Let X be the class of graphs of vertex degree at most 2.

- What can you say about the structure of graphs in X? In particular, how a connected graph in this class looks like?
- Is X a hereditary class?
- Find the set of minimal forbidden induced subgraphs for X.

4 Classes with small forbidden induced subgraphs

As we have seen already, $Free(\overline{K}_2)$ is the class of complete graphs. Similarly, $Free(K_2)$ is the class of empty (edgeless) graphs. In this section, we will describe some other simple classes of graphs defined by small forbidden induced subgraphs.

Lemma 20. *A graph G is P_3-free if and only if it is a disjoint union of cliques, i.e. every connected component of G is a clique.*

Proof. Clearly, if G is a disjoint union of cliques, then G is P_3-free. Conversely, let G be a P_3-free graph and assume it contains a connected component which is not a clique. Consider any two non-adjacent vertices in this component and a shortest (i.e. chordless) path connecting them. This path contains at least 2 edges, and therefore, G contains a P_3. This contradiction completes the proof. □

From Lemma 20 it follows that the complement of a P_3-free graph is a graph whose vertices can be partitioned into independent sets so that any two vertices belonging to different independent sets are adjacent. Such graphs are called *complete multipartite*.

Corollary 21. *A graph is complete multipartite if and only if it is $\overline{P}_3$-free.*

Exercises

- Which of the following graphs are complete multipartite: K_n, P_n, C_n, Q_n?

Definition 22. If the number of parts (independent sets) in a complete multipartite graph is at most two, the graph is called *complete bipartite*. A complete bipartite graph with two parts of size n and m is denoted $K_{n,m}$. Graphs of the form $K_{1,n}$ are called *stars*.

Corollary 23. *A graph is complete bipartite if and only if it is $(\overline{P}_3, K_3)$-free.*

Proof. Neither $\overline{P}_3$ nor K_3 is a complete bipartite graph, therefore a complete bipartite graph is $(\overline{P}_3, K_3)$-free. Conversely, if a graph is $\overline{P}_3$-free, then it is complete multipartite, and if it is K_3-free, then number of parts cannot be larger than 2, since otherwise a K_3 arises. $\square$

Corollary 24. *A graph G is (P_3, K_3)-free if and only if the maximum vertex degree in G is at most 1.*

Proof. If the maximum vertex degree in G is at most 1, then clearly it is (P_3, K_3)-free, since P_3 and K_3 have vertices of degree 2.

Conversely, if G is P_3-free, then every connected component of G is a clique (Lemma 20) and if G is K_3-free, then every component has size at most 2. Therefore, the maximum vertex degree in G is at most 1. $\square$

Corollary 25. *A graph G is $(P_3, 2K_2)$-free if and only if its vertex set can be partitioned into two subsets C and I so that C is a clique and I is the set of isolated vertices.*

Proof. If G is P_3-free, then every connected component of G is a clique, and if G is $2K_2$-free, then at most one of the components of G has more than one vertex. Therefore, the vertices of a $(P_3, 2K_2)$-free graph can be partitioned into a clique and the set of isolated vertices. The converse statement follows from the obvious fact that neither P_3 nor $2K_2$ admits such a partition. $\square$

Definition 26. A *paw* is the unique (up to isomorphism) graph with the degree sequence $(1, 2, 2, 3)$.

EXERCISES

- Show that $Free(K_3) \cup Free(P_3) \cup Free(\overline{P}_3) \subseteq Free(paw)$.
- Is the inverse inclusion $Free(paw) \subseteq Free(K_3) \cup Free(P_3) \cup Free(\overline{P}_3)$ true?

Lemma 27. *Every connected paw-free graph is either K_3-free or $\overline{P}_3$-free.*

Proof. Assume a connected *paw*-free graph G contains a triangle a, b, c. Then no vertex $x \in V(G) - \{a, b, c\}$

(1) has 1 neighbour in $\{a, b, c\}$, since otherwise $G[a, b, c, x] = paw$;

(2) has 0 neighbours in $\{a, b, c\}$. Indeed, assume x has no neighbours in $\{a, b, c\}$. Since G is connected, there are must exist a path connecting x to $\{a, b, c\}$. Without loss of generality we may assume that x is a vertex closest to $\{a, b, c\}$ which has no neighbours in this set. Then x is adjacent to a vertex y which has a neighbour in $\{a, b, c\}$. According to (1), y has at least 2 neighbours in $\{a, b, c\}$. But then two neighbours of y together with y and x induce a *paw* in G, a contradiction.

We denote by V_{ab}, V_{ac}, V_{bc} the subsets of $V(G)$ with exactly 2 neighbours in $\{a, b, c\}$, and by V_3 the remaining vertices of G, i.e. those adjacent to each vertex of $\{a, b, c\}$. Then

- each of V_{ab}, V_{ac}, V_{bc} is an independent set. Indeed, if, say, V_{ab} contains two adjacent vertices x, y, then a, c, x, y induce a *paw* in G.

- any two vertices belonging to different sets V_{ab}, V_{ac}, V_{bc}, V_3 are adjacent. Indeed, if, say, $x \in V_{a,b} \cup V_3$ is not adjacent to $y \in V_{b,c}$, then a, b, x, y induce a *paw* in G.
- $G[V_3]$ is $\overline{P}_3$-free. Indeed, if $G[V_3]$ contains a $\overline{P}_3$ induced by x, y, z, then $G[a, x, y, z]$ is an induced *paw*.

Therefore, G is complete multipartite, i.e. G is $\overline{P}_3$-free. □

5 On the speed of graph properties

How many n-vertex graphs with a property P are there? An answer to this question depends on whether we count unlabelled (i.e. up to isomorphism) or labelled graphs in P. In labelled graphs, the vertices are assigned numbers $1, 2, \ldots, n$ and two labelled graphs are counted different if they have different sets of edges, i.e. if there is at least one pair ij which creates an edge in one of the graphs and a non-edge in the other.

In most cases, counting unlabelled graphs is much harder. But even for labelled graphs the above question remains highly non-trivial for most of the properties. Let us denote by

- $P(n)$ the set of n-vertex labelled graphs in P.

The value of $|P(n)|$, as a function of n, is known as the *speed* of P [1]. The exact value of $|P(n)|$ is available only for very few properties. The following set of exercises reveals some of them.

EXERCISES

- What is the number of n-vertex labelled graphs?
- What is the number of complete n-vertex labelled graphs?
- What is the number of n-vertex labelled paths?
- What is the number of n-vertex labelled stars?
- What is the number of n-vertex labelled graphs with one edge? with k edges?
- What is the number of n-vertex labelled graphs in the class $Free(\overline{P}_3, K_3)$?
- What is the number of n-vertex labelled graphs in the class $Free(P_3, 2K_2)$?
- What is the number of n-vertex labelled graphs in the class $Free(P_3, K_3)$?

For most properties, only asymptotic values of $|P(n)|$ or bounds are available. We will return to this topic, when more results are proved.

6 Graphs without cycles

Definition 28. A connected graph without cycles is called a *tree*.

EXERCISES

- Which of the following graphs are trees: P_n, C_n, K_n, *paw*, $\overline{P}_n$, $\overline{C}_n$, $\overline{K}_n$, co-*paw*?
- Is the class of trees hereditary?

Definition 29. A graph every connected component of which is a tree is called a *forest*. In other words, a forest is a graph without cycles. Forests are also known as *acyclic graphs*.

> EXERCISES
>
> - Which of the following graphs are forests: P_n, C_n, K_n, *paw*, $\overline{P}_n$, $\overline{C}_n$, $\overline{K}_n$, co-*paw*?
> - Is the class of forests hereditary?

Theorem 30. *The following statements are equivalent for a graph T:*

(1) T is a tree.

(2) Any two vertices in T are connected by a unique path.

(3) T is minimally connected, i.e. T is connected but $T - e$ is disconnected for any edge e of T.

(4) T is maximally acyclic, i.e. T is acyclic but $T + uv$ contains a cycle for any two non-adjacent vertices u, v of T.

(5) T is connected and $|E(T)| = |V(T)| - 1$.

(6) T is acyclic and $|E(T)| = |V(T)| - 1$.

Proof. (1)→ (2): Since a tree is a connected graph, any two vertices must be connected by at least one path. If there would be two paths connecting two vertices, then a cycle could be easily found.

(2)→ (1): If any two vertices in T are connected by a path, T is connected. Since the path is unique, T is acyclic. Therefore, T is a tree.

(2)→ (3): Let $e = ab$ be an edge in T and assume $T - e$ is connected. Then a and b are connected in $T - e$ by a path P. But then P and the edge e create two different paths connecting a to b in T, contradicting (2).

(3)→ (2): Since T is connected, any two vertices in T are connected by a path. This path is unique, since otherwise there would exist an edge e in T (belonging to one of the paths but not to the other), such that $T - e$ is connected, contradicting (3).

(1)→ (4): By (1), T is acyclic. For the same reason, T is connected and therefore any two non-adjacent vertices u, v of T are connected by a path. This path together with uv create a cycle in $T + uv$.

(4)→ (1): To prove this implication, we only have to show that T is connected. Assume it is not, and let u and v be two vertices of T in different connected components. Obviously $T + uv$ has no cycles, which contradicts (4).

(1,2,3,4)→ (5): We prove $|E(T)| = |V(T)| - 1$ by induction on $n = |V(T)|$. For $n = 1, 2$, true. Assume true for less than n vertices and let T be a tree with n vertices and let $e = ab$ be an edge in T. By (3), $T - e$ is disconnected. Let T_1 and T_2 be connected components of $T - e$. Obviously, T_1 and T_2 are trees. Since each of them has fewer vertices than T, we know that $|E(T_1)| = |V(T_1)| - 1$ and $|E(T_2)| = |V(T_2)| - 1$. We also know that $|V(T)| = |V(T_1)| + |V(T_2)|$ and $|E(T)| = |E(T_1)| + |E(T_2)| + 1$. Therefore, $|E(T)| = |(V(T_1)| - 1 + |V(T_2)| - 1) + 1 = |V(T)| - 2 + 1 = |V(T)| - 1$, as required.

(5) $\to$ (1,2,3,4): We only need to show that T is acyclic. As long as T contains a cycle, delete an edge from the cycle and denote the resulting graph by T' (i.e. the graph obtained from T by destroying all cycles). Clearly, T' is a tree, since deletion of an edge from a cycle cannot destroy the connectivity. Then from the previous paragraph we know that $|E(T')| = |V(T')| - 1$. On the other hand, we did not delete any vertex of T, i.e. $|V(T')| = |V(T)|$. Therefore, $|E(T')| = |V(T')| - 1 = |V(T)| - 1 = |E(T)|$ and hence $E(T') = E(T)$, i.e. no edge has been deleted from T. In other words, T is acyclic.

(1,2,3,4,5)$\to$ (6): obvious.

(6) $\to$ (1,2,3,4,5): We only need to show that T is connected. Denote by $T_1, \ldots, T_k$ the connected components of T. Each of them is a tree, and hence for each of them $|E(T_i)| = |V(T_i)| - 1$. Therefore,

$$|E(T)| = \sum_{i=1}^{k} |E(T_i)| = \sum_{i=1}^{k} (|V(T_i)| - 1) = \sum_{i=1}^{k} |V(T_i)| - k = |(V(T)| - k.$$

By (6), $|E(T)| = |V(T)| - 1$ and hence $k = 1$, i.e. T is connected. □

Corollary 31. *Every tree T with at least 2 vertices has at least 2 vertices of degree 1.*

Proof. By Handshake Lemma, $\sum deg(v)_{v \in V(T)} = 2|E(T)|$ and by Theorem 30, $|E(T)| = |V(T)| - 1$. Therefore, $\sum deg(v)_{v \in V(T)} = 2|V(T)| - 2$. Therefore, T must contain at least 2 vertices of degree 1. □

Definition 32. In a tree a vertex of degree 1 is called a *leaf*.

EXERCISES

- The *mean degree* of a graph G with n vertices is the value $d_{mean} = \frac{1}{n} \sum_{i=1}^{n} deg(v_i)$. Express the number of vertices of a tree in terms of the mean degree.
- Let T be a tree with n vertices. Assume each vertex of T has degree 1 or $d > 1$. Find a formula for the number of vertices of degree 1 in T.
- Show that a tree has exactly two vertices of degree 1 if and only if it is a path.
- Characterize the class of forests every connected component of which is a star in terms of minimal forbidden induced subgraphs.
- Characterize the class of forests every connected component of which is a path in terms of minimal forbidden induced subgraphs.

EXERCISES

- Let $T_1, T_2, \dots, T_k$, $k > 1$, be subtrees of a tree T such that any two of them intersect each other (i.e. have a vertex in common). Show that there is a vertex of T common to all subtrees $T_1, T_2, \dots, T_k$.

  Solution.

  Induction on k. For $k = 2$, this follows by assumption. Suppose the statement is true for $k-1$ trees. Then

  $$V_1 = \bigcap_{i=1}^{k-1} V(T_i) \neq \emptyset \text{ and } V_2 = \bigcap_{i=2}^{k} V(T_i) \neq \emptyset.$$

  If the intersection $V_1 \cap V_2$ is not empty, then every vertex of it belongs to all subtrees. So, assume $V_1 \cap V_2 = \emptyset$. By assumption $V_3 = V(T_1) \cap V(T_k)$ is not empty. If V_3 intersects V_1 or V_2, then this intersection belongs to all subtrees. Therefore, we assume that V_3 is disjoint both from V_1 and V_2.

  Let us consider three vertices $v_1 \in V_1$, $v_2 \in V_2$ and $v_3 \in V_3$. Any pair of them is connected by a unique path in T. We denote them by P_{12}, P_{13}, P_{23}. Every subtree contains at least two vertices in the triple $\{v_1, v_2, v_3\}$. Therefore, every subtree contains at least one of the paths P_{12}, P_{13}, P_{23}.

  If there is no vertex common to all three paths, then there must exist a cycle, which is impossible. Therefore, there exists a vertex u belonging to all three paths. Since each subtree contains at least one of the paths P_{12}, P_{13}, P_{23}, vertex u belongs to all subtrees.

6.1 Prüfer code. The number of labelled trees

Every labelled graph can be described by listing all its edges. Every edge is a pair of vertices. Therefore, every tree with n vertices can be described by listing $2(n-1)$ labels. In 1918, Prüfer proposed a more compact way of describing a labelled tree.

Let T be a tree with vertices $\{1, 2, \dots, n\}$. Let a_1 be the smallest leaf in T (i.e. the leaf with smallest label) and let b_1 be the only neighbour of a_1. By deleting a_1 from T we obtain a new tree that will be denoted T_1. Let a_2 be the smallest leaf in T_1 and b_2 the only neighbour of a_2. By deleting a_2 from T_1 we obtain a new tree denoted T_2. Proceeding in this way, in $n-2$ steps we obtain the tree T_{n-2} with one edge denoted $a_{n-1}b_{n-1}$. Observe that the sequence $a_1b_1, a_2b_2, \dots, a_{n-1}b_{n-1}$ contains all edges of T. To the tree T we associate the ordered sequence $P(T) = b_1, b_2, \dots, b_{n-2}$, which is called the Prüfer code of T.

Let us show that the Prüfer code $P(T)$ allows to compute the numbers $a_1, a_2, \dots, a_{n-1}, b_{n-1}$, and therefore, to restore the tree T. Consider first the number a_1. Due to the choice of a_1, we conclude the following.

(1) *No vertex smaller than a_1 is a leaf in T.*

Also, it is not difficult to see that

(2) *Every vertex which is not a leaf in T appears in $P(T)$.* Indeed, every vertex x which is not a leaf has at least 2 neighbours in T. At least one of these neighbours must be deleted

by the end of the process of constructing the code, because at the end we are left with 2 adjacent vertices. If a neighbour of x is deleted at step i, then $b_i = x$.

From (1) and (2) it follows that

(3) *a_1 is the smallest natural number not in $P(T)$.*

Assume now that we have found the numbers $a_1, \ldots, a_{i-1}$. Consider the tree T_{i-1}. This tree corresponds to the subsequence (suffix) $b_i, b_{i+1}, \ldots, b_{n-2}$ of the code. In T_{i-1} vertex a_i is the smallest leaf. Arguing similarly as before, we conclude that

(4) *a_i is the smallest number in the set $\{1, 2, \ldots, n\} - \{a_1, \ldots, a_{i-1}\}$ which does not belong to $\{b_i, b_{i+1}, \ldots, b_{n-2}\}$. Alternatively, a_i is the smallest natural number which does not belong to $\{a_1, \ldots, a_{i-1}, b_i, b_{i+1}, \ldots, b_{n-2}\}$.*

This rule allows us to find the numbers $a_1, a_2, \ldots, a_{n-2}$. These numbers are pairwise distinct. The remaining two numbers a_{n-1} and b_{n-1} are the numbers in the set $\{1, 2, \ldots, n\} - \{a_1, \ldots, a_{n-2}\}$.

Theorem 33. (Cayley's formula) *The number of labelled trees with n vertices is n^{n-2}.*

Proof. The Prüfer code of a tree with n vertices is a word (ordered sequence) of length $n-2$ in the alphabet of n letters. The total number of such words is n^{n-2}. To prove that this also is the number of labelled trees, we will show that the mapping P that maps a tree T to its code $P(T)$ is a bijection (between trees and sequences). We have seen already that the tree can be restored from its code in a unique way. Therefore, P is an injective mapping. Now let us show that P is a surjection, i.e. for every sequence β of length $n-2$ in the alphabet of n letters, there is a tree T such that $P(T) = \beta$.

The decoding procedure (i.e. the procedure for restoring the tree from its code) described above is applicable to *any* sequence $b_1, b_2, \ldots, b_{n-2}$ of numbers chosen from the set $\{1, 2, \ldots, n\}$. This procedure produces numbers $a_1, a_2, \ldots, a_{n-1}, b_{n-1}$ from the same set. From these numbers we can create a graph with edges $a_1b_1, a_2b_2, \ldots, a_{n-1}b_{n-1}$. Therefore, all we have to do is to show that this graph is a tree. Let us denote by G_i the graph vertices $\{1, 2, \ldots, n\}$ and edges $\{a_ib_i, \ldots, a_{n-1}b_{n-1}\}$. The graph G_{n-1} has one edge and therefore has no cycles. The number a_i is different from the numbers $a_{i+1}, \ldots, a_{n-1}$ and from the numbers $b_{i+1}, \ldots, b_{n-1}$ (this follows from the decoding algorithm). Therefore, vertex a_i has degree 1 in the graph G_i, i.e. the transformation of G_{i+1} into G_i does not create cycles. As a result, the graph G_1 has no cycles. Since this graph has n vertices and $n-1$ edges, we conclude by Theorem 30 that it is a tree. □

Exercises

- Reconstruct trees from their Prüfer codes:
  - $(3, 3, 5, 5, 6, 6)$
  - $(1, 5, 1, 5, 9, 8, 2)$
  - $(4, 1, 6, 2, 2, 2, 7)$
  - $(2, 2, 2, \ldots, 2)$
  - $(1, 2, 3, \ldots, n-2)$

6.2 Rooted trees

- A *rooted tree* is a tree with a designated vertex v_0 called the *root*. According to Theorem 30, each vertex v_i is connected to the root by a unique path $v_i, v_{i-1}, \ldots, v_1, v_0$, in which case v_{i-1} is called *father* (or *parent*) of v_i, and v_i is a *child* (or *son*) of v_{i-1}.
- The *height* of a rooted tree is the distance from the root to the farthest leaf.
- A *binary* tree is a rooted tree in which every vertex has at most two children.
- A *full binary tree* is a binary tree in which every non-leaf vertex has exactly two children.
- A *complete binary tree* is a full binary tree in which all leaves are of the same distance from the root.

EXERCISES

- What is the number of leaves in a complete binary tree of height h?
- What is the maximum number of vertices in a binary tree of height h?
- What is the minimum height of a binary tree with n leaves?
- What is the number of vertices of degree 3 in a binary tree with t leaves?
- What is the number of *rooted* labelled trees with n vertices?

7 Modular decomposition and cographs

From the exercises in Section 2, we know that P_4 (the path on 4 vertices) is neither disconnected nor the complement to a disconnected graph. What can we say about the connectivity of graphs that do not contain P_4 as an induced subgraph?

7.1 P_4-free graphs – Cographs

Definition 34. A graph G is called *complement reducible* (or *cograph* for short) if every induced subgraph of G with at least two vertices is ether disconnected or the complement to a disconnected graph.

Theorem 35. *A graph G is a cograph if and only if it is P_4-free.*

Proof. Since neither P_4 nor its complement is disconnected, every cograph is P_4-free.

Conversely, let G be a P_4-free graph. We will show by induction on $n = |V(G)|$ that G is ether disconnected or the complement to a disconnected graph. Let $a \in V(G)$ and $G' = G - a$. Without loss of generality we can assume, by the induction hypothesis, that G' is disconnected (otherwise we can consider the complement of G'). If a is adjacent to every vertex in G', then the complement of G is disconnected. Consider a connected component H of G' which has a non-neighbour x of a. On the other hand, a must have a neighbour y in H, else H is a connected component of G. Since x and y belong to the same connected component of G', there is a chordless path connecting them. Without loss of generality we may assume that x and y are adjacent. Let z be a neighbour of a in another component of G'. Then z, a, y, x induce a P_4

in G. This contradiction shows that either there is a connected component in G' which does not have any neighbour of a or a is adjacent to every vertex in G'. In the first case G is disconnected and in the second, it is the complement to a disconnected graph. □

This theorem suggests a recursive method of decomposing a cograph G, which can be described by a rooted tree T as follows.

Step 0: Create the roof x of T.

Step 1: If G has just one vertex, then label the node x of T by that vertex and stop. Otherwise,

- if G is disconnected, then
  - label the node x of T by 0,
  - split G into connected components $G_1, \ldots, G_k$,
  - add to the node x k child vertices $x_1, \ldots, x_k$.
- if the complement of G is disconnected, then
  - label the node x of T by 1,
  - split G into co-components $G_1, \ldots, G_k$,
  - add to the node x k child vertices $x_1, \ldots, x_k$.

Step 2: For each $i = 1, \ldots, k$, go to Step 1 with $G := G_i$ and $x := x_i$.

A labelled tree constructed in this way is called a *co-tree* representing G.

Every leaf node of this tree (i.e. a node without a child) corresponds to a vertex of G and is labelled by this vertex. Every internal node x (i.e. a node with a child) corresponds to the subgraph G_x of G induced by the leaves of the tree rooted at x. The label of x shows how G_x is composed of the subgraphs of G corresponding to the children of x. If the label of x is 0, then G_x is the union of the subgraphs defined by the children. If the label of x is 1, then G_x is the join of the subgraphs defined by the children: that is, we form the union and add an edge between every two vertices belonging to different subgraphs.

An equivalent way of describing the cograph formed from a cotree is that two vertices are connected by an edge if and only if the lowest common ancestor of the corresponding leaves is labeled by 1.

If we require the labels on any root-leaf path of this tree to alternate between 0 and 1, this representation is unique [4].

The co-tree representation of a cograph has many interesting and important applications. Let us briefly discuss two of them.

The first of them deals with *coding of graphs*, i.e. representing graphs by words in a finite alphabet, which is important for representing graphs in computer memory. For general labelled graphs, we need one bit of information for each pair of vertices, which requires $n(n-1)/2$ bits for n-vertex graphs. In case of cographs, the label of each internal node of a co-tree represents the adjacency of more than just one pair of vertices, which allows a more compact coding of cographs.

Exercises

- Show that a labelled P_4-free graph with n vertices can be represented by a binary word that uses $O(n \log_2 n)$ bits of information, i.e. at most $cn \log_2 n$ bits for a constant c.

The second example, deals with the independent set problem, i.e. the problem of finding in a graph an independent set of maximum cardinality. In general, this problem is computationally hard. For cographs, it allows an efficient implementation which is based on the following facts.

Exercises

- $\alpha(G + H) = \alpha(G) + \alpha(H)$.
- $\alpha(G \times H) = \max(\alpha(G), \alpha(H))$.

If the label of an internal node x of the co-tree T is 0, then we know that a maximum independent set in G_x (the subgraph of G corresponding to x) is the union of maximum independent sets of $G_{x_1}, \ldots, G_{x_k}$ (the subgraphs of G corresponding to the children of x). If the label of x is 1, then a maximum independent set in G_x is a maximum independent set in one of the graphs $G_{x_1}, \ldots, G_{x_k}$.

7.2 Modules and modular decomposition

Let G be a graph, U a subset of its vertices and $v \in V(G)$ a vertex not in U. We will say that v *distinguishes* U if it has both a neighbour and a non-neighbour in U.

Definition 36. A *module* in a graph is a subset of vertices indistinguishable by the vertices outside of the subset.

In other words, a module is a subset U of vertices with the property that every vertex v outside of U is either adjacent to all vertices of U, in which case we say that v is *complete* to U, or non-adjacent to all vertices of U, in which case we say that v is *anticomplete* to U.

By definition, every vertex is a module and the set of all vertices of the graph is a module. We call such modules trivial.

Definition 37. A module is *trivial* if it consists of a single vertex or includes all the vertices of the graph.

Some graphs may also have non-trivial modules. For instance, every connected component with at least two vertices in a disconnected graph is a non-trivial module. Similarly, if $\overline{G}$ is disconnected, then every co-component with at least two vertices is a non-trivial module of G.

Definition 38. A graph in which every module is trivial is called *prime*.

Exercises

- Describe prime P_4-free graphs.
- Which of the following graphs are prime: C_n, P_n, K_n, *paw*?

Let us establish several important properties of modules.

Lemma 39. *The intersection of two modules is a module.*

Proof. Let U and W be two modules and x a vertex outside U. By definition x does not distinguish U. Therefore, it does not distinguish $U \cap W$. Similarly, no vertex outside W distinguishes $U \cap W$. Therefore, no vertex outside $U \cap W$ distinguishes $U \cap W$, i.e. $U \cap W$ is a module. $\square$

Lemma 40. *If two modules U and W have a non-empty intersection, then their union is a module too.*

Proof. Consider a vertex $x \notin U \cup W$, and assume x is anticomplete to U. Then x is anticomplete to the intersection $U \cap W$, and therefore, to W and hence to $U \cup W$. Similarly, if x is complete to U, then it is complete to W and to $U \cup W$. Thus $U \cup W$ is a module. $\square$

Definition 41. A module U in G is *proper* if $U \neq V(G)$.

Theorem 42. *Let G be a graph which is connected and co-connected. Then any two maximal (with respect to set inclusion) proper modules of G are disjoint. In other words, G admits a unique partition into maximal proper modules.*

Proof. Let U and W be two maximal proper modules in G. Assume they have non-empty intersection. Then their union is a module by Lemma 40. Therefore, $U \cup W = V(G)$, since otherwise U, W are not maximal proper modules. We observe that neither $U - W$ nor $W - U$ is empty, since otherwise U, W are not maximal.

Let $x \in U - W$ and assume x is complete to W. Then $W - U$ must be complete to U, since otherwise a vertex $y \in W - U$ which has a non-neighbour $z \in U$ does not distinguish U (as having the neighbour $x \in U$ and the non-neighbour $z \in U$), contradicting the fact that U is a module. But then the complement of G is disconnected (there are no edges between U and $W - U$). Similarly, if x is anticomplete to W, then G is disconnected. This contradiction shows that U and W are disjoint. As a result, every vertex of G is contained in a unique maximal proper module, and therefore, a partition of G into maximal proper modules is unique. $\square$

Lemma 43. *Any two disjoint modules U and W are either complete or anticomplete to each other.*

Proof. Let $x \in U$ and assume first x is complete to W. Then every vertex of W is complete to U (since it has a neighbour in U). Therefore, W and U are complete to each other. Similarly, if x is anticomplete to W, then W and U are anticomplete to each other. $\square$

The above results suggest a recursive method of decomposing a graph G into modules, known as *modular decomposition*. This method generalizes the decomposition of a cograph into components and co-components and can also be described by a decomposition tree. In this tree, every leaf also represents a vertex of G, and every internal node x corresponds to the subgraph of G induced by the leaves of the tree rooted at x. Also, every internal vertex x has a label showing how the subgraph of G associated with x is composed of the subgraphs associated with the children of x.

Modular decomposition

Input: a graph G
Output: a modular decomposition tree T of G.

0. Create the root x of T
1. If $|V(G)| = 1$, then label x by the only vertex of $V(G)$ and stop.
2. If G is disconnected, partition it into connected components $M_1, \ldots, M_k$ and go to 5.
3. If co-G is disconnected, partition G into co-components $M_1, \ldots, M_k$ and go to 5.
4. If G and co-G are connected, partition G into maximal modules $M_1, \ldots, M_k$ and go to 5.
5. Construct the graph G^0 from G by contracting each M_j $(j = 1, \ldots, k)$ to a single vertex, and label x by G^0.
6. Add to the node x k children $x_1, \ldots, x_k$ and for each $j = 1, \ldots, k$, implement steps 1–5 with $G := G[M_j]$ and $x := x_j$.

7.3 Prime extensions

Every graph which is not prime is contained, as an induced subgraph, in some prime graphs. In algorithmic graph theory, it is important to find the set all minimal prime graphs containing a given non-prime graph. Sometimes this set is finite, sometimes it is infinite. Below we illustrate the process of finding the set of all minimal prime extensions of a graph G with $G = 2K_2$ and show that in this case this set is finite.

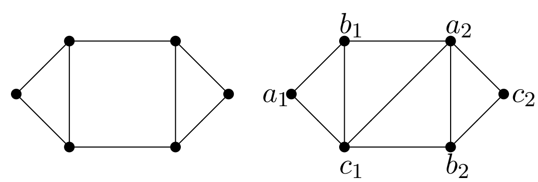

Figure 4: The graphs H_1 (left) and H_2 (right)

Theorem 44. *If a prime graph G contains a $2K_2$, then G contains a P_5 or an H_1 or an H_2.*

Proof. Let a_1, a_2, b_1, b_2 be four vertices that induce in G a $2K_2$ with edges a_1a_2 and b_1b_2. Define $A = \{a_1, a_2\}$ and $B = \{b_1, b_2\}$. As long as possible apply the following two operations:

- If G has a vertex $x \notin A \cup B$ that has a neighbour in A and no neighbour in B, add x to A.
- If G has a vertex $x \notin A \cup B$ that has a neighbour in B and no neighbour in A, add x to B.

Assume now that none of the two operations is applicable to G. Observe that both $G[A]$ and $G[B]$ are connected by construction.

Since G is prime, there must be a vertex x outside A that distinguishes A. Clearly, x is not in B. Without loss of generality, we may assume that x distinguishes two adjacent vertices $a \in A$ and $a' \in A$, since otherwise two adjacent vertices distinguished by x can be found on a path connecting a to a' in $G[A]$. We know that x must have a neighbour in B, since otherwise it must belong to A by assumption. If x distinguishes B, then a, a', x together with any two adjacent vertices of B distinguished by x induce a P_5.

Now assume that x is complete to B. Similarly, we conclude that there must be a vertex y outside $A \cup B$ that distinguishes two adjacent vertices $b \in B$ and $b' \in B$ and which is complete to A. Now if x is not adjacent to y, then a, a', b, b', x, y induce an H_1, and if x is adjacent to y, then a, a', b, b', x, y induce an H_2. $\square$

Exercises • Find all minimal prime extensions of P_3.

8 Separating cliques and chordal graphs

Definition 45. In a connected graph G, a *separator* is a subset S of vertices such that $G - S$ is disconnected.

It is not difficult to see that every connected graph G which is not complete has a separator. Indeed, if G is not complete, it must have two non-adjacent vertices, say x and y. Then obviously the set $S = V(G) - \{x, y\}$ is a separator, since $G - S$ consists of two nonadjacent vertices. Of special interest in graph theory are clique separators, also known as *separating cliques*, i.e. separators that induce a complete graph. Not every connected graph has a separating clique.

Graphs that have no separating cliques are called *irreducible*. Every graph G can be decomposed into irreducible induced subgraphs as follows. If G has a separating clique S, then decompose it into proper induced subgraphs $G_1, \ldots, G_k$ with $G_1 \cup \ldots \cup G_k = G$ and $G_1 \cap \ldots \cap G_k = S$. Then decompose $G_1, \ldots, G_k$ in the same way, and so on, until all the graphs obtained are irreducible.

Definition 46. A graph is *chordal* (or *triangulated*) if it has no chordless cycles of length at least 4.

Theorem 47. *A connected graph G which is not complete is chordal if and only if each of its minimal (with respect to set inclusion) separators is a clique.*

Proof. Assume first that G is a chordal graph. If G is not complete, it has a couple of vertices a, b that are not adjacent. Let $X \subset V(G)$ be a minimal set separating a from b, i.e. a minimal set such that in the graph $G - X$ vertices a and b belong to different connected components. Let A denote the component of $G - X$ containing a and B the component of $G - X$ containing b. Also, denote $G_A := G[A \cup X]$ and $G_B := G[B \cup X]$.

Let us show that X is a clique. Assume X contains two non-adjacent vertices s and t. By the minimality of X, both s and t have neighbours in A. Therefore, there is a (chordless) path in G_A connecting s to t all of whose vertices, except s and t, belong to A. Similarly, there is a chordless path connecting s to t in G_B with all vertices, except s and t, outside X. These two paths together create a chordless cycle of length at least 4, which is a contradiction to the fact that G is chordal. Therefore, X is a clique.

Suppose now that every minimal separator in G is a clique. Assume G has a chordless cycle $(v_1, \ldots, v_k)$ with $k \geq 4$. Let X be a minimal set separating v_1 from v_3. Obviously, X must contain v_2 and at least one vertex from $v_4, \ldots, v_k$, since otherwise $G - X$ contains a path connecting v_1 to v_3. By our assumption X is a clique. On the hand, v_2 has no neighbours among $v_4, \ldots, v_k$. This contradiction shows that G has no chordless cycles of length 4 or more, i.e. G is triangulated. □

To establish an important property of chordal graphs, let us introduce the following definition.

Definition 48. A vertex in a graph is *simplicial* if its neighbourhood is a clique.

Theorem 49. *Every chordal graph G has a simplicial vertex. Moreover, if G is not complete, then it has at least two non-adjacent simplicial vertices.*

Proof. We prove the theorem by induction on $n = |V(G)|$. The statement is obviously true for complete graphs and graphs with at most 3 vertices. Assume that it is also true for graphs with less than n vertices and let G be a non-complete graph with $n > 3$ vertices. Consider two non-adjacent vertices a and b in G and let X be a minimal (a, b)-separator. Let A denote the component of $G - X$ containing a and B the component of $G - X$ containing b. Also, denote $G_A := G[A \cup X]$ and $G_B := G[B \cup X]$.

If G_A is a complete graph, then any of vertex of A is simplicial in G. If G_A is not complete, then by the induction hypothesis it must contain at least two non-adjacent simplicial vertices. Since X is a clique (by Theorem 47), at most one of these vertices belongs to X. Therefore, at least one of them belongs to A. Obviously, this vertex is also simplicial in G. Similarly, B contains a vertex which is simplicial in G. Thus, G contains two non-adjacent simplicial vertices. $\square$

The class of chordal graphs has several important subclasses. Let us consider some of them.

8.1 Split graphs

In this section, we study the intersection of the class of chordal graphs and the class of their complements. This intersection contains neither cycles $C_4, C_5, C_6 \ldots$ nor their complements. Since the complement of C_4 is $2K_2$ and every cycle of length at least 6 contains a $2K_2$, we conclude that there are only three minimal graphs that do not belong to the intersection. In other words, the intersection coincides with the class $Free(C_4, C_5, 2K_2)$. Graphs in this class have an interesting structural property.

Lemma 50. *The vertex set of every graph G in the class $Free(C_4, C_5, 2K_2)$ can be partitioned into a clique and an independent set.*

Proof. Consider a maximal clique C in G such that $G - C$ has as few edges as possible. Our goal is to show that $S = V(G) - C$ is an independent set. To prove this, assume by contradiction that S contains an edge ab. Then either $N_C(a) \subseteq N_C(b)$ or $N_C(b) \subseteq N_C(a)$, otherwise G contains a C_4. Suppose $N_C(a) \subseteq N_C(b)$. Clearly $N_C(b)$ is a proper subset of C, because of the maximality of C. On the other hand, if $C - N_C(b)$ contains at least two vertices, say x and y, then $G[a, b, x, y] = 2K_2$. Hence we conclude that $C = N_C(b) \cup \{z\}$.

Assume there is a vertex $x \in V(G) - C$ which is adjacent to z. Then x is not adjacent to b. Indeed, if x would be adjacent to b, then we would have $N_C(b) \subseteq N_C(x)$. But then $N_C(x) = C$ that contradicts the maximality of C. It follows that x must be adjacent to a, otherwise $G[a, b, x, z] = 2K_2$. Consider a vertex y in C which is not adjacent to x (which exists because C is maximal). If y is adjacent to a, then $G[x, z, y, a] = C_4$. And if y is not adjacent to a, then $G[x, z, y, b, a] = C_5$. In both cases, we have a contradiction. Therefore, z does not have neighbours outside C. But then $C' = (C - \{z\}) \cup \{b\}$ is a maximal clique and the subgraph $G - C'$ has fewer edges than $G - C$, which contradicts the choice of C. $\square$

Definition 51. A graph G is a *split* graph if the vertices of G can be partitioned into a clique and an independent set.

From Lemma 50 we know that all $(C_4, C_5, 2K_2)$-free graphs are split. On the other hand, none of the graphs C_4, C_5 and $2K_2$ is a split graph, which can be easily seen. Therefore, the following conclusion holds.

Theorem 52. *A graph is a split graph if and only if it is $(C_4, C_5, 2K_2)$-free.*

Definition 53. Hereditary classes of graphs that can be characterized by finitely many forbidden induced subgraphs will be called *finitely defined*.

Finitely defined classes are of special interest in graph theory for many reasons. For instance, graphs in such classes can be recognized efficiently, i.e. in polynomial time. The following theorem, proved in [9], provides a sufficient condition for a hereditary class to be finitely defined.

Theorem 54. *Let P and Q be two hereditary classes of graphs and let $P \cdot Q$ denote the class of graphs whose vertices can be partitioned into a set inducing a graph from P and a set inducing a graph from Q. If both P and Q are finitely defined and there is a constant bounding the size of a maximum clique for all graphs in P and the size of a maximum independent set for all graphs in Q, then $P \cdot Q$ is finitely defined.*

The finiteness of the number of forbidden induced subgraphs for the class of split graphs follows directly from this theorem with P being the set of empty (edgeless) graphs and Q being the set of complete graphs.

EXERCISES

- Show that the class of graphs whose vertices can be partitioned into two parts, one inducing a graph of bounded vertex degree and the other inducing a graph whose complement is of bounded vertex degree, is finitely defined.

In addition to a nice characterization in terms of forbidden induced subgraphs, the class of split graphs also admits an interesting characterization via degree sequences. For a non-increasing degree sequence $d = (d_1, \dots, d_n)$ let us define $m(d) = \max\{i : d_i \geq i - 1\}$. For instance, for the sequence $d = (4, 4, 2, 2, 1, 1)$, we have $m(d) = 3$.

Theorem 55. *Let $d = (d_1, \dots, d_n)$ be a non-increasing degree sequence of a graph G and $m = m(d)$. Then G is a split graph if and only if*

$$\sum_{i=1}^{m} d_i - \sum_{i=m+1}^{n} d_i = m(m-1).$$

Proof. Let G be a split graph and let (A, B) be a partition of its vertex set into a clique A and an independent set B which maximizes the size of A. If A or B is empty, the result is obvious. Now let $a \in A$, $b \in B$ and denote $k = |A|$. Then clearly $deg(a) \geq k - 1$ and $deg(b) < k$. Therefore, $m = k$. Since A is a clique and B is an independent set, $\sum\limits_{i=1}^{m} d_i - \sum\limits_{i=m+1}^{n} d_i$ must be equal twice the number of edges in $G[A]$ (the subgraph of G induced by A), i.e. $m(m-1)$, which proves the theorem in one direction.

Conversely, let G be a graph with vertex set $\{v_1, \dots, v_n\}$ such that $deg(v_i) = d_i$ (for all $i = 1, \dots, n$) and $\sum\limits_{i=1}^{m} d_i - \sum\limits_{i=m+1}^{n} d_i = m(m-1)$. Let us define $A = \{v_1, \dots, v_m\}$ and $B = V(G) - A$. We split the sum $\sum\limits_{i=1}^{m} d_i$ into two parts $\sum\limits_{i=1}^{m} d_i = C + D$, where C is the contribution of the edges with both endpoints in A, while D is the the contribution of the edges exactly one endpoint of which belongs to A. Obviously, $C \leq m(m-1)$ and $D \leq \sum_{i=m+1}^{n} d_i$. Moreover, the equality $\sum_{i=1}^{m} d_i - \sum_{i=m+1}^{n} d_i = m(m-1)$ is valid if and only if $C = m(m-1)$ and $D = \sum_{i=m+1}^{n} d_i$. The first of the two last equalities means that A is a clique, while the second means that B is an independent set. Therefore, G is a split graph. □

EXERCISES

- Determine if the graphs given by the following degree sequences are split: $(4,3,3,2,1,1)$, $(3,3,3,3)$, $(2,2,2,2)$, $(1,1,1,1)$.
- Prove that a graph G is split if and only if its complement $\overline{G}$ is split.
- Let $G = (A, B, E)$ be a split graph with clique A and independent set B. Is it true that if G self-complementary, then $|A|$ necessarily equals $|B|$?
- Find all regular split graphs.
- Find all split graphs that are trees.
- Which number is larger: the number of n vertex labelled trees or the number of n vertex split graphs?
- What is the relationship between the class of split graphs and $Free(P_4, 2K_2, C_4)$?

8.2 Threshold graphs

An important subclass of split graphs is known in the literature under the name *threshold graphs*. To define this class, let us introduce the following notion. Given a graph G with n vertices $V = \{v_1, \ldots, v_n\}$ and a subset $U \subseteq V$, the *characteristic vector* of U is a binary vector $(\beta_1, \ldots, \beta_n)$ such that $\beta_i = 1$ if $v_i \in U$ and $v_i = 0$ otherwise.

Definition 56. A graph G with n vertices $V = \{v_1, \ldots, v_n\}$ is a *threshold graph* if there exist real numbers $a_1, \ldots, a_n$ and b such that the 0-1 solutions of $\sum_i a_i x_i \leq b$ are precisely the characteristic vectors of independent sets in G.

Example: let G be a graph with the degree sequence $(3,2,2,1)$. With each vertex v_i of G we associate the number a_i equal the degree of the vertex, and let $b = 3$. It is not difficult to verify that the 0-1 solutions of $\sum_i a_i x_i \leq b$ are precisely the characteristic vectors of independent sets in G. Therefore, G is threshold.

Claim 57. *The graphs C_4, P_4 and $2K_2$ are not threshold.*

Proof. Let us denote the vertices of a C_4 by v_1, v_2, v_3, v_4 listed along the cycle. By deleting the edge v_1v_2 we obtain a P_4 and by deleting from this P_4 the edge v_3v_4 we obtain a $2K_2$. Assume there is an inequality

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 \leq b$$

such that the 0-1 solutions of this inequality are precisely the independent sets of one of these graphs. Then, since the sets $\{v_1, v_3\}$ and $\{v_2, v_4\}$ are independent and the sets $\{v_1, v_4\}$ and $\{v_2, v_3\}$ are not independent in all three graphs, we have

$$a_1 + a_3 \leq b, \quad a_2 + a_4 \leq b, \quad a_1 + a_4 > b, \quad a_2 + a_3 > b.$$

By adding the first two of these inequalities we obtain $\sum_i a_i \leq 2b$ and by adding the last two we obtain $\sum_i a_i > 2b$. This contradiction shows that the graphs C_4, P_4 and $2K_2$ are not threshold. □

Claim 58. *If G is a threshold graph, then any induced subgraph of G is a threshold graph.*

Proof. Let $a_1, \ldots, a_n$ and b be real numbers such that the 0-1 solutions of $\sum_i a_i x_i \leq b$ are precisely the characteristic vectors of independent sets in G, and let H be a subgraph of G induced by a set $S \subseteq V(G)$. Then the independent sets of H are precisely the 0-1 solutions of

$$\sum_{i:v_i \in S} a_i x_i \leq b.$$

□

From the above two claims we conclude that the class of threshold graphs is hereditary and it is a subclass of $Free(C_4, P_4, 2K_2)$. In fact, the inverse inclusion is also valid, i.e. every graph in $Free(C_4, P_4, 2K_2)$ is threshold. We leave the proof of this fact as an exercise below and summarize the above discussion in the following theorem, which was originally proved by Földes and Hammer [6].

Theorem 59. *A graph G is a threshold graph if and only if G is $(C_4, P_4, 2K_2)$-free.*

Exercises

- Show that every graph in the class $Free(C_4, P_4, 2K_2)$ has either a dominating vertex (i.e. a vertex which is adjacent to all other vertices of the graph) or an isolated vertex (i.e. a vertex of degree 0).
- Show that if G is a threshold graph then by adding to it either a dominating vertex or an isolated vertex we obtain again a threshold graph.
- Show that there are 2^{n-1} pairwise non-isomorphic threshold graphs on n vertices. Hint: establish a bijection between threshold graphs and 0-1 sequences.
- Prove that a graph G is threshold if and only if its complement $\overline{G}$ is threshold.
- Find all self-complementary threshold graphs.
- Show that the graph H_n shown in Figure 5 is an n-universal threshold graph, i.e. it contains every threshold graph with n vertices as an induced subgraph.

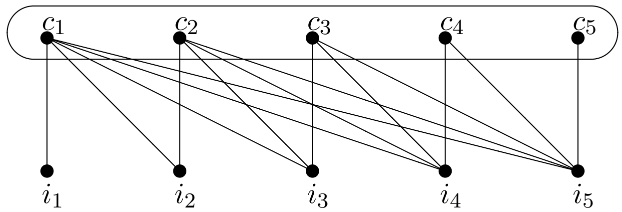

Figure 5: The H_n with $n = 5$ (the oval represents a clique)

9 Bipartite graphs

Definition 60. A graph G is *bipartite* if the vertex set of G can be partitioned into at most 2 independent sets.

EXERCISES

- Is the class of bipartite graphs hereditary?
- Which of the following graphs are bipartite: P_n, K_n, C_n, Q_n?

Bipartite graphs are also called 2-colorable graphs, in which case the two parts (independent sets) are called *color classes*. It is not difficult to see that for a connected bipartite graph there is a unique partition of its vertices into two independent sets (parts, color classes). If a bipartite graph is not connected, it admits more than one bipartition. A bipartite graph G with a given bipartition $A \cup B$ (i.e. a graph whose vertices are partitioned into two color classes A and B) will be denoted $G = (A, B, E)$.

In many respects, bipartite graphs are similar to split graphs. Indeed, the vertices of graphs from both classes can be partitioned into two subsets, and what is important is the adjacency of vertices belonging to *different* parts of the graph. This implies in particular that the number of n-vertex labeled graphs in these classes is asymptotically the same. On the other hand, there are several fundamental differences between these two classes. One of them is that the class of split graphs is characterized by finitely many forbidden induced subgraphs (Theorem 52), while the class of bipartite graphs has an infinite forbidden induced subgraph characterization. The latter fact is due to König.

Theorem 61. *A graph G is bipartite if and only if it contains no cycles of odd length.*

Proof. Let G be a bipartite graph with a bipartition $V_1 \cup V_2$ and assume $C = (v_1, v_2, \ldots, v_k)$ is a cycle, i.e., v_i is adjacent to v_{i+1} $(i = 1, \ldots, k-1)$ and v_k is adjacent to v_1. Without loss of generality let $v_1 \in V_1$. Vertex v_2 is adjacent to v_1 and hence it cannot belong to V_1, i.e., $v_2 \in V_2$. Similarly, $v_3 \in V_1$, $v_4 \in V_2$ and so on. In general, v_j with odd j belongs to V_1 and v_j with even j belongs to V_2. Since v_k is adjacent to v_1, it must be that k is even. Hence C is an even cycle.

Conversely, let G be a graph without odd cycles. Without loss of generality we assume that G is connected, for if not, we could treat each of its connected components separately. Let v be a vertex of G and define V_1 to be the set of vertices of G of an odd distance from v, i.e., $V_1 = \{u \in V(G) \mid dist(u, v) \text{ is odd}\}$. Also, let V_2 be the set of remaining vertices, i.e., $V_2 = V(G) - V_1$.

Consider two vertices x, y in V_1, and let P^x and P^y be two shortest paths connecting v to x and y respectively. Obviously, P^x and P^y have v in common. If it is not the only common vertex for P^x and P^y, denote by v' the common vertex of these two paths which is closest to x and y (the "last" common vertex on the paths). Since P^x and P^y have the same parity (they both are odd), the paths connecting v' to x and y are also of the same parity. Therefore, x is not adjacent it y, since otherwise the paths connecting v' to x and y together with the edge xy would create an odd cycle.

We proved that no two vertices of V_1 are adjacent. Similarly, one can prove that no two vertices of V_2 are adjacent. Thus, $V_1 \cup V_2$ is a bipartition of G. $\square$

EXERCISES

- What is the intersection of the class of bipartite graphs and the class of chordal graphs?
- What is the intersection of the class of bipartite graphs and the class of P_3-free graphs?
- What is the intersection of the class of bipartite graphs and the class of P_4-free graphs?

Definition 62. A bipartite graph $G = (A, B, E)$ is *complete bipartite* if every vertex of A is adjacent to every vertex of B. A compete bipartite graph with parts of size n and m is denoted $K_{n,m}$.

EXERCISES

- Prove that an n-vertex graph with more than $n^2/4$ edges is not bipartite.
- Find all n-vertex bipartite graphs with $n^2/4$ edges.
- Prove that for a non-empty regular bipartite graph the number of vertices in both parts is the same.
- Does there exist a bipartite graph G with $\delta(G) + \Delta(G) > |V(G)|$?
- Characterize the bipartite graphs with the following property: for every pair of non-adjacent vertices there is a vertex adjacent to both of them.

Exercises: Let X be the class of P_4-free graphs whose complement is bipartite.

(i) Show that X is a hereditary class.

(ii) Find all minimal forbidden induced subgraphs for X.

(iii) Describe the structure of graphs in X whose complement is connected.

Solution.

(i) We know that both P_4-free graphs and bipartite graphs are hereditary classes. Therefore, the class of complements of bipartite graphs is hereditary and hence P_4-free graphs whose complement is bipartite form a hereditary class.

(ii) We know that bipartite graphs are precisely the graphs containing no odd cycles as induced subgraphs. Therefore, the class of their complements is the class of graphs containing no complements of odd cycles as induced subgraphs. Thus, the class X can be characterized by forbidding P_4 and the complements of odd cycles as induced subgraphs. However, not all of them are minimal forbidden graphs for this class. Indeed, since P_4 is a self-complementary graph, the complement of each odd cycle of length at least 5 contains a P_4. Therefore, the set of minimal forbidden induced graphs for the class X consists of P_4 and the complement of C_3.

(iii) Let G be a graph in X with $\overline{G}$ being connected. We know that $\overline{G}$ is a bipartite graph. We also know that $\overline{G}$ is P_4-free, since P_4 is self-complementary. Then $\overline{G}$ is a complete bipartite graph. Indeed, if $\overline{G}$ contains two non-adjacent vertices in opposite parts of $\overline{G}$, then the shortest path between them must contain at lest 3 edges, in which case a P_4 arises. Therefore, $\overline{G}$ is a complete bipartite graph, and hence G is the disjoint union of two cliques.

9.1 Chain graphs

Let us say that a set of vertices in a graph forms a chain if for any two vertices x, y in this set either $N(x) \subseteq N(y)$ or $N(y) \subseteq N(x)$. In other words, a set A is a chain if the vertices of this set can be linearly ordered under inclusion of their neighbourhoods, i.e. $A = \{a_1, a_2, \ldots, a_k\}$ with $N(a_1) \subseteq N(a_2) \subseteq \ldots \subseteq N(a_k)$.

Definition 63. A bipartite graph $G = (A, B, E)$ is a *chain* graph if the vertices in each part of the graph form a chain.

Exercises

- Is the class of chain graphs hereditary?
- Show that the class of chain graphs is precisely the class of $2K_2$-free bipartite graphs.
- Show that a connected P_5-free graph is $2K_2$-free.
- Show that if in a chain graph $G = (A, B, E)$ we replace one of the parts with a clique, then the resulting graph will be a threshold graph, and vice versa.

Theorem 64. *A chain graph $G = (A, B, E)$ with at least three vertices is prime if and only if $|A| = |B|$ and for each $i = 1, \ldots, |A|$ each part of the graph contains exactly one vertex of degree i.*

Proof. The sufficiency of the statement can be easily checked by inspection. Now assume that G is prime and has at least 3 vertices. Then it must be connected. As a result, each vertex of A has degree at least 1. Also, it is obvious that every vertex of A has degree at most $|B|$. This implies that if $|A| < |B|$, then A has a couple of vertices of the same degree, and therefore, of the same neighbourhood. But then these two vertices create a non-trivial module. However, this is not possible, since G is prime. Therefore, $|A| \geq |B|$. Similarly, $|B| \geq |A|$. As a result $|A| = |B|$. As before, from the primality of G we know that no two vertices of A have the same degree, and similarly for B. Therefore, for each $i = 1, \ldots, |A|$ each part of the graph contains exactly one vertex of degree i. $\square$

EXERCISES: Show that the only prime graphs with odd number of vertices in the class $Free(P_5, K_3)$ are K_1 and C_5.

Solution:

First, we show that prime graphs in this class which are C_5-free are chain graphs, i.e. $2K_2$-free bipartite graphs. Let G be a (P_5, K_3)-free graph which does not contain a C_5. Then G is bipartite, since by forbidding P_5 we also forbid all cycles of length at least 6. Since G is prime, it must be connected. A connected P_5-free bipartite graphs must be $2K_2$-free. Indeed, if G is connected and contains a $2K_2$, then a path connected the two edges of the $2K_2$ must have at least 2 edges. The two edges of the path together with the two edge of the $2K_2$ create an chordless path on P_5 vertices, which is forbidden. We know (from the lectures) that a prime $2K_2$-free bipartite graph with at least three vertices have the same number of vertices in both parts. Therefore, the only prime graph in this class which is C_5-free and has an odd number of vertices is K_1.

Second, we show that the only prime graph in this class containing a C_5 is the C_5 itself. Assume that G is a prime (P_5, K_3)-free graph containing a $C_5 = (1, 2, 3, 4, 5)$ and suppose G is not equal to the C_5. Therefore, it must contain a vertex v that has a neighbour in the C_5 (since a prime graph is necessarily connected). Suppose, without loss of generality, that 3 is a neighbour of v. Then v must have at least one more neighbour on the C_5, since otherwise a P_5 arises. Clearly, v is adjacent neither to 2 nor to 4, because G is K_3-free. For the same reason, v cannot be adjacent to both 1 and 5. Therefore, we conclude that v has exactly two neighbours, say 1 and 3. Now we have two C_5's: $(1, 2, 3, 4, 5)$ and $(1, v, 3, 4, 5)$. This implies that there must exist a vertex, say u, distinguishing v and 2, since otherwise $\{2, v\}$ is a nontrivial module. Suppose u is adjacent to 2 but not to v. Applying the arguments similar to those above, we conclude that u must have exactly one more neighbour on the cycle $(1, 2, 3, 4, 5)$. But then u has exactly one neighbour on the cycle $(1, v, 3, 4, 5)$, which is impossible due to P_5-freeness of G. This contradiction shows that $G = C_5$.

Denote by B_n the prime chain graph with n vertices in each part. An example of this graph for $n = 5$ is given in Figure 6.

Theorem 65. *The graph B_n is an n-universal chain graph, i.e. it contains all n-vertex graphs with n vertices as induced subgraphs.*

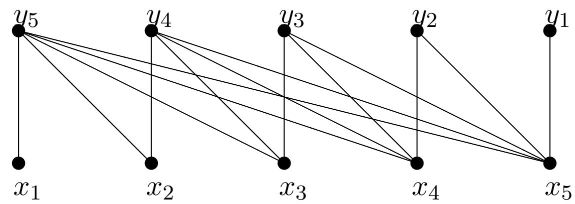

Figure 6: The graph B_5

Proof. We prove the theorem by induction on n. Obviously, K_1 is an induced subgraph of B_1. Assume the theorem is true for every chain graph with less than n vertices and let $G = (A, B, E)$ be a chain graph with n vertices. We denote the vertices of G by $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ and assume $N(x_i) \subseteq N(x_{i+1})$ and $N(y_i) \subseteq N(y_{i+1})$ (see Figure 6 for an illustration). Consider a vertex $v \in A$ with a largest neighbourhood.

If v is adjacent to all vertices of B, then G can be embedded into B_n as follows. Vertex v is mapped to x_n, while the rest of G can be embedded into the B_{n-1} obtained from B_n by deleting x_n and y_1 (which follows by the induction assumption).

If v has a non-neighbour in B, say v is not adjacent to $u \in B$, then u must be isolated in G. Then G can be embedded into B_n as follows. Vertex u is mapped to y_1, while the rest of G can be embedded into the B_{n-1} obtained from B_n by deleting x_n and y_1 (by induction). □

9.2 Bipartite complement and related classes

Definition 66. The *bipartite complement* of a bipartite graph $G = (A, B, E)$ is the bipartite graph $G = (A, B, (A \times B) - E)$.

EXERCISES

- Show that the bipartite complement of a chain graph is again a chain graph.
- Show that the bipartite complement of a P_7-free bipartite graph is again a P_7-free bipartite graph.
- Find the bipartite complement of C_6, $3K_2$.

Definition 67. A bipartite graph G is a *bi-complement reducible graph* if any induced subgraph of G with at least two vertices is either disconnected or the bipartite complement of a disconnected graph.

The class of bi-complement reducible graphs is a bipartite analog of cographs (complement reducible graphs). This class has been characterized in [7] by means of forbidden induced subgraphs as follows.

Theorem 68. *A bipartite graph is bi-complement reducible if and only if it is* $(P_7, S_{1,2,3}, Sun_4)$*-free, where* $S_{1,2,3}$ *and* Sun_4 *are the graphs represented in Figure 7.*

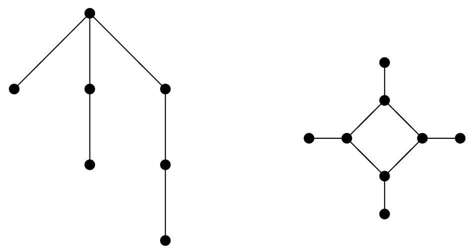

Figure 7: Graphs $S_{1,2,3}$ (left) and Sun_4 (right)

EXERCISES

- Let Sun_1 be the graph obtained from Sun_4 by deleting 3 vertices of degree 1. Show that every connected Sun_1-free bipartite graph is either complete bipartite or C_4-free.

(*) Let Sun_2 be the graph obtained from Sun_4 by deleting 2 vertices of degree 1 of distance 3 from each other. Show that every prime Sun_2-free bipartite graph is the bipartite complement of a regular graph of degree 1.

EXERCISES

- Show that if one part in a bipartite graph forms a chain, then the other forms a chain too.
- If one part of a bipartite graph can be partitioned into two chains, does it mean that the other admits a partition into two chains too?

Definition 69. A *bichain* graph is a bipartite graph in which the vertices in each part can be partition into at most two chains.

EXERCISES

- Is the class of bichain graphs hereditary?
- Show that the bipartite complement of a bichain graph is again a bichain graph.

Theorem 70. *A bipartite graph is a bichain graph if and only if it is* $(3K_2, C_6, P_7)$*-free.*

Proof. To prove the theorem, we will show that a bipartite graph $G = (U, V, E)$ with a bipartition $U \cup V$ is $(3K_2, C_6, P_7)$-free if and only if the vertices in each part of the graph can be partition into at most two chains.

One direction of the proof is simple, because at least one part in each of the graphs P_7, C_6 and $3K_2$ contains three vertices whose neighbourhoods are incomparable.

Now assume that G is $(3K_2, C_6, P_7)$-free. Suppose, for contradiction, that in one part of G, say U, there exist three vertices a, b, c whose neighbourhoods are incomparable. Then, in the part V, there exists a vertex d which is adjacent to a but not b, and a vertex e which is adjacent to b but not a. We will split the proof into three cases:

Case 1. Suppose c is adjacent to both d and e. Then there must exist a vertex f which is adjacent to a but not c. Vertex b must be non-adjacent to f, otherwise $afbecd$ would form an induced C_6. So there must exist a vertex g which is adjacent to b but non-adjacent to c. Again, to avoid an induced C_6, the vertex g must also be non-adjacent to a. But then $fadcebg$ forms an induced P_7, which is a contradiction.

Case 2. Now suppose that c is adjacent to exactly one of d and e, say e. Then there must exist a vertex g which is adjacent to b but not c, and a vertex h which is adjacent to c but not b. If a were adjacent to neither of g and h, then $adbgch$ would form an induced $3K_2$. If a were adjacent to exactly one of g and h, say g, then $dagbech$ would form an induced P_7. Finally, if a is adjacent to both g and h, then $agbech$ would form an induced C_6. Each of these possibilities is a contradiction.

Case 3. Finally, suppose c is non-adjacent to d and e. Then there must exist a vertex h which is adjacent to c but not b, and a vertex i which is adjacent to c but not a. Note that h and i must not be the same vertex, since otherwise $adbech$ would form an induced $3K_2$. The vertex a must be adjacent to h, since otherwise $adbech$ would form an induced $3K_2$. Similarly, b must be adjacent to i. Now $dahcibe$ forms an induced P_7, which is a contradiction.

We have exhausted all possible cases, each leading to a contradiction. Thus our proof is complete. □

9.3 Chordal bipartite graphs

Definition 71. A bipartite graph G is *chordal bipartite* if it contains no chordless cycles of length at least 6.

The class of chordal bipartite graphs is a bipartite analog of chordal (triangulated) graphs, and it inherits many properties of chordal graphs. For instance, the fact that every minimal separator in a chordal graphs is a clique can be rephrased in the case of chordal bipartite graphs as follows.

A subset $S \subset V(G)$ of a connected bipartite graph G is an *edge separator* if there are edges $e, e' \in E(G)$ that lie in different connected components of $G - S$. A bipartite graph is *separable* if it contains an edge separator.

Theorem 72. *A bipartite graph G is chordal bipartite if and only if every minimal edge separator of G induces a complete bipartite graph.*

Also, an edge $xy \in E(G)$ of a bipartite graph G is *bisimplicial* if $N(x) \cup N(y)$ induces a complete bipartite graph. The notion of a bisimplicial edge is a bipartite analog of the notion of a simplicial vertex.

Theorem 73. *Every separable chordal bipartite graph contains at least two separable bisimplicial edges.*

This theorem provides an analog of the fact that every non-complete chordal graph contains at least two non-adjacent simplicial vertices.

Exercises

- Let G be a graph with vertices $v_1, \ldots, v_n$. Define $B(G)$ be the bipartite graph with $2n$ vertices $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ such that x_i is adjacent to y_j if and only if either $i = j$ or $v_i v_j \in E(G)$. Show that if $B(G)$ is chordal bipartite, then G is chordal. Is the converse true?

10 Graph parameters

10.1 Graph lettericity

Definition 74. A *k-letter graph* G is a graph defined by a finite word $x_1x_2\dots x_n$ over an alphabet X of size k together with a subset $S \subseteq X^2$ such that:

- $V(G) = \{x_1, x_2, \dots, x_n\}$
- $E(G) = \{x_ix_j : i \leq j \text{ and } (x_i, x_j) \in S\}$

EXERCISES

- Let k be a fixed constant. Is the class of k-letter graphs hereditary? Is it finitely defined?
- Show that chain graphs and threshold graphs are 2-letter graphs.
- Show that every graph is a k-letter graph for some $k \leq |V(G)|$.

The *lettericity* of a graph G, denoted $\ell(G)$, is the minimum k such that G is a k-letter graph.

EXERCISES

- Show that $\ell(G) = \ell(\overline{G})$.

Theorem 75. $\ell(nK_2) = n$.

Proof. First, it is not difficult to see that $\ell(nK_2) \leq n$, since n letters suffices (one letter per edge). Assume $\ell(nK_2) < n$, then there must exist a letter a representing at least 3 vertices of the graph. Clearly, $(a, a) \notin S$, since otherwise a triangle arises. Then the neighbour of the middle a is different from a, say b. If this neighbour appears before the middle a, it must also be adjacent to the last a. If it appears after the middle a, it must also be adjacent to the first a. In both case, b has at least two neighbours. Therefore, $\ell(nK_2) \geq n$. □

10.2 Tree-width

As we have seen, trees have many important properties. It is therefore natural to ask to what degree these properties can be transferred to more general graphs. This question is partially answered with the help of the notion of tree-width. To define this notion, let us first define the notion of tree-decomposition.

Let G be a graph and T a tree. Let $\mathcal{V} = (V_t)_{t\in T}$ be a family of vertex sets $V_t \subseteq V(G)$ indexed by the vertices t of T. The pair $(T, \mathcal{V})$ is called a tree decomposition of G if it satisfies the following three conditions:

(T1) $V(G) = \cup_{t\in T} V_t$,

(T2) for every edge e of G, there exists a $t \in T$ such that both endpoints of e lie in V_t,

(T3) for any tree nodes t_1, t_2, t_3 of T such that t_2 lies on the unique path connecting t_1 to t_3 in T, we have $V_{t_1} \cap V_{t_3} \subseteq V_{t_2}$.

Conditions (T1) and (T2) together say that G is the union of the subgraphs $G[V_t]$ (i.e. subgraphs of G induced by V_t); we call these subgraphs and the sets V_t themselves the *parts* (also called *bags*) of the decomposition and we say that $(T, \mathcal{V})$ is a tree-decomposition of G into these parts. Condition (T3) implies that the parts of $(T, \mathcal{V})$ are organized roughly like a tree.

EXERCISES

- Show that if H is a subgraph of G and $(T, (V_t)_{t\in T})$ is a tree-decomposition of G, then $(T, (V_t \cap V(H))_{t\in T})$ is a tree-decomposition of H.
- Show that for any clique in a graph G and any of its tree-decompositions, there is a bag containing the clique.

Definition 76. The *width* of a decomposition $(T, \mathcal{V})$ is $\max\{|V_t| - 1 \ : \ t \in T\}$. The *tree-width* of a graph G, denoted $tw(G)$, is the least width of any tree-decomposition of G.

The notion of tree-width can also be defined in a completely different way as follows. Let us call a graph H a *triangulation* of a graph G if H is chordal (triangulated) and G is a spanning subgraph of H, i.e. $V(G) = V(H)$ and $E(G) \subseteq E(H)$. Then

Theorem 77. *The tree-width of G is $\min\{\omega(H) - 1 \ : \ H$ is a triangulation of $G\}$, where $\omega(H)$ is the size of a maximum clique in H.*

Proof. The proof of this result is based on the following claim.

> *G is a chordal graph if and only if it has a tree-decomposition with each bag being a clique.*
>
> We prove this claim by induction on $n = |V(G)|$. Assume first that G has a tree-decomposition with each bag being a clique, and let $(T, \mathcal{V})$ be such a decomposition with minimum number of bags (vertices of T). If this number is at most 1, then G is complete, and hence chordal. So, assume T has an edge t_1t_2. The deletion of this edge splits T into two subtrees T_1 and T_2 (with $t_i \in T_i$). For $i = 1, 2$, we denote $U_i := \cup_{t\in T_i} V_t$ and $G_i := G[U_i]$. Clearly $(T_i, (V_t)_{t\in T_i})$ is a tree-decomposition of G_i with each bag being a clique. Therefore, by the induction hypothesis, we know that G_i is chordal. Also, according to conditions (T1) and (T2) of the definition of tree-decomposition, $G = G_1 \cup G_2$ and $V(G_1 \cap G_2) = V_{t_1} \cap V_{t_2}$ Since V_{t_1} and V_{t_2} are cliques, $G_1 \cap G_2$ is a complete graph. Together with the fact that G_1 and G_2 are chordal this implies that G is chordal too.
>
> Conversely, assume G is chordal. If it is complete, there is nothing to prove. Otherwise, by Theorem 47, it has a separating clique X. Let $A_1, \ldots, A_k$ be the connected components of $G - X$. We denote $G_1 = G[X \cup A_1]$ and $G_2 = G[X \cup A_2 \cup \ldots \cup A_k]$. By induction hypothesis, each of G_i has a tree-decomposition $(T_i, \mathcal{V}_i)$ with each bag being a clique. We know (from one of the previous exercises) that each of the tree-decompositions has a bag containing X. Let t_i be a vertex of T_i $(i = 1, 2)$ representing the bag containing X. Then it is not difficult to check that $(T_1 \cup T_2) + t_1t_2, \mathcal{V}_1 \cup \mathcal{V}_2)$ is a tree-decomposition of G with each bag being a clique. This completes the proof of the claim.

Now we turn to the proof of the theorem. Let H be a triangulation of G. Since G is a subgraph of H, any tree-decomposition of H is also a tree-decomposition of G, and hence $tw(G) \leq tw(H)$. Since every clique of H is contained in one of the bags of its tree-decomposition, we know that $tw(H) \leq \omega(H) - 1$. Therefore, $tw(G) \leq \omega(H) - 1$ for any triangulation H of G. In particular,

$$tw(G) \leq \min\{\omega(H) - 1 \ : \ H \text{ is a triangulation of } G\}.$$

Conversely, let $(T, \mathcal{V})$ be a tree-decomposition of G of width $tw(G)$. If this tree-decomposition has a bag, which is not a clique, we complete it to a clique. This transforms G into a new graph H with the same tree-decomposition. In this decomposition each bag of H is a clique, and hence, by the above claim, H is a chordal graph. Therefore, H is a triangulation of G. Since $(T, \mathcal{V})$ is a tree decomposition of H and every clique of H is contained in some bag of the decomposition, $\omega(H)$ is not larger than the size of a maximum bag in $(T, \mathcal{V})$, i.e. $\omega(H) \leq tw(G) + 1$. Thus,

$$tw(G) \geq \omega(H) - 1 \geq \min\{\omega(H) - 1 \ : \ H \text{ is a triangulation of } G\}.$$

□

EXERCISES

- Determine the tree-width of a complete bipartite graph $K_{n,n}$.
- Show that in the class of chordal graphs of bounded vertex degree the tree-width is bounded by a constant.

10.3 Clique-width

The *clique-width* of a graph G is the minimum number of labels needed to construct G using the following four operations:

(i) Creation of a new vertex v with label i (denoted by $i(v)$).

(ii) Disjoint union of two labeled graphs G and H (denoted by $G \oplus H$).

(iii) Joining by an edge each vertex with label i to each vertex with label j ($i \neq j$, denoted by $\eta_{i,j}$).

(iv) Renaming label i to j (denoted by $\rho_{i \to j}$).

Every graph can be defined by an algebraic expression using these four operations. For instance, a chordless path on five consecutive vertices a, b, c, d, e can be defined as follows:

$$\eta_{3,2}(3(e) \oplus \rho_{3\to2}(\rho_{2\to1}(\eta_{3,2}(3(d) \oplus \rho_{3\to2}(\rho_{2\to1}(\eta_{3,2}(3(c) \oplus \eta_{2,1}(2(b) \oplus 1(a))))))))).$$

Such an expression is called a *k-expression* if it uses at most k different labels. Thus the clique-width of G, denoted $\mathrm{cw}(G)$, is the minimum k for which there exists a k-expression defining G. For instance, from the above example we conclude that $\mathrm{cw}(P_5) \leq 3$. A k-expression defining a graph G can also be represented in a natural way by a tree the leaves of which correspond to the creation of vertices of G (see Figure 8 for an illustration).

Theorem 78. $cw(\overline{G}) \leq 2cw(G)$.

Theorem 79. $cw(G) \leq 2^{2tw(G)+2} + 1$.

Theorem 80. $cw(G) = \max\{cw(H) \ : \ H$ *is a prime induced subgraph of* $G\}$.

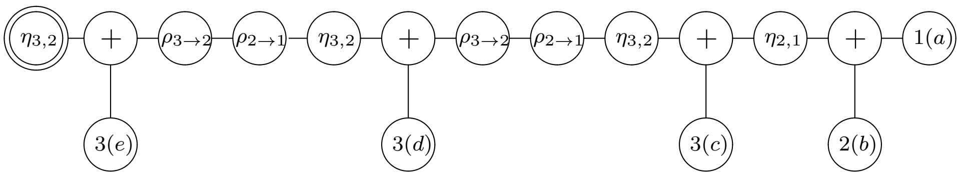

Figure 8: The tree representing the expression (1) defining a P_5

Proof. Clearly, for any induced subgraph H of G, we have $cw(H) \leq cw(G)$, because a k-expression defining H can be obtained from a k-expression defining G by omitting every operation which is not relevant for the vertices of H.

To prove the inverse inequality, we will use induction on $n = |V(G)$. If G is prime, then together with the previous paragraph we conclude that

$$cw(G) = \max\{cw(H) \ : \ H \text{ is a prime induced subgraph of } G\}.$$

So, let G be non-prime, and let $M_1, \ldots, M_p$ be maximal non-trivial modules of G. By contracting each M_i into a single vertex m_i we obtain the characteristic graph G_0 of G. We separately construct expressions representing the graphs G_0 and $G[M_i]$ for each i, assuming by induction that each expression uses at most $\max\{cw(H) \ : \ H$ is a prime induced subgraph of $G\}$ different labels. If in the expression defining G_0 the vertex m_i is created with label j, we finish the construction of the graph $G[M_i]$ by renaming all labels to j. Then in the tree describing G_0 we replace the node creating m_i with the root of the tree creating $G[M_i]$. The resulting tree represents G and uses at most $\max\{cw(H) \ : \ H$ is a prime induced subgraph of $G\}$ labels, as required. □

Exercises

- Show that the clique-width of any cograph is at most 2.
- Show that the clique-width of any forest is at most 3.
- Show that the clique-width of a cycle is at most 4.
- Show that the clique-width of the complement of a cycle is at most 5.
- Show that for any bipartite graph G the clique-width of its bipartite complement is at most $4cw(G)$ (hint: use Theorem 79).
- Show that the clique-width of a chain graph is at most 3 (hint: use Theorems 64 and 80.

Theorem 81. *Let G be a square $n \times n$ grid with $n \geq 3$. Then $cw(G) \geq n$.*

Proof. Let A be a k-expression defining G and T the tree representing A. For each node a of T, we denote by $T(a)$ the tree rooted at a. This tree represents a subgraph of G, not necessarily induced.

Let a be a lowest $\oplus$ node of T such that the graph represented by $T(a)$ contains a full row and a full column of G (a is lowest in the sense that in the graph $T(a)$ no other node possesses this property). Let b and c be two children of a. We color all vertices of G in $T(b)$ by blue and all vertices of G in $T(c)$ by red. All the remaining vertices of G are colored white.

We may assume, by the symmetric role of rows and columns, that G contains neither blue nor red column. Indeed, if a column is blue, then no row is red (since it intersects the blue column) and no row is blue (since otherwise a is not minimal), and similarly if there would be a red column.

By the choice of a, there is a row r. For each j, we denote by $v_j := v_{i,j}$ the vertex in column j closest to $v_{r,j}$ and with $\mathit{color}(v_{i,j}) \neq \mathit{color}(v_{r,j})$. Such a vertex exists for each j, because $v_{r,j}$ is either blue or red and no column of G is completely blue or completely red. Also, for each j, we define the unique vertex u_j as follows: if $i < r$, then $u_j := v_{i+1,j}$, otherwise $u_j := v_{i-1,j}$.

By definition all vertices $u_1, \ldots, u_n$ are non-white (i.e. either red or blue). Let us show that no two of them have the same label. Consider two vertices u_{j_1} and u_{j_2}. Assume without loss of generality that v_{j_1} lies above u_{j_1}. We also may assume, without loss of generality, that the v_{j_1} is not adjacent to u_{j_2}, since if this is the case, then v_{j_1} and u_{j_2} lie in the same row, while v_{j_2} is above u_{j_2}, in which case v_{j_2} is not adjacent to u_{j_1}. Thus we have that v_{j_1} is adjacent to u_{j_1} (as being in the same column strictly above u_{j_1}) and is not adjacent to u_{j_2}. Also, by the choice of the vertices, v_{j_1} and u_{j_1} have different colors, which means that the edge $v_{j_1}u_{j_1}$ is not present in the subgraph of G represented by the tree $T(a)$. Therefore, if u_{j_1} and u_{j_2} have the same label at $T(a)$, then the creation of the edge $v_{j_1}u_{j_1}$ also creates the edge $v_{j_1}u_{j_2}$. Since the latter pair is not adjacent in G, u_{j_1} and u_{j_2} must have different labels at $T(a)$. □

11 Coverings and Matchings

11.1 Covering

Given a graph G, a vertex $v \in V(G)$ and an edge e incident to v, we will say that v and e *cover* each other.

Definition 82. For a graph $G = (V, E)$ a set $U \subseteq V$ is called a *vertex cover* if it covers all the edges of G, i.e. if every edge of G is incident to at least one vertex in U. The minimum number of vertices in a vertex cover of a graph G is called the *vertex cover number* of G and is denoted by $\beta(G)$.

Example.

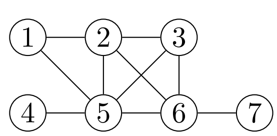

Figure 9:

Let G be a graph on Figure 9. Then

1. $\{1, 3, 4, 7\}$ is a maximum independent set of G, i.e. $\alpha(G) = 4$.
2. $\{2, 3, 5, 6\}$ is a maximum clique of G, i.e. $\omega(G) = 4$.
3. $\{2, 5, 6\}$ is a minimum vertex cover of G, i.e. $\beta(G) = 3$.

EXERCISES

- Does there exist a connected graph $G = (V, E)$ whose minimum vertex cover is V?

There is a simple relationship between independent sets and vertex covers in a graph.

Theorem 83. *Let $G = (V, E)$ be a graph. A subset U of vertices of G is a vertex cover if and only if $\overline{U} = V \setminus U$ is an independent set.*

Proof. If U is a vertex cover, then every edge is incident to (is covered by) at least one vertex of U, and therefore there are no edges connecting two vertices in $\overline{U}$, i.e. $\overline{U}$ is an independent set. Conversely, if $\overline{U}$ is an independent set, then there are no edges between vertices in $\overline{U}$, and hence each edge is covered by (incident to) at least one vertex of U, i.e. U is a vertex cover. $\square$

Corollary 84. *$\alpha(G) + \beta(G) = n$ for any n-vertex graph G.*

11.2 Matching

Let us say that two edges in a graph are *adjacent* if they have a vertex in common, i.e. if they are incident to the same vertex.

Definition 85. A set of pairwise non-adjacent edges in a graph is called a *matching*. The maximum number of edges in a matching in a graph G is called the *matching number* of G and denoted by $\mu(G)$.

EXERCISES

- Show that $\mu(G) \leq \beta(G) \leq 2\mu(G)$ for any graph $G = (V, E)$.
- Give examples of graphs where the two inequalities (above) become equalities.

How can we find a matching of maximum size in a graph G? To answer this question, let us consider an arbitrary matching M in G. We will say that the edges of M are *M-strong*, while all the other edges of G are *M-weak*. Also, we will say that a vertex $a \in V$ is *matched* (or *saturated*) (by M) if a is incident to an edge in M. Otherwise, a is *unmatched* (*unsaturated*).

Definition 86. An *alternating path* (with respect to M) is a path in which M-strong edges alternate with M-weak edges.

Definition 87. An *augmenting path* (with respect to M) is an alternating path that starts and ends in unmatched vertices.

If M is a matching and P is an augmenting path with respect to M, then it is easy to see that $M \oplus E(P) := (M \setminus E(P)) \cup (E(P) \setminus M)$ is also a matching and $|M \oplus E(P)| = |M| + 1$.

Theorem 88 (Berge, 1957)**.** *Let $G = (V, E)$ be a graph and M be a matching in G. Then M is a maximum matching if and only if there are no augmenting paths with respect to M.*

Proof. If M is a maximum matching, then there are no augmenting paths with respect to M. Indeed, if P is an augmenting path, then $M \oplus E(P)$ is a matching of larger size $|M \oplus E(P)| = |M| + 1$.

In order to prove the converse statement, consider a matching M in G and suppose that M is not maximum. We show that there is an augmenting path in G with respect to M. Let M^* be another matching in G such that $|M^*| > |M|$. Let H be a subgraph of G formed by the edges of $M \oplus M^*$, i.e. by the edges that belong to exactly one of the matchings M or M^*. Note that in the graph H every vertex is incident to at most two edges (at most one edge from M and at most one edge from M^*). In other words, H has vertex degree at most two. Therefore every connected component in H is a cycle or a path. In each of these paths and cycles edges alternate between M and M^*. Since $|M^*| \geq |M|$, there is a connected component which contains more edges from M^* than from M. Clearly, this connected component is a path in which the first and the last edges belong to M^*. Therefore, this path is augmenting with respect to M. □

11.3 Matchings vs independent sets

Definition 89. The line graph of a graph G is the graph whose vertex set is $E(G)$ with two vertices being adjacent if and only if the corresponding edges of G have a vertex in common. The line graph of G is denoted $L(G)$.

EXERCISES

- Is the class of line graphs hereditary (i.e. closed under deletion of vertices)? Is it monotone (i.e. closed under deletion of vertices and edges)?
- Show that $K_{1,3}$ is a minimal (with respect to vertex deletion) non-line graph.
- Find the number of vertices and edges of the line graph $L(G)$ of a graph G with the degree sequence $(d_1, d_2, \ldots, d_n)$.
- The wheel W_k is the graph obtained from a chordless cycle C_k ($k \geq 3$) by adding a vertex which dominates the cycle (i.e. a vertex which is adjacent to every vertex of the cycle). Determine the values of k for which W_k a line graph. In those cases where W_k is a line graph, find a graph G such that $L(G) = W_k$. In the cases where W_k is not a line graph, provide a proof.
- Derive Berge's Theorem in the terminology of independent sets.

11.4 Matchings in bipartite graph

Theorem 90 (König, 1931)**.** *If G is a bipartite graph, then $\mu(G) = \beta(G)$*

Proof. Let $G = (A, B, E)$ be a bipartite graph and M a maximum matching in G. Since every vertex cover of G must cover M, every vertex cover has at least $|M|$ vertices. Therefore every vertex cover of cardinality $|M|$ is a minimum vertex cover.

We show how to construct a vertex cover U of cardinality $|M|$. For every edge $ab \in M$, $a \in A, b \in B$, we add to U

- vertex b, if there exists an alternating path which starts in A at an unmatched vertex and ends in b;
- vertex a, otherwise.

Now let $ab \in E$ be an arbitrary edge of the graph with $a \in A$ and $b \in B$. We show that the edge ab is covered by at least one vertex of U. If $ab \in M$, then either a or b belongs to U by the definition of U. Assume now that $ab \notin M$. Since M is maximum, the case that a and b are both unmatched is impossible. So, we are left with the following three cases.

Case 1. Assume both a and b are matched: a is matched with b', while b is matched with a'. If $a \in U$, we are done. Otherwise, $b' \in U$, in which case there is an alternating path from an unmatched vertex of A to b'. Extending this path by two edges $b'a$ and ab, we conclude that the same is true for b, and hence $b \in U$.

Case 2. If a is unmatched and b is matched, then ab is an alternating path from an unmatched vertex of A to b. Therefore, $b \in U$.

Case 3. Let b be unmatched and let a be matched with a vertex b'. If $a \in U$, we are done. Otherwise, $b' \in U$, in which case there is an alternating path from an unmatched vertex of A to b'. Extending this path by two edges $b'a$ and ab, we conclude that there is an alternating path between two unmatched vertices, i.e. an augmenting path. But this contradicts the fact that M is maximum.

□

Let $G = (A, B, E)$ be a bipartite graph. Clearly, if there exists a matching covering A (i.e. covering each vertex of A), then

$$|N(S)| \geq |S| \text{ for all } S \subseteq A. \tag{1}$$

In other words, (1) is a necessary condition for the existence of a matching covering A. Philip Hall proved in 1935 that this condition is also sufficient.

Theorem 91 (Hall, 1935)**.** *G contains a matching covering A if and only if (1) holds.*

Proof. In order to prove the theorem we assume that G satisfies condition (1) and find a matching covering A. To this end, we show that for every matching M in G that leaves a vertex $a \in A$ unmatched there is an augmenting path with respect to M.

Let A^* be the set of vertices in A that can be reached from a by a non-trivial alternating path, and $B^* \subseteq B$ the set of all penultimate vertices of such paths. The last edges of these paths lie in M, so $|A^*| = |B^*|$. Hence by condition (1), there is an edge from a vertex v in $S = A^* \cup \{a\}$ to a vertex b in $B \setminus B^*$.

As $v \in A^* \cup \{a\}$, there is an alternating path P from a to v. Then there is an alternating path P^* from a to b. If b is matched, say by $ba' \in M$, then $P^* \cup \{ba'\}$ is an alternating path, and hence a' must belong to A^* and b must belong to B^*. But $b \notin B^*$, so b is unmatched, and therefore P^* is the desired augmenting path. □

Exercises
(*) Derive Hall's theorem from König's theorem.

Definition 92. A matching is called *perfect* if it covers all vertices of the graph.

Corollary 93. *Every regular bipartite graph has a perfect matching.*

Proof. Let $G = (A, B, E)$ be a bipartite graph every vertex of which has degree d. Consider an arbitrary subset $S \subseteq A$ and its neighbourhood $N(S)$. Then $|S|d \leq |N(S)|d$, because $|S|d$ is the number of edges between S and $N(S)$ (i.e. the number of edges incident to the vertices of S), while $|N(S)|d$ counts all edges between S and $N(S)$ and possibly some other edges incident to the vertices of $N(S)$ (if such edges exist). Therefore, (1) holds for G and hence there is a matching covering A. This matching is perfect, because $|A|d = |B|d$ (which is the total number of edges counted twice from each side of the graph) implying $|A| = |B|$. □

12 Eulerian tours and Hamiltonian cycles

Definition 94. A *walk* (of length k) in a graph G is a non-empty alternating sequence

$$v_0 e_0 v_1 e_1 \ldots e_{k-1} v_k$$

of vertices and edges in G such that $e_i = v_i v_{i+1}$ for all $i < k$. If $v_0 = v_k$, the walk is closed.

A path is a walk in which all vertices are distinct. A cycle is a closed walk in which all internal vertices are distinct.

12.1 Eulerian tours

Definition 95. A closed walk in a graph is called an *Eulerian tour* if it contains every edge of the graph exactly once.

Theorem 96 (Euler, 1736)**.** *A connected graph has an Eulerian tour if and only if every vertex has even degree.*

Proof. The degree condition is clearly necessary, since each time the Euler tour passes a vertex it uses exactly two edges: one is used to enter the vertex and one to leave it.

Conversely, let G be a connected graph with all degrees even, and let

$$W = v_0 e_0 \ldots e_{l-1} v_l$$

be a longest walk in G using no edge more than once. Since W cannot be extended, it already contains all the edges incident to v_l. By assumption, the number of such edges is even. Hence $v_l = v_0$ (why?), so W is a closed walk.

Suppose W is not an Eulerian tour. Then G has an edge e outside of W. Moreover, since G is connected, we may assume that e incident to a vertex of W, say $e = uv_i$. Then the walk

$$u e v_i e_i \ldots e_{l-1} v_l e_0 \ldots e_{i-1} v_i$$

is longer than W, a contradiction. □

12.2 Hamiltonian cycles

Definition 97. Let G be a graph with $n \geq 3$ vertices. A cycle that contains every vertex of G is called *Hamiltonian.*

> EXERCISES
>
> - Let G be a graph with n vertices. Show that if the degree of each vertex of G is at least $n/2$, then G is connected.

Theorem 98 (Dirac, 1952)**.** *Every graph with $n \geq 3$ vertices and minimum degree at least $n/2$ has a Hamiltonian cycle.*

Proof. Let $G = (V, E)$ be a graph with $n \geq 3$ vertices and minimum degree at least $n/2$. By the above exercise, G is connected.

Let $P = v_0 \dots v_k$ be a longest path in G. By the maximality of P, all the neighbours of v_0 and all the neighbours of v_k lie on P. Hence for at least $n/2$ indices $0 \leq i \leq k-1$, we have $v_i v_k \in E$ and for at least $n/2$ indices $1 \leq i \leq k$, we have $v_0 v_i \in E$. The second part of the last sentence can be rephrased as follows: for at least $n/2$ indices $0 \leq i \leq k-1$, we have $v_0 v_{i+1} \in E$. Since $k < n$, by the pigeon hole principle, there is an index $0 \leq i \leq k-1$ such that both $v_i v_k \in E$ and $v_0 v_{i+1} \in E$.

We claim that the cycle

$$C = v_0 v_{i+1} v_{i+2} \dots v_{k-1} v_k v_i v_{i-1} \dots v_1 v_0$$

is Hamiltonian. Assume the contrary. Then, since G is connected, there exists vertex v_j of C adjacent to a vertex y not in C. But then the path from y to v_j and then around C is longer than P. This contradicts the maximality of P and hence proves the theorem. □

Theorem 99. *Let G be a connected graph G with at least 3 vertices. If for every induced path abc,*

$$deg(a) + deg(c) \geq |N(a) \cup N(b) \cup N(c)|,$$

then G has a Hamiltonian cycle.

Proof. Consider any induced path abc in G. Since $deg(a)+deg(c) = |N(a) \cup N(c)| + |N(a) \cap N(c)|$, our degree assumption implies that

$$|N(a) \cap N(c)| \geq |N(a) \cup N(b) \cup N(c)| - |N(a) \cup N(c)| = |N(b) \setminus N(\{a, c\})| \geq |\{a, c\}| \geq 2. \quad (2)$$

In particular, G contains a cycle.

Let C be a longest cycle in G. Assume that G has no Hamiltonian cycle and let $u \notin C$ be a vertex that has a neighbour on C. Denote $M = N(u) \cap V(C)$. For a vertex $v \in M$ let v^+ denote the successor of v on C along some fixed orientation of C, and let $M^+ = \{v^+ \mid v \in M\}$.

Since C is a longest cycle, we have

(*) $M \cap M^+ = \emptyset$ (i.e. u is not adjacent to any two consecutive vertices of C);

(**) no two vertices of $M^+ \cup \{u\}$ are adjacent or have a common neighbour outside C.

In particular, the paths uvv^+ are induced. Therefore, every $v \in M$ satisfies

$$|N(u) \cap N(v^+)| \geq^{(2)} |N(v) \setminus N(\{u, v^+\})| \geq |N(v) \cap M^+| + 1.$$

The last inequality comes from the fact that, by (**), both u and the vertices of M^+ lie outside of $N(\{u, v^+\})$. Let K be the number of edges between M and M^+. This number must satisfy

$$K = \sum_{v \in M} |N(v) \cap M^+| \leq \sum_{v \in M} (|N(u) \cap N(v^+)| - 1) =^{(**)} K - |M|.$$

For the last equality note that, by (**), v^+ has all its common neighbours with u in M. By assumption, M is not empty and hence $K \leq K - |M|$ is a contradiction, which proves the theorem. □

Exercises

- Let $\kappa(G)$ denotes the number of connected components of G. Prove that if a connected graph G has a vertex set X such that $\kappa(G - X) > |X|$, then G has no Hamiltonian cycles.
- Prove that if a graph G with n vertices has a Hamiltonian cycle, then $\alpha(G) \leq n/2$.

13 Perfect Graph Theorem and related results

Definition 100. A *vertex coloring* of a graph G is an assignment of colors to the vertices of G in such a way that no two adjacent vertices receive the same color. A graph is *k-colorable* if it admits a vertex coloring with at most k colors. The minimum number of colors in a proper vertex coloring of G is the *chromatic number* of G, denoted $\chi(G)$.

In other words, vertex coloring can be viewed as partitioning the vertex set of a graph into independent sets, also called *color classes*. The minimum number of color classes is the chromatic number of the graph.

Trivially, graphs of chromatic number 1 are empty (edgeless) graphs. Also, it is not difficult to see that graphs of chromatic number at most 2 are bipartite.

Exercises

- What is the chromatic number and the clique number of C_{2k+1}?
- What is the chromatic number and the clique number of $\overline{C}_{2k+1}$?
- Show that $|E(G)| \geq \chi(G)(\chi(G) - 1)/2$ for any graph G.
- Show that $|V(G)| \leq \chi(G)\alpha(G)$.

It is not difficult to see that the chromatic number of a graph is never smaller than its clique number, i.e $\chi(G) \geq \omega(G)$ for any graph G.

13.1 Perfect graphs and Berges conjectures

Definition 101. A graph G such that $\chi(H) = \omega(H)$ for every induced subgraph H of G is called PERFECT.

By definition, the class of perfect graphs is hereditary. This class is very important in graph theory because it contains many interesting subclasses. Let us consider some simple examples.

EXERCISES

- Show that complete graphs are perfect.
- Show that empty (edgeless) graphs are perfect.
- Show that cographs are perfect.
- Show that bipartite graphs are perfect.
- Show that split graphs are perfect.
- Show that co-bipartite graphs are perfect.
- Characterize the set of K_3-free perfect graphs.

In the next section, we will consider more subclasses of perfect graphs. For the time being, let us look at the graphs which are not perfect with the aim of identifying minimal non-perfect graphs. It is not difficult to see that every cycle of odd length at least 5 is a minimal non-perfect graph. With a bit of work (based on the two sets of exercises above), one can also show that the complement of cycles of odd length at least 5 are minimal non-perfect (this will also follow from a more general result stated as Theorem 103).

Are there any other minimal non-perfect graphs? Berge conjectured in 1963 that there are none, and this was known as *strong perfect graph conjecture*. This conjecture was proved in 2006 and is now known as the Strong Perfect Graph Theorem.

Theorem 102. *A graph is perfect if and only if it contains neither odd cycles of length at least 5 nor their complements as induced subgraphs.*

This theorem was proved in [3]. Its proof is long and technical, and it would not be too illuminating to attempt to sketch it. Instead, we prove a related result, which was formerly known as *weak perfect graph conjecture*, and now is known as Perfect Graph Theorem.

Theorem 103. *A graph is perfect if and only if its complement is perfect.*

To prepare the proof of this result, let us introduce the following graph operation. Given a graph G and a vertex x in G, we will say that we *expand* x to an edge xx' if we add a new vertex x' and connect it to x and to all neighbours of x.

Lemma 104. *Every graph obtained from a perfect graph by expanding a vertex is again a perfect graph.*

Proof. We use induction on the number of vertices. It is not difficult to see that all graphs with at most 3 vertices are perfect, which establishes the basis of the induction. To make the induction step, consider a perfect graph G with at least 4 vertices and let G' be obtained from G by expanding a vertex $x \in V(G)$ to an edge xx'.

Let H be an induced subgraph of G'. If H is a *proper* induced subgraph, i.e. $H \neq G'$, then either H is an induced subgraph of G, in which case H is perfect by definition, or H is obtained from an induced subgraph of H by expanding x, in which case H is perfect by induction assumption, as being strictly smaller than G'. In either case, $\chi(H) = \omega(H)$.

It remains to show that $\chi(G') = \omega(G')$. Clearly, $\omega(G) \leq \omega(G') \leq \omega(G) + 1$ and $\chi(G') \leq \chi(G) + 1$. If $\omega(G') = \omega(G) + 1$, then $\chi(G') \leq \chi(G) + 1 = \omega(G) + 1 = \omega(G')$ and we are done. So, assume $\omega(G) = \omega(G')$. Then no maximum clique of G contains x (since otherwise such a clique together with x' would create a clique of size $\omega(G) + 1$ in G').

Let us color the vertices of G with $\omega(G)$ colors and let X be the color class containing x. Any maximum clique K of G meets X, but (as we have seen earlier) does not meet x. Therefore, the clique number of the graph $H = G - (X - \{x\})$ is strictly less than $\omega(G)$. Since G is perfect, we may color H with at most $\omega(G) - 1$ colors. The graph G' can be obtained from H by adding the set $(X - \{x\}) \cup \{x'\}$, which is independent (since X is independent, and x and x' are twins). Therefore, the coloring of H with at most $\omega(G) - 1$ colors can be extended to a coloring of G' with at most $\omega(G)$ colors, i.e. $\chi(G') \leq \omega(G) = \omega(G')$. $\square$

With the help of Lemma 104, we prove the following important result.

Lemma 105. *Every perfect graph G has a clique intersecting all maximum independent sets of G.*

Proof. Let $\alpha := \alpha(G)$ be the independence number of G (i.e. the size of a maximum independent set in G), let $\mathcal{A}$ denote the family of all independent sets of size α in G, and $\mathcal{K}$ the family of all cliques in G.

Assume by contradiction that for every clique $K \in \mathcal{K}$, there exists a set $A_K \in \mathcal{A}$ with $K \cap A_K = \emptyset$. For a vertex x of G, we denote

$$k(x) := |\{K \in \mathcal{K} \ : \ x \in A_K\}|,$$

i.e. $k(x)$ is the number of sets A_K containing x. Let us replace in G every vertex x by a complete graph G_x with $k(x)$ vertices, joining all the vertices of G_x to all the vertices of G_y whenever x is adjacent to y in G. Denote the resulting graph by G'. In other words, G' can be obtained by repeated vertex expansion from the graph $G[\{x \ : \ k(x) > 0\}]$. The latter graph is an induced subgraph of G and hence perfect by assumption. Therefore, G' is perfect by Lemma 104.

By construction, every clique in G' has the form $\cup_{x \in K} G_x$ for some clique $K \in \mathcal{K}$. Let $X \in \mathcal{K}$ be a clique maximizing $\sum_{x \in X} k(x)$. Then we have

$$\omega(G') = \sum_{x \in X} k(x).$$

Each set A_k has at most 1 common vertex with X, as X is a clique and A_K is an independent set. If A_k has a vertex in X, then it contributes exactly 1 to the sum $\sum_{x \in X} k(x)$, and if A_k does not have vertices in X, then it contributes nothing to this sum. Therefore, this some can be re-written as

$$\sum_{x \in X} k(x) = \sum_{K \in \mathcal{K}} |X \cap A_K|.$$

Since X intersects not all sets A_K (by definition it does not intersect A_X), we conclude that the above sum is at most $|\mathcal{K}|-1$, i.e. $\omega(G') \leq |\mathcal{K}|-1$. On the other hand,

$$|V(G')| = \sum_{x\in V(G)} k(x) = \sum_{K\in\mathcal{K}} |A_K| = \alpha|\mathcal{K}|.$$

We know that $|V(G')| \leq \chi(G')\alpha(G')$ (which is true for any graph). Also, by construction of G', we have $\alpha(G') \leq \alpha$. Therefore,

$$\chi(G') \geq \frac{|V(G')|}{\alpha(G')} \geq \frac{|V(G')|}{\alpha} = |\mathcal{K}|.$$

Summarizing the above discussion we obtain

$$\chi(G') \geq |\mathcal{K}| > |\mathcal{K}|-1 \geq \omega(G'),$$

contradicting the fact that G' is perfect. This contradiction shows G has a clique intersecting all maximum independent sets of G. □

Proof of Theorem 103 will be given by induction on $|V(G)|$. For $|V(G)|=1$ this trivial. So, let G be a graph with at least two vertices.

Every proper induced subgraph of $\overline{G}$ (i.e. an induced subgraph different from $\overline{G}$) is the complement of a proper induced subgraph of G, and is hence perfect by induction. Therefore, to prove that $\overline{G}$ is perfect it suffices to show that $\chi(\overline{G}) \leq \omega(\overline{G})$. To this end, we find, by Lemma 105, a clique K intersecting all maximum independent sets in G, and conclude that

$$\omega(\overline{G}-K) = \alpha(G-K) < \alpha(G) = \omega(\overline{G}),$$

so by the induction hypothesis

$$\chi(\overline{G}) \leq \chi(\overline{G}-K)+1 = \omega(\overline{G}-K)+1 \leq \omega(\overline{G}),$$

as desired. This completes the proof of Theorem 103.

13.2 Subclasses of perfect graphs

From the Strong Perfect Graph Theorem (Theorem 102) and the definition of chordal graphs, we can easily conclude that chordal graphs are perfect.

EXERCISES

- Use Theorem 102 to show that chordal graphs are perfect.

However, the fact that chordal graphs are perfect was known well before the proof of Theorem 102. It is based on the existence in chordal graphs separating clique and the following proposition.

Proposition 106. *Let G be a graph containing a separating clique X, and let $C^1,\dots,C^k$ be the connected components of $G-X$. Then G is perfect if and only if $G[X\cup C^i]$ is perfect for each $i=1,\dots,k$.*

Proof. One direction of the statement is obvious, since an induced subgraph of a perfect graph is perfect by definition.

Assume now that $G[X \cup C^i]$ is perfect for each $i = 1, \dots, k$. Let H be an induced subgraph of G. To prove the result, we need to show that $\chi(H) \leq \omega(H)$. Let $X_H = V(H) \cap X$ and $C^i_H = V(H) \cap C^i$, and $H_i = H[X_H \cup C^i_H]$. It is not difficult to see that

$$\chi(H) = \max\{\chi(H_i) \ : \ i = 1, \dots, k\} \text{ and } \omega(H) = \max\{\omega(H_i) \ : \ i = 1, \dots, k\}.$$

Also, for each $i = 1, \dots, k$, we have $\chi(H_i) = \omega(H_i)$, since H_i is an induced subgraph of the perfect graph $G[X \cup C^i]$. Suppose that H_{i_m} has the maximum chromatic number among all graphs H_i. Then,

$$\chi(H) = \chi(H_{i_m}) = \omega(H_{i_m}) \leq \max\{\omega(H_i) \ : \ i = 1, \dots, k\} = \omega(H),$$

as required. □

Below we mention some more important subclasses of perfect graphs.

13.2.1 Comparability Graphs

A binary relation on a set A is a subset of $A^2 = A \times A$.

Definition 107. A binary relation $\mathcal{R}$ on A is a strict partial order if it is

- asymmetric, i.e., $(a, b) \in \mathcal{R}$ implies $(b, a) \notin \mathcal{R}$;
- transitive, i.e., $(a, b) \in \mathcal{R}$ and $(b, c) \in \mathcal{R}$ implies $(a, c) \in \mathcal{R}$.

It is not difficult to see that asymmetry implies that the relation is irreflexive, i.e., $(a, a) \notin \mathcal{R}$ for all $a \in A$.

A strict partial order $\mathcal{R}$ on a set A can be represented by an oriented graph G with vertex set A and edge set $\mathcal{R}$. By forgetting (ignoring) the orientation of G, we obtain a graph which is known in the literature as a *comparability graph*. In other words, an indirected graph is a comparability graph if it admits a transitive orientation, i.e., the edges of the graph can be oriented in such a way that the existence of an arc directed from a to b and an arc directed from b to c implies the existence of an arc directed from a to c. Sometimes comparability graphs are also called *transitively orientable graphs*. The example in Figure 2 shows that P_5 (a path on 5 vertices) is a transitively orientable graph.

EXERCISES

- Is the class of comparability graphs hereditary?
- Show that C_{2k+1} with $k > 1$ is not transitively orientable.
- Show that the complement of C_k ($k \geq 6$) is not a comparability graph.
- Show that bipartite graphs are transitively orientable, i.e. find a transitive orientation of bipartite graphs.

13.2.2 Permutation Graphs

Let π be a permutation of $\{1, 2, \ldots, n\}$. A pair (i, j) is called an *inversion* if $(i-j)(\pi(i)-\pi(j)) < 0$. The graph of permutation π, denoted $G[\pi]$, has $\{1, 2, \ldots, n\}$ as its vertex set with i and j being adjacent if and only if (i, j) is an inversion.

Definition 108. A graph G is said to be a permutation graph if there is a permutation π such that G is isomorphic to $G[\pi]$.

Lemma 109. *The complement of a permutation graph is a permutation graph.*

Proof. In order to prove the lemma, let us represent the permutation graph of a permutation $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ as follows. Consider two parallel lines each containing n points numbered consecutively from 1 to n. Then for each $i = 1, \ldots, n$ connect the point i of the first line to the point $\pi(i)$ on the second line. Two segments $[i, \pi(i)]$ and $[j, \pi(j)]$ cross each other if and only if (i, j) is an inversion of π. Therefore, the permutation graph $G[\pi]$ of π is a graph whose vertices correspond to the n segments connecting the two parallel lines with two vertices being adjacent if and only if the respective segments cross each other.

Now consider a new permutation π^r obtained from π by reversing the order of the points on the second line, i.e., $\pi^r(i) = \pi(n-i+1)$. Then clearly two segments $[i, \pi(i)]$ and $[j, \pi(j)]$ cross each other if and only if the respective segments $[i, \pi^r(i)]$ and $[j, \pi^r(j)]$ do not. Therefore, $G[\pi^r]$ is the complement of $G[\pi]$. $\square$

Lemma 110. *Every permutation graph is a comparability graph.*

Proof. Let G be the permutation graph of a permutation π on $\{1, 2, \ldots, n\}$. To show that G is a comparability graph, we will find a transitive orientation of G. To this end, we orient each edge ij with $i < j$ from i to j. Let us show that this orientation is transitive. Assume there is an arc $i \to j$ and an arc $j \to k$. Therefore, $i < j < k$ and $\pi(k) < \pi(j) < \pi(i)$ and hence ik is an inversion. Thus, ik is an edge of G and this edge is oriented from i to k. This proves that the proposed orientation is transitive. $\square$

Theorem 111. *A graph G is a permutation graph if and only if both G and $\overline{G}$ are comparability graphs.*

EXERCISES

- Show that every chain graph is a permutation graph.

13.2.3 Interval graphs

Definition 112. A graph G is an interval graph if it is the intersection graph of intervals on the real line.

EXERCISES

- Is the class of interval graphs hereditary?
- Show that every interval graph is chordal, i.e. show that C_k with $k \geq 4$ is not an interval graph.

Theorem 113. *A graph G is an interval graph if and only if G is chordal and $\overline{G}$ is transitively orientable.*

Exercises
• Show that the complement of a chain graph is an interval graph. Moreover, show that the complement of a chain graph has an interval representation in which all intervals are of the same length.
• Show that threshold graphs are interval graphs. Is it possible to represent a threshold graph by intervals of the same length?

13.3 χ-bounded classes of graphs

Let us repeat that the clique number is a lower bound for the chromatic number of any graph. For perfect graphs, the clique number is also an upper bound for the chromatic number. However, in general, there is no upper bound on the chromatic number of a graph in terms of its clique number. In other words, the difference $\chi(G) - \omega(G)$ can be arbitrarily large.

Theorem 114. *There exist K_3-free graphs with arbitrarily large chromatic number.*

Proof. To prove the theorem, we will inductively construct an infinite sequence of K_3-free graphs $G_2, G_3, \ldots, G_i, \ldots$ such that $\chi(G_i) = i$. Let $G_2 = K_2$. Now assume we have constructed the graph G_i and let $V = V(G_i) = \{v_1, \ldots, v_n\}$ be its vertex set. Then the graph G_{i+1} is defined as follows: $V(G_{i+1}) = V \cup V' \cup \{v\}$, with $V' = \{v'_1, \ldots, v'_n\}$. We let V induce G_i in G_{i+1}. Also, for each $i = 1, \ldots, n$, we connect vertex v'_i to those vertices of V which are adjacent to v_i in G_i, and we connect vertex v to all vertices of V'.

Assume G_{i+1} contains a triangle. Then all its vertices must belong to $V \cup V'$, since the neighbourhood of v is an independent set. Exactly one vertex of this triangle belongs to V', since V induces a K_3-free graph, while V' is an independent set. However, if v'_i, v_j, v_k is a triangle in G_{i+1}, then v_i, v_j, v_k is a triangle in G_i, which is impossible. Therefore, G_{i+1} is K_3-free.

Now let us show that $\chi(G_{i+1}) = i + 1$. First, we observe that $\chi(G_{i+1}) \leq i + 1$, because any coloring of G_i with i colors can be extended to a coloring f of G_{i+1} with $i+1$ colors by defining $f(v'_i) = f(v_i)$ for $i = 1, \ldots, n$ and by coloring v with a new color. To show that $\chi(G_{i+1}) \geq i+1$, assume to the contrary that G_{i+1} admits a coloring with i colors. Then the vertices of V' use at most $i - 1$ of these colors (since the i-th color is needed for vertex v). But then the vertices of V could also be colored with $i - 1$ colors, by defining the color of v_i to be equal to the color of v'_i. Since this is not possible, we conclude that $\chi(G_{i+1}) \geq i + 1$. □

Paul Erdős proved an important generalization of Theorem 114.

Theorem 115. *For any fixed $k \geq 3$, there exist $(C_3, \ldots, C_k)$-free graphs with arbitrarily large chromatic number.*

From the above two theorems it follows that the chromatic number of a graph is generally not upper bounded by any function of its clique number. However, for graphs in some special classes this may be the case, and such classes are known as χ-bounded.

Definition 116. A class of graphs is called *χ-bounded* if there is a function f such that for every graph G in this class, $\chi(G) \leq f(\omega(G))$.

There are several interesting conjectures about χ-bounded classes. One of them deals with classes defined by a single forbidden induced subgraph G. A necessary condition for a class $Free(G)$ to be χ-bounded is that G must be acyclic, which follows from Theorem 115.

Corollary 117. *Given a graph G, the class $Free(G)$ is χ-bounded only if G is a forest (i.e. a graph without cycles).*

Proof. Assume G is not a forest, and C_k be a chordless cycle contained in G. Then $Free(C_3, \ldots, C_k) \subseteq Free(C_k) \subseteq Free(G)$. By Theorem 115, the class $Free(C_3, \ldots, C_k)$ is not χ-bounded, and hence so is $Free(G)$. $\square$

An interesting conjecture about χ-bounded classes states that the above necessary condition is also sufficient.

Conjecture 118. *A class of graph defined by a single forbidden induced subgraph G is χ-bounded if and only if G is a forest.*

The conjecture was proved for several important cases. Below we verify it for one them.

Theorem 119. *If G is a $2K_2$-free graph, then $\chi(G) \leq \binom{\omega(G)+1}{2}$.*

Proof. Let $\omega = \omega(G)$ and A be a clique of size ω in G. For every pair a, b of distinct vertices of A, let C_{ab} consist of those vertices of G that are adjacent neither to a nor to b. Then C_{ab} is an independent set, since otherwise a $2K_2$ arises. Therefore, defining $C = \cup C_{ab}$, we conclude that $\chi(G[C]) \leq \binom{\omega}{2}$.

Consider now a vertex v not in $A \cup C$. Then v must have exactly one non-neighbour in A (why?). For each vertex $a \in A$, let I_a be the set of vertices of G for which a is the only non-neighbour in A. Then $\{a\} \cup I_a$ is an independent set, since otherwise any two adjacent vertices of I_a together with $A - \{a\}$ would create a clique large than A. Therefore, the vertices not in C can be colored with $|A| = \omega$ colors (ω colors for the vertices of A; the vertices of I_a use the color of a). As a result, $\chi(G) \leq \binom{\omega}{2} + \omega = \binom{\omega(G)+1}{2}$. $\square$

EXERCISES

- Prove that the class of mK_2-free graphs is χ-bounded for each fixed $m \geq 2$.

The K_3-free graphs with large chromatic number revealed in Theorem 114 are rather exotic, because it is known that *for almost all graphs*, if $\omega(G) \leq k$, then $\chi(G) \leq k$. In the next section, we give a formal definition of the notion of "almost all graphs" and describe several results related to this notion.

14 Properties of almost all graphs

We recall from Section 5 that for a graph property P we denote by $P(n)$ the set of labelled graphs with n vertices in P.

EXERCISES

- Find the number of labelled complete bipartite graphs with n vertices.
- Find the number of labelled graphs of degree 1 with n vertices.
- Let x_n be the number of n-vertex labelled graphs in the class $Free(P_3)$ and let y_n be the number of n-vertex labelled graphs in the class $Free(K_{1,3}, C_3, C_4, C_5, \ldots)$. Determine the values of n for which

  i) $x_n > y_n$,

  ii) $x_n < y_n$,

  iii) $x_n = y_n$.

Denote by Γ the set of all labelled graphs. Clearly $|\Gamma(n)| = 2^{\binom{n}{2}}$.

Definition 120. We will say that *almost all graphs* have property P if

$$\lim_{n\to\infty} |P(n)|/|\Gamma(n)| = 1.$$

Theorem 121. *Almost all graphs are connected.*

Proof. Let $S(n)$ denote the set of all connected graphs from $\Gamma(n)$, and $S_t(n)$ the set of graphs from $\Gamma(n)$ containing at least one component of size t. Then

$$|S(n)| \geq |\Gamma(n)| - \sum_{t=1}^{\lfloor n/2 \rfloor} |S_t(n)|.$$

To find an upper bound on the number of graphs in $S_t(n)$, consider n-vertex graphs whose vertices can be partitioned into two subsets V_1 of size t and V_2 of size $n-t$ in such a way that there no edges between the subsets. There

$$\binom{n}{t} 2^{\binom{t}{2}+\binom{n-t}{2}} = \binom{n}{t} 2^{\binom{n}{2}-t(n-t)}$$

such graphs. Clearly, every graph with a connected component of size t belongs to the set of such graphs, i.e.

$$|S_t(n)| \leq \binom{n}{t} 2^{\binom{n}{2}-t(n-t)}.$$

Summarizing, we obtain

$$\frac{|S(n)|}{|\Gamma(n)|} \geq 1 - \sum_{t=1}^{\lfloor n/2 \rfloor} \binom{n}{t} 2^{-t(n-t)}.$$

Denoting $f(t) = \binom{n}{t} 2^{-t(n-t)}$ and comparing $f(t)$ and $f(t+1)$, we conclude that in the interval $[1, \ldots, \lfloor n/2 \rfloor]$ the function $f(t)$ is decreasing. Therefore,

$$\sum_{t=1}^{\lfloor n/2 \rfloor} f(t) < \frac{n}{2} f(1) = \frac{n}{2} \frac{n}{2^{n-1}} = \frac{n^2}{2^{n-2}} \to 0.$$

Therefore, the fraction $|S(n)|/|\Gamma(n)|$ tends to 1. $\square$

Now let us generalize the above result from connected graphs to k-connected graphs.

Definition 122. A graph G is k-connected if it has no set W of at most $k-1$ vertices such that $G-W$ is disconnected. The larges number k such that G is k-connected is the connectivity number of G. Alternatively, the connectivity number of G is the smallest k such that the deletion of k vertices disconnects the graph.

The graphs of connectivity number 0 are precisely disconnected graphs. Every 1-connected is connected. The k-connected graphs are graphs of connectivity number at least k.

In order to show that almost all graphs are k-connected we will prove a more general result. Let $P_{i,j}$ be the property (set) of graphs in which for *any* two disjoint subsets U and W of vertices with $|U| \leq i$ and $|W| \leq j$, there is a vertex $v \notin U \cup W$ which is complete to U and anticomplete to W. For instance,

- every graph G in $P_{2,k-1}$ is k-connected. Indeed, assume G has a subset W of at most $k-1$ vertices such that $G-W$ is disconnected. Consider two vertices u_1 and u_2 from different connected components of $G-W$. By definition of $P_{2,k-1}$, there must exist a vertex $v \notin \{u_1, u_2\} \cup W$ which is adjacent to both u_1 and u_2, contradicting to the assumption that $G-W$ is disconnected.

- every prime graph belongs to $P_{1,1}$, since in a prime graph for any two vertices u and w there is a vertex v distinguishing them, i.e. adjacent to one of them and non-adjacent to the other.

Theorem 123. *For any fixed natural numbers i and j, almost all graphs have property $P_{i,j}$.*

Proof. We will prove the theorem in the framework of random graphs, in which case the probability that a graph G has a property P is the ration $|P(n)|/|\Gamma(n)|$, To show that almost all graphs have property P we have to show that this ratio tends to 1.

We will show that a random graph G has property $P_{i,j}$ with probability tending to 1 assuming that any two vertices of G are adjacent with probability $p = 1/2$ (although the result holds for any value p with $0 < p < 1$). More precisely, will show that with probability tending to 0 G does not belong $P_{i,j}$.

Let us fix two disjoint sets of vertices U and W in G. Then the probability that a vertex $v \notin U \cup W$ is complete to U is $(1/2)^{|U|} = 2^{-|U|}$. Similarly, the probability that v is anticomplete to W is $2^{-|W|}$. Thus, the probability that v is complete to U and anticomplete to W is $2^{-|U|-|W|}$. Therefore, the probability that v is not of this type is $1 - 2^{-|U|-|W|}$, and the probability that no vertex in $V(G) - (U \cup W)$ is of this type is

$$(1 - 2^{-|U|-|W|})^{n-|U|-|W|} \leq (1 - 2^{-i-j})^n,$$

since the events that different vertices $v \notin U \cup W$ are not of the type are independent. Also, there are

$$\sum_{i_1 \leq i, j_1 \leq j} \binom{n}{i_1} \binom{n-i_1}{j_1} \leq \sum_{i_1 \leq i, j_1 \leq j} n^{i_1} n^{j_1} \leq ijn^{i+j} \leq n^{i+j+2}$$

ways to choose U and W of size at most i and j respectively. Therefore, the probability that G has a pair U, W with no suitable vertex $v \notin U \cup W$ is at most

$$n^{i+j+2}(1 - 2^{-i-j})^n.$$

This probability tends to 0, since $(1 - 2^{-i-j})$ is less than 1. $\square$

Theorem 124. *Diameter of almost all graphs is 2.*

Proof. Let $\Gamma^2(n)$ the set of n-vertex labeled graphs of diameter 2. We will show that $\lim_{n\to\infty} |\Gamma^2(n)|/|\Gamma(n)| = 1$. Clearly, the diameter of almost all graphs is at least 2, since there is just one graph with n vertices of diameter 1 (the complete graph K_n). Now let us show that the diameter of almost all graphs is at most 2. To this end, let us denote by $\Gamma'(n)$ the set of n-vertex labeled graphs of diameter more than 2. We will show that $\lim_{n\to\infty} |\Gamma'(n)|/|\Gamma(n)| = 0$.

In a graph from $\Gamma'(n)$, consider two vertices u and v of distance more than 2, and let U be the neighbourhood of u. Denote $k = |U|$. We know that $v \notin U$ (otherwise $dist(u,v) = 1$) and v has no neighbours in U (otherwise $dist(u,v) = 2$). Let $\Gamma'_{u,v,U}(n)$ denote the set of graphs from $\Gamma'(n)$ with fixed vertices u, v and a fixed set U as above. By fixing u, v and U, we fix the adjacency value for $k+n-1$ pairs of vertices (k pairs create edges and $n-1$ pairs create non-edges). Therefore,

$$|\Gamma'_{u,v,U}(n)| = 2^{\binom{n}{2}-k-n+1}.$$

Let $\Gamma'_{u,U}(n) = \bigcup_{v\notin(U\cup\{u\})} \Gamma'_{u,v,U}$. Then

$$|\Gamma'_{u,U}(n)| \le (n-k-1)|\Gamma'_{u,v,U}| < n2^{\binom{n}{2}-k-n+1}.$$

Next, let $\Gamma'_u(n) = \bigcup_U \Gamma'_{u,U}(n)$. For a fixed k, there are $\binom{n-1}{k}$ ways to choose U. Also, $k \le n-2$ (since $u, v \notin U$). Therefore,

$$|\Gamma'_u(n)| \le \sum_{k=0}^{n-2} \binom{n-1}{k} |\Gamma'_{u,U}(n)| < \sum_{k=0}^{n-2} \binom{n-1}{k} n2^{\binom{n}{2}-k-n+1} = n2^{\binom{n}{2}-n+1} \sum_{k=0}^{n-2} \binom{n-1}{k} 2^{-k}$$

In order to estimate $\sum_{k=0}^{n-2} \binom{n-1}{k} 2^{-k}$, apply the Binomial Theorem:

$$\sum_{k=0}^{n-2} \binom{n-1}{k} 2^{-k} = (3/2)^{n-1} - \frac{1}{2^{n-1}} < (3/2)^{n-1}$$

Therefore,

$$|\Gamma'_u(n)| < n2^{\binom{n}{2}-n+1}(3/2)^{n-1}$$

Since $\Gamma'(n) = \bigcup_u \Gamma'_u(n)$, we conclude that

$$|\Gamma'(n)| = n|\Gamma'_u(n)| < n^2 2^{\binom{n}{2}-n+1}(3/2)^{n-1}.$$

Thus,

$$\lim_{n\to\infty} \frac{\Gamma'(n)}{\Gamma(n)} = \lim_{n\to\infty} n^2(3/4)^{n-1} = 0,$$

as required. $\square$

EXERCISES

- Show that if a graph has more than $\binom{n-1}{2}$ edges, then it is connected.
- Show that if almost all graphs have property P_1 and almost all graphs have property P_2, then almost all graphs have property $P_1 \cap P_2$.

15 Turán graphs

Let us denote by $H \subseteq G$ the fact that G contains H as a subgraph, not necessarily induced. Clearly, K_n contains all graphs with at most n vertices as subgraphs. Now let us ask the following question: given a graph H with at most n vertices, how many edges an n-vertex graph G should have to contain H as a subgraph? Alternatively, what is the maximum possible number of edges that a graph with n vertices can have *without containing* a copy of H as a subgraph?

Definition 125. A graph G on n with the maximum possible number of edges containing no copy of H as a subgraph is called *extremal* for n and H; the number of its edges is denoted $ex(n, H)$.

Let us emphasize that if G is extremal for n and H, then $E(G)$ is also a maximal (with respect to set inclusion) set such that $H \nsubseteq G$, i.e. adding *any edge* to G results in a copy of H. However, the converse is generally not true: a graph G can be edge-maximal (i.e. adding *any edge* to G results in a copy of H) but not extremal (i.e. with fewer than $ex(n, H)$ edges). Consider, for instance, $2K_2$. It is not difficult to verify that it is an edge-maximal graph without P_4 as a subgraph (i.e. adding any edge to $2K_2$ gives rise to a P_4). However, this graph is not extremal, because there is a graph with 3 edges containing no P_4 as a subgraph, namely, $K_{1,3}$.

In what follows we analyze the case of $H = K_r$. A special role in our analysis will be given to so called complete multipartite graphs.

Definition 126. A *complete multipartite graph* is a graph whose vertices can be partitioned into independent sets (also called partition sets) with all possible edges between any two different sets.

In other words, a graph G is complete multipartite if and only if its complement $\overline{G}$ is a graph in which every every connected component is a clique, or simply, $\overline{G}$ is a disjoint union of cliques.

A complete multipartite graph with r partition sets is called complete r-partite. For $r = 2$, this is a complete bipartite graph.

For each r and $n \geq r$, there is a unique (up to isomorphism) complete r-partite graph with n vertices whose partition sets differ in size by at most 1. This graph is called the *Turán graph* and is denoted by $T^r(n)$.

Definition 127. For natural numbers r and $n \geq r$, the *Turán graph* $T^r(n)$ is the unique complete r-partite graph whose partition sets differ in size by at most 1. The number of edges in the Turán graph $T^r(n)$ is denoted $t_r(n)$.

For convenience, we extend the notion of the Turán graph to values of n smaller than r by defining $T^r(n) = K_n$ in this case.

Lemma 128. *For natural numbers n and r with $n \geq r$,*

$$t_r(n) = t_r(n-r) + (n-r)(r-1) + \binom{r}{2}.$$

Proof. Let $G = T^r(n)$ and let K be an induced subgraph of G containing exactly one vertex in each partition set of G. Then $G - K$ is again a Turán graph $T^r(n-r)$ and hence the number of its edges is $t_r(n-r)$. Each vertex of $G-K$ has exactly $r-1$ neighbours in K (it is adjacent to all vertices of K except the vertex from the same partition set). Therefore, there are $(n-r)(r-1)$ edges between K and the rest of the graph. Since K contains $\binom{r}{2}$ edges, the result follows. □

Clearly, any complete k-partite graph with $k < r$ is K_r-free.

Lemma 129. *Among complete multipartite graphs, the Turán graph $T^{r-1}(n)$ is the unique n-vertex graph with maximum number of edges that does not contain K_r.*

Proof. For $n < r$, the statement is obvious, i.e. K_n is the unique n-vertex graph with maximum number of edges that does not contain K_r.

Let $n \geq r$ and G be a complete multipartite n-vertex graph with maximum number of edges that does not contain K_r. Then the number of partition sets in G is at most $r-1$, else G contains K_r. Also, the number of partition sets is at least $r-1$. Indeed, if the number of partition sets is less than $r-1$, then G contains a partition set U with at least 2 vertices. By splitting this set arbitrarily into two non-empty subsets, say U_1 and U_2, and adding all possible edges between U_1 and U_2 we obtain another complete multipartite graph, which has more edges than G and which also does not contain K_r. This contradiction shows that G is a complete $(r-1)$-partite graph.

Finally, assume that V_1 and V_2 are two partition sets of G with $|V_1| - |V_2| \geq 2$. Then we can increase the number of edges in G by moving a vertex from V_1 to V_2. Therefore, partition sets of G differ in size by at most 1, i.e. G is the Turán graph $T^{r-1}(n)$ □

Theorem 130. *For all natural numbers $r > 1$ and n, every graph $G \not\supseteq K_r$ with n vertices and $ex(n, K_r)$ edges is the Turán graph $T^{r-1}(n)$.*

Proof. We apply induction on n. For $n \leq r-1$, we have $G = K_n = T^{r-1}(n)$. For $n = r$, $G = K_r - e$ (i.e. the graph obtained from K_r by deleting an edge) and $K_r - e = T^{r-1}(r)$, as required.

Suppose now that $n > r$. Since G is an edge-maximal graph without a K_r subgraph, G has a subgraph $K = K_{r-1}$. By the induction hypothesis, $G - K$ has at most $t_{r-1}(n-r+1)$ edges, and each vertex of $G - K$ has at most $r-2$ neighbours in K (since otherwise G contains K_r). Therefore,

$$|E(G)| \leq t_{r-1}(n-r+1) + (n-r+1)(r-2) + \binom{r-1}{2} = t_{r-1}(n).$$

The equality on the right follows from Lemma 128.

Since G is extremal for K_r, we must have equality in the above inequality. Thus, every vertex of $G - K$ has exactly $r-2$ neighbours in K. Let $x_1, \ldots, x_{r-1}$ be the list of vertices of K. For $i = 1, \ldots, r-1$, we denote by V_i the set of vertices v of G with $N(v) \cap K = K - \{x_i\}$. Since each vertex of G has exactly $r-2$ neighbours in K (including the vertices of K themselves), the sets $V_1, \ldots, V_{r-1}$ form a partition of $V(G)$ (i.e. a collection of disjoint subsets containing collectively all vertices of G). Also, since G does not contain K_r, each of the sets V_i is independent. Hence, G is a $(r-1)$-partite graph. As $T^{r-1}(n)$ is the unique $(r-1)$-partite graph with n vertices and the maximum number of edges, our claim that $G = T^{r-1}(n)$ follows from the extremality of G.

One more proof. Let G be a graph with maximum number of edges that does not contain K_r. Assume G contains a $K_1 + K_2$ induced by vertices x, y_1, y_2 with $y_1 y_2$ being an edge. Deleting x and duplicating y_1 (i.e. creating a new vertex with the same neighbourhood as y_1) transforms G into a new graph with the same number of vertices which again contains no copy of K_r. If $deg(y_1) > deg(x)$, then this transformation increases the number of edges, which is not possible since G is extremal for n and K_r. Therefore, $deg(y_1) \leq deg(x)$. Similarly, $deg(y_2) \leq deg(x)$.

But now deleting y_1 and y_2 and duplicating x twice transforms G into a new graph with the same number of vertices and containing no copy of K_r. Moreover, the new graph has strictly more edges then G. This is impossible, since G is extremal for n and K_r, and hence G is (K_1+K_2)-free. Therefore, by Lemmas 20 and 129 G is the Turán graph $T^{r-1}(n)$. $\square$

Let us now discuss some properties of Turán graphs.

Observation 131. *The sizes of partition sets of the Turán graph $T^r(n)$ are given by*

$$\lfloor \frac{n}{r} \rfloor, \lfloor \frac{n+1}{r} \rfloor, \lfloor \frac{n+2}{r} \rfloor, \ldots, \lfloor \frac{n+r-1}{r} \rfloor.$$

Lemma 132. *Among complete r-partite graphs with n vertices, the Turán graph $T^r(n)$ is the graph maximizing the minimum vertex degree. Moreover, if $n \geq r+2$, the $T^r(n)$ contains at least 3 vertices of minimum degree.*

Proof. *Exercise.* $\square$

Lemma 133. $t_{r-1}(n) \approx \frac{1}{2}n^2\frac{r-2}{r-1}$. *In particular, if $r-1$ divides n, then* $t_{r-1}(n) = \frac{1}{2}n^2\frac{r-2}{r-1}$.

Proof. In $T_{r-1}(n)$ there are $(r-1)(r-2)/2$ pairs of partition sets. Each set is of size $n/(r-1)$ and therefore, each pair of sets is joint by $\frac{n^2}{(r-1)^2}$ edges. $\square$

Corollary 134. $\lim_{n\to\infty} \frac{t_{r-1}(n)}{\binom{n}{2}} = \frac{r-2}{r-1}$.

Lemma 135. *Let G be a graph with vertex set $V = \{v_1, v_2, \ldots, v_n\}$. If G is K_r-free, then there is an $(r-1)$-partite graph G' with vertex set V such that $deg_G(v_i) \leq deg_{G'}(v_i)$ for all $i = 1, \ldots, n$.*

Proof. We apply induction on r. For $r = 2$, the theorem is obvious. Suppose now that $r > 2$ and the result holds for values smaller than r. Pick a vertex of maximum degree in G and let W be the set of its neighbours. Denote by H the subgraph of G induced by W. Since G is K_r-free, H is K_{r-1}-free. By the induction hypothesis, H can be replaced by a $(r-2)$-partite graph H' in such a way that we do not decrease the degrees of vertices. Let G' be the graph obtained by adding to H' the vertices of $V-W$ and connecting every vertex of $V-W$ to every vertex of W. Then $deg_G(v_i) \leq deg_{G'}(v_i)$. This is true for the vertices of W by assumption. For any vertex in $v_i \in V-W$, this is true because $deg_G(v_i) \leq \Delta(G) = |W| = deg_{G'}(v_i)$, where $\Delta(G)$ denotes the maximum vertex degree in G. $\square$

Theorem 130 tells us that every graph with n vertices and $t_{r-1}(n)+1$ edges contains a K_r. Moreover, it "almost" contains a K_{r+1}.

Theorem 136. *If $n \geq r+1$, then every graph G with n vertices and $t_{r-1}(n)+1$ edges contains a $K_{r+1}-e$ as a subgraph.*

Proof. Apply induction on n. For $n = r+1$, $G = K_{r+1}-e$. Now assume that the result follows for every graph with less than $r+2$ vertices, and consider a graph G with $n \geq r+2$ vertices. Let us show that

Claim 137. $\delta(G) \leq \delta(T^{r-1}(n))$.

Proof. In $(T^{r-1}(n))$ either all vertices have the same degree, say d, (if $r-1$ divides n), or the degree sequence consists of two different values, say d and $d+1$. If all vertices of $(T^{r-1}(n))$ have degree d and $\delta(G) \geq d+1$, then, taking into account that $n \geq r+2 \geq 3$,

$$\sum_{x \in V(G)} deg(x) \geq \sum_{i=1}^{n} d + n = 2t_{r-1}(n) + n \geq 2t_{r-1}(n) + 3 > 2t_{r-1}(n) + 2 = |E(G)|.$$

If $r-1$ does not divides n, then the number of vertices belonging to the biggest partition sets, and hence the number of vertices of degree d, is at least 3 (since otherwise $n = r$). Thus with $\delta(G) \geq d+1$ we would have

$$\sum_{x \in V(G)} deg(x) - \sum_{x \in V(T^{r-1}(n))} deg(x) \geq 3,$$

while we know that

$$\sum_{x \in V(G)} deg(x) - \sum_{x \in V(T^{r-1}(n))} deg(x) = 2t_{r-1}(n) + 2 - 2t_{r-1}(n) = 2.$$

□

Let x be a vertex of minimum degree in G. Since $deg(x) \leq \delta(T^{r-1}(n))$, the graph $G - x$ has at least $t_{r-1}(n-1)+1$ edges. Therefore, by the induction hypothesis, $G - x$ contains a $K_{r+1} - e$ as a subgraph. □

According to the above theorem, increasing $t_{r-1}(n)$ by at least one edge leads to the appearance of K_r (and even $K_{r+1} - e$). More interestingly, increasing $t_{r-1}(n)$ by an arbitrarily small fraction of n^2 leads to the appearance of $T_r(rs)$ (also denoted by K_r^s; this is the complete r-partite graph with partition sets of size s) for arbitrary values of s. This result is known as Erdős-Stone Theorem.

Theorem 138. *For every $r \geq 2$ and $s \geq 1$, and every $\epsilon > 0$, there is a number n_0 such that every graph with $n \geq n_0$ vertices and at least*

$$t_{r-1}(n) + \epsilon n^2$$

edges contains K_r^s as a subgraph.

We do not prove this theorem, because it is beyond the scope of the module. Instead, we present an important result which follows from this theorem.

Theorem 139. *For any graph H with at least one edge,*

$$\lim_{n \to \infty} \frac{ex(n, H)}{\binom{n}{2}} = \frac{\chi(H) - 2}{\chi(H) - 1}.$$

Proof. Denote $r := \chi(H)$. Since H cannot be colored with $r-1$ colors, we have $H \not\subseteq T^{r-1}(n)$ for all values of n. Therefore,

$$t_{r-1}(n) \leq ex(n, H).$$

Indeed, if $ex(n,H) < t_{r-1}(n)$, then $ex(n,H)$ is not a maximum number of edges in a graph containing no H subgraph. On the other hand, $H \subseteq K_r^s$ for all sufficiently large s, and hence

$$ex(n,H) \leq ex(n,K_r^s)$$

for all those s. Let us fix such an s. For every $\epsilon > 0$, Theorem 138 implies that for large enough n

$$ex(n,K_r^s) < t_{r-1}(n) + \epsilon n^2.$$

Therefore, for large enough n we have

$$\begin{aligned} t_{r-1}(n)/\binom{n}{2} &\leq ex(n,H)/\binom{n}{2} \\ &\leq ex(n,K_r^s)/\binom{n}{2} \\ &< t_{r-1}(n)/\binom{n}{2} + \epsilon n^2/\binom{n}{2} \\ &= t_{r-1}(n)/\binom{n}{2} + 2\epsilon/(1-1/n) \\ &\leq t_{r-1}(n)/\binom{n}{2} + 4\epsilon \text{ assuming } n \geq 2 \end{aligned}$$

Since $t_{r-1}(n)/\binom{n}{2}$ converges to $(r-2)/(r-1)$, we conclude that so does $ex(n,H)/\binom{n}{2}$. □

16 Ramsey's Theorem with variations

Pigeonhole Principle: *If $n+1$ letters are placed in n pigeonholes, then some pigeonhole must contain more than one letter.*

More generally: Let r and p be positive integers. Then there is an $n = n(r,p)$ such that for any coloring of n objects with r different colors there exist p objects of the same color.

Exercises

- Find the minimum value of this number $n = n(r,p)$.

The Pigeonhole Principle has an important generalization proved in the beginning of 20-th century by British mathematician Frank Ramsey at the age of 26 and known as Ramsey's Theorem.

Theorem 140. *Let k, r, p be positive integers. Then there is a positive integer $n = n(k,r,p)$ with the following property. If the k-subsets of an n-set are colored with r colors, then there is a monochromatic p-set, i.e., a p-set all of whose k-subsets have the same color.*

Before we prove this theorem, let us consider some particular cases. For $k=1$, the theorem coincides with the The Pigeonhole Principle.

For $k=2$, coloring 2-subsets can be viewed as coloring the edges of a complete graph, i.e. for $k=2$, the theorem can be reformulated as follows: For any positive integers r and p, there is a positive integer $n = n(n,p)$ such that if the edges of a n-vertex graph are colored with r colors, then there is a monochromatic clique of size p, i.e., a clique all of whose edges have the same color.

16.1 Two colors

In the case of $r = 2$ colors, the statement of Ramsey's Theorem can be further rephrased as follows: for any positive integer p, there is a positive integer $n = n(p)$ such that every graph with at least n vertices has either a clique of size p or an independent set of size p.

Definition 141. The minimum n such that every graph with n vertices contains either a clique of size p or an independent set of size p is the SYMMETRIC RAMSEY NUMBER $R(p)$.

Definition 142. The minimum n such that every graph with n vertices contains either a clique of size p or an independent set of size q is the RAMSEY NUMBER $R(p, q)$.

EXERCISES

- Show that $R(p, p) = R(p)$.
- Show that $R(p, q) = R(q, p)$.

Let us consider Ramsey numbers (symmetric and non-symmetric) for some small values of p and q.

EXERCISES

- Show that $R(2, k) = k$.

Claim 143. $R(3, 3) = 6$

Proof. Since C_5 is a self-complementary graph and C_5 contains no triangle (i.e. no clique of size 3), $R(3, 3) \geq 6$. Consider now an arbitrary graph G with 6 vertices. Our purpose is to show that either G or its complement contains a triangle. Let v be a vertex of G. Then either the degree of v in G is at least 3 or the degree of v in the complement of G is at least 3.

Assume that v has at least three neighbours a, b, c in G. If at least two of the neighbours of v are adjacent, say a is adjacent to b, then a, b, v form a triangle in G. If a, b, c are pairwise non-adjacent, then they form a triangle in the complement of G.

If v has at least three neighbours in the complement of G, the arguments are similar. □

Claim 144. $R(3, 4) = 9$

Proof. Let G be a graph on 8 vertices obtained from a cycle C_8 by joining pairs of vertices that are of distance 4 in the cycle. Clearly, this graph contains no triangle. For each vertex, the set of its non-neighbours induces a P_4. Since P_4 contains no independent set of size 3, the graph G has no independent set of size 4. Therefore, $R(3, 4) \geq 9$.

Now consider a graph G on 9 vertices and let v be a vertex of G.

Assume first that v has at least four neighbours a, b, c, d in G. Since $R(2, 4) = 4$, the graph induced by a, b, c, d contains either an edge, which together with v creates a triangle, or an independent set of size 4.

Now assume that v has at least 6 non-neighbours (the vertices non-adjacent to v). Then the subgraph induced by these non-neighbours contains either an independent set of size 3, which together with v form an independent set of size 4, or a clique of size 3.

If v has less than 4 neighbours and less than 6 non-neighbours, then the degree of v is exactly 3. Since v was chosen arbitrarily, we must conclude that the degree of every vertex of G is 3, which is not possible by Corollary 13. □

Below is the list of all known Ramsey numbers:

$R(1,k) = 1$, $R(2,k) = k$, $R(3,3) = 6$, $R(3,4) = 9$, $R(3,5) = 14$, $R(3,6) = 18$, $R(3,7) = 23$, $R(3,8) = 28$, $R(3,9) = 36$, $R(4,4) = 18$, $R(4,5) = 25$.

For some numbers, only bounds are known. For instance, $40 \leq R(3,10) \leq 43$, $43 \leq R(5,5) \leq 49$, $102 \leq R(6,6) \leq 165$. It is probable that the exact value of $R(6,6)$ will remain unknown forever. [1]

16.1.1 Bounds on Ramsey numbers

Theorem 145. *The number $R(p,q)$ exists and for $p,q > 1$ it satisfies $R(p,q) \leq R(p-1,q) + R(p,q-1)$.*

Proof. For $p = 1$ or $q = 1$, we have $R(p,q) = 1$. Assume now that $p,q > 1$ and let G be a graph with $R(p-1,q) + R(p,q-1)$ vertices and v a vertex of G. Denote by G_1 the subgraph of G induced by the neighbours of v, and by G_2 the subgraph induced by the non-neighbours of v. Since $|V(G_1)| + |V(G_2)| + 1 = R(p-1,q) + R(p,q-1)$, we have either $|V(G_1)| \geq R(p-1,q)$ or $|V(G_2)| \geq R(p,q-1)$. In the former case, G_1 contains either an independent set of size q, in which case we are done, or a clique of size $p-1$. This clique together with v create a clique of size p, and we are done again. The case $|V(G_2)| \geq R(p,q-1)$ is similar. $\square$

This theorem does more than giving bounds on Ramsey numbers. It actually proves Ramsey's Theorem for $k = 2$ and $r = 2$. We know that Ramsey numbers exist for small values of p and q. This theorem shows that the existence of Ramsey numbers for small values of p and q implies the existence of Ramsey numbers for larger values of p and q.

On the other hand, this theorem only shows the *existence* of Ramsey numbers without giving any explicit bound. The next result overcomes this difficulty.

Theorem 146. *If $p,q \geq 2$, then $R(p,q) \leq \binom{p+q-2}{p-1}$*

Proof. We prove the result by induction on p and q. For $p = 2$, we have $R(2,q) = q = \binom{2+q-2}{2-1}$ and similarly for $q = 2$.

From Theorem 145 we know that $R(p,q) \leq R(p-1,q) + R(p,q-1)$. Together with inductive assumption and Pascal's triangle for binomial coefficients, this implies

$$R(p,q) \leq R(p-1,q) + R(p,q-1) \leq \binom{p+q-3}{p-2} + \binom{p+q-3}{p-1} = \binom{p+q-2}{p-1}.$$

$\square$

In addition to an upper bound, it would be interesting to know any lower bound on the Ramsey number.

Theorem 147. $R(p,p) \geq 2^{(p-2)/2}$

[1] "Imagine an alien force, vastly more powerful than us landing on Earth and demanding the value of $R(5,5)$ or they will destroy our planet. In that case, we should marshal all our computers and all our mathematicians and attempt to find the value. But suppose, instead, that they asked for $R(6,6)$, In that case, we should attempt to destroy the aliens before they destroy us." – Paul Erdős

Proof. There are $2^{\binom{n}{2}} = 2^{n(n-1)/2}$ labelled graphs on n vertices. Among them, not more than $2\binom{n}{p}2^{\binom{n}{2}-\binom{p}{2}}$ graphs contain a monochromatic set on p vertices. Therefore, the proportion of such graphs does not exceed

$$\frac{2 \cdot \binom{n}{p} \cdot 2^{\binom{n}{2}-\binom{p}{2}}}{2^{\binom{n}{2}}} = \binom{n}{p} 2^{1-\binom{p}{2}}.$$

It is known (and not difficult to see) that $\binom{n}{p} < n^p$. Also, it is not difficult to check that $1 - \binom{p}{2} = 1 - p(p-1)/2 < -p(p-2)/2$. Therefore, $\binom{n}{p}2^{1-\binom{p}{2}} < n^p 2^{-p(p-2)/2}$, and hence for $n = 2^{p-2}/2$, the proportion of graphs containing a monochromatic set of size p is strictly less than 1. In other words, there are graphs with $n = 2^{p-2}/2$ vertices containing neither a clique of size p nor an independent set of size p, i.e. $R(p,p) \geq 2^{(p-2)/2}$. $\square$

16.2 More colors

Let $R_k(s_1, \ldots, s_k)$ be the smallest n such that whenever the edges of a K_n are colored with k colours, then we can find a K_{s_i} of colour i for some $1 \leq i \leq k$.

The existence of numbers $R_r(s_1, \ldots, s_k)$ can be shown by analogy with Theorem 145, i.e. by analogy with the proof of Theorem 145 one can show that

Theorem 148. *The number $R_r(s_1, \ldots, s_r)$ exists and satisfies $R_r(s_1, \ldots, s_r) \leq R_r(s_1 - 1, \ldots, s_r) + \cdots + R_r(s_1, \ldots, s_r - 1)$.*

Alternatively, the existence of numbers $R_r(s_1, \ldots, s_k)$ can be shown by induction on the number of colours: $r = 1$ is trivial.

Given $s_1, \ldots, s_r$ $(r \geq 2)$, let $n = R(s, R_{r-1}(s_2, \ldots, s_r))$. Then for any r-colouring of K_n, view it as a 2-colouring of K_n with the colours "1" and "2 or 3 or $\cdots$ or r". So, by choice of n we have *either* a K_{s_1} coloured with colour 1 (in which case we are done) *or* a $K_{R_{r-1}(s_2,\ldots,s_r)}$ coloured with colours $2, \ldots, r$, in which case we are done by inductive assumption.

16.3 Infinite version

Theorem 149. *Let k and r be positive integers and X is an infinite set. If the k-subsets of X are colored with r colors, then it contains a monochromatic infinite subset.*

Proof. We prove the theorem by induction on k. For $k = 1$, the result is obvious ("infinite Pigeonhole Principle"). So, assume $k > 1$ and let the k subsets of X to be colored with r colors.

We will construct an infinite sequence $X_0, X_1, \ldots$ of infinite subsets of X as follows. For $i = 0$, we let $X_0 := X$. Now assume that the sets $X_0, X_1, \ldots, X_i$ have been constructed. We arbitrarily choose $x_i \in X_i$ and color the $k-1$-subsets of $X_i - \{x_i\}$ with r colors by assigning to each subset $Z \subseteq X_i - \{x_i\}$ of $k-1$ elements the color that is assigned to $\{x_i\} \cup Z$ in the coloring of X. By the induction hypothesis, $X_i - \{x_i\}$ has an infinite monochromatic subset all of whose $k-1$-subsets have the same color. We define this subset to be X_{i+1} and denote the color of all of its $k-1$-subsets by c_i.

The above procedure produces an infinite sequence of elements $x_0, x_1, \ldots$ together with an infinite sequence of colors $c_0, c_1, \ldots$ such that for any k indices $i_1, \ldots, i_k$, the color of the k-set $\{x_{i_1}, \ldots x_{i_k}\}$ is c_{i_1}. Since the number of colors used is finite, in the sequence $c_0, c_1, \ldots$ there must be an infinite subsequence of indices $\{i_1, i_2, \ldots\}$ such that $c_{i_1} = c_{i_2} = \ldots$. Then all the k subsets of the infinite set $\{x_{i_1}, x_{i_2}, \ldots\}$ have the same color. $\square$

16.4 Proof of Ramsey Theorem

Lemma 150. *Let $V_0, V_1, \ldots$ be an infinite sequence of disjoint non-empty finite sets, and let G be a graph with the vertex set $V_0 \cup V_1 \cup \ldots$. Assume that each vertex $v \in V_i$ with $i \geq 1$ has a neighbour $f(v)$ in V_{i-1}. Then G contains an infinite path $v_0, v_1, \ldots$ such that $v_i = f(v_{i+1})$.*

Proof. Let $\mathcal{P}$ be the set of all finite paths of the form $v, f(v), f(f(v)), \ldots$ ending on V_0. Since G is infinite, $\mathcal{P}$ is infinite too, and since V_0 is finite, infinitely many paths in $\mathcal{P}$ end at the same vertex $v_0 \in V_0$. Of these paths, infinitely many also agree on their penultimate vertex $v_1 \in V_1$, because V_1 is finite, and so on. Although the set of paths considered decreases from step to step, it is infinite after any finite number of steps. Therefore, v_n is defined for any $n = 0, 1, 2 \ldots$. $\square$

Theorem 151. *Let k, r, p be positive integers. Then there is a positive integer $n = n(k, r, p)$ with the following property. If the k-subsets of an n-set are colored with r colors, then there is a monochromatic p-set, i.e., a p-set all of whose k-subsets have the same color.*

Proof. Let us denote by $[n]$ the set $\{1, \ldots, n\}$ and by $[n]^k$ the set of all k subsets of $[n]$.

Assume to the contrary that the theorem fails for some k, r, p, i.e. for each n, there is a coloring of $[n]^k$ with r colors such that $[n]$ contains no monochromatic p-set. Let us call such a coloring *bad*. Our aim is to combine these bad colorings into a bad coloring of the set N^k, where $N = \{1, 2, \ldots\}$.

For each n, let V_n denote the set of all bad coloring of $[n]^k$. For an arbitrary $c \in V_n$, let $f(c)$ be the restriction of c to $[n-1]^k$. Clearly, $f(c)$ is bad. Therefore, by Lemma 150 there exists an infinite sequence of bad colorings $c_1, c_2, \ldots$ such that $c_n = f(c_{n+1})$. For every m, all colorings c_n with $n \geq m$ agree on $[m]^k$. Therefore, for each k-subset $Y = \{i_1, \ldots, i_k\}$ of N with $i_1 < \ldots < i_k$, the value of $c_n(Y)$ coincide for all $n \geq i_k$. Let us define $c(Y)$ as this common value $c_n(Y)$. According the infinite version of Ramsey's Theorem, there exists an infinite subsets N' of N all of whose k subsets have the same color. Let's take any subset $S \subset N'$ with p elements, say $i_1 < i_2 < \ldots < i_p$. By the choice of S all k subsets of S have the same coloring in c. On the other hand, c assigns to the k subsets of S the same colors as any coloring c_n with $n \geq i_p$, and we know that c_n is bad, i.e. S is not monochromatic in c_n. This contradiction shows that initial assumption was wrong, and hence the result holds. $\square$

16.5 Ramsey-type results

Ramsey's Theorem with $k = r = 2$ tells us that if the number of vertices in a graph is sufficiently large, then it contains either a big clique or a big independent set, i.e. big cliques and independent sets form a set "unavoidable configurations" in large graphs. Since many graph problems can be reduced to connected graphs, it is natural to ask what are the big "unavoidable *connected* configurations" in large graphs. An answer to this question is given in Theorem 153. We need the following helpful lemma to prove the theorem.

Lemma 152. *A graph G of diameter D and of maximum vertex degree Δ has less than*

$$\frac{\Delta}{\Delta - 2}(\Delta - 1)^D$$

vertices.

Proof. Let v_0 be an arbitrary vertex of G and V_i be the set of vertices of distance i from v_0 ($V_0 = \{v_0\}$). Then V_{D+1} is empty, since the diameter of G is D. Also, $|V_0| = 1$ and $|V_1| = \Delta$. For

$2 \le i \le D$, we have $|V_i| \le (\Delta-1)|V_{i-1}|$, since each vertex v of V_{i-1} has at most Δ neighbours in V_i (remember that v must have a neighbour in V_{i-2}). Thus, by induction, $|V_i| \le \Delta(\Delta-1)^{i-1}$, and hence

$$|V(G)| \le 1 + \Delta \sum_{i=0}^{D-1} (\Delta-1)^i = 1 + \frac{\Delta}{\Delta-2}((\Delta-1)^D - 1) < \frac{\Delta}{\Delta-2}(\Delta-1)^D,$$

as required. □

Theorem 153. *For every natural numbers ℓ, s, t, there is a number $n = n(\ell, s, t)$ such that every connected graph with at least n vertices contains either K_ℓ or $K_{1,s}$ or P_t as an induced subgraph.*

Proof. Let $d = R(\ell-1, s)$ and $n = \frac{d}{d-2}(d-1)^t$. Consider a connected graph G with at least n vertices. If G has a vertex v of degree at least d, then the neighbourhood of v contains either a clique K of size $\ell-1$ (in which case $K \cup \{v\}$ is a clique of size ℓ) or an independent set S of size s (in which case $S \cup \{v\}$ induces a star $K_{1,s}$).

If the maximum vertex degree is bounded in G by d, then the diameter of G must be at least t by Lemma 152. But then any path between any two vertices of distance $t-1$ forms an induced P_t. □

It is known that many graph problems can be reduced to graphs that are not only connected but also co-connected (i.e. whose complement is connected), also known as doubly connected. Therefore, it would be interesting to find large unavoidable doubly connected graphs. An answer to this question was found in [5]. To formulate this result, let us introduce the following notations.

Let

- $K'_{1,s}$ be the graph obtained from $K_{1,s}$ by subdividing exactly one edge exactly once.
- $K_{2,s} - e$ be the graph obtained from $K_{2,s}$ by deleting one edge.
- $K^+_{2,s}$ be the graph obtained from $K_{2,s}$ by connecting the two vertices of degree s and by adding to each of them a pendant edge.

Theorem 154. *For every s, there is an $n = n(s)$ such that every doubly connected graph on at least n vertices contains one of the following graphs as an induce subgraph: P_s, $K'_{1,s}$, $K_{2,s} - e$, $K^+_{2,s}$, $\overline{P}_s$, $\overline{K'}_{1,s}$, $\overline{K_{2,s} - e}$, $\overline{K^+_{2,s}}$.*

Connected and co-connected graphs are examples of *prime* graphs (with respect to modular decomposition). Therefore, it is natural to ask about big unavoidable *prime* graphs. Only a partial answer to this question is available, namely, the list of big unavoidable *prime* graphs has been found only for permutation graphs. This result was obtained in [2] in the terminology of permutations.

Ramsey theory in graphs also has a "bipartite" analog stating that every complete bipartite graph with sufficiently many vertices in each part whose edges are colored with two colors has a big "monochromatic" biclique (i.e. a complete bipartite subgraph). Alternatively,

Theorem 155. *For every s, there is an $n = n(s)$ such that every bipartite graph G with at leas n vertices in each part contains either $K_{s,s}$ or the bipartite complement of $K_{s,s}$.*

Proof. Let $n = s2^{2s}$ and let $G = (A, B, E)$ be a bipartite graph with $|A| \geq n$ and $|B| \geq n$. Consider an arbitrary subset $A' \subseteq A$ with $2s$ vertices. We split the vertices of B into at most 2^{2s} subsets in accordance with their neighbourhood in A'. Since $|B| \geq s2^{2s}$, there must exist a subset $B' \subseteq B$ with at least s vertices. By definition all vertices of B' have the same neighbourhood in A', say A''. If $|A''| \geq s$, then $A'' \cup B'$ is a biclique with at least s vertices in each part. Otherwise, $(A' - A'') \cup B'$ is the bipartite complement of a biclique with at least s vertices in each part. $\square$

The above theorem was proved in the symmetric case. In a similar way, one can show the existence of a non-symmetric bipartite Ramsey number $Rb(s, p)$, i.e. the number with the property that every bipartite graph with at least $Rb(s, p)$ vertices in each part contains either a $K_{s,s}$ or the bipartite complement of $K_{p,p}$. We leave this as an exercise.

Exercises

- Show the existence of the non-symmetric bipartite Ramsey number $Rb(s, p)$.

Theorem 156. *For any natural t and p, there is a number $N = N(t, p)$ such that every bipartite graph with a matching of size at least $N(t, p)$ has either a bi-clique $K_{t,t}$ or an induced matching of size p.*

Proof. For $p = 1$ and arbitrary t, we can define $N(t, p) = 1$. Now, for each fixed t, we prove the lemma by induction on p. Without loss of generality, we prove it for values of the form $p = 2^s$. Suppose we have shown the lemma for $p = 2^s$ for some $s \geq 0$. Let us now show that it is sufficient to set $N(t, 2p) = Rb(t, Rb(t, N(t, p)))$, where Rb is the non-symmetric bipartite Ramsey number.

Consider a graph G with a matching of size at least $Rb(t, Rb(t, N(t, p)))$. Without loss of generality, we may assume that G contains no vertices outside of this matching. We also assume that G does not contain an induced $K_{t,t}$, since otherwise we are done. Then G must contain the bipartite complement of a $K_{Rb(t,N(t,p)),Rb(t,N(t,p))}$ with vertex classes, say, A and B. Now let C and D consist of the vertices matched to vertices in A and B respectively in the original matching in G.

Note that A, B, C, D are pairwise disjoint. $G[A \cup C]$ and $G[B \cup D]$ now each contain a matching of size $Rb(t, N(t, p))$. There are no edges between A and B. However there may exist edges between C and D. By our assumption, $G[C \cup D]$ is $K_{t,t}$-free, therefore it must contain the bipartite complement of $K_{N(t,p),N(t,p)}$, with vertex sets $C' \subset C$, $D' \subset D$. Let $A' \subset A$ and $B' \subset B$ be the set of vertices matched to C' and D' respectively in the original matching in G. Now there are no edges in $G[A' \cup B']$ and none in $G[C' \cup D']$, but $G[A' \cup C']$ and $G[B' \cup D']$ both contain a matching of size $N(t, p)$. Since G is $K_{t,t}$-free, by the induction hypothesis, we conclude that they both contain an induced matching of size p. Putting these together we find that G contains an induced matching of size $2p$. $\square$

Exercises

- Extend Theorem 156 to a non-bipartite case as follows: show that every graph containing a sufficiently large matching contains either a big clique or a big induced biclique or a large induced matching.

17 On the speed of hereditary graph properties

The question of deciding whether a certain graph property P is valid for almost all graphs or not, is based on estimating the size of the property, i.e. the number of n-vertex graphs in P. Recently, a considerable attention has been given to the study of the size (also known as the speed) of hereditary graph properties [1]. In this section we present some of the results on this topic.

As before, we denote by $\Gamma(n)$ the set of all n-vertex labeled graphs and by $P(n)$ the set of graphs from $\Gamma(n)$ that possess property P. We assume that P is infinite, i.e. $|P(n)| > 0$ for all $n > 0$. One of such properties is the set of all complete graphs $\mathcal{K} = \{K_1, K_2, \ldots\}$. Moreover, $\mathcal{K}$ is a minimal infinite hereditary class of graphs. Indeed, for every *proper hereditary* subclass X of $\mathcal{K}$ at least one of the graph of $\mathcal{K}$ must be excluded (forbidden). Clearly, by forbidding K_n we exclude from $\mathcal{K}$ all graphs with at least n vertices.

Similarly, the set $\overline{\mathcal{K}} = \{\overline{K}_1, \overline{K}_2, \ldots\}$ of all edgeless graphs is a minimal infinite hereditary class. From Ramsey's theorem it follows that these are the *only* two minimal infinite hereditary classes of graphs.

Exercises

- What is the number of n-vertex labelled graphs with one edge? two edges?
- What is the number of n-vertex labelled paths?
- What is the number of n-vertex labelled stars, i.e. graphs of the form $K_{1,n-1}$?
- What is the number of n-vertex labelled graphs in the class $Free(P_3, K_2)$?
- What is the number of n-vertex labelled graphs in the class $Free(K_3, \overline{K}_3)$?

Each of the classes $\mathcal{K}$ and $\overline{\mathcal{K}}$ contains exactly one n-vertex graph for each value of n. Let us call a hereditary property P *constant* if there exist a constant c such that $|P(n)| \leq c$ for all $n > 0$. The following theorem characterizes the family of constant hereditary properties.

Theorem 157. *The following statements are equivalent for any hereditary property P:*

(1) P is constant;

(2) there exists an n_0 such that $P(n) - (\mathcal{K}(n) \cup \overline{\mathcal{K}}(n))$ is empty for all $n > n_0$;

(3) none of the following classes is a subclass of P:

- $\mathcal{S}$ *the class of graphs each of which is either an edgeless graph or a star (i.e. a graph of the form $K_{1,n}$ for some n),*
- $\mathcal{E}$ *the class of graphs with at most one edge,*
- $\overline{\mathcal{S}}$ *the class of complements of graphs in $\mathcal{S}$,*
- $\overline{\mathcal{E}}$ *the class of complements of graphs in $\mathcal{E}$.*

Proof. Clearly, (2) implies (1). Also, it is not difficult to see that $|\mathcal{S}(n)| = |\overline{\mathcal{S}}(n)| = n + 1$ and $|\mathcal{E}(n)| = |\overline{\mathcal{E}}(n)| = \binom{n}{2} + 1$. Therefore, (1) implies (3). It remains to show that (3) implies (2).

Suppose to the contrary that the set $Q = P - (\mathcal{K} \cup \overline{\mathcal{K}})$ contains infinitely many graphs. By Ramsey's Theorem, Q contains graphs with arbitrarily large clique or arbitrarily large independent set. Assume Q contains graphs with arbitrarily large independent set. Then for each k there exist a graph G on Q with an independent set of size k. Since G contains an edge (no edgeless graph belong to Q), then G contains an induced subgraph of the form $K_{1,s} + \overline{K}_{k-s}$ with $s \geq 1$. All these graphs also belong to Q (since P is hereditary). Among these graphs, there exist graphs with arbitrarily large value of s or $k - s$. In the first case, P contains all graphs in $\mathcal{S}$, while in the second case it contains all graphs in $\mathcal{E}$.

If Q contains graphs with arbitrarily large clique, then P contains either all graphs in $\overline{\mathcal{S}}$ or all graphs in $\overline{\mathcal{E}}$. $\square$

Determining exact values of $|P_n|$ is possible only for very simple classes of graphs. More frequently, the question of interest is the asymptotic behaviour of $|P(n)|$. In particular, the following value, known as the entropy of P, has been studied for hereditary classes:

$$Entropy(P) = \lim_{n \to \infty} \frac{\log_2 |P(n)|}{\binom{n}{2}}.$$

In other words, it is the limit of the ratio of the logarithm of the number of n-vertex labelled graphs in P to the logarithm of the number of all n-vertex labelled graphs. To explain the role of logarithm in this formula, let us observe that the logarithm of the number of all n-vertex labelled graphs can be viewed as the length of a binary word representing an arbitrary labelled graph with n vertices. Moreover, in case of arbitrary graphs, this is the minimum length of such a word, because for each pair of vertices we need exactly one bit of information to describe the adjacency of these vertices. There are $\binom{n}{2}$ pairs of vertices, therefore we need $\binom{n}{2}$ bits of information.

However, if we know that our graph G belongs to a particular class P, we may need fewer bits than $\binom{n}{2}$ to describe G. In this case, $\log_2 |P(n)|$ gives a low bound on the number of bits, because the number of different binary words needed to describe graphs in $P(n)$ cannot be smaller than the number of graphs. Since there are 2^k binary words of length k, we need at least $\log_2 |P(n)|$ bits to describe an arbitrary graph in $P(n)$.

For an arbitrary graph H, let $Free_m(H)$ denote the class of graphs containing no subgraph (not necessarily induced) isomorphic to H. Then

$$Entropy(Free_m(H)) = \frac{\chi(H) - 2}{\chi(H) - 1}.$$

This intriguing similarity between the entropy of $Free_m(H)$ and Theorem 139 has a simple explanation. Since $ex(n, H)$ is the maximum number of edges in an n-vertex graph in the class $Free_m(H)$, and this class is closed under deletion of edges, $ex(n, H)$ can also be viewed as the minimum number of bits needed to describe an arbitrary graph G from $Free_m(H)$, i.e. for each of the $ex(n, H)$ possible edges of G we need to describe whether it is present in G or not.

In general, the entropy of a hereditary class P can be described as follows. Let $\mathcal{E}_{i,j}$ denote the class of graphs whose vertices can be partitioned into at most i clique and j independent sets. The *index* of P, denoted $k(P)$, is the maximum k such that P contains a class $\mathcal{E}_{i,j}$ with $i + j = k$. Then

$$Entropy(P) = \frac{k-1}{k}.$$

18 Minors and minor-closed graph classes

In addition to the subgraph and induced subgraph relation, one more partial order on graphs is important in graph theory.

Definition 158. Let xy be an edge in a graph G. The CONTRACTION of the edge xy is the operation consisting in deleting the vertices x and y from G and adding a new vertex which is adjacent to every vertex of $G - \{x, y\}$ that is adjacent to x or y in G.

Definition 159. A graph H is said to be a MINOR of a graph G, denoted $H \preccurlyeq G$, if H can be obtained from G by a (possibly empty) sequence of vertex deletions, edge deletions and edge contractions.

Definition 160. A class X of graphs is said to be MINOR-CLOSED if $G \in X$ implies $H \in X$ for every minor H of G.

EXERCISES

- Is a minor-closed class of graphs hereditary?
- Show that every minor-closed class of graphs can be described by a set of minimal forbidden minors.
- Can you characterize the class of K_3-minor-free graphs, i.e. graphs containing no K_3 as a minor?
- Is the class of bipartite graphs minor-closed?
- Is the class of graphs of vertex degree at most 3 minor-closed?
- Is the class of graphs of vertex degree at most 2 minor-closed?

18.1 Hadwiger's conjecture

The following conjecture was proposed by Hadwiger.

Conjecture 161. *If $\chi(G) \geq r$, then $G \succcurlyeq K_r$.*

For any fixed r, the conjecture is equivalent to saying that every graph without a K_r minor can be colored with $r - 1$ colours.

EXERCISES

- Show that Hadwiger's conjecture holds for $r \leq 3$.

Hadwiger's conjecture was also proved for $k = 4, 5, 6$ and still open for $k = 7$. Below we discuss some results about K_4-minor-free and K_5-minor-free graphs.

18.1.1 $r = 4$

For $r = 4$, the graphs without a K_4 minor can be characterized as follows.

Proposition 162. *A graph with at least 3 vertices is edge-maximal without a K_4 minor if and only if it can be constructed recursively from triangles by pasting along K_2s.*

One of the interesting consequences of Proposition 162 is that all edge-maximal K_4-minor-free graphs have the same number of edges.

Corollary 163. *Every edge-maximal graph G without a K_4 minor has $2|V(G)| - 3$ edges.*

<table><tr><td>

Exercises

- Prove corollary 163 by induction on $|V(G)|$.

</td></tr></table>

One more important consequences of Proposition 162 is that Hadwiger's conjecture holds for $k = 4$.

Corollary 164. *Hadwiger's conjecture holds for $k = 4$.*

<table><tr><td>

Exercises

- Prove corollary 164. Hint: use the fact that if G arises from G_1 and G_2 by pasting along a complete graph then $\chi(G) = \max\{\chi(G_1), \chi(G_2)\}$ (see the proof of Proposition 106).

</td></tr></table>

18.1.2 $r = 5$ and planar graphs

An important subclass of K_5-minor-free graphs is the class of planar graphs.

Definition 165. A graph is said to be PLANAR if it can be drawn in the plane in such a way that no two edges intersect each other. Drawing a graph in the plane without edge crossing is called EMBEDDING the graph in the plane (or PLANAR EMBEDDING or PLANAR REPRESENTATION).

<table><tr><td>

Exercises

- Is the class of planar graphs minor-closed?

</td></tr></table>

Given a planar representation of a graph G, a *face* (also called a *region*) is a maximal section of the plane in which any two points can be joint by a curve that does not intersect any part of G. When we trace around the boundary of a face in G, we encounter a sequence of vertices and edges, finally returning to our final position. Let $v_1, e_1, v_2, e_2, \ldots, v_d, e_d, v_1$ be the sequence obtained by tracing around a face, then d is the degree of the face. Some edges may be encountered twice because both sides of them are on the same face. A tree is an extreme example of this: each edge is encountered twice. The following result is known as Euler's Formula.

Theorem 166 (Euler's Formula)**.** *If G is a connected planar graph with n vertices, m edges and f faces, then*

$$n - m + f = 2.$$

Proof. We prove by induction on m. If $m = 0$, then $G = K_1$, a graph with 1 vertex and 1 face. The formula is true in this case. Assume it is true for all planar graphs with fewer than m edges and suppose G has m edges.

Case 1: G is a tree. Then $m = n - 1$ and obviously $f = 1$. Thus $n - m + f = 2$, and the result holds.

Case 2: G is not a tree. Let C be a cycle in G, and e an edge in C. Consider the graph $G - e$. Compared to G this graph has the same number of vertices, one edge fewer, and one face fewer (since removing e coalesces two faces in G into one in $G - e$). By the induction hypothesis, in $G - e$ we have $n - (m - 1) + (f - 1) = 2$. Therefore, in G we have $n - m + f = 2$, which completes the proof. □

Corollary 167. *If G is a connected planar graph with $n \geq 3$ vertices and m edges, then $m \leq 3n - 6$. If additionally G has no triangles, then $m \leq 2n - 4$.*

Proof. If we trace around all faces, we encounter each edge exactly twice. Denoting the number of faces of degree k by f_k, we conclude that $\sum_k kf_k = 2m$. Since the degree of any face in a (simple) planar graph is at least 3, we have

$$3f = 3\sum_{k\geq 3} f_k \leq \sum_{k\geq 3} kf_k = 2m.$$

Together with the Euler's formula, this proves that $m \leq 3n - 6$. If additionally, G has no triangles, then

$$4f = 4\sum_{k\geq 4} f_k \leq \sum_{k\geq 4} kf_k = 2m.$$

and hence $m \leq 2n - 4$. □

Corollary 168. *K_5 and $K_{3,3}$ are not planar.*

Proof. For K_5 we have $n = 4$, $m = 10$ and $m > 3n - 6$. Therefore, K_5 is not planar by Corollary 167. For $K_{3,3}$ we have $n = 6$, $m = 9$, and $m > 2n - 4$. Noticing that $K_{3,3}$ is triangle-free, we conclude, again by Corollary 167, that $K_{3,3}$ is not planar. □

Corollary 169. *Every planar graph has a vertex of degree at most five.*

Proof. Suppose a planar graph G has n vertices and m edges. If $n \leq 6$, the statement is obvious. So suppose $n > 6$. If we let D be the sum of the degrees of the vertices of G. If each vertex of G had degree at least 6, then we would have $D \geq 6n$. On the other hand, since G is planar, we have

$$D = 2m \leq 2(3n - 6) = 6n - 12.$$

Therefore, G must have a vertex of degree at most 5. □

Corollary 170. *The chromatic number of any planar graph is at most 5.*

Proof. We use induction on the number of vertices. For graphs with at most 5 vertices, this is trivial. Let G be a planar graph with $n > 5$ vertices, x a vertex of degree (at most) 5 in G and $y_1, \ldots, y_5$ the neighbours of x. We know at least two of the neighbours of x, say y_1 and y_2, must be non-adjacent, since otherwise G is not planar (contains a K_5). Let us delete the edges xy_3, xy_4, xy_5 and contract the edges xy_1 and xy_2 into a single vertex z and denote the resulting graph by G'. By induction we know that there is a 5-coloring $c : V(G') \to \{1, 2, 3, 4, 5\}$ of G'. We extend this coloring to G as follows. Every vertex of G' different from z is also a vertex of G and we keep the same color for it. For y_1 and y_2 we assign $c(y_1) = c(y_2) = c(z)$. Since in the neighbourhood of x only 4 colors are used, we may assign the remaining color to x. $\square$

Corollary 170 has an important generalization known as the Four Colour Theorem.

Theorem 171 (Four Colour Theorem)**.** *The chromatic number of any planar graph is at most 4.*

This theorem confirms Hadwiger's conjecture for planar graphs. Moreover, it also proves Hadwiger's conjecture for all K_5-minor-free graphs, because for $r = 5$ Hadwiger's conjecture is equivalent to the Four Colour Theorem. It is interesting that K_5-minor-free graphs inherit many more properties of planar graphs. In particular, similarly to planar graphs a K_5-minor-free graph with n vertices has at most $3n - 6$ edges.

As we have seen, K_5 or $K_{3,3}$ are not planar. Also, it is not difficult to check that every proper minor of K_5 or $K_{3,3}$ is a planar graph. Therefore, K_5 and $K_{3,3}$ are two minimal forbidden minors of the class of planar graphs. The following theorem, which is due to Kuratowski, states that K_5 and $K_{3,3}$ are the *only* minimal forbidden minors for the class of planar graphs.

Theorem 172 (Kuratowski's Theorem)**.** *A graph is planar if and only if it does not contain K_5 and $K_{3,3}$ as minors.*

In the next section we will show that *every* minor-closed class of graphs can be described by *finitely* many forbidden minors.

18.2 On the speed of hereditary graph properties (continued)

As we have seen earlier, K_3-minor-free graphs are forests and we known that edge-maximal n-vertex graphs in this class have $n-1$ edges. Also, K_4-minor-free n-vertex graphs have at most $2n - 3$ edges, and K_5-minor-free n-vertex graphs have at most $3n - 6$ edges. More generally, it is known that for every minor closed class X of graphs there is a constant c such that every n-vertex graph in X has at most cn edges. This implies the following upper bound on the number of n-vertex labelled graphs in minor-closed classes.

Theorem 173. *For every minor-closed class X of graphs there is a constant c such that the number of n-vertex labelled graphs in X is at most n^{cn} for all $n \geq 1$.*

18.3 Minors and Well-Quasi-Ordering

One of the most fundamental results about the graph minor relation is that the set of all simple graph is well-quasi-ordered with respect to this relation. To state this formally, we need to recall and introduce some notions from the theory of partially ordered sets.

A binary relation on a set X is a subset of X^2, where X^2 denotes the set of all ordered pairs of elements of X. If Q is a binary relation and $(x, y) \in Q$, we will also write $x \leq_Q y$. If $x \leq_Q y$ and $x \neq y$, we will write $x <_Q y$.

A *quasi-order* on X is a binary relation Q which reflexive ($(x, x) \in Q$ for each $x \in X$) and transitive ($(x, y) \in Q$ and $(y, z) \in Q$ imply $(x, z) \in Q$). If the relation is additionally antisymmetric, then it is a *partial order*.

Exercises

- Show that subgraph, induced subgraph and minor relation on graphs is a partial order (and hence a quasi-order).
- Show that the inclusion relationship on the set of all hereditary classes is a partial order.
- Given a graph G, define a binary relation Q on $V(G)$ by $(x, y) \in Q$ if and only if $N[x] \subseteq N[y]$, where $N[x]$ is the closed neighbourhood of x, i.e. $N[x] = N(x) \cup \{x\}$. Show that Q is a quasi-order, but not a partial order.

Two elements x, y are *comparable* with respect to a quasi-order $Q \subseteq X^2$ if either $(x, y) \in Q$ or $(y, x) \in Q$. Otherwise, they are *incomparable*.

Exercises

- Determine whether P_4, C_4 and C_5 are comparable (in pairs) with respect to subgraph, induced subgraph and minor relation.

A set of pairwise comparable elements is called a *chain* and a set of pairwise incomparable elements is called an *antichain*.

Definition 174. A quasi-order Q on a set X is a *well-quasi-order* (wqo) if it contains neither infinite antichains nor infinite strictly decreasing chains (i.e. sequences $x_1 >_Q x_2 >_Q x_3 >_Q \ldots$).

Exercises

- Does the set of all simple graphs contains infinite strictly decreasing chains with respect to subgraph, induced subgraph or minor relation? Does it contain infinite antichains?
- Does the family of all hereditary classes contain infinite strictly decreasing chains with respect to the inclusion relation? Does it contain infinite antichains?

Theorem 175. *The set of all simple graph is well-quasi-ordered by the minor relation.*

The prove of this theorem is long and complicated and is beyond the scope of the module. Instead, we will present some important corollaries from this result and will prove an important special case of the theorem.

The following corollary from Theorem 175 can be viewed as a generalization of Kuratowski's Theorem for planar graphs.

Corollary 176. *Every minor-closed class of graphs can be described by finitely many forbidden minors.*

To prove an important special case of Theorem 175, we need a number of auxiliary results. We also need to update the terminology.

Let Q be a quasi-order on a set X and $x_1, x_2, \ldots$ an infinite sequence of elements of X. Any pair (x_i, x_j) with $i < j$ such that $x_i \leq_Q x_j$ will be called a *good* pair of this sequence. Any infinite sequence containing a good pair is called *good*, otherwise it is bad.

The terminology of good and bad sequences, provides an alternative definition of well-quasi-ordering: a quasi-order Q on a set X is a well-quasi-order if and only if every infinite sequence $x_1, x_2, \ldots$ of elements of X is good. Moreover, if Q is wqo then every infinite sequence contains an infinite chain, not just a good pair.

Lemma 177. *Let X be an infinite set and Q a well-quasi-order on X. Then every infinite sequence $x_1, x_2, \ldots$ of elements of X contains an infinite increasing subsequence, i.e. a subsequence $x_{i_1}, xi_2, x_{i_3} \ldots$ with $i_1 < i_2 < i_3 < \ldots$ such that*

$$x_{i_1} \leq_Q x_{i_2} \leq_Q x_{i_3} \leq_Q \ldots.$$

.

Proof. We color the pairs $(x_i x_j)$ with $i < j$ with 3 colors as follows: color 1 if x_i and x_j are incomparable, color 2 if $x_i \leq_Q x_j$, and color 3 if $x_i >_Q x_j$. By (infinite) Ramsey's Theorem, there must be an infinite monochromatic subset of X. This subset cannot be of color 1 or 3, since otherwise X contains either an infinite antichain or an infinite strictly decreasing chain. Therefore, the given sequence contains an infinite increasing subsequence. $\square$

Let $\leq$ be a quasi-ordering on a set X. For finite subsets $A, B \subseteq X$, we write $A \leq B$ if there is an injective mapping $f : A \to B$ (i.e. a mapping that maps different elements of A to different element of B) such that $a \leq f(a)$ for all $a \in A$. This naturally extends the quasi-order $\leq$ to a quasi-order on the set of all finite subsets of X, denoted X^*.

Lemma 178. *If X is well-quasi-ordered by $\leq$, then so is X^*.*

Proof. Assume X^* is not wqo. We construct inductively an infinite bad sequence $A_0, A_1, A_2, \ldots$ in X^* as follows. We start with the empty sequence. Given a natural n, assume that A_i has been constructed for all $i < n$, and that there exists a bad sequence in X^* starting with $A_0, A_1, \ldots, A_{n-1}$.

Choose $A_n \in X^*$ so that some bad sequence starts with $A_0, A_1, \ldots, A_n$ and A_n is as small as possible. Clearly, the sequence $A_0, A_1, A_2, \ldots$ constructed in this way is a bad sequence in X^*; in particular $A_n \neq \emptyset$ for all n. For each n, pick an element $a_n \in A_n$ and define $B_n := A_n - \{a_n\}$. Since X is wqo by $\leq$, the sequence $a_0, a_1, a_2, \ldots$ has an infinite increasing subsequence $a_{i_0}, a_{i_1}, a_{i_2}, \ldots$, by Lemma 177. By the minimal choice of A_{i_0}, the sequence

$$A_0, A_1, \ldots, A_{i_0-1}, B_{i_0}, B_{i_1}, B_{i_2}$$

is not bad, i.e. it contains a good pair. This pair cannot be of the form (A_i, A_j), since all A-sets are incomparable. Also, this pair is not of the form (A_i, B_j), since we know that $B_j \leq A_j$, and if $A_i \leq B_j$ then by transitivity $A_i \leq A_j$, which is impossible. Therefore, we assume that a good pair has the form (B_i, B_j) with $B_i \leq B_j$. Then by extending the injection $B_i \to B_j$ with $a_i \to a_j$ we conclude that $A_i \leq A_j$, which contradicts the fact that A_i and A_j are incomparable. This final contradiction shows that our assumption that X^* is not wqo was wrong. $\square$

18.3.1 Topological minors

A subdivision of an edge is the operation of creation of a new vertex on the edge. In a sense, this is an operation opposite to edge contraction, i.e. if an xy has been subdivided by a new vertex z, then by contracting the edge xz in the new graph, we again obtain the original graph.

A graph H is a subdivision of a graph G if H is obtained from G by a sequence of edge subdivision.

Definition 179. A graph H is a TOPOLOGICAL MINOR of a graph G if a subdivision of H is a subgraph of H.

EXERCISES
• Show that if H is a topological minor of G, then H is a minor of G.

Theorem 180. *The set of finite trees is well-quasi-ordered by the topological minor relation.*

Proof. We will prove a slightly stronger result, where each tree is equipped with a root and for two rooted trees T and T' we will write $T \sqsubseteq T'$ if there is an isomorphism ϕ which maps a subdivision of T into a subtree of T' and preserves the parent-child relation. It is not difficult to see that the relation $\sqsubseteq$ is stronger than the topological minor relation.

Assume to the contrary that $\sqsubseteq$ is not a well-quasi-order on the set of finite trees. By analogy with Lemma 178 we construct a bad sequence of trees by choosing a segment of the first i trees $T_0, T_1, \dots, T_{n-1}$ and inductively assuming that there is a bad sequence starting with this segment. Then T_n is chosen so that it has minimum number of vertices among all trees for which $T_0, T_1, \dots, T_n$ starts a bad sequence.

For each n, we denote by r_n the root of T_n and by A_n the set of trees obtained by removing r_n from by T_n. Let is show that the union $A := \cup_{n\geq 0} A_n$ of all these threes is wqo by $\sqsubseteq$.

Let $T^{(0)}, T^{(1)}, T^{(2)}, \dots$ be any infinite sequence of trees in A. For every natural k, let $n = n(k)$ by such that $T^{(k)} \in A_n$. Pick k with the smallest $n(k)$. Then the sequence

$$T_0, T_1, \dots, T_{n(k)-1}, T^{(k)}, T^{(k+1)}, \dots$$

must be good, because $T^{(k)}$ has strictly less vertices than $T_{n(k)}$. Therefore, this sequence has a good pair (T, T'). Assume T is one of the first $n(k)$ members of the sequence. Since these members are pairwise incomparable, T' is not one of them. Therefore, T' is a subtree of T_ℓ with $\ell \geq n(k)$. But then $T \sqsubseteq T' \sqsubseteq T_{ell}$, in which case (T, T_ℓ) is a good pair. Since this is impossible, we conclude that both T and T' belong to A and hence the sequence $T^{(0)}, T^{(1)}, T^{(2)}, \dots$ is good. This shows that A is well-quasi-ordered by $\sqsubseteq$.

Therefore, A^* is wqo and hence the sequence $A_0, A_1, \dots$ has a good pair, say (A_i, A_j) with $i < j$. This means that every tree of A_i can be embedded (with respect to $\sqsubseteq$) into a tree of A_j so that different trees of A_i are embedded into different trees of A_j. Now we extend the union of these embeddings into a mapping $\phi : T_i \to T_j$ by letting $\phi(r_i) = r_j$. This map defined an embedding of T_i into T_j showing that $T_i \sqsubseteq T_j$. Therefore, (T_i, T_j) is a good pair in our initial bad sequence. This contradiction shows that the set of finite trees is wqo by $\sqsubseteq$ and hence by the topological minor relation. □

References

[1] J. Balogh, B. Bollobás and D. Weinreich, The speed of hereditary properties of graphs, *J. Combin. Theory Ser. B* 79 (2000) 131–156.

[2] R. Brignall, S. Huczynska, and V. Vatter, Decomposing simple permutations, with enumerative consequences. *Combinatorica*, 28 (4) (2008) 385–400.

[3] M. Chudnovsky, N. Robertson, P. Seymour, R. Thomas, The strong perfect graph theorem, *Annals of Mathematics* 164 (1) (2006) 51-229.

[4] D.G. Corneil, H. Lerchs, Bham, L. Stewart Burlingham, Complement reducible graphs, *Discrete Applied Mathematics*, 3 (1981) 163-174.

[5] G. Ding and P. Chen, Unavoidable doubly connected large graphs, *Discrete Mathematics*, 280 (2004) 1-12.

[6] S. Földes, P.L. Hammer, Split graphs, *Congr. Numer.* 19 (1977) 311-315.

[7] V. Giakoumakis and J.-M. Vanherpe, Bi-complement reducible graphs, *Advances in Appl. Math.* 18 (1997) 389-402.

[8] H. Prüfer, Neuer Beweis eines Satzes über Permutationen, *Arch. Math. Phys.* 27 (1918) 742-744.

[9] I.E. Zverovich, r-bounded k-complete bipartite bihypergraphs and generalized split graphs. *Discrete Math.* 247 (2002) 261–270.