Lecture Notes on Graph Theory

Vadim Lozin

1 Introductory concepts

A graph $G = (V, E)$ consists of two finite 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, $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$.

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

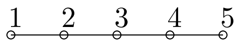

Figure 1: Graph $G = (V, E)$ with $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}\}$

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, we obtain *hypergraphs*.
- by allowing V and E to be infinite sets, we obtain *infinite graphs*.

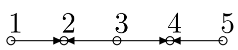

Figure 2: An oriented graph $G = (V, E)$ with $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1, 2), (3, 2), (3, 4), (4, 5)\}$

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

All graphs in these notes will be simple. For notational convenience, instead of representing an edge by $\{a, b\}$ we shall denote it by ab. The following is a list of basic graph notions and notations.

- The sets of vertices and edges of a graph G will be denoted $V(G)$ and $E(G)$, respectively.
- Two vertices x, y of a graph G are said to be *adjacent* if $xy \in E(G)$, in which case we also say that x is *connected* to y or x is a *neighbor* of y. If $xy \notin E$, then x and y are *nonadjacent* (*not connected*, *non-neighbors*).
- If $e = xy$ is an edge of G, then e is incident with x and y. We also say that x and y are the endpoints of e.
- The *neighborhood* of a vertex $v \in V(G)$, denoted $N(v)$, is the set of vertices adjacent to v, i.e., $N(v) = \{u \in V(G) \mid vu \in E(G)\}$. For a subset $U \subseteq V(G)$, $N(U) = \bigcup_{v \in U} N(v)$.
- The *degree* of a vertex $v \in V(G)$, denoted $deg(v)$, is the number of edges incident with v. Alternatively, $deg(v) = |N(v)|$.
- 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.
- A graph $G_1 = (V_1, E_1)$ is 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 (with all incident edges) and some edges. G_1 is a *spanning subgraph* of G_2 if $V_1 = V_2$, i.e., G_1 can be obtained from G_2 by deleting some edges but not vertices. G_1 is an induced subgraph of G_2 if 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 (with all incident edges) but not edges.
- A *path* in a graph is a sequence of 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 graph is *connected* if any two vertices of the graph are joint by a path. If a graph G is *disconnected* (i.e., not connected), then every maximal (with respect to inclusion) connected subgraph of G is called a *connected component* of G.
- 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 $\max\limits_{a,b \in V(G)} dist(a, b)$.
- 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 $f(u)f(v) \in E_2$.
- 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 $\overline{G}$.
- A graph G is *self-complementary* if G is isomorphic to its complement.
- The *adjacency matrix* of a graph G with vertex set $V = \{1, 2, \ldots, n\}$ is a binary $n \times n$ matrix $A = (a_{i,j})$ such that $a_{i,j} = 1$ if i is adjacent to j, and $a_{i,j} = 0$ otherwise.
- A set of pairwise adjacent vertices in a graph is called a clique, and a set of pairwise non-adjacent vertices is an independent set.

2 Graphs and graph operations

- C_n is a chordless cycle on n vertices,
- P_n is a chordless path on n vertices,
- K_n is complete graph with n vertices,
- O_n is an empty (edgeless) graph with n vertices,
- $K_{n,m}$ is a complete bipartite graph with parts of size n and m.

For two graphs H and G with $V(H) \cap V(G) = \emptyset$, we denote by $H + G$ the disjoint union of H and G, i.e. the graph of form $V(H+G) = (V(H) \cup V(G), E(H+G) = E(H) \cup E(G))$. In particular, if $G = H$, we shall write $2G$ instead of $G + G$. Also, nG will denote n disjoint copies of a graph G.

The *k-th power* of a graph G, denoted G^k, is defined as follows: $V(G^k) = V(G)$ and $u, v \in V(G^k)$ are adjacent if and only if the distance $dist(u, v)$ in G is at most k.

Given an edge $e \in E(G)$, we shall write $G - e$ to denote the graph obtained from G by removal of edge e. Similarly, given a vertex x of G, we shall denote by $G - x$ the graph obtained from G by removal of x. Obviously, along with the vertex we delete all the edges incident with it.

3 Some results

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

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

Corollary 3 *In every graph, the number of vertices of odd degree is even.*

Theorem 4 *Diameter of almost all graphs is 2.*

Proof. Let $\Gamma(n)$ denote the set of all labeled graphs on n vertices and $\Gamma^2(n)$ 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 neighborhood of u. Denote $k = |U|$. We know that $v \notin U$ (otherwise $dist(u, v) = 1$) and v has no neighbors 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 $\Gamma'_{u,U}(n) = \bigcup\limits_{v \notin (U \cup \{u\})} \Gamma'_{u,v,U}$. Then

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

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

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

In order to estimate $\sum\limits_{k=0}^{n-2} \binom{n-1}{k} 2^{-k}$, 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 $\Gamma'(n) = \bigcup\limits_u \Gamma'_u(n)$, 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. ∎

4 Classes of graphs

For a set of graphs X, we denote by $\overline{X} = \{G \in X \mid \overline{G} \in X\}$ the set of complementary graphs, and by $X_n = \{G \in X \mid |V(G)| = n\}$ the set of n-vertex graphs in X.

4.1 Simple classes

- The class of all graphs. It will be denoted by Γ.
- The class of empty (edgeless) graphs. Clearly for every $n \geq 1$, there exists only one empty graph on n vertices. The class of all empty graphs will be denoted by O.
- The class complete graphs, i.e., graphs containing all possible edges, will be denoted by K. Clearly $\overline{K} = O$.

4.2 Comparability Graphs

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

Definition 5 *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.

Exercise: show that C_5 is not transitively orientable.

4.3 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.

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 6 *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 $G[\pi]$. ■

Theorem 7 *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. ■

4.4 Bipartite Graphs

Definition 8 *A graph $G = (V, E)$ is bipartite if the vertices of G can be partitioned into two subsets V_1 and v_2 in such a way that no two vertices in the same subset are connected by an edge.*

Exercise: Show that every bipartite graph is a comparability graph. (Hint: find a transitive orientation of a bipartite graph).

Theorem 9 *A graph G is bipartite if and only if G has no cycles of odd length.*

Proof. Let G be a bipartite graph with partite sets V_1 and 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. ■

4.5 Trees

Definition 10 *A tree is a connected graph without cycles.*

Theorem 11 *If $G = (V, E)$ is a connected graph, then the following are equivalent.*

(a) G is a tree.

(b) G has no cycles.

(c) For every pair of distinct vertices u and v in G, there is exactly one path from u to v.

(d) $G - e$ is disconnected for any edge $e \in E(G)$.

(e) $|E| = |V| - 1$.

Proof. See [1], page 137. ■

Corollary 12 *Every tree with $n \geq 2$ vertices contains at least 2 vertices of degree 1.*

Proof. Let $(d_1, d_2, \ldots, d_n)$ be the degree sequence of a tree T with $n \geq 2$ vertices. By Lemma 2 and Theorem 11(e), $\sum_{i=1}^{n} d_i = 2n - 2$. Taking into account that d_i is strictly positive for each i, we conclude that at least two d_i's must be equal 1. ■

4.5.1 Counting Trees

Theorem 13 *The number of labeled trees on n vertices is n^{n-2}.*

There are many proofs of this formula. The first proof belongs to Cayley, so sometimes this formula is called *Cayley's Tree Formula.* We will present an elegant proof which is due to Prüfer, which establishes a bijection between the set of all labeled trees and the set of all sequences of length $n-2$ composed of the elements of $\{1, 2, \ldots, n\}$.

Prüfer's Method for Assigning a Sequence to a Labeled Tree T

1. Let $i = 0$, and let $T_0 = T$.
2. Find a vertex of degree 1 in T_i with the smallest label and call it v.
3. Record in the sequence the label of the v's neighbor.
4. Remove v from T_i to create a new tree T_{i+1}.
5. If T_{i+1} has two vertices, then stop. Otherwise, increment i by 1 and go back to step 2.

Notice that none of the vertices of degree 1 of the original tree T appears in the sequence. More generally, each vertex v appears in the sequence exactly $deg(v) - 1$ times. Since each tree leads to a unique sequence, the number of trees is at most the number of sequences, i.e. at most n^{n-2}. In order to show that the number of sequences is at most the number of trees, we will assign to each sequence a unique tree.

Prüfer's Method for Assigning a Labeled Tree with vertex set $\{1, \ldots, n\}$ to a Sequence $\sigma = (j_1, j_2, \ldots, j_{n-2})$, where $j_i \in \{1, \ldots, n\}$ for each $i = 1, \ldots, n-2$.

1. Let $i = 0$, let $\sigma_0 = \sigma$ and let $S_0 = \{1, 2, \ldots, n\}$.
2. Let j be the smallest number in S_i that does not appear in the sequence σ_i.
3. Connect vertex j to vertex j_{i+1} (the first element of σ_i).
4. Remove the first element of σ_i to create a new sequence σ_{i+1}, and remove j from S_i to create a new set S_{i+1}.
5. If the sequence σ_{i+1} is empty, connect the last two vertices in S_{i+1} and stop. Otherwise, increment i by 1 and go back to step 2.

The fact that this procedure results in a tree can be proved by induction on n. Notice that if T_{n-1} is a tree with $n-1$ vertices, then adding to it a vertex and connecting this vertex to any vertex of T_{n-1} result in a tree with n vertices.

4.6 Planar Graphs

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*).

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 14 *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 15 *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 $\sum_k k f_k = 2m$. Since the degree of any face in a (simple) planar graph is at least 3, we have

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

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

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

and hence $m \leq 2n - 4$. ■

Corollary 16 *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 15. 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 15, that $K_{3,3}$ is not planar. ■

Corollary 17 *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. ■

A graph H is said to be a minor of a graph G if H can be obtained from G by a sequence of vertex deletions, edge deletions and edge contractions (contraction of an edge is the operation of substituting an edge uv with a new vertex adjacent to every neighbor of u and v). It is not difficult to see that the class of planar graphs is minor-closed, i.e., if G is a planar graph then every minor of G also is a planar graph. Therefore, every graph containing K_5 or $K_{3,3}$ as a minor is not a planar graph. The converse statement is also true, which is due to Kuratowski:

Theorem 18 *A graph is planar if and only if it does not contain K_5 and $K_{3,3}$ as minors.*

4.6.1 Coloring planar graphs

Given a graph G, a *k-coloring* of the vertices of G is a partition of the vertex set into k independent sets. If such a partition exists, G is said to be *k-colorable.* Another way of viewing a k-coloring is as follows. Instead of assigning vertices to k partitioning sets, assign each vertex one of k colors in such a way that no two vertices of the same color are adjacent. The *chromatic number* of G, denoted $\chi(G)$, is the minimum k such that G is k-colorable.

The famous *Four Color Problem* asks whether each planar graph is 4-colorable. In 1976, Appel and Haken announced that they have found a proof. To complete their proof, they verified thousands of cases with computers using over 1000 hours of computer time. While the Appel-Haken proof is accepted as being valid, mathematicians still search for alternative proofs. The proof of a weaker version of the problem (5-Color Theorem) can be found in [1], page 166. A weaker statement of 6-colorability of planar graphs follows trivially from Corollary 17 by induction.

4.6.2 Graph parameters measuring "non-planarity"

The *thickness* of a graph G, denoted $t(G)$, is the minimum number of planar subgraphs P_i of G such that $\bigcup_i P_i = G$. Obviously $t(G) = 1$ if and only if G is planar. From Corollary 15 it follows that

$$t(G) \geq \lceil \frac{m}{3n-6} \rceil.$$

Moreover, taking into account that for any positive integer numbers a, b

$$\lceil a/b \rceil = \lfloor (a+b-1)/b \rfloor,$$

we have

$$t(G) \geq \lfloor \frac{m+3n-7}{3n-6} \rfloor.$$

Exercise. Show that $t(K_n) \geq \lfloor \frac{n+7}{6} \rfloor$.

The *skewness* of a graph G is the minimum number of edges we have to remove from G in order to obtain a planar graph. From Corollary 15 it follows that

$$sk(K_n) = \binom{n}{2} - 3n + 6, \quad n \geq 3.$$

The *crossing number* $cr(G)$ of a graph G is the minimum number of pairwise intersections of edges possible when drawing the graph on the plane.

Lemma 19 *For any graph with n vertices and m edges, $cr(G) \geq m - 3n$.*

Proof. Consider a drawing of G which has exactly $cr(G)$ crossings. Each of these crossings can be removed by removing an edge from G. Thus we can find a graph with at least $m - cr(G)$ edges and n vertices which is planar. From the Euler's formula we have $m - cr(G) \leq 3n$, and the claim follows. ■

Theorem 20 *If $m > 4n$, then $cr(G) \geq \frac{m^3}{64n^2}$.*

Proof. We will use a probabilistic argument. Let p be a probability parameter to be chosen later, and construct a random induced subgraph H of G by allowing each vertex of G to lie in H independently with probability p.

Now consider a drawing of G with $cr(G)$ crossings. Without loss of generality we may assume that every crossing in this drawing involves four distinct vertices of G.

Since H is an induced subgraph of G, this drawing contains a drawing of H. By Lemma 19, $cr(H) \geq |E(H)| - 3|V(H)|$. Taking mathematical expectation, we obtain $\mathcal{E}(cr(H)) \geq \mathcal{E}(|E(H)|) - 3\mathcal{E}(|V(H)|)$. Since each of the n vertices in G had a probability p of being in H, we have $\mathcal{E}(|V(H)|) = np$. Similarly, since each of the edges in G has a probability p^2 of remaining in H (since both endpoints need to stay in H), then $\mathcal{E}(|E(H)|) = p^2 m$. Finally, every crossing in the drawing of G has a probability p^4 of remaining in H, since every crossing involves four vertices, and so $\mathcal{E}(cr(G)) = p^4 cr(G)$. Thus we have $p^4 cr(G) \geq p^2 m - 3pn$. If we now set p to equal $4n/m$ (which is less than one, since we assume that m is greater than $4n$), we obtain after some algebra $cr(G) \geq \frac{m^3}{64n^2}$. ■

4.7 Hereditary classes of graphs

Definition 21 *A class of graphs is called hereditary if $G \in X$ implies $H \in X$ for any induced subgraph H of G.*

In other words, a class of graphs is hereditary if it is closed under deletion of vertices.

Examples. The set of all graphs is obviously hereditary. Classes of empty (edgeless) and complete graphs are hereditary. The classes of permutation, comparability and bipartite graphs are also hereditary, but the trees is not a hereditary class. The minimal hereditary class containing all trees is the class of forests, i.e., graphs every connected component of which is a tree. Alternatively, a forest is a (not necessarily connected) graph without cycles.

For a set of graphs M, let us denote by $Free(M)$ the class of all graphs containing no graphs from M as induced subgraphs. If $X = Free(M)$, we say that M is the set of forbidden induced subgraphs for X, and any graph G in X is said to be M-free. With this notation, the class of bipartite graphs is $Free(C_3, C_5, C_7, \ldots)$ and the class of forests is $Free(C_3, C_4, C_5, \ldots)$.

Theorem 22 *A class of graphs X is hereditary if and only if there is a set M such that $X = Free(M)$. Moreover, there is only one minimal set M with this property.*

Proof. The sufficiency is obvious. To prove the necessity, let X be a hereditary class. If X contains all graphs, then $M = \emptyset$. Suppose now X does not contain all graphs, and let M be the set of minimal graphs that are not in X (i.e., $H \in M$ if and only if $H \notin X$ but $H - v \in X$ for each vertex $v \in V(H)$). Let G be a graph in X. By definition, for any induced subgraph H of G, we have $H \in X$. Therefore, $G \in Free(M)$. On the other hand, if $G \in Free(M)$, then

$G \in X$. Indeed, if $G \notin X$, then G contains an induced subgraph from M, which contradicts the assumption. Thus, $X = Free(M)$.

Now let us show that that M is the only minimal set such that $X = Free(M)$. Suppose $X = Free(N)$ for some N other than M. Then $M \subseteq N$. Indeed, assume to the contrary that $H \in M - N$. By minimality, any proper induced subgraph of H is in X, and hence is in $Free(N)$. Thus, H contains no induced subgraph in N, and therefore $H \in X = Free(M)$. But this contradicts $H \in M$. ■

Theorem 23 *A hereditary class of graphs $X = Free(M)$ is finite if and only if M contains K_m and $\overline{K}_n$ for some m and n.*

Proof. If M does not contain any K_m, then X includes all complete graphs and hence is infinite. Similarly if M does not contain any $\overline{K}_n$, X includes all empty (edgeless) graphs. Suppose now M contains both K_m and $\overline{K}_n$ for some m and n. Then due to Ramsey Theorem, there is a k such that any graph with at least k vertices contains either K_m or $\overline{K}_n$. Hence X is finite in this case. ■

Theorem 24 *If $X = Free(M_1)$ and $Y = Free(M_2)$, then $X \cup Y$, $X \cap Y$ and $\overline{X}$ are hereditary classes. Moreover, $\overline{X} = Free(\overline{M_1})$, $X \cap Y = Free(M_1 \cup M_2)$.*

Theorem 25 *A class of graphs $Free(M_1)$ is a subclass of a class of graphs $Free(M_2)$ if and only if for every $G \in M_2$, there is $H \in M_1$ such that H is an induced subgraph of G.*

Proof. Assume $Free(M_1) \subseteq Free(M_2)$, and suppose to the contrary 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 $Free(M_2)$, G contains an induced subgraph $H \in M_2$. Due to the assumption, H contains an induced subgraph $H' \in M_1$. Clearly H' is an induced subgraph of G which contradicts the fact that $G \in Free(M_1)$. ■

4.7.1 Examples of hereditary classes

We have seen several examples of hereditary classes: bipartite, permutation, comparability graphs. Below we will list some more important classes.

1. *Chordal graphs.* $CH = Free(C_4, C_5, C_6, \ldots)$.
2. *Co-chordal graphs.* The class of complements to chordal graphs: $\overline{CH} = Free(\overline{C_4}, \overline{C_5}, \ldots)$.
3. *Split graphs.* $CH \cap \overline{CH} = \Gamma(C_4, \overline{C_4}, C_5, \overline{C_5}, \ldots) = \Gamma(C_4, C_5, 2K_2)$.

   Theorem 26 . *The vertex set of every graph G in $Free(C_4, C_5, 2K_2)$ 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 (a, b). 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 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. 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 case, we have a contradiction. Therefore, z does not have neighbors outside C. But than $C' = (C - \{z\}) \cup \{b\}$ is a maximal clique and subgraph $G - C'$ has fewer edges than $G - C$ that contradicts the choice of C. ■

The graphs that admit decomposition into a clique and an independent set are called *split* graphs. The above theorem states that any $(C_4, C_5, 2K_2)$-free graph is a split graph. Clearly none of graphs $C_4, C_5, 2K_2$ is split, hence any split graphs is $(C_4, C_5, 2K_2)$-free. The two propositions proved an induced subgraph characterization of split graphs. Note that this characterization contains a finite number of graphs.

4. $2K_2$-free graphs. This class is of interest because the number of maximal independent sets is bounded by a polynomial in this class. More generally, we will show that

Theorem 27 *If G is a $(p+1)K_2$-free graph with m edges, then*

$$T(G) \leq \sum_{i=0}^{p} \binom{m}{i}.$$

To prove the theorem, we will need a number of intermediate results.

A vertex of a graph is called *dominating* if it is adjacent to any other vertex of the graph. If a and b are vertices of a graph G, then $G_{a,b}$ will denote the subgraph obtained from G by deleting a, b and all the vertices adjacent to a or to b.

Lemma 28 . *If a is not a dominating vertex in a graph G, then*

$$T(G) \leq T(G - a) + \sum_{b \in N(a)} T(G_{a,b}).$$

Proof. Let H_1 be the set of those maximal independent sets (MIS) in G that include any MIS of the graph $G - a$, and let H_2 be the set of all the other MIS in G. If $X \in H_1$, then either X does not contain a and hence it is a MIS in the graph $G - a$ or X contains a and then $X - a$ is a MIS in $G - a$. Conversely, with each MIS Y in the graph $G - a$ one can associate some $X \in H_1$ as follows: if a is adjacent to at least one vertex in Y, then set $X := Y$, otherwise set $X := Y \cup a$. Thus, there is a bijection between H_1 and the family of all MIS in $G - a$. Therefore $T(G) = T(G - a) + |H_2|$.

Clearly, each independent set in the family H_2 contains vertex a. Let $X \cup \{a\} \in H_2$. Notice that X is not empty because a is not dominating. Since X is not a MIS in the graph $G - a$ (by definition of H_2), there is a vertex $b \in N(a)$ such that X is a MIS in the subgraph $G_{a,b}$. Consequently, $|H_2| \leq \sum_{b \in N(a)} T(G_{a,b})$. ■

Lemma 29 *For any vertex a of a graph G, the equality*

$$M(G) = M(G - a) + \sum_{b \in N(a)} (M(G_{a,b} + 1))$$

holds.

Proof. The proposition is a consequence of the fact that each induced matching either is included in $G - a$ or contains an edge incident with the vertex a. ■

Lemma 30 *For arbitrary graph G,*

$$T(G) \leq M(G) + 1.$$

Proof. For the complete graph with n vertices, we have $T(K_n) = n$, $M(K_n) = n(n-1)/2$ and the inequality is true for all n. If G is not a complete graph, then it has a non-dominating vertex a. The statement of the theorem is proved by induction on the number of vertices with the help of Lemmas 28 and 29:

$$T(G) \leq T(G - a) + \sum_{b \in N(a)} T(G_{a,b}) \leq M(G - a) + 1 + \sum_{b \in N(a)} (M(G_{a,b}) + 1) = M(G) + 1.$$

■

5. *Claw-free graphs* — $K_{1,3}$-free graphs.
6. *Line graphs.* Let $G = (V, E)$ be a graph. A graph H is called the line graph of if $V(H) = E$ and two vertices are adjacent in H if they have a vertex in common.

   Exercise. Show that $K_{1,3}$ is not a line graph.
7. *Cographs* — P_4-free graphs.

Theorem 31 *A graph is a cograph if and only if any of its induced subgraphs with at least two vertices is ether disconnected or the complement to a disconnected graph.*

Proof. Since neither P_4 nor its complement is disconnected, the "if" part is trivial.

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-neighbor x of a. On the other hand, a must have a neighbor 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 connected them. Without loss of generality we mat assume that x and y are adjacent. Let z be a neighbor of a in another component of G'. Then z, a, y, x is an induced P_4 in G. This contradiction shows that either there is a connected component in G' which does not have any neighbor 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. ■

8. *Perfect graphs.* A graph G is said to be perfect if for every induced subgraph H of G, the chromatic number of H equals the size of a maximum clique in H. Strong Berge Theorem states that a graph is perfect if and only if it is $(C_5, \overline{C_5}, C_7, \overline{C_7}, \ldots)$-free.

9. *Graphs of bounded clique-width.*

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

   – Creation of a new vertex v with label i (denoted $i(v)$).
   – Disjoint union of two labeled graphs G and H (denoted $G \oplus H$).
   – Joining by an edge each vertex with label i to each vertex with label j (denoted $\eta_{i,j}$).
   – Renaming label i to j (denoted $\rho_{i \to j}$).

   In general, clique-width can be arbitrarily large. Moreover, it is unbounded in many restricted graph families. On the other hand, in some specific classes of graphs the clique-width is bounded by a constant. Consider, for instance, a chordless path P_5 on five consecutive vertices a, b, c, d, e. By means of the four operations described above this graph can be constructed as follows:

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

   This expression uses only three different labels. Therefore, the clique-width of P_5 is at most 3.

10. *Threshold graphs.* Threshold graphs can be characterized in terms of forbidden induced subgraphs as $(2K_2, C_4, P_4)$-free graphs. This implies that threshold graphs are split graphs. Therefore, the vertices of a threshold graph can be partitioned into a clique C and an independent set I. Moreover, for any two vertices $a, b \in I$, either $N(a) \subseteq N(b)$ or $N(b) \subseteq N(a)$, since otherwise an induced P_4 arises. Therefore, the vertices of I can ordered under inclusion of their neighborhoods, i.e., a precedes b implies $N(a) \subseteq N(b)$. Similarly, the vertices of C can be ordered under inclusion of their neighborhoods.

    Exercise: Show that every threshold graph has either a dominating vertex or an isolated vertex.

    Exercise: Show that the clique-width of a threshold graph is at most 2.

5 Matchings and Hall's Theorem

Definition 32 *In a graph, a matching is a set of edges no two of which have a vertex in common.*

We say that a matching M covers a vertex v if v is an endpoint of an edge from M. A matching M is maximum if it has the largest possible cardinality. A matching M is maximal if no other matching contains M.

If G is a bipartite graph with partite sets X and Y, we say that X *can be matched into* Y if there is a matching covering all the vertices of X.

Theorem 33 *Let G be a bipartite graph with partite sets X and Y. Then X can be matched into Y if and only if $|N(S)| \geq |S|$ for each $S \subseteq X$.*

Proof. First, suppose that X can be matched into Y and let S be a subset of X. Since S itself is also matched into Y, we see immediately that $|S| \leq |N(S)|$.

Conversely, assume $|S| \leq |N(S)|$ for each subset $S \subseteq X$ and let M be a maximum matching in G. Denote by X_M and Y_M the vertices of X and Y that are covered by M. We claim that $X_M = X$. By contradiction, suppose $X_0 = X - X_M$ is the set of vertices of X that are not covered by M and let Y_1 be the neighborhood of X_0. If Y_1 contains a vertex which is not covered by M, then obviously M is not a maximum matching. Therefore, $Y_1 \subseteq Y_M$ and we may define the set X_1 of vertices of X_M that are matched with the vertices of Y_1. Since $|X_0 \cup X_1| > |X_0|$, the neighborhood of $X_0 \cup X_1$ must be strictly greater than Y_1. Denoting $i = 1$, we proceed according to the following procedure: as long as $N(X_i) \subseteq Y_M$, we increment i by one, define $Y_i := N(X_{i-1}) - (Y_1 \cup \ldots \cup Y_{i-1})$ and X_i to be the set of vertices of X_M that are matched with the vertices of Y_i. The procedure stops when we reach a set X_k that has a neighbor in Y which is not covered by M. Notice that at each step of the procedure, $|X_0 \cup \ldots \cup X_i| > |X_0 \cup \ldots \cup X_{i-1}|$, and hence the sets $X_1, X_2, \ldots X_k \subseteq X_M$ and $Y_1, Y_2, \ldots Y_k \subseteq Y_M$ are non-empty. We now define a sequence of vertices $y_{k+1}, x_k, y_k, x_{k-1}, y_{k-1}, \ldots, y_1, x_0$ according to the following procedure:

- let y_{k+1} be a neighbor of X_k which is not covered by M, x_k be a vertex of X_k adjacent to y_{k+1} and $i = k$,
- as long as $i > 0$, define y_i to be the vertex of Y_i which is matched with x_i, decrease i by one, and let x_i be a vertex of X_i which is adjacent to y_{i+1}.

In the sequence $y_{k+1}, x_k, y_k, x_{k-1}, y_{k-1}, \ldots, y_1, x_0$ neither x_0 nor y_{k+1} are covered by M and for each $i = 1, \ldots, k$ the edge $x_i y_i$ belongs to M. But then by replacing in M the edges $x_i y_i$ by the edges $x_i y_{i+1}$ $(i = 0, \ldots, k)$ we obtain a matching which is strictly greater than M. This contradiction shows that $X_M = X$. ∎

5.1 System of distinct representatives

Definition 34 *Let $A = \{A_1, A_2, \ldots, A_n\}$ be a finite collection of finite sets. A list of distinct elements $x_1, x_2, \ldots, x_n$ is called a* SYSTEM OF DISTINCT REPRESENTATIVES *(SDR) for sets in A if $x_i \in A_i$ for each $i = 1, 2, \ldots, n$.*

Example: Let $A_1 = \{a, b\}$, $A_2 = \{b\}$, $A_3 = \{c, d\}$, $A_4 = \{a, c, b\}$. Then (a, b, d, c) is an SDR for the family $A = \{A_1, A_2, A_3, A_4\}$. However, if we replace A_3 by $A_3' = \{a, c\}$, then it is not difficult to verify that the collection $A = \{A_1, A_2, A_3', A_4\}$ does not have an SDR. This example raises the following question: under what conditions a finite collection of finite sets has an SDR? An answer to this question is given by the following theorem.

Theorem 35 *Let $A = \{A_1, A_2, \ldots, A_n\}$ be a finite collection of finite non-empty sets. Then A has an SDR if and only if for each set of indices $I \subseteq \{1, 2, \ldots, n\}$,*

$$|I| \leq |\bigcup_{i \in I} A_i|.$$

Proof. Let $X = \bigcup_{i=1}^{n} A_i$ be the union of all subsets of A and $x_1, x_2, \ldots, x_p$ be the elements of X. Define G to be the bipartite graph with partite sets A and X in which $A_i x_j$ is an edge if and only if $x_j \in A_i$. In other words, the neighborhood of A_i is the set of elements belonging to A_i, and for any $I \subseteq \{1, 2, \ldots, n\}$, the neighborhood of the set $\{A_i \mid i \in I\}$ is $\bigcup_{i \in I} A_i$.

Obviously, A has an SDR if and only if A can be matched into X. By Theorem 33, A can be matched into X if and only if for each $I \subseteq \{1, 2, \ldots, n\}$, we have $|I| \leq |\bigcup_{i \in I} A_i|$. ∎

6 Ramsey Theory

Definition 36 *For positive integers p, q, the Ramsey number $R(p, q)$ is the smallest integer n such that every graph on n vertices contains either a clique of size p or an independent set of size q.*

Exercise: show that $R(2, k) = k$.

Theorem 37 $R(3, 3) = 6$

Proof. Since C_5 is a self-complementary graph and C_5 contains no triangle (a 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 neighbors a, b, c in G. If at least two of the neighbors 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 neighbors in the complement of G, the arguments are similar. ∎

Theorem 38 $R(3, 4) = 9$

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 non-neighbors induce 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 neighbors 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-neighbors (the vertices non-adjacent to v). Then the subgraph induced by these non-neighbors 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 neighbors and less than 6 non-neighbors, 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 3.

∎

Theorem 39 *If $p, q > 1$, then $R(p, q) \leq R(p - 1, q) + R(p, q - 1)$.*

Proof. 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 neighbors of v, and by G_2 the subgraph induced by the non-neighbors 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. ∎

6.1 Known Ramsey Numbers

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]

References

[1] Edward E. Bender and S. Gill Williamson, Foundations of Combinatorics with Applications, *Dover Publications* 2006 (available online at http://www.math.ucsd.edu/~ebender/CombText/)

[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 $R(6,6)$, we should attempt to destroy the aliens" – Paul Erdős