## 6.1 Theory and examples

You have already seen quite a few strategies and ideas, and you might say: "Enough with these tricks! When will we go to serious facts?" We will try to convince you that the following results are more than simple tools or tricks. They help to create a good base, which is absolutely indispensable for someone who enjoys mathematics, and moreover, they are the first steps to some really beautiful and difficult theorems or problems. And you must admit that the last problems discussed in the previous units are quite serious facts. It is worth mentioning that these strategies are not a panacea. This assertion is proved by the fact that every year problems that are based on well-known tricks prove to be very difficult in contests.

We will "disappoint" you again in this unit by focusing on a very familiar theme: graphs without complete subgraphs. Why do we say familiar? Because there are hundreds of problems proposed in different mathematics competitions around the world and in professional journals that deal with this subject. And each such problem seems to add something. Before passing to the first problem, we will assume that the basic knowledge about graphs is known and we will denote by $d(V)$ and $C(V)$ the number, and the set of vertices adjacent to $V$, respectively. Also, we will say that a graph has a complete $k-$subgraph if there are $k$ vertices any two of which are connected. For simplicity, we will say that $G$ is $k$-free if it does not contain a complete $k-$subgraph. First we will discuss one famous classical result about $k$-free graphs, namely Turan's theorem. Before that, though, we prove a useful lemma, also known as Zarankiewicz's lemma, which is the main step in the proof of Turan's theorem.

**Example 1.** If $G$ is a $k$-free graph, then there exists a vertex having degree at most $\left\lfloor \frac{k-2}{k-1}n \right\rfloor$.

[Zarankiewicz]

**Solution.** Suppose not and take an arbitrary vertex $V_1$. Then

$$|C(V_1)| > \left\lfloor \frac{k-2}{k-1} n \right\rfloor,$$

so there exists $V_2 \in C(V_1)$. Moreover,

$$\begin{aligned} |C(V_1) \cap C(V_2)| &= d(V_1) + d(V_2) - |C(V_1) \cup C(V_2))| \\ &\geq 2\left(1 + \left\lfloor \frac{k-2}{k-1} n \right\rfloor\right) - n > 0. \end{aligned}$$

Pick a vertex $V_3 \in C(V_1) \cap C(V_2)$. A similar argument shows that

$$|C(V_1) \cap C(V_2) \cap C(V_3)| \geq 3\left(1 + \left\lfloor \frac{k-2}{k-1} n \right\rfloor\right) - 2n.$$

Repeating this argument, we find

$$V_4 \in C(V_1) \cap C(V_2) \cap C(V_3)$$

$$V_{k-1} \in \bigcap_{i=1}^{k-2} C(V_i).$$

Also,

$$\left|\bigcap_{i=1}^{j} C(V_i)\right| \geq j\left(1 + \left\lfloor \frac{k-2}{k-1} n \right\rfloor\right) - (j-1)n.$$

This can be proved easily by induction. Thus

$$\left|\bigcap_{i=1}^{k-1} C(V_i)\right| \geq (k-1)\left(1 + \left\lfloor \frac{k-2}{k-1} n \right\rfloor\right) - (k-2)n > 0,$$

and, consequently, we can choose

$$V_k \in \bigcap_{i=1}^{k-1} C(V_i).$$

But it is clear that $V_1, V_2, \ldots, V_k$ form a complete $k$ graph, which contradicts the assumption that $G$ is $k$-free.

We are now ready to prove Turan's theorem.

**Example 2.** The greatest number of edges of a $k$-free graph with $n$ vertices is

$$\frac{k-2}{k-1}\cdot\frac{n^2-r^2}{2}+\binom{r}{2},$$

where $r$ is the remainder left by $n$ when divided to $k-1$.

[Turan]

**Solution.** We will use induction on $n$. The first case is trivial, so let us assume the result true for all $k$-free graphs having $n-1$ vertices. Let $G$ be a $k$-free graph with $n$ vertices. Using Zarankiewicz's lemma, we can find a vertex $V$ such that

$$d(V)\leq\left\lfloor\frac{k-2}{k-1}n\right\rfloor.$$

Because the subgraph determined by the other $n-1$ vertices is clearly $k$-free, using the inductive hypothesis we find that $G$ has at most

$$\left\lfloor\frac{k-2}{k-1}n\right\rfloor+\frac{k-2}{k-1}\cdot\frac{(n-1)^2-r_1^2}{2}+\binom{r_1}{2}$$

edges, where $r_1=n-1 \pmod{k-1}$.
Let $n=q(k-1)+r=q_1(k-1)+r_1+1$. Then $r_1\in\{r-1,r+k-2\}$ (this is because $r-r_1\equiv 1 \pmod{k-1}$) and it is easy to check that

$$\left\lfloor\frac{k-2}{k-1}n\right\rfloor+\frac{k-2}{k-1}\cdot\frac{(n-1)^2-r_1^2}{2}+\binom{r_1}{2}=\frac{k-2}{k-1}\cdot\frac{n^2-r^2}{2}+\binom{r}{2}$$

The inductive step is proved. Now, it remains to construct a $k$-free graph with $n$ vertices and $\frac{k-2}{2}\cdot\frac{n^2-r^2}{k-1}+\binom{r}{2}$ edges. This is not difficult. Just consider $k-1$ classes of vertices, $r$ of them having $q+1$ elements and the rest $q$ elements, where $q(k-1)+r=n$ and join the vertices situated in different groups. It is immediate that this graph is $k$-free, has $\frac{k-2}{k-1}\cdot\frac{n^2-r^2}{2}+\binom{r}{2}$

edges and also the minimal degree of the vertices is $\left\lfloor \frac{k-2}{k-1} n \right\rfloor$. This graph is called Turan's graph and is denoted by $T(n,k)$.
These two theorems generate numerous beautiful and difficult problems. For example, using these results yields a straightforward solution for the following Bulgarian problem.

**Example 3.** There are 2001 towns in a country, each of which is connected with at least 1600 towns by a direct bus line. Find the largest $n$ for which it must be possible to find $n$ towns, any two of which are connected by a direct bus line.

Spring Mathematics Tournament 2001

**Solution.** Practically, the problem asks to find the greatest $n$ such that any graph $G$ with 2001 vertices and minimum degree at least 1600 is not $n$-free. But Zarankiewicz's lemma implies that if $G$ is $n$-free, then at least one vertex has degree at most $\left\lfloor \frac{n-2}{n-1} 2001 \right\rfloor$. So, we need the greatest $n$ for which $\left\lfloor \frac{n-2}{n-1} 2001 \right\rfloor < 1600$. It is immediate to see that $n = 5$. Thus for $n = 5$ any such graph $G$ is not $n$-free. It suffices to construct a graph with all degrees of the vertices at least 1600, which is 6-free. We will take of course $T(2001, 6)$, whose minimal degree is $\left\lfloor \frac{4}{5} 2001 \right\rfloor = 1600$ and which is (as shown before) 6-free. Thus, the answer is $n = 5$.

Here is a beautiful application of Turan's theorem in combinatorial geometry.

**Example 4.** Consider 21 points on a circle. Show that at least 100 pairs of points subtend an angle less than or equal to $120°$ at the center.

Tournament of the Towns 1986

**Solution.** In such problems, it is more important to choose the right graph than to apply the theorem, because as soon as the graph is appropriately chosen, the solution is more or less straightforward. Here we will consider the graph with vertices at the given points and we will connect two points if they subtend an angle less than or equal to 120° at the center. Therefore we need to prove that this graph has at least 100 edges. It seems that this is a reversed form of Turan's theorem, which maximizes the number of edges in a $k$-free graph. Yet, the reversed form of the reversed form is the natural one. Applying this principle, let us look at the "reversed" graph, the complementary one. We must show that it has at most $\binom{21}{2} - 100 = 110$ edges. But this is immediate, since it is clear that this new graph does not have triangles and so, by Turan's theorem, it has at most $\frac{21^2 - 1}{4} = 110$ edges, and the problem is solved.

At first glance, the following problem seems to have no connection with the previous examples, but, as we will immediately see, it is a simple consequence of Zarankiewicz's lemma. It is an adaptation of an USAMO 1978 problem. Anyway, this is trickier than the actual contest problem.

**Example 5.** There are $n$ delegates at a conference, each of them knowing at most $k$ languages. Among any three delegates, at least two speak a common language. Find the least number $n$ such that for any distribution of the languages satisfying the above properties, it is possible to find a language spoken by at least three delegates.

**Solution.** We will prove that $n = 2k+3$. First, we prove that if there are $2k+3$ delegates, then the conclusion of the problem holds. The condition "among any three of them there are at least two who can speak the same language" suggests taking the 3-free graph with vertices the persons and whose edges join persons that do not speak a common language. From Zarankiewicz's lemma, there exists a vertex whose degree is at most $\left\lfloor \frac{n}{2} \right\rfloor = k+1$. Thus, it is not connected with at least $k+1$ other vertices. Hence there exists a person $A$ and $k+1$ persons $A_1, A_2, \dots, A_{k+1}$ that can communicate with $A$. Because $A$ speaks at most $k$ languages, there are two persons among $A_1, A_2, \dots, A_{k+1}$ that speak with $A$ in the same language. But that language is spoken by at least three delegates and we are done. It remains to prove now that we can create a situation in which there are $2k+2$ delegates, but no language is spoken by more than two delegates. We use again Turan's graph, by creating two groups of $k+1$ delegates. Assign to each pair of persons in the first group a common language, so that the language associated is different for any two pairs in that group. Do the same for the second group, taking care that no language associated with a pair in the second group is identical to a language associated with a pair in the first group. Persons in different groups do not communicate. Then it is clear that among three persons, two will be in the same group and therefore will have a common language. Of course, any language is spoken by at most two delegates.

The following problem turned out to be an upset at one of the Romanian Team Selection Tests for 2004 IMO, being solved by only four contestants. The idea is even easier than in the previous problems, but this time we need a little observation that is not so obvious.

**Example 6.** Let $A_1, A_2, \dots, A_{101}$ be different subsets of the set $\{1, 2, \dots, n\}$. Suppose that the union of any 50 subsets has more than $\frac{50}{51}n$ elements. Prove that among them there are three any two of which having common elements.

[Gabriel Dospinescu] Romanian TST 2004

__Solution.__ As the conclusion suggests, we should take a graph with vertices the subsets, connecting two subsets if they have common elements. Let us assume that this graph is 3-free. The main idea is not to use Zarankiewicz's lemma, but to find many vertices with small degrees. In fact, we will prove that there are at least 51 vertices all of them having degree at most 50. Suppose this is not the case, so there are at least 51 vertices whose degrees are greater than

51. Let us pick such a vertex $A$. It is connected with at least 51 vertices, so it must be adjacent to a vertex $B$ whose degree is at least 51. Because $A$ and $B$ are each connected with at least 51 vertices, there is a vertex adjacent to both, so we have a triangle, contradicting our assumption. Therefore, we can find $A_{i_1}, \ldots, A_{i_{51}}$, all of them having degrees at most 50. Consequently, $A_{i_1}$ is disjoint from at least 50 subsets. Because the union of these fifty subsets has more than $\frac{50}{51}n$ elements, we infer that $|A_{i_1}| < n - \frac{50}{51}n = \frac{n}{51}$. In a similar way, we obtain $|A_{i_j}| \leq \frac{n}{51}$ for all $j \in \{1, 2, \ldots, 51\}$ and so

$$|A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_{50}}| \leq |A_{i_1}| + \cdots + |A_{i_{50}}| < \frac{50}{51}n,$$

which contradicts the hypothesis.

We continue with an adaptation of a very nice and quite challenging problem from the American Mathematical Monthly.

**Example 7.** Prove that the complement of any 3-free graph with $n$ vertices and $m$ edges has at least

$$\frac{n(n-1)(n-5)}{24} + \frac{2}{n}\left(m - \frac{n^2 - n}{4}\right)^2$$

triangles.

[A.W Goodman] AMM

**Solution.** Believe it or not, the number of triangles from the complementary graph can be expressed in terms of the degrees of the vertices of the graph only. More precisely, if $G$ is a 3−free graph, then the number of triangles from the complementary graph is

$$\binom{n}{3} - \frac{1}{2}\sum_{x \in X} d(x)(n - 1 - d(x)),$$

where $X$ is the set of vertices of $G$. Indeed, consider all triples $(x, y, z)$ of vertices of $G$. We will count the triples that do not form a triangle in the complementary graph $\overline{G}$. Indeed, consider the sum $\sum_{x \in X} d(x)(n-1-d(x))$. It counts twice every triple $(x, y, z)$ in which $x$ and $y$ are connected, while $z$ is not adjacent to any of $x$ and $y$: once for $x$ and once for $y$. But it also counts twice every triple $(x, y, z)$ in which $y$ is connected with both $x$ and $z$: once for $x$ and once for $z$. Therefore, $\frac{1}{2} \sum_{x \in X} d(x)(n-1-d(x))$ is exactly the number of triples $(x, y, z)$ that do not form a triangle in the complementary graph. (Here we have used the fact that $G$ is 3-free.) Now, it is enough to prove that

$$\binom{n}{3} - \frac{1}{2} \sum_{x \in X} d(x)(n-1-d(x)) \geq \frac{n(n-1)(n-5)}{24} + \frac{2}{n}\left(m - \frac{n^2-n}{4}\right)^2.$$

Because $\sum_{x \in X} d(x) = 2m$, after a few computations the inequality reduces to

$$\sum_{x \in X} d^2(x) \geq \frac{4m^2}{n} \tag{6.1}$$

But this is the Cauchy-Schwarz inequality combined with $\sum_{x \in X} d(x) = 2m$. Finally, two chestnuts. The following problem is not directly related to our topic at first glance, but it gives a very beautiful proof of Turan's theorem:

**Example 8.** Let $G$ be a simple graph. To every vertex of $G$ one assigns a nonnegative real number such that the sum of the numbers assigned to all vertices is 1. For any two vertices connected by an edge, compute the product of the numbers associated to these vertices. What is the maximal value of the sum of these products?

Solution. The answer is not obvious at all, so let us start by making a few remarks. If the graph is complete of order $n$ then the problem reduces to finding the maximum of $\sum_{1\le i<j\le n} x_ix_j$ knowing that $x_1+x_2+\cdots+x_n=1$. This is easy, since

$$\sum_{1\le i<j\le n} x_ix_j=\frac{1}{2}(1-\sum_{i=1}^{n}x_i^2)\le\frac{1}{2}\left(1-\frac{1}{n}\right).$$

The last inequality is just the Cauchy-Schwarz inequality and we have equality when all variables are $\frac{1}{n}$ . Unfortunately, the problem is much more difficult in other cases, but at least we have an idea of a possible answer: indeed, it is easy now to find a lower bound for the maximum: if $H$ is the complete subgraph with maximal number of vertices $k$, then by assigning these vertices $\frac{1}{k}$, and to all other vertices 0, we find that the desired maximum is at least $\frac{1}{2}(1-\frac{1}{k})$. We still have to solve the difficult part: showing that the desired maximum is at most $\frac{1}{2}(1-\frac{1}{k})$. Let us proceed by induction on the number $n$ of vertices of $G$. If $n=1$ everything is clear, so assume the result true for all graphs with at most $n-1$ vertices and take a graph $G$ with $n$ vertices, numbered $1,2,...,n$. Let $A$ be a set of vectors with nonnegative coordinates and whose components add up to 1 and $E$ the set of edges of $G$. Because the function $f(x_1,x_2,...,x_n)=\sum_{(i,j)\in E} x_ix_j$ is continuous on the compact set $A$, it attains its maximum in a point $(x_1,x_2,...,x_n)$. Denote by $f(G)$ the maximum value of this function on $A$. If at least one of the $x_i$ is zero, then $f(G)=f(G_1)$ where $G_1$ is the graph obtained by erasing vertex $i$ and all edges that are incident to this vertex. It suffices to apply the induction hypothesis to $G_1$ (clearly, the maximal complete subgraph of $G_1$ has at most as many vertices as the maximal complete subgraph of $G$ ). So, suppose that all $x_i$ are positive. We may assume that $G$ is not complete, since this case has already been discussed. So, let us assume for example that vertices 1 and 2 are not connected. Choose any number $0<a\le x_1$ and assign to vertices $1,2,...,n$ of $G$ the numbers $x_1-a, x_2+a, x_3,...,x_n$. By maximality of $f(G)$, we must have

$$\sum_{i\in C_1} x_i\le\sum_{i\in C_2} x_i,$$

where $C_1$ is the set of vertices that are adjacent to vertex 2 and not adjacent to vertex 1 (the definition of $C_2$ being clear). By symmetry, we deduce that we must actually have

$$\sum_{i \in C_1} x_i = \sum_{i \in C_2} x_i,$$

which shows that $f(x_1, x_2, ..., x_n) = f(0, x_1 + x_2, x_3, ..., x_n)$. Hence we can apply the previous case and the problem is solved. Observe that the inequality in Turan's theorem follows by taking all $x_i$ to be $\frac{1}{n}$.

The final problem is a very beautiful result on the number of complete subgraphs of a graph:

**Example 9.** What is the maximal number of complete maximal subgraphs that a graph on $n$ vertices can have?

[Leo Moser, J. W. Moon]

**Solution.** Let us suppose that $n \geq 5$, the other cases being easy to check. Let $f(n)$ be the desired number and $G$ a graph for which this maximum is attained. Clearly, this graph is not complete, so there are two vertices $x$ and $y$ not connected by an edge. In order to simplify the solution, we need several notations. Let $V(x)$ be the set of vertices that are adjacent to $x$, $G(x)$ the subgraph obtained by erasing vertex $x$ and $G(x, y)$ the graph obtained by erasing all edges incident to $x$ and replacing them with edges from $x$ to any vertex in $V(y)$ . Finally, let $a(x)$ be the number of complete subgraphs with vertices in $V(x)$, maximal with respect to $G(x)$ and let $c(x)$ be the number of complete maximal subgraphs of $G$ that contain $x$.

Now, we pass to serious things: by erasing edges incident to $x$, exactly $c(x) - a(x)$ complete maximal subgraphs vanish, and by joining $x$ with all vertices of $V(y)$, exactly $c(y)$ complete maximal subgraphs appear. So, if $c(G)$ is the number of complete maximal subgraphs in the graph $G$, then we have the relation

$$c(G(x, y)) = c(G) + c(y) - c(x) + a(x).$$

By symmetry, we can assume that $c(y) \geq c(x)$. By maximality of $c(G)$, we must have $c(G(x,y)) \leq c(G)$, which is the same as $c(y) = c(x)$ and $a(x) = 0$. Therefore $G(x,y)$ also has $f(n)$ complete maximal subgraphs. In the same way, we deduce that $c(G(x,y)) = c(G(y,x)) = c(G)$. Now take a vertex $x$ and let $x_1, x_2, ..., x_k$ be the vertices not adjacent to $x$. By performing the previous operations, we change $G$ into $G_1 = G(x_1, x)$, then into $G_2 = G_1(x_2, x)$ and so on until $G_k = G_{k-1}(x_k, x)$, by conserving the number $f(n)$ of maximal complete subgraphs. Observe now that $G_k$ has the property that $x, x_1, ..., x_k$ are not joined by edges, yet $V(x_1) = V(x_2) = \cdots = V(x_k) = V(x)$. Now, we know what to do: if $V(x)$ is void, we stop the process. Otherwise, consider a vertex of $V(x)$ and apply the previous transformation. In the end, we obtain a complete multipartite graph $G'$ whose vertices can be partitioned into $r$ classes with $n_1, n_2, ..., n_r$ vertices, two vertices being connected by an edge if and only if they do not belong to the same class. Because $G'$ has $f(n)$ maximal complete subgraphs, we deduce that

$$f(n) = \max_r \max_{n_1+n_2+\cdots+n_r=n} n_1 n_2 ... n_r. \tag{6.2}$$

(6.2) can be easily computed. Indeed, let $(n_1, n_2, ..., n_r)$ the $r$-tuple for which the maximum is attained. If one of these numbers is at least equal to 4, let us say $n_1$, we consider $(2, n_1 - 2, n_3, ..., n_r)$ for which the product of the components is at least the desired maximum. So none of the $n_i$ exceed 3. Even more, since $2 \cdot 2 \cdot 2 < 3 \cdot 3$, there are at most two numbers equal to 2 among $n_1, n_2, ..., n_r$. This shows that $f(n) = 3^{\frac{n}{3}}$ if $n$ is a multiple of 3, $f(n) = 4 \cdot 3^{\frac{n-4}{3}}$ if $n-1$ is a multiple of 3 and $f(n) = 2 \cdot 3^{\frac{n-2}{3}}$ otherwise.