NDMI012: Combinatorics and Graph Theory 2

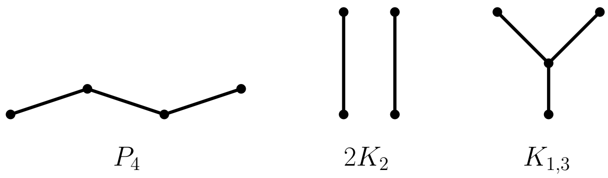
Lecture #13
Extremal combinatorics

Irena Penev

1 Turán's theorem

Given a positive integer n and a graph H, an n-vertex graph G without an H subgraph is *extremal* (for the property of not containing H as a subgraph) if it has the largest possible number of edges among all n-vertex graphs without an H subgraph; $\mathrm{ex}(n, H)$ is the number of edges of an extremal n-vertex graph without an H subgraph. In other words, $\mathrm{ex}(n, H)$ is the maximum number of edges that an n-vertex graph that does not contain H as a subgraph can have.

Obviously, any extremal graph G without an H subgraph is "edge-maximal" without an H subgraph, i.e. any graph obtained from G by adding one or more edges to it, contains H as a subgraph. The converse, however, does not hold in general: it is possible that a graph is edge-maximal without an H subgraph, without being extremal. For example, $2K_2$ is a four-vertex edge-maximal graph without a P_4 subgraph,[1] but it is not extremal: indeed, $K_{1,3}$ also has four vertices and no P_4 subgraph, and it has more edges than $2K_2$.



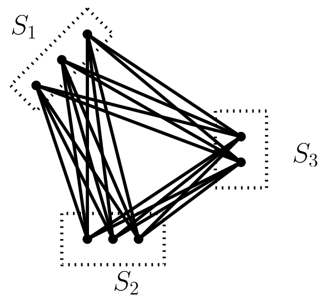
The following was proven in Combinatorics & Graphs Theory 1.

Mantel's theorem. *For any positive integer n, we have that $ex(n, K_3) = \lfloor \frac{n^2}{4} \rfloor$, and moreover, $K_{\lfloor \frac{n}{2} \rfloor, \lceil \frac{n}{2} \rceil}$ is an extremal n-vertex graph without a K_3 subgraph.*

[1] As usual, P_4 is the four-vertex (and three-edge) path.

Mantel's theorem is a special case of Turán's theorem, to which we now turn.

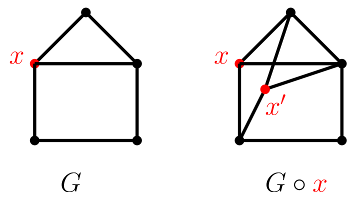
For a positive integer r, a *complete r-partite graph* is a graph G whose vertex set can be partitioned into r (possibly empty) stable sets (called *parts*), pairwise complete to each other. For example, the graph below is complete 3-partite, with parts S_1, S_2, S_3.



A *complete multipartite graph* is any graph that is complete r-partite for some r.

The *r-partite Turán graph on n vertices*, denoted by $T_r(n)$, is the complete r-partite graph on n vertices, in which the sizes of any two parts differ by at most one (so, each part is of size $\lfloor \frac{n}{r} \rfloor$ or $\lceil \frac{n}{r} \rceil$); $t_r(n)$ is the number of edges of $T_r(n)$. We note that the complete 3-partite graph above is in fact the graph $T_3(8)$.

Recall that *duplicating a vertex x* of a graph G produces a supergraph $G \circ x$ by adding to G a vertex x' and making it adjacent to all the neighbors of x in G, and to no other vertices of G (in particular, x and x' are nonadjacent in $G \circ x$). An example is shown below.



Obviously, $\omega(G \circ x) = \omega(G)$, i.e. G contains K_{r+1} is a subgraph if and only if $G \circ x$ does.

Turán's theorem. *Let n and r be positive integers. Then $ex(n, K_{r+1}) = t_r(n)$, and furthermore, $T_r(n)$ is the unique (up to isomorphism) extremal n-vertex graph without a K_{r+1} subgraph.*

Proof. We may assume that $r < n$, for otherwise, $T_r(n) \cong K_n$, and the result is immediate.

It is clear that $T_r(n)$ is an n-vertex graph without a K_{r+1} subgraph. Now, let G be any n-vertex extremal graph without a K_{r+1} subgraph. We must show that $G \cong T_r(n)$.

Claim. G is a complete multipartite graph.

Proof of the Claim. Suppose otherwise. Then there exist pairwise distinct vertices $y_1, x, y_2 \in V(G)$ such that $y_1x, xy_2 \notin E(G)$, but $y_1y_2 \in E(G)$.[2] If $d_G(y_1) > d_G(x)$, then $G_1 := (G \setminus x) \circ y_1$ is an n-vertex graph that does not contain K_{r+1} as a subgraph, and $|E(G_1)| > |E(G)|$, contrary to the fact that G is extremal.

So, $d_G(y_1) \leq d_G(x)$, and similarly, $d_G(y_2) \leq d_G(x)$. Now, let G' be the graph obtained from $G \setminus \{y_1, y_2\}$ by duplicating x twice. Then G' is an n-vertex graph with no K_{r+1} subgraph, and (since $y_1y_2 \in E(G)$) we have that $|E(G')| = |E(G)| - (d_G(y_1) + d_G(y_2)) + 1 + 2d_G(x) \geq |E(G)| + 1$, contrary to the fact that G is extremal. This proves the Claim. ♦

Now, using the Claim, we fix a partition $(S_1, \dots, S_k)$ of $V(G)$ into non-empty stable sets, pairwise complete to each other. Clearly, G contains K_k is a subgraph,[3] and so $k \leq r$. Suppose that $k < r$. Then since $r < n$, at least one of the sets $S_1, \dots, S_k$ has more than one vertex; by symmetry, we may assume that $|S_k| \geq 2$. Fix $a \in S_k$. Then consider the graph G' obtained from G by adding edges between a and all vertices of $S_k \setminus \{a\}$; then G' is a complete $(k+1)$-partite graph, it does not contain K_{r+1} as a subgraph (because $k < r$), and it has more edges than G, contrary to the fact that G is extremal. So, $k = r$.

It remains to show that any two of $S_1, \dots, S_r$ differ in size by at most one (this will imply that $G \cong T_r(n)$). Suppose otherwise. By symmetry, we may assume that $|S_1| \geq |S_2| + 2$. Now, fix a vetex $a \in S_1$, and let G' be the graph obtained by first deleting all edges between a and S_2, and then adding all edges between a and $S_1 \setminus \{a\}$. (This effectively "moves" a into S_2.) Now G is still a complete r-partite graph on n vertices, and it does not contain K_{r+1} as a subgraph. Furthermore, since $|S_1| \geq |S_2| + 2$, we see that $|E(G')| \geq |E(G)| + 1$. But this contradicts the fact that G is extremal. □

2 The Erdős-Ko-Rado theorem

Suppose we are given positive integers r and n, and we want to select a maximum number of pairwise intersecting r-element subsets of $\{1, \dots, n\}$. What is this maximum number? For $r > \frac{n}{2}$, any two r-element subsets of $\{1, \dots, n\}$ intersect, and there are $\binom{n}{r}$ many such subsets. How about if $r \leq \frac{n}{2}$? In that case, we can fix any $x \in \{1, \dots, n\}$, and consider all r-element

[2]Let us justify this. Since G is not complete multipartite, $\overline{G}$ is not the disjoint union of complete graphs. Consequently, some component C of $\overline{G}$ is not a complete graph. Since C is not complete, we see that C contains some two distinct, non-adjacent vertices, call them a and b. Since C is connected, there is an induced path $p_1, \dots, p_t$ in C, with $p_1 = a$ and $p_t = b$; since a and b are non-adjacent in C, we see that $t \geq 3$. We now set $y_1 := p_1$, $x := p_2$, and $y_2 := p_3$, and we observe that $y_1x, y_2x \notin E(G)$ and $y_1y_2 \in E(G)$.

[3]Indeed, we just take one vertex from each S_i, and we obtain a clique of size k.

subsets of $\{1,\dots,n\}$ that contain x; there are $\binom{n-1}{r-1}$ many such subsets, and obviously, they pairwise intersect.[4] As the following theorem shows, this is in fact best possible.

The Erdős-Ko-Rado theorem. *Let r and n be positive integers such that $r \leq \frac{n}{2}$. Then there are at most $\binom{n-1}{r-1}$ many pairwise distinct and pairwise intersecting r-element subsets of $\{1,\dots,n\}$.*

Proof. Let $A_1,\dots,A_m$ be pairwise distinct and pairwise intersecting r-element subsets of $\{1,\dots,n\}$. We must show that $m \leq \binom{n-1}{r-1}$.

Let c be the number of ordered pairs (C,A), where

- C is a directed cycle with vertex set $\{1,\dots,n\}$;[5]
- A is an r-vertex directed subpath of C;
- $V(A) = A_i$ for some $i \in \{1,\dots,m\}$.

Now we count in two ways, as follows. On the one hand, we can form an ordered pair (C,A) by first selecting one of the sets $A_1,\dots,A_m$ (we have m choices), then ordering its vertices to form a directed path (there are $r!$ choices), and then ordering the remaining $n-r$ vertices to complete the cycle C (there are $(n-r)!$ choices). So,

$$c \quad = \quad mr!(n-r)!.$$

We now count in another way. First, there are $(n-1)!$ ways of ordering $\{1,\dots,n\}$ to obtain a directed cycle C. Next, we claim that for fixed C, there are at most r directed subpaths of C that correspond to one of A_i's. Indeed, suppose the subpath $a_1,a_2,\dots,a_r$ of C corresponds to one of $A_1,\dots,A_m$. Then for any other subpath of C corresponding to one of $A_1,\dots,A_m$ (and therefore containing at least one of $a_1,\dots,a_r$), there exists some $i \in \{1,\dots,r-1\}$ such that either a_i is the terminal vertex of the path, or a_{i+1} is the initial vertex of the path; but since $r \leq \frac{n}{2}$, the r-vertex subpath terminating at a_i and the r-vertex subpath starting at a_{i+1} have no vertices in common, and so at most one of them can correspond to one of $A_1,\dots,A_m$. Thus, in addition to $a_1,\dots,a_r$, there are at most $r-1$ subpaths of C corresponding to one of $A_1,\dots,A_m$; in total, at most r subpaths of C correspond to one of $A_1,\dots,A_m$. This proves that

$$c \quad \leq \quad (n-1)!r.$$

[4]For the case when $r = \frac{n}{2}$, here is another construction. Fix any $x \in \{1,\dots,n\}$, and consider all r-element subsets of $\{1,\dots,n\} \setminus \{x\}$. Since $r = \frac{n}{2}$, all these subsets pairwise intersect, and there are $\binom{n-1}{r} = \binom{n-1}{(n-1)-r} = \binom{n-1}{(n-1)-\frac{n}{2}} = \binom{n-1}{\frac{n}{2}-1} = \binom{n-1}{r-1}$ many of them.

[5]Vertics $1,\dots,n$ need **not** appear in that order on the cycle.

We now have that

$$mr!(n-r)! \;=\; c \;\leq (n-1)!r,$$

and so

$$m \;\leq\; \tfrac{(n-1)!r}{r!(n-r)!} \;=\; \tbinom{n-1}{r-1},$$

which is what we needed to show. □

3 The Sunflower lemma

A *sunflower* is a family (i.e. collection) $\mathscr{S}$ of sets (called *petals*) such that there exists a set S (called a *kernel*) with the property that for all distinct $S_1, S_2 \in \mathscr{S}$, we have that $S_1 \cap S_2 = S$.[6]

The Sunflower lemma (Erdős-Rado)**.** *Let ℓ and p be positive integers, and let $\mathscr{A}$ be a family of sets such that*

- $|A| \leq \ell$ *for all* $A \in \mathscr{A}$, *and*
- $|\mathscr{A}| > (p-1)^\ell \ell!$.

Then there exists a sunflower $\mathscr{S} \subseteq \mathscr{A}$ with p petals.

Proof. We assume inductively that the lemma is true for smaller values of ℓ. More precisely, we assume that for all positive integers $\ell' < \ell$, and all families $\mathscr{A}'$ of sets such that

- $|A| \leq \ell'$ for all $A \in \mathscr{A}'$, and
- $|\mathscr{A}'| > (p-1)^{\ell'} \ell'!$,

there exists a sunflower $\mathscr{S}' \subseteq \mathscr{A}'$ with p petals.

Note that $|\mathscr{A}| \geq p$; so, if $p \leq 2$, then any p elements of $\mathscr{A}$ form a sunflower with p petals, and we are done. So, we may assume that $p \geq 3$. Next, suppose that $\ell = 1$. Then $|A| \leq 1$ for all $A \in \mathscr{A}$ and $|\mathscr{A}| > p - 1$. We then take any p elements of $\mathscr{A}$, and we observe that they form a sunflower (with an empty kernel). So, from now on, we assume that $\ell \geq 2$.

Let $\mathscr{D} \subseteq \mathscr{A}$ be a collection of pairwise disjoint sets, with $|\mathscr{D}|$ chosen maximum. If $|\mathscr{D}| \geq p$, then any p elements of $\mathscr{D}$ form a sunflower (with an empty kernel), and we are done. So assume that $|\mathscr{D}| < p$. Let $D = \bigcup \mathscr{D}$; then $|D| \leq |\mathscr{D}|\ell \leq (p-1)\ell$. Furthermore, since $|\mathscr{A}| \geq \ell \geq 2$, the maximality of $\mathscr{D}$ guarantees that $\mathscr{D}$ contains at least one non-empty set,[7] and so $D \neq \emptyset$.

[6] It is possible that $S = \emptyset$.

[7] Let us justify this. Suppose that $\mathscr{D}$ contains no non-empty sets. Then either $\mathscr{D} = \emptyset$ or $\mathscr{D} = \{\emptyset\}$. Since $|\mathscr{A}| \geq 2$, we see that $\mathscr{A}$ contains at least one non-empty set, say A. But then $\mathscr{D} \cup \{A\}$ is a set of pairwise disjoint elements of $\mathscr{A}$, contrary to the maximality of $\mathscr{D}$.

Claim. There exists some $d \in D$ such that d belongs to more than $(p-1)^{\ell-1}(\ell-1)!$ many elements of $\mathscr{A}$.

Proof of the Claim. We consider two cases: when $\emptyset \in \mathscr{A}$, and when this is not the case.

Suppose first that $\emptyset \notin \mathscr{A}$ (and consequently, $\emptyset \notin \mathscr{D}$). Then every element of $\mathscr{A}$ intersects D: indeed, since $\emptyset \notin \mathscr{D}$, we know that every element of $\mathscr{D}$ intersects D, and by the maximality of $\mathscr{D}$, every element of $\mathscr{A} \setminus \mathscr{D}$ intersects D. But then by the Pigeonhole Principle, some element of D belongs to at least

$$\left\lceil \frac{|\mathscr{A}|}{|D|} \right\rceil > \frac{(p-1)^{\ell}\ell!}{(p-1)\ell} = (p-1)^{\ell-1}(\ell-1)!$$

many elements of $\mathscr{A}$, which is what we needed.

Suppose now that $\emptyset \in \mathscr{A}$. Then the maximality of $\mathscr{D}$ guarantees that $\emptyset \in D$. Since $D \neq \emptyset$, it follows that $|\mathscr{D}| \geq 2$. Since $\emptyset \in \mathscr{D}$, we see that $|D| \leq (|\mathscr{D}|-1)\ell \leq (p-2)\ell$. Now by the maximality of $\mathscr{D}$, every element of $\mathscr{A} \setminus \{\emptyset\}$ intersects D. But then by the Pigeonhole Principle, some element of D belongs to at least

$$\begin{aligned} \left\lceil \frac{|\mathscr{A} \setminus \{\emptyset\}|}{|D|} \right\rceil &\geq \frac{(p-1)^{\ell}\ell!}{(p-2)\ell} \\ &= (p-1)^{\ell-1}(\ell-1)!\frac{p-1}{p-2} \\ &> (p-1)^{\ell-1}(\ell-1)! \end{aligned}$$

many elements of $\mathscr{A}$, which is what we needed. This proves the Claim. ⧫

Let $d \in D$ be as in the Claim, and set $\mathscr{A}' := \{A \setminus \{d\} \mid A \in \mathscr{A}, d \in A\}$. Then $|\mathscr{A}'| > (p-1)^{\ell-1}(\ell-1)!$; furthermore, $|A| \leq \ell - 1$ for all $A \in \mathscr{A}'$. So, by the induction hypothesis, there exists a sunflower $\mathscr{S}' \subseteq \mathscr{A}'$ with p petals. Now, set $\mathscr{S} := \{A \cup \{d\} \mid A \in \mathscr{S}'\}$; then $\mathscr{S} \subseteq \mathscr{A}$ is a sunflower with p petals, and we are done. □