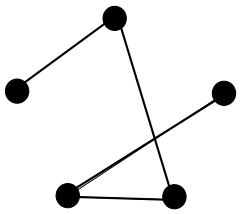
EXTREMAL COMBINATORICS

1. Introduction

Combinatorics is a subfield of mathematics which studies discrete structures. Today I will give a brief introduction of extremal combinatorics which studies the relation of the size of a discrete structure to various other properties.

A graph $G = (V, E)$ consists of a vertex set V and an edge set E, where V is an aribtrary finite set, and E is a subset of (non-ordered) pairs of elements of V. We say that two vertices $v, w \in V$ are *adjacent* if $\{v, w\} \in E$.

We may graphically represent a graph by representing each vertex as a dot, and connect two dots if and only if the pair consisting of the two vertices that the dots represent form an edge. The following figure is a representation of a graph with five vertices and four edges.



However, try to avoid restricting thinking of a graph only through its graphical representation, since the abstract definition is much more powerful. Remember that the strength of mathematics comes form abstraction.

A *complete graph* is a graph where every pair of vertices form an edge, and an *empty graph* is a graph where no pair of vertices form an edge.

2. Graph Properties

Graph theory studies relations between various graph properties. Extremal graph theory focuses on the cases when one of the properties is the "size" of the graph.

Following is a collection of some graph properties:

- $|V|$ or $v(G)$: Number of vertices
- $|E|$ or $e(G)$: Number of edges

A *clique* is a set of vertices which form a complete graph, and an *independent set* is a set of vertices which form an empty graph.

- $\kappa(G)$: Clique number, the number of vertices in the largest clique
- $\alpha(G)$: Independence number, the number of vertices in the largest independence set.

Example. Suppose that we are transmitting a message over some channel. A message consists of bits of zeros or ones. If a channel adds a lot of noise, then it might be better to send 00 instead of 0 and 11 instead of 1.

In general, a *code* $\mathcal{C}$ is a set of 0-1 vectors of length n. A code has minimum distance d if for every two vectors $v, w \in \mathcal{C}$, v and w differs in at least d of its coordinates. Note that in a code of minimum distance d, we can recover at most $\frac{d}{2}$ errors. To be efficient, it is best to have as much vectors in the code as possible.

Given n and d, what is the largest possible number of vectors one can have in a code of minimum distance d? Let V be the vertex set of all possible 0-1 vectors of length n. Two vertices are adjacent if and only if they have distance at most $d-1$. The problem is equivalent to finding the largest independent set of the graph.

A *proper coloring* of a graph is a coloring of its vertex set where for every edge in the graph, the colors on its two endpoints are different. Note that a proper coloring of a graph partitions the vertex set of the graph into independent sets.

- $\chi(G)$: Chromatic number, the minimum number of colors that are required for a proper coloring of G.

At first glance, chromatic number might look like an arbitrary concept to study. However, chromatic number is a very interesting concept which appears in many different unexpected context throughout mathematics.

Example. The following problem can be modeled using graphs and chromatic number.

Consider the courses offered in MIT on Spring 2013. Suppose that the schedule of each course is predetermined. We would like to assign lecture rooms for each course so that there are no conflict. Is this possible?

Let G be the following graph. The vertex set is the set of courses, where two vertices (courses) are adjacent if and only if they overlap in time. Now consider each lecture room as a color. Then it is possible to assign lecture rooms so that there are no conflict if and only if we can properly color G with the given set of colors. Thus such assignment exists if and only if the number of lecture rooms are larger than $\chi(G)$.

3. Turán Theorem : Number of edges versus subgraphs

For two graphs $G = (V, E)$ and $G' = (V', E')$, we say that G' is a *subgraph* of G and denote it as $G' \subset G$ if $V' \subset V$ and $E' \subset E$. In other words, G' is a subgraph of G if one can obtain G' from G by removing some vertices and edges.

For a vertex $v \in V$, the *degree* $d(v)$ of a vertex is the number of edges which have v as one of its endpoints.

Proposition 3.1. $\sum_{v \in V} d(v) = 2e(G)$.

Proof. The left hand side counts each edge exaclty twice, thus the equation follows. □

Theorem 3.2. *(Mantel, 1907) Suppose that we are given a graph G satisfying $e(G) > \frac{v(G)^2}{4}$. Then G contains triangle as a subgraph.*

Proof. Suppose that we are given a graph G satisfying $e(G) > \frac{v(G)^2}{4}$ with no triangles.

Note for a fixed edge $e = \{v, w\}$, there is no vertex x adjacent to both v and w, since otherwise G contains a triangle. Therefore,

$$(d(v) - 1) + (d(w) - 1) \leq v(G) - 2.$$

Sum the equation above for all edges over the graph and we have,

$$\sum_{e=\{v,w\}\in E} \big(d(v) + d(w)\big) \leq e(G) \cdot v(G).$$

In the left hand side of the equation above, the degree $d(v)$ of a vertex v is added exactly $d(v)$ times. Hence the left hand side equals to

$$\sum_{v\in V} d(v)^2,$$

which by Cauchy-Schwartz inequality is at least

$$\frac{1}{v(G)}\Big(\sum_{v\in V} d(v)\Big)^2 = \frac{4}{v(G)}e(G)^2,$$

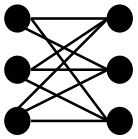
where we used the proposiiton above. Therefore,

$$\frac{4}{v(G)}e(G)^2 \leq e(G) \cdot v(G),$$

from which we conclude that $e(G) \leq \frac{v(G)^2}{4}$.

However, this is a contradiction. Hence G contains a triangle. □

The following graph, known as the *complete bipartite graph,* shows that the inequality given in the theorem above is tight.



More generally, let $K_{n,n}$ be the graph over a vertex set $V = V_1 \cup V_2$, where $V_1 = V_2$. Edge set consists of the pairs which have one vertex in V_1 and the other in V_2. This graph has $n^2 = \frac{v(G)^2}{4}$ edges and contains no triangle. (the figure above depicts $K_{3,3}$).

Here is a generalization to larger cliques.

Theorem 3.3. *(Turán, 1941) If a graph G has more than $(1 - \frac{1}{r})\frac{v(G)^2}{2}$ edges, then it contains a complete graph on $r + 1$ vertices as a subgraph.*

For more general graphs, we have the following theorem.

Theorem 3.4. *(Erdős, Stone, Simonovits) For any fixed graph H and any fixed $\varepsilon > 0$, there exists an integer n_0 such that for every graph with at least n_0 vertices and at least*

$$(1 - \frac{1}{\chi(H) - 1} + \varepsilon)\frac{v(G)^2}{2}$$

edges contains H as a subgraph.

Note the sudden appearance of the chromatic number in this theorem!

It is easy to see that cycles of even length have chromatic number 2 and cycles of odd length have chromatic number 3. Thus the theorem above asserts that one only needs tiny amount of edges to force a cycle of even length, but needs at least about $\frac{v(G)^2}{4}$ edges to force a cycle of odd length.

4. Hall's theorem

A *bipartite graph* is a graph of chromatic number 2. Equivalently, it is a graph over a vertex set $V = V_1 \cup V_2$, where all the edges have one vertex in V_1 and the other in V_2.

For bipartite graphs with parts of equal size, i.e., $|V_1| = |V_2|$, a *perfect matching* is a collection of edges such that every vertex is contained in exactly one of them.

For a vertex subset X of a graph G, let $N(X)$ be the set of *neighbors* of X, i.e., the set of vertices not in X which are adjacent to some vertex in X.

Theorem 4.1. *(Hall, 1935) Let G be a bipartite graph with parts V_1 and V_2 of equal size. G contains a perfect matching if and only if $|N(X)| \geq |X|$ for all $X \subset V_1$.*

Proof. It is clear that if G contains a perfect matching, then $|N(X)| \geq |X|$ for all $X \subset V_1$. We now prove the converse.

Let $|V_1| = |V_2| = n$. We prove the statement by induction on n. The statement is clearly true for $n = 1$. Suppose that we are proving it for n, while it is true for all smaller values.

If a stonger inequality $|N(X)| \geq |X|+1$ holds for every non-empty proper subset X of V_1, then pick an arbitrary edge edge $\{v_1, v_2\}$ for $v_1 \in V_1$ and $v_2 \in V_2$. Let $V_1' = V_1 \setminus \{v_1\}$ and $V_2' = V_2 \setminus \{v_2\}$. Note that for all $X \subset V_1'$, we have

$$|N(X) \setminus \{v_2\}| \geq |N(X)| - 1 \geq |X|.$$

Hence the graph induced on the vertex set $V_1' \cup V_2'$ satisfies Hall's condition for $n - 1$. Therefore it contains a perfect matching. This perfect matching together with $\{v_1, v_2\}$ forms a perfect matching of the original graph.

Otherwise, assume that there exists a proper subset $X \subset V_1$ such that $|N(X)| = |X|$. By Hall's condition, there exists a perfect matching between X and $N(X)$.

Moreover, we claim that the subgraph induced on $V_1 \setminus X$ and $V_2 \setminus N(X)$ also satisfies Hall's condition. Suppose not. Then there exists a subset $U \subset V_1 \setminus X$ such that $|N(U) \setminus N(X)| < |U|$. However, this implies that

$$|N(U \cup X)| \leq |N(X)| + |N(U) \setminus N(X)| < |X| + |U|,$$

and is a contradiction. Therefore, the claim is true. Hence there exists a perfect matching between $V_1 \setminus X$ and $V_2 \setminus N(X)$.

By combining the two matchings, we obtain a perfect matching of the original graph. □

Hall's theorem is also refered to as Hall's marriage theorem for the following reason. Suppose that there are n boys and n girls. We place a bipartite graph over the vertex set of boys and girls where a boy and a girl forms an edge if and only if they are interested in each other. Then Hall's theorem gives a condition on which we can match boys and girls so that each boy and girl forming a couple, the two are interested in each other.

5. Dirac's Theorem

A *Hamilton cycle* is a cycle which passes through each vertex of the graph exactly once.

Theorem 5.1. *(Dirac) If a vertex has minimum degree at least $\frac{v(G)}{2}$, then it contains a Hamilton cycle.*

Proof. Suppose that $v_1, v_2, \cdots, v_\ell$ forms a longest path P in a given graph. all the neighbors of v_ℓ and v_1 must lie in the set $\{v_1, \cdots, v_\ell\}$ as otherwise there exists a path longer than P in the given graph.

Since

$$d(v_1) + d(v_\ell) \geq v(G) \geq \ell,$$

there exists an index i such that v_1 is incident to v_i and $v_\ell$is incident to v_{i-1}. Using these edges, we can find a cycle $v_1, v_2, \cdots, v_{i-1}, v_\ell, v_{\ell-1}, \cdots, v_i, v_1$.

If $\ell = n$, then we just found a Hamilton cycle.

If $\ell \leq \frac{n}{2}$, then every vertex v_j on the cycle which has a neighbor in $V \setminus \{v_1, \cdots, v_\ell\}$. Otherwise, if $\ell > \frac{n}{2}$, then every vertex $w \in V \setminus \{v_1, \cdots, v_\ell\}$ has a neighbor on the cycle. In both cases, we can find a path longer than P in the given graph, thus obtain contradiction. Therefore, the graph contains a Hamilton cycle. □

6. Erdős-Szekeres theorem

Extremal combinatorics studies various discrete structures. We conclude this lecture with one theorem on permutations which can be understood through analyzing graphs.

Theorem 6.1. *Let $a_1, \cdots, a_{n^2+1}$ be a permutation of the integers $1, 2, \cdots, n^2+1$. Show that there exists a monotone (increasing or decreasing) subsequence of length $n+1$.*

Proof. Consider a graph over the vertex set $1, 2, \cdots, n^2+1$, where we place an edge between i and j for $i < j$ if and only if $a_i > a_j$.

For each i let $f(i)$ be the length of the longest increasing subsequence ending at a_i. If $f(i) \geq n+1$ for some i, then we are done. Otherwise, color the vertex i using color $f(i)$. Thus $1 \leq f(i) \leq n$ for all i, and this is a vertex coloring using at most n colors.

We claim that each color class forms an independent set. Suppose for the contrary that there exist indices i and j such that $f(i) = f(j)$ and $i < j$ for which $\{i, j\}$ forms an edge. Then by definition, we have $a_i < a_j$. However, then we can find an increasing subsequence of length $f(i)+1$ ending at a_j, contradicting $f(i) = f(j)$. This is a contradiction. Therefore, each color class forms an independent set, and we have a proper coloring.

Since $\chi(G) \cdot \alpha(G) \geq n^2 + 1$ (homework problem) and $\chi(G) \leq n$, we see that $\alpha(G) \geq n+1$. $\square$