1 Minimal Examples and Extremal Problems

Minimal and extremal problems are really variations on the same question: what is the largest or smallest graph you can find which either avoids or satisfies some graph property? "Size" in these contexts is usually "number of vertices" but in some contexts is "number of edges" on a graph with a predetermined number of vertices.

1.1 Minimal examples

If we are trying to find the smallest graph satisfying a certain property, it's generally called a *minimal example*. For most properties, the minimal examples are pretty dull: whether we are minimizing on number of vertices or number of dges among a fixed set of vertices, most of our graph paremeters have simple minimal examples. The smallest graph of diameter d is a path on $d+1$ vertices; the smallest graph of chromatic number k or clique number k is a connected graph on k vertices, and the smallest graph of independence number k is k non-adjacent vertices. These are self-evident and not terribly interesting.

The one parameter we have seen so far with an interesting and not too difficult minimal example is the parameter of connectivity. Finding the fewest number of vertices on which we may connect a graph of connectivity k is not too interesting: once again, a complete graph K_{k+1} is best, but if we fix a number of vertices $n \geq k+1$, we might ask how many edges a graph on n vertices must have in order to have connectivity k? It is easy to find a necessary number of edges.

Proposition 1. *If $|G| = n$ and $\kappa(G) = k$, then $\|G\| \geq \lceil \frac{nk}{2} \rceil$.*

Proof. It has been seen previously, when first introducing the concept of connectivity, that $\kappa(G) \geq \delta(G)$, so $\delta(G) \geq k$. Then

$$2\|G\| = \sum_{v \in V(G)} d(v) \geq \sum_{v \in V(G)} \delta(G) = |G|\delta(G) \geq nk$$

and thus $\|G\| \geq \frac{nk}{2}$, and since $\|G\|$ must be an integer, we may take the ceiling of this lower bound. □

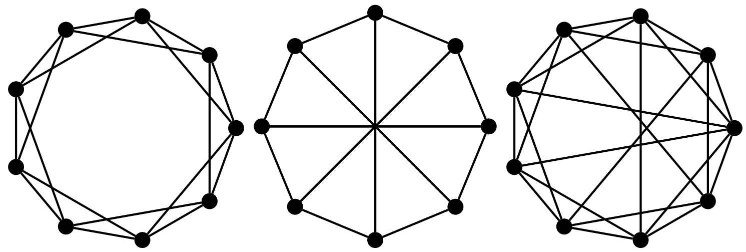
This almost-trivial bound will in fact completely satisfy our purposes, because we can show that specific graphs achieve it.

Definition 1. The *Harary graph* $H_{n,k}$ is a graph on the n vertices $\{v_1, v_2, \ldots, v_n\}$ defined by the following construction:

- If k is even, then each vertex v_i is adjacent to $v_{i\pm 1}$, $v_{i\pm 2}, \ldots$, $v_{i\pm\frac{k}{2}}$, where the indices are subjected to the wraparound convention that $v_i \equiv v_{i+n}$ (e.g. v_{n+3} represents v_3).
- If k is odd and n is even, then $H_{n,k}$ is $H_{n,k-1}$ with additional adjacencies between each v_i and $v_{i+\frac{n}{2}}$ for each i.
- If k and n are both odd, then $H_{n,k}$ is $H_{n,k-1}$ with additional adjacencies

$$\{v_1, v_{1+\frac{n-1}{2}}\}, \{v_1, v_{1+\frac{n+1}{2}}\}, \{v_2, v_{2+\frac{n+1}{2}}\}, \{v_3, v_{3+\frac{n+1}{2}}\}, \cdots, \{v_{\frac{n-1}{2}}, v_n\}$$

Below, the graphs $H_{9,4}$, $H_{8,3}$, and $H_{9,5}$ are shown as examples of these three classes.



Note that these graphs are (except when n and k are both odd) highly symmetrical, and that each contains $\lceil \frac{nk}{2} \rceil$ edges. We thus have an explicit construction of minimal k-connected graphs, following our proof of the Haray graphs' connectivity:

Theorem 1. *$H_{n,k}$ is k-connected.*

Proof. We shall start by proving this in the symmetric cases. If $k = 2r$ for some r, then let us consider some subset S of $V(H_{n,k})$, with $|S| < k$; wlog let us assume $v_1 \notin S$, and there is some other $v_i \notin S$. Let us consider the sets $\{v_2, \ldots, v_{i-1}\}$ and $\{v_{i+1}, \ldots, v_n\}$. S is drawn from these two sets, and $|S| < 2r$, so one or the other of these sets contains fewer than r elements of S; without loss of generality, let $\{v_2, \ldots, v_{i-1}\} \cap S$ have fewer than S elements. Now, either $i \leq k+1$, in which case $v_1 \sim v_i$, or there is some element v_{i_1} of $\{v_2, v_3, \ldots, v_{k+1}\}$ which is not in S, since $|S \cap \{v_2, v_3, \ldots, v_{k+1}\}| < k$. So $v_1 \sim v_{i_1}$ Now either $i \leq k + i_1$ in which case $v_{i_1} \sim v_i$, or there is some element v_{i_2} of $\{v_{i_1+1}, v_{i_1+2}, \ldots, v_{i_1+k}\}$ which is not in S, since $|S \cap \{v_2, v_3, \ldots, v_{k+1}\}| < k$, so $v_{i_1} \sim v_{i_2}$. Continuing this procedure as far as necessary, we will construct a path from v_1 to v_i. Thus, since v_i was an arbitrary vertex of $H_{n,k}$, and by symmetry v_1 is equivalent to any other choice of vertex, we have shown connectivity between arbitrary vertices of $H_{n,k} - S$, so $H_{n,k}$ is k-connected.

The other two cases will be left as exercises to the reader! □

1.2 Turán-type Extremal Problems

Extremal problems deal with one of two identical questions: how large must a graph be before certain substructures become inevitable, and how large can a graph be without achieving certain substructures. These are really the same question, posed as an existential negative (there exists a graph G without property P) or a universal positive (every graph G has property P). We might start with a simple example to illuminate this logical equivalency: how many edges can we put among n vertices and keep our chromatic number below 2?

We know the structure of graphs with $\chi(G) \leq 2$ very well: they are exactly the bipartite graphs, since every bipartite graph is 2-colorable and no non-bipartite graph is 2-colorable. If we were to divide the n vertices into parts A and B, then the maximum number of edges we could put among the parts would of course be $|A| \cdot |B|$ to form the complete bipartite graph $K_{|A|,|B|}$. If $|A| = k$ and $|B| = n - k$, then such a construction contains $k(n-k)$ edges; considered as a function of k, this has a unique maximum at $k = \frac{n}{2}$, so the optimal choice of $|A|$ is either $\lceil \frac{n}{2} \rceil$ or $\lfloor \frac{n}{2} \rfloor$ (which is a single value when n is even. Thus, we may with confidence say that it is possible to produce a graph of chromatic number less than or equal to 2 with m edges if and only if $m \leq \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$. Conversely, if $m > \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$, it is guaranteed that $\chi(G) \geq 3$. $\lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$ is thus the *extremal number* for the condition

$\chi(G) \leq 2$, since it represents the most edges that can be put into a graph while retaining that condition. A graph with an extremal number of edges is often called an *extremal graph*.

In fact, this argument can be extended to other chromatic numbers. A graph is 3-colorable if and only if its vertices can be partitioned into "color classes" A, B, and C such that no vertices within a single part are adjacent. This could be thought of as a "tripartite" graph, after a fashion, and the "complete tripartite" graph $K_{a,b,c}$ could be constructed by placing mutual adjacencies among all 3 parts, using $ab + ac + bc$ edges in total. Multivariate calculus (or a symmetry argument) will establish that, given $a + b + c = n$ this sum is largest when a, b, and c are as close to equal in size as possible, so the most edges that a 3-colorable graph on n vertices can have will be:

$$\begin{cases} \lfloor \frac{n}{3} \rfloor^2 + 2\lceil \frac{n}{3} \rceil \lfloor \frac{n}{3} \rfloor \text{ if } n \equiv 1 \bmod 3 \\ \lceil \frac{n}{3} \rceil^2 + 2\lceil \frac{n}{3} \rceil \lfloor \frac{n}{3} \rfloor \text{ if } n \equiv 2 \bmod 3 \end{cases}$$

(Note that when $n \equiv 0 \bmod 3$, these two are equal, and simplify to $\frac{1}{3}n^2$) In fact, this same line of argument applies to k-colorability in general, to the extent that we might want to describe this class of graphs:

Definition 2. The *Turán graph* $T_{n,k}$ is a complete k-partite graph with n vertices and parts $S_1, S_2, \ldots, S_k$ such that each $|S_i|$ is either $\lfloor \frac{n}{k} \rfloor$ or $\lceil \frac{n}{k} \rceil$.

Note that the Turán graph, despite the appearance of choice in size for each part, is really uniquely defined, since if $n \cong r \bmod k$, then exactly r of the parts will be of size $\lceil \frac{n}{k} \rceil$ and $k - r$ parts will be of size $\lfloor \frac{n}{k} \rfloor$.

Finding the number of edges in an arbitrary Turán graph is a combinatorial chore, because of the mismatched sizes and the potential for three different sizes of sets for edges to traverse between. Using r defined as above, it's tedious but not difficult to determine that

$$t_{n,k} = \|T_{n,k}\| = \binom{r}{2} \left\lceil \frac{n}{k} \right\rceil^2 + r(k-r) \left\lceil \frac{n}{k} \right\rceil \left\lfloor \frac{n}{k} \right\rfloor + \binom{k-r}{2} \left\lfloor \frac{n}{k} \right\rfloor^2$$

It is worth noting that in the special case where n is a multiple of k, this will be just $\binom{k}{2}\left(\frac{n}{k}\right)^2 = \frac{k-1}{2k}n^2$, and even when n is not a multiple of k, it is very close to $\frac{k-1}{2k}n^2$. This lends itself nicely to a concept of *extremal density*: since a complete graph contains $\binom{n}{2}$ edges, we may, after a fashion, drop n itself out of consideration by asking what fraction of the total possible edges an extremal graph would have; in this case:

$$\frac{t_{n,k}}{|K_n|} \approx \frac{k-1}{k}$$

Frequently we go one step further to file the edges off of this approximation and seek the tendency of the density for massive graphs:

Definition 3. The *density* of a graph G is $\frac{\|G\|}{\|K_{|G|}\|} = \frac{\|G\|}{\binom{n}{2}}$. For an infinite family of graphs $G_1, G_2, G_3, \ldots$ with $|G_i| = i$, the *asymptotic density* of the family is $\lim_{n\to\infty} \frac{\|G_n\|}{\binom{n}{2}}$.

For instance, the asymptotic density of P_n, C_n, or $H_{n,k}$ for fixed k is 0; the asymptotic density of $H_{n,\epsilon n} = \epsilon$, and the asymptotic density of $K_n = 1$. And, more to our point here, the asymptotic density of $T_{n,k}$ is $\frac{k-1}{k}$.

Thus far, however, we have only looked at extremal properties of chromatic nubmer. What is far more common in discussing extremal properties is forbidden subgraphs.

Definition 4. A graph G is *H-free* if H is not a subgraph of G. The *extremal number* $\mathrm{ex}(n, H)$ is the largest value of m such that there exists an H-free graph G with n vertices and m edges. Such a G is called an *extremal* or *maximal H-free graph.*

We can see from our previous investigation a fairly simple bound on the extremal number of a clique-free graph.

Proposition 2. $\mathrm{ex}(n, K_r) \geq t_{n,r-1}$.

Proof. The Turán graph $T_{n,r-1}$ is $(r-1)$-partite, so any selection of r vertices must have 2 from the same part which are thus non-adjacent, so these vertices do not form a K_r and thus $T_{n,r-1}$ is K_r-free, so the largest K_r-free graph on n vertice must have at least as many edges as $T_{n,r-1}$. □

In fact, we can do better than this:

Theorem 2 (Turán '41)**.** *$\mathrm{ex}(n, K_r) = t_{n,r-1}$, and, in fact, every extremal K_r-free graph on n vertices is isomorphic to $T_{n,r-1}$.*

Proof. We use induction on n. Trivially, if $n < r$, $T_{n,r-1}$ is a K_n, which is K_r-free and uniquely maximal. Let us consider G to be a maximal K_r-free graph on $n \geq r$ vertices. Since it is maximally K_r-free, adding an absent edge $\{u, v\}$ to it induces a K_r, so either u or v already lies in a K_{r-1}, which we shall call H. By the inductive hypothesis, $G - H$ has at most $t_{n-r+1,r-1}$ edges. In addition, since G is K_r-free, every vertex of $G - H$ must be nonadjacent to some vertex in H, so each element of $G - H$ has at most $r - 2$ neighbors in H. And thus:

$$\begin{aligned}
\|G\| &\leq \|G - H\| + (n - r + 1)(r - 2) + \|H\| \\
&\leq t_{n-r+1,r-1} + (n - r + 1)(r - 2) + \binom{r-1}{2} \\
&\leq t_{n,r-1}
\end{aligned}$$

The final line follows from the rather reliable number of edges produced by adding exactly $r - 1$ vertices to a balanced $(r-1)$-partite graph: namely, adding one vertex to each part, so that every single vertex already in the graph gets $r - 2$ new neighbors, and the $r - 1$ newly added vertices are mutually adjacent.

We know that $T_{n,r-1}$ achieves equality on the above bound, so since G is maximal and the above upper bound is attainable, it is attained by a maximal graph and $\|G\| = t_{n,r-1}$.

Furthermore, the equality (and maximality) above is only attainable if every single vertex in $G - H$ has $r - 2$ neighbors in H. Thus, in order for G to be maximal, it must (by induction hypothesis) consist of a $T_{n-r+1,r-1}$, which is $(r-1)$-partite, together with a K_{r-1} and edges from each vertex of the $T_{n-r+1,r-1}$ to $r-2$ elements of the K_{r-1}. This structure can easily be shown to be $(r-1)$-partite, since each vertex of the K_{r-1} can be assigned to a part whose elements it is not a neighbor of. Since G is $(r-1)$-partite, we already know edge-maximality to be achieved uniquely by $T_{n,r-1}$. □

We thus know the exact structure of maximal K_r-free graphs, and their asymptotic density:

$$\lim_{n \to \infty} \frac{\mathrm{ex}(n, K_r)}{\binom{n}{2}} = \frac{r-2}{r-1}$$

and this is necessarily an upper bound on the size of arbitrary H-free graphs.

Corollary 1. *If* $|H| = r$, *then* $\mathrm{ex}(n, H) \leq t_{n,r-1}$, *so* $\lim_{n\to\infty} \frac{\mathrm{ex}(n,H)}{\binom{n}{2}} \leq \frac{r-2}{r-1}$.

Proof. Since $H \subset K_r$, any H-free graph is K_r-free, so the largest H-free graphon n vertices is K_r-free; thus $\mathrm{ex}(n, H) \leq \mathrm{ex}(n, K_r)$. □

Note that for many graphs this is an unacceptably inaccurate bound on asymptotic density: for instance, the asymptotic density of T-free graphs for any tree T is in fact zero, regardless of the number of vertices.

The exact value of the asympotic density is in fact determined solely by the chromatic number: unfortunately, the heavy lifting for this result is a bit beyond the scope of this class.

1.3 Ramsey numbers: extremality on vertices

An interesting fact about graph parameters, somewhat alluded to by previous investigations, is that it's quite difficult to keep both $\omega(G)$ and $\alpha(G)$ small: adjacencies tend to drive up $\omega(G)$, while nonadjacencies increase $\alpha(G)$. In fact, we can see that it's actually impossible to keep from having either a moderately large clique or a moderately large independent set:

Proposition 3. *If* $|G| \geq 6$, *then either* $\alpha(G) \geq 3$ *or* $\omega(G) \geq 3$.

Proof. Let v be a vertex of G, chosen arbitrarily. If $d(v) \geq 3$, then let a, b, and c be neighbors of v. If any of $a \sim b$, $b \sim c$, or $c \sim a$ are true, then v and the two adjacent vertices form a K_3, so $\omega(G) \geq 3$. On the other hand, if $a \not\sim b$, $b \not\sim c$, and $c \not\sim a$, then $\{a, b, c\}$ is an independent set of size 3, so $\alpha(G) \geq 3$.

Conversely, if $d(v) < 3$, then v has at least 3 non-neighbors a, b, and c. if any two of them are non-adjacent, then they form an independent set with v; if all three of them are mutually adjacent, then they form a K_3.

Note as a counterexample for $|G| < 6$ that $\omega(C_5) = \alpha(C_5) = 2$. □

Speaking in terms of independent sets and cliques is a traditional and useful way to express this result, however, in terms of proof accessibility, this argument is usually phrased in terms of edge-coloring, where we might associate one color with a "present edge" and another color with an "absent edge". Note that this concept of coloring is distinct from that used in the previously presented edge- and vertex-coloring problems; here we place no restrictions on using the same color in adjacent structures. So, the same proof might be presented in the following form:

Proposition 4. *If every edge of a* K_6 *is colored either red or blue, there is a monochromatic (entirely red or entirely blue)* K_3.

Proof. Consider a vertex v; it has 5 incident edges to be colored red or blue. 3 edges must be the same color, and without loss of generality we can say this color is blue, and name these edges' terminal vertices a, b, and c. If $\{a, b\}$ is blue, then $\{v, a, b\}$ is a blue K_3; likewise for $\{a, c\}$ and $\{b, c\}$. However, if these three are all red instead, we have a red K_3. □

This little elegant proof leads to a definition for a more general sort of problem:

Definition 5. The *Ramsey number* $R(k, \ell)$ is the least value of n such that a K_n whose edges are colored red and blue must contain either a red K_k or a blue K_ℓ; alternatively, it is the least n such that if $|G| = n$, then either $\alpha(G) \geq k$ or $\omega(G) \geq \ell$.

So the above proof is essentially that $R(3, 3) = 6$. There are a couple of simple, obvious properties of Ramsey numbers:

- $R(k, \ell) = R(\ell, k)$.
- $R(k, \ell) \leq R(\ell + 1, k)$, and $R(k, \ell) \leq R(\ell, k + 1)$.
- $R(k, 2) = k$.