1 Ramsey Theory

1.1 Classical Ramsey numbers

Furthermore, there is a beautiful recurrence to give bounds on Ramsey numbers, but we will start with a simple but distinctly nontrivial example, to set the stage:

Proposition 1. *$R(3,4) \leq 10$.*

Proof. Consider a vertex v in a coloring of a K_{10}. v has 9 incident edges, in some variety of colors. It must either have at least 6 incident blue edges, or 4 incident red edges (since otherwise it would have fewer than 8 incident edges total).

If v is incident on 6 blue edges, consider the endpoints $\{v_1, \ldots, v_6\}$ of those edges. Among these 6 vertices, a K_6 is colored red and blue; by our previous proposition, this K_6 contains either a red K_3 or a blue K_3. A blue K_3 among vertices $\{v_i, v_j, v_k\}$, together with the blue edges to v from each of these, will form a blue K_4. Thus, if v has 6 blue edges incident on it, we are guaranteed either a blue K_4 or a red K_3.

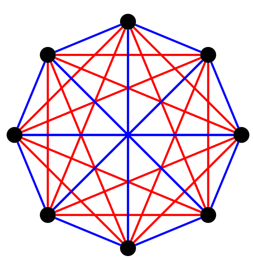
On the other hand, if v is incident on 4 red edges, then either there is a red edge between the endpoints of these edges, producing a red K_3, or the edges among all four of them are blue, producing a blue K_4. □

We can actually improve this result slightly by maknig use of a parity argument.

Corollary 1. *$R(3,4) \leq 9$.*

Proof. The above argument demonstrated that if any vertex is incident on 6 blue edges or 4 red edges, then the graph must contain a red K_3 or a blue K_4. In order to prevent this from happening in a coloring of K_9, in which each vertex is incident on 8 edges, every vertex must be incident specifically on 5 blue edges and 3 red edges. However, the total number of pairs (v, e) of vertices and incident blue edges would then be $9 \cdot 5 = 45$, but the number of such pairs must be even, since for each blue edge e, there are exactly two vertices which are its endpoints. Thus, a coloring with the incidence properties described above is impossible, and some vertex has either 6 incident blue edges or 4 incident blue edges; thus any coloring of K_9 contains a red K_3 or a blue K_4. □

This bound is in fact sharp, as can be seen by this example of a coloring of a K_8 without a red K_3 or blue K_4.



The above argument doesn't really use anything special about $R(3,4)$, and can in fact be generalized to give a nice upper bound on $R(k, \ell)$:

Theorem 1. *$R(k,\ell) \leq R(k-1,\ell) + R(k,\ell-1)$. Furthermore, if $R(k-1,\ell)$ and $R(k,\ell-1)$ are both even, this bound can be reduced by 1.*

Proof. Let $n = R(k-1,\ell) + R(k,\ell-1)$, for brevity. Consider a vertex v in a coloring of a K_n. v has $n-1$ incident edges, and must either have at least $R(k-1,\ell)$ incident red edges, or $R(k,\ell-1)$ incident blue edges (since otherwise it would have fewer than $R(k-1,\ell)+R(k,\ell-1)$ incident edges total).

If v is incident on $R(k-1,\ell)$ red edges, consider the induced subgraph H among the $R(k-1,\ell)$ endpoints of those edges. Since this subgraph is a complete graph on $R(k-1,\ell)$ vertices, it must contain either a red K_{k-1} or a blue K_ℓ; since the vertices of such a red K_{k-1} are all adjacent via red edges to v, a red K_{k-1} together with v forms a K_k, so that the graph as a whole must contain either a red K_k or a blue K_ℓ.

If, on the other hand, v is incident on $R(k,\ell-1)$ blue edges, a nearly identical argument can be made for the inevitability of either a red K_k or a blue $K_{\ell-1}$ among its neighbors, which implies either a red K_k or a blue K_ℓ in the larger graph. Thus, a K_n colored in red and blue must contain a red K_k or a blue K_ℓ.

Furthermore, if $R(k-1,\ell)$ and $R(k,\ell-1)$ are both even, it can be shown that any coloring of a K_{n-1} must also contain a red K_k or a blue K_ℓ. As seen above, if any vertex v has $R(k-1,\ell)$ incident red edges or $R(k,\ell-1)$ incident blue edges, then the graph must contain a red K_k or a blue K_ℓ. The only way to avoid such an occurrance in coloring the $n-2$ edges incident on a single vertex of K_{n-1} is to color exactly $R(k-1,\ell)-1$ of them red and $R(k,\ell-1)-1$ of them blue. This must be the case at every vertex of the K_{n-1}, resulting in $\frac{(R(k-1,\ell)-1)(n-1)}{2}$ red edges and $\frac{(R(k,\ell-1)-1)(n-1)}{2}$ blue edges. However, since $R(k-1,\ell)$, $R(k,\ell-1)$, and n are all even, these expressions are not integers, and thus describe impossible situations, so some vertex of K_{n-1} must b3e colored in such a way as to make the argument for K_n apply. □

This gives a moderately acceptable bound: using this approach and the known values $R(k,2) = R(2,k) = k$, it is not too hard to show that $R(k,\ell) \leq \binom{k+\ell-2}{\ell-1}$, even without using the parity improvement.

Unfortunately, this bound is exact in only a few cases. It yields the correct values $R(3,5) = 14$ and $R(4,4) = 18$, but is otherwise a little above the correct values: $R(3,6) = 18 < 19 = R(3,5) + R(2,6) - 1$.

Surprisingly few nontrivial Ramsey numbers are known exactly. We know $R(3,k)$ through $k = 9$, $R(4,4)$, and $R(4,5)$. The bulk of ongoing research into two-color Ramsey numbers is focused on $R(5,5)$, which is known to be between 43 and 49 inclusive.

One might wonder why our modern computational power isn't very effective, given how small these numbers are. Even though these nubmers are small, the sample space is enormous! The number of different colorings of a K_{43}, for instance, is $2^{\|K_{43}\|} = 2^{\binom{43}{2}} = 2^{903} > 10^{270}$; even with some clever sample-space-reducing tricks such as eliminating symmetries and rejecting obviously K_5-containing colorings, this is still a set of colorings far too large to be brute-forced to find an upper bound.

Lower bounds, by contrast, are easy to demonstrate: a single sample coloring of a complete graph will suffice. However, except when looking at a value of n significantly smaller than the correct Ramsey number, finding such an example is difficult.

In fact, a general rule or method for finding lower bounds on Ramsey numbers eluded researchers for a long time, until the following extraordinary result with an unusual proof:

Theorem 2 (Erdős, '47). *If* $\binom{n}{k}2^{1-\binom{k}{2}} < 1$*, then* $R(k,k) > n$*.*

Proof. Let the edges of K_n be colored independently and at random, i.e., for each edge, flip a fair coin, and color it red or blue depending on the result. For a subset S of $V(K_n)$ with $|S| = k$, let A_S represent the event that the elements of S are the vertices of a monochromatic K_k. Since each of the $\binom{k}{2}$ edges among vertices of S has exactly a $\frac{1}{2}$ probability of being red, then the probability that the edges among the elements of S are all red (forming a red clique) is $\left(\frac{1}{2}\right)^{\binom{k}{2}}$. Likewise, the probability that they are all blue is $\left(\frac{1}{2}\right)^{\binom{k}{2}}$. So the probability of the event A_S is $2 \cdot \left(\frac{1}{2}\right)^{\binom{k}{2}} = 2^{1-\binom{k}{2}}$. Now, the event that there is a monochromatic clique anywhere in the K_n occurs only if some A_S occurs: in other words, this event occurs only if at least one A_S occurs. The probability that at least one of a family of events occurs is no more than the sum of their individual probabilities (by inclusion-exclusion, essentially), and thus:

$$\begin{aligned} Pr(\text{monochromatic } k\text{-clique}) &\leq \sum_{|S|=k} Pr(A_S) \\ &\leq \sum_{|S|=k} 2^{1-\binom{k}{2}} \\ &\leq \binom{n}{k} 2^{1-\binom{k}{2}} < 1 \end{aligned}$$

so the probability that a random coloring contains a monochromatic k-clique is less than one, so the probability that a random coloring does not contain a monochromatic k-clique is greater than zero: that is to say, an actual possibility! So, some coloring of K_n contians no monochromatic k-clique. □

The above inequality condition for a lower bound on $R(k,k)$ is messy, so it's generally bounded by the following:

Corollary 2. $R(k,k) \geq \lfloor 2^{k/2} \rfloor$.

Proof. This follows from a moderately involved application of Stirling's formula to show that $\binom{2^{k/2}}{k} < 2^{\binom{k}{2}-1}$. □

The stunning thing about this result is that it was a shift in the way lower bounds were constructed: traditionally, you needed an example, and this was one of the earliest examples of a *non-constructive bound* developed: it asserted, with strong support, that a specific graph with specific properties existed, but didn't furnish the actual graph in question.

This also marked the first appearance of what came to be known as the *probabalisitic method* in combinatorics: using probability to assert that something can or must happen. This argument is not actually essentially probabalistic: multiply every quantity appearing in this argument by $\binom{n}{2}$, and replace "events" with "sets of graphs", and it's a counting argument instead. However, probability allows for some refinements on this argument: for instance, when looking at an asymmetric Ramsey number $R(k,\ell)$ with $k > \ell$, the optimal randomization technique might not be to use a fair coin with even chances of creating a red and blue edge, but to use an unfair coin which is more likely to add red edges. Using this modification of the random approach, fairly good lower bounds for the Ramsey numbers can be attained (the definitive text for randomization techniques as applied to combinatorics and graph theory is *The Probabalistic Method*, by Noga Alon and Joel Spencer, full of exotic techniques using probability).

1.2 Multicolor Ramsey problems

The question of finding monochromatic cliques after coloring a K_n edges with 2 colors is of traditional interest for two reasons: first, it's the simplest nontrivial edge-coloring problem, and second, it has a conenction, explored even before we introduced the coloring conceit, to clique numbers and independence numbers of arbitrary graphs. Adding a third color breaks the underlying paradigm of "existence-vs.-nonexistence", but it presents a theoretical richness of its own. Not much has been done with multicolor Ramsey problems, but the definition, and a simple example of the one known case, will serve to illustrate the concept:

Definition 1. The *Ramsey number* $R(k_1, k_2, k_3, \dots, k_r)$ is the least value of n such that a K_n whose edges are colored in r colors must contain, for some i, a K_{k_i} in color i.

Proposition 2. $R(3, 3, 3) = 17$.

Proof. For clarity and brevity, let us call the colors used red, blue, and green. Pick a vertex v of a tricolored K_{17}. It has 16 incident edges, so at least 6 of them will be the same color; without loss of generality, assume that color was green. Then, among these six vertices, we either have a green edge, which together with the green edges emenating to v form a green K_3, or all edges among these 6 vertices are red and blue, in which case, since $R(3, 3) = 6$, there must be a red or a blue K_3 among these vertices.

Thus, $R(3, 3, 3) \leq 17$. We could show that $R(3, 3, 3) > 16$ with an example of a coloring of K_{16} with 3 colors such that it contains no K_3. Doing so is not terribly difficult but the construction is moderately tedious.

There is, however, a most elegant construction to show that $R(3, 3, 3) > 8$: let each vertex v be represented by a unique string of 3 bits. If two vertices differ in their first bit, color the edge between them green; if two vertices have the same first bit and a different second bit, color their edge red; finally, if they differ only in the last bit, color their edge blue. It is easy to show that the edges in each color form a bipartite graph, so there are no monochromatic K_3s. □

Other Ramsey-type problems

Structures akin to the Ramsey problem abound. We shall not focus on these in any depth, but here is a list of interesting theorems and concepts which have given rise to related research:

Definition 2. The *generalized Ramsey number* $R(H_1, H_2, \dots, H_k)$ is the least n such that a coloring of the edges of K_n with r colors must contain an H_i in color i, for some i.

Theorem 3 (van der Waerden, '27)**.** *For any value of r and k, there is a value of N such that if the natural numbers $1, 2, \dots, N$ are colored in r colors, there is a monochromatic arithmetic sequence of length k.*

Theorem 4 (Hales and Jewett, '63). *For any value of n and r, there is a value of d such that if the cells of a d-dimensional lattice are colored with r colors, there will be a line of monochromatic cells of length n.*

Theorem 5 (Graham and Rothschild, '71). *There is a value of d such that, if the vertices of K_{2^d} are associated with the vertices of a d-dimensional hypercube and the edges are colored in two colors, there is a monochromatic K_4 whose vertices all lie in the same plane.*

The Graham-Rothschild result is particularly noteworthy in that it has the worst known bounds on a specific number. It has been varified experimentally that $d > 10$. The upper bound, however, is so massive that Ron Graham had to invent a new notation to even describe it (the new notation defines operations which are an extension of the logical progression of addition, multiplication, and exponentiation, and then defines g_i as the performance of the g_{i-1}th operation in the series on 3 and 3; the number required is g_{64}.

2 Hamiltonian cycles

We now return to a concept introduced early and not much revisited. The presence and identification of Euler tours in graphs turned out to be an exceedingly simple matter: we might reasonably hope that the vertex-based analogue, the Hamiltonian cycle, is as easily found and identified. Unfortunately, this seems not to be the case: there are many open problems relating to the identification of Hamiltonian graphs. Even though a full characterization of the class of Hamiltonian graphs does not exist, however, there are a number of conditions we can look at which are necessary or sufficient for a graph to be Hamiltonian.

One of our necessary conditions is vaguely reminiscent of Tutte's Theorem:

Proposition 3. *G is Hamiltonian only if for every nonempty $S \subset V$, $G - S$ has at most $|S|$ components.*

Proof. Consider a graph G and Hamiltonian cycle thereon; note that G must be connected, since it has a cycle through every vertex. Given a set S, the Hamiltonian cycle can only pass from one component of $G - S$ to another through an element of S (since if elements of C_1 and C_2 were adjacent to each other, they would not be distinct components in S). Since the Hamiltonian cycle passes through each component of $G - S$, it must pass from one component to another a number of times greater than or equal to the number of components; thus $|S|$ is less than or equal to the number of components. □

One might note a corollary to this, since it is, for the most part, a stronger condition than Tutte's Theorem:

Corollary 3. *If G has an even number of vertices and is Hamiltonian, then G has a perfect matching.*

Proof. Since $|G|$ is even and G is connected, Tutte's criterion is satisfied for $S = \emptyset$. For nonempty S, the above Hamiltonicity condition is stronger than the Tutte criterion, so Tutte's Theorem holds. □

Of course, there is a much simpler proof of this same result: take a Hamiltonian cycle of even length, and select every second edge to get a matching.

Theorem 6 (Dirac '52)**.** *If $|G| \geq 3$ and $\delta(G) \geq \frac{|G|}{2}$, then G is Hamiltonian.*

Proof. Let us consider G satisfying the above conditions. Note that G must be connected: if it had 2 or more components, some component would contain no more than $\frac{|G|}{2}$ vertices, and vertices in that component would have degree of less than $\frac{|G|}{2}$. Let us look at a longest path $P = v_0 \sim v_1 \sim \cdots \sim v_k$ in G. If $v_0 \sim v_k$, then we may form a cycle from P, to be inspected later.

Otherwise, since this is the longest path possible, no simple extension from v_0 or v_k is a path, which is to say, every neighbor of v_k and every neighbor of v_0 is in $P - \{v_0, v_k\}$. Thus we know that v_0 has at least $\frac{|G|}{2}$ neighbors in $P - \{v_0, v_k\}$, as does v_k. Now consider the *successors* of v_k's neighbors; that is to say, points v_i such that $v_{i-1} \sim v_k$. There are $\frac{|G|}{2}$ of these in $P - \{v_k\}$. Since $|P - \{v_k\}| \leq |G| - 1$, and we have two sets of size $\frac{|G|}{2}$ on these poitns, they must intersect somewhere: thus, there is an i such that $v_0 \sim v_i$ and $v_{i-1} \sim v_k$. Using this value, we can produce a cycle

$$v_0 \sim v_1 \sim \cdots \sim v_{i-1} \sim v_k \sim v_{k-1} \sim v_{k-2} \sim \cdots \sim v_i \sim v_0$$

Thus, in each case we have produced a cycle on the same vertex-set as P. If these vertices do not form a component, then there is some outside vertex u adjacent to some v_i; however, then we could find a longer path than path P by breaking the cycle at v_i and continuing to u. Thus, $\{v_0, \ldots, v_k\}$ form a component. But since G is connected, they must thus be all the vertices of G. □

This same basic argument can be used to prove a strong, peculiarly constructive statement about Hamiltonian graphs:

Proposition 4 (Ore '60)**.** *For a graph G with nonadjacent vertices u and v such that $d(u) + d(v) \geq |G|$, it follows that G is Hamiltonian if and only if $G + e$ is Hamiltonian, for $e = \{u, v\}$.*