Ramsey Theory

We begin with a simple question: How many people do we need to have at a party to ensure that we can find either 3 of them who all know each other, or 3 of them who are all strangers to each other? [We assume that knowing someone is symmetric – i.e. if A knows B then B knows A.]

We abstract the problem to one in graph theory in the following way: consider the people at the party as vertices of a graph, and draw an edge between any two of the people. We can now assign a colour to each edge – red if the two people know each other, and blue if they do not. A set of 3 people who all know each other then corresponds to a triangle all of whose edges are red, and a set of 3 people who are mutual strangers corresponds to a triangle all of whose edges are blue. Therefore, we are colouring each edge of a complete graph on n vertices either red or blue, and looking for a triangle all of whose edges are the same colour, and we are looking for the smallest value of n that we can choose to ensure that this happens, or asking whether there is necessarily any value of n that ensures that this happens.

To express this and similar questions more easily, we make the following definitions:

Definition 1. A *2-colouring* of the edges of a graph G is a function assigning to each edge of G one of two colours. More generally, an *n-colouring* is a function assigning to each edge of G, one of a fixed set of n colours.

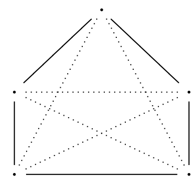
Definition 2. Given a colouring of the edges of a graph G, a subgraph G' is *monochromatic* if there is some colour such that every edge in G' is that colour.

solution:

The solution to this problem is that $n = 6$ is the smallest value of n such that however we 2-colour the edges of a K_n, we always get a monochromatic triangle. To show this, we need to show that however we 2-colour the edges of a K_6, we can always find a monochromatic triangle, and that there is a way of 2-colouring the edges of a K_5 so that we cannot find a monochromatic triangle.

For the first part, pick a vertex v in a K_6, and consider the edges incident with v. There are 5 of them, and they are divided into two sets – red edges and blue edges. Therefore, by the generalised pigeon-hole principle, there must either be 3 red edges or 3 blue edges. W.L.O.G., suppose there are 3 red edges, to vertices w_1, w_2, w_3. If any of the three edges between two of w_1, w_2 and w_3 is red, then it completes a red triangle. On the other hand, if none of those three edges is red, then they are all blue, and so w_1, w_2 and w_3 form a blue triangle. Therefore, in either case, there is either a red triangle or a blue triangle.

For the second part, we simply exhibit a 2-colouring of K_5 with no monochromatic triangles:



Having solved the problem, we look for ways to make it more general. The choice of 3 people all of whom know each other, or all of whom are mutual strangers was arbitrary. We could instead have asked how large n should be in order that whenever we 2-colour the edges of a K_n, we always get a monochromatic K_4, or in general, we could look for a monochromatic K_m for any choice of m.

Another way to generalise is to invest in a larger set of coloured pens – what if we 3-colour the edges of a K_n? What if we k-colour them for an arbitrary k.

Finally, we can generalise by noticing that it is not necessary to look for the same size of monochromatic set in each colour. We could for example ask how large n has to be to ensure either a red K_4 or a blue triangle. It turns out that we will need to study this generalisation in order to get results about the first generalisation.

First, we need some notation.

Definition 3. We define $R(k,m)$ to be the smallest n such that whenever we 2-colour a K_n red and blue, we always get either a red K_k or a blue K_m.

In fact this definition is not yet satisfactory, since we do not yet know that there is any value of n that we can choose so that we ensure a red K_k or a blue K_m. The first thing we should do is to prove that we can always find an n with this property.

Theorem 1 (Ramsey's Theorem)**.** *For any positive integers k, m, there is an n such that however we 2-colour (red and blue) the edges of a K_n, we can always find either a red K_k or a blue K_m. (So $R(k,m)$ is defined for every k, m. Furthermore, we have that $R(k,m) \leqslant 2^{k+m}$.*

Proof. We prove this by induction on k and m. The base case is easy – if $k = m = 1$ the result is trivially true, and indeed we have shown the result for $k = m = 3$. Now suppose that $R(k, m-1)$ and $R(k-1, m)$ are both defined. Let $n = R(k, m-1) + R(k-1, m)$. Then let v be a vertex of K_n. v has $n-1$ neighbours. This means that however we have 2-coloured the K_n, v has either $R(k-1, m)$ red neighbours or $R(k, m-1)$ blue neighbours. In the first case, we now look at the subgraph on the set of red neighbours of v. If this has a red K_{k-1}, then this K_{k-1}, with v added, forms a red K_k in the whole graph. On the other hand, if it has a blue K_m, then this is a K_m in the whole graph. Therefore, the whole graph has either a red K_k or a blue K_m. A similar argument applies if v has $R(k, m-1)$ blue neighbours.

This argument has shown that $R(k,m) \leqslant R(k-1,m) + R(k,m-1)$. From this it is an easy induction to show that $R(k,m) \leqslant 2^{k+m}$. $\square$

Remark 1. Clearly, we can improve the upper bound given in this theorem by looking more closely at the small values of $R(k, m)$, but the upper bound will still be an exponential, and indeed the best known upper bound is an exponential. On the other hand, we will shortly see some lower bounds for $R(k, m)$, and see that there is a very significant gap between the upper bounds and the lower bounds.

Theorem 2. *For any sequence $m_1, m_2, \ldots, m_k$ of positive integers, there is some n such that whenever we k-colour the edges of a K_n with colours $\{c_1, c_2, \ldots, c_k\}$, we get a K_{m_i} all of whose edges are of colour c_i for some $i \in \{1, \ldots, k\}$.*

Proof. We will prove this by induction on k. We proved the $k = 2$ case above. Now suppose we know the result for $k - 1$, and for 2. Imagine becoming colour-blind between colours c_1 and c_2. We now think that the graph is $k-1$-coloured, and we can choose n large enough that whenever we $k - 1$-colour the K_n, we get either a K_{m_i} all of whose edges are of colour c_i for some $i \in \{3, \ldots, k\}$, or a $K_{R(m_1,m_2)}$ all of whose edges are of the colour which is either c_1 or c_2. In the first case, we are done. In the second case, we have a $K_{R(m_1,m_2)}$, which in the original colouring was 2-coloured with c_1 and c_2. By definition of $R(m_1, m_2)$, this means that we can find either a K_{m_1} of colour c_1 or a K_{m_2} of colour c_2. $\square$

Lower Bounds for Ramsey Numbers

So far, we have produced upper bounds on how large these $R(m, n)$ are (and we can use the proof of the previous theorem to produce upper bounds for the larger numbers of colours. However, we do not yet know whether these upper bounds are good ones – i.e. is the actual value of $R(k, m)$ close to the upper bound we have? Might we be able to choose a much smaller value of n and still be guarenteed either a red K_k or a blue K_m?

In a sense, finding lower bounds should be easier – to show that $R(k, m) > N$, we only need to find one 2-colouring of K_N that does not have either a red K_k or a blue K_m. For small values of k and m, this is indeed the best way to find Ramsey numbers. However, for larger values of k, m it isn't so efficient, because the colourings we can describe are well organised – i.e. they have a lot of structure, and this often means that there will be large monochromatic subgraphs.

In fact, the best known approach to this problem is by taking a random colouring if K_N, and showing that the probability of getting either a red K_k or a blue K_m is less than 1. If the probability is less than one, then there must be at least one colouring in which there is no red K_k or blue K_m.

Of course, the probability of getting a red K_k or a blue K_m is very difficult to calculate. However, a quantity which is much easier to calculate is the expected number of red K_k and blue K_ms. This gives an upper bound for the probability, since the expected number of them is at least the probability that there is at least one. We exemplify this technique in the case where $k = m$, but similar methods could be applied to get lower bounds on more general Ramsey numbers.

Theorem 3. $R(m,m) > 2^{\frac{m-1}{2}}$.

Proof. 2-colour the edges of a K_n red with probability $\frac{1}{2}$, and blue with probability $\frac{1}{2}$ (colour each edge either red or blue). Given m vertices of the graph, the probability that the K_m they form has all its edges red is $\left(\frac{1}{2}\right)^{\binom{m}{2}}$. Therefore, the expected number of red K_m is $\frac{\binom{n}{m}}{2^{\binom{m}{2}}}$. Similarly, the expected number of blue K_m is $\frac{\binom{n}{m}}{2^{\binom{m}{2}}}$, so the expected number of monochromatic K_m is $\frac{2\binom{n}{m}}{2^{\binom{m}{2}}}$. We therefore have that there is some colouring with no monochromatic K_m as long as $\frac{2\binom{n}{m}}{2^{\binom{m}{2}}} \leqslant 1$. (it is easy to show that if equality holds, then there must be a colouring with no monochromatic K_m, since it is easy to produce colourings with more than one.) This gives

$$2\binom{n}{m} \leqslant 2^{\binom{m}{2}}$$

If we have

$$n \leqslant 2^{\frac{m-1}{2}}$$

then we can deduce

$$\begin{aligned} n^m &\leqslant 2^{\binom{m}{2}} \frac{m!}{2} \\ 2\frac{n^m}{m!} &\leqslant 2^{\binom{m}{2}} \\ 2\binom{n}{m} &\leqslant 2^{\binom{m}{2}} \end{aligned}$$

□

Remark 2. We have shown that Ramsey numbers grow exponentially, but the exact rate is not known. The number to be raised to the power m is somewhere between $\sqrt{2}$ and 4. The lower bounds that have so far been obtained by constructing explicit colourings of a K_n are significantly smaller.

Infinite Ramsey Theory

Having shown that for arbitrarily large finite numbers k and m, we can always find an n such that whenever we 2-colour a K_n, we get either a red K_k or a blue K_m, we are naturally led to ask whether if we 2-colour an infinite complete graph, we will always get an infinite monochromatic set. The answer is yes, and the proof is almost exactly the same as the proof for the finite case, but simpler in some ways, since we don't have to keep track of the sizes of things so carefully.

Theorem 4. *If the complete graph on $\mathbb{N}$ is 2-coloured, then there is a monochromatic infinite subset (i.e. a subset X such that for any $i, j \in X$, the edges ij are the same colour).*

Proof. We will form a sequence $x_0, x_1, \ldots$ of natural numbers such that the colour of the edge $x_i x_j$ with $i < j$ depends only on i. We will then call this colour c_i. Once we have done this, we know that the c_i are all either red or blue, so either infinitely many are red or infinitely many are blue. If infinitely many are red, then $\{x_i | c_i = \text{red}\}$ is an infinite red set. If infinitely many are blue, then $\{x_i | c_i = \text{red}\}$ is an infinite blue set, so once we have found the sequence $x_0, x_1, \ldots$, we will be done.

Let $X_0 = \mathbb{N}$, and let $x_0 = 0$. Start by considering 0. It has either an infinite number of red neighbours, or an infinite number of blue neighbours. If it has an infinite number of red neighbours, let X_1 be the set of its red neighbours, otherwise, let X_1 be the set of its blue neighbours. If we choose all subsequent x_i from the set X_1, then every edge from x_0 to x_i will be the same colour.

We choose x_1 to be the smallest element in X_1. Consider the edges from x_1 to other elements of X_1. Either infinitely many of them are red or infinitely many are blue. If infinitely many are red, let $X_2 \subset X_1$ be the set of red neighbours of x_1 in X_1. Otherwise, let $X_2 \subset X_1$ be the set of blue neighbours of x_1 in X_1. Now we will choose all subsequent values for our sequence to be in X_2. We let x_3 be the smallest element of X_2, and continue in the same way to get the required sequence $x_0, x_1, \ldots$. □

We can use this infinite version of Ramsey's theorem to prove the finite version without getting any upper bounds.

Corollary 1 (Ramsey's Theorem). *For any positive integers k, m, there is an n such that however we 2-colour (red and blue) the edges of a K_n, we can always find either a red K_k or a blue K_m.*

Proof. Suppose there are values of k, m for which there is no such n. Then for every n, there is a colouring of the edges of a complete graph on $\{0, 1, \ldots, n-1\}$ with no red K_k or blue K_m. Furthermore, if we take a colouring of the edges of the complete graph on $\{0, 1, \ldots, N-1\}$ for some $N > n$, with no red K_k or blue K_m, we can restrict to a colouring of $\{0, 1, \ldots, n-1\}$ with no red K_k or blue K_m. Therefore, given a colouring of $\{0, 1, \ldots, n-1\}$ with no red K_k or blue K_m, we can ask how much we can extend it – i.e. for what N is our colouring the restriction of a colouring of the edges of the complete graph on $\{0, 1, \ldots, N-1\}$ with no red K_k or blue K_m. Since there are colourings of arbitrarily large complete graphs with no red K_k or blue K_m, and there are only finitely many colourings of $\{0, 1, \ldots, n-1\}$, we can find some colouring which can be extended to arbitrarily large N. Furthermore, there is an extension of this colouring to $n+1$ that can also be extended to arbitrarily large N, and so on.

By picking these extensions, we get a sequence of colourings $c_1, c_2, \ldots$ such that c_i is a colouring of K_i, and if $i < j$ then c_j is an extension of c_i. We can combine these to get a colouring of the edges of the complete graph on $\mathbb{N}$ by colouring the edge ij with $i < j$ the colour it is given by c_{j+1} (since the later colourings extend the earlier ones, this is the same as the colouring it is given by every c_l for $l > j$. No initial segment $\{0, 1, \ldots, N-1\}$ contains a red

K_k or a blue K_m, since the initial segment is coloured by c_N; but if the whole graph contains a red K_k or a blue K_m, then since k and m are finite, it would have to be contained in some initial segment $\{0, 1, \ldots, N-1\}$, so the whole graph doesn't contain a red K_k or a blue K_m, and therefore does not contain a monochromatic infinite subset. This contradicts the previous theorem, so our original assertion that there were values k, m for which no such n exists must have been impossible. $\square$

One advantage of this proof is that the same argument can be used to prove:

Corollary 2. *For any $m \in \mathbb{N}$ there is an n such that whenever the complete graph on $\{0, 1, \ldots, n-1\}$ is 2-coloured, there is a monochromatic subset X whose size is at least as big as the smallest number in X, and is also at least m.*