Mathellaneous

Norman Do

Party Problems and Ramsey Theory

1 Problems Pertaining to People at Parties

Suppose that you are at a party and you notice that there are three people, all of whom know each other — hardly a surprising observation, one must admit. But at another party the following night, you happen to notice that there are three people, all of whom do not know each other. Following that, you wonder whether it might always be the case that a party must include either three mutual friends or three mutual strangers. Of course, this statement would not apply to a party of one or two people, but perhaps there is a certain critical mass, so that parties with enough people do possess this property.

> How many people do you need at a party to guarantee that there are three people all of whom know each other or three people all of whom do not know each other?

Of course, we can pictorially represent our party by replacing each person with a point in the plane and using a red line segment to join people who are acquainted with each other and a blue line segment to join people who are not. Thus, we find ourselves within the realm of graph theory, and our party problem can be rephrased in the following less social, though more colourful, terminology.

> What is the smallest value of N such that if the edges of K_N are coloured red or blue, then the resulting graph must contain a red K_3 or a blue K_3?

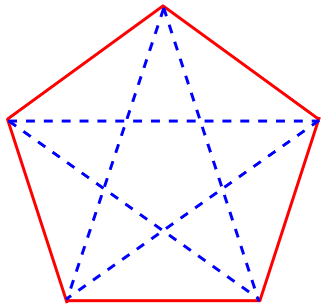
Here, we have used the notation K_N to represent the complete graph on N vertices — that is, the graph with N vertices and an edge between every pair of them. For example, K_1 represents a vertex, K_2 represents an edge between two vertices and K_3 represents a triangle. The answer to our problem, can now be stated as follows.

> If the edges of K_6 are coloured red or blue, then the resulting graph must contain a red K_3 or a blue K_3. Furthermore, it is possible to colour the edges of K_5 red or blue so that the resulting graph does not contain a red K_3 or a blue K_3.

The proof of this statement is delightfully simple and elegant. Consider any vertex V in the coloured K_6 and the five adjacent edges. By the pigeonhole principle, at least three of these edges, VA, VB, VC are of the same colour and, without loss of generality, we may assume that they are red. If our graph is to avoid red triangles, then the edge AB is forced to be blue. Similarly, the edges BC and CA are forced to be blue, thereby creating the blue triangle ABC. So any K_6 whose edges have been coloured red or blue must contain a red triangle or a blue triangle.

To complete the proof, it suffices to demonstrate a colouring of the edges of K_5 which is devoid of red or blue triangles — an example is pictured below[1].

[1] Due to the monochromatic nature of the hardcopy *Gazette*, the blue lines are dashed.



Problem: Prove that if the edges of K_6 are coloured red or blue, then the resulting graph must actually contain two distinct, though not necessarily disjoint, monochromatic copies of K_3.

2 Ramsey's Theorem and Ramsey Theory

Of course, there is no need for us to restrict our attention to trios of friends or strangers. More generally, we can ask the following question.

> How many people do you need at a party to guarantee that there are m people all of whom know each other or n people all of whom do not know each other?

Once again, we can state the problem in graph theoretic terms.

> What is the smallest value of N such that if the edges of K_N are coloured red or blue, then the resulting graph must contain a red K_m or a blue K_n?

Such problems are central to the domain of mathematics known as Ramsey theory[2]. In keeping with the accepted notation, let us denote the answer to this problem by $R(m, n)$. Note that in asking for the smallest such value of N, we are presuming that there is indeed a value of N in the first place. A priori, it is not at all clear that this is the case. For example, are all parties with a sufficiently large attendance guaranteed to have either a million people all of whom know each other or a million people all of whom do not know each other? That this is indeed the case is the conclusion of Ramsey's theorem.

Ramsey's theorem:
For every pair of positive integers m and n, the value of $R(m, n)$ is finite. In other words, there is a positive integer N such that if the edges of K_N are coloured red or blue, then there exists a red K_m or a blue K_n.

Proof. It requires only a moment's thought to conclude that $R(2, n) = R(n, 2) = n$. We will prove by induction on $m + n$ that

$$R(m, n) \leq R(m - 1, n) + R(m, n - 1)$$

[2]Ramsey theory is named after Frank Plumpton Ramsey (1903–1930), a most remarkable man who made significant contributions not only to mathematics, but also to economics and philosophy. After learning to read German in little over a week, he was handed the task of translating the text of Wittgenstein's *Tractatus Logico Philosophicus* while still at the tender age of 19. John Maynard Keynes himself encouraged Ramsey to try his hand at economics, resulting in him contributing three important papers. In a biographical article on Ramsey [3], Mellor indicates that "Ramsey's enduring fame in mathematics... rests on a theorem he didn't need, proved in the course of trying to do something we now know can't be done!" Despite this, the evidence clearly indicates that Ramsey was a first-rate mathematician. However, as is so often the case in mathematics, Ramsey's career shone brightly yet all too briefly. He died at the age of 26 after complications from an abdominal operation.

which will provide us not only with a proof that $R(m,n)$ exists, but also with a computable upper bound for its value.

Consider a complete graph on $R(m-1,n)+R(m,n-1)$ vertices and pick any vertex v from the graph. Let V_R denote the set of vertices which are connected to v by a red edge and let V_B denote the set of vertices which are connected to v by a blue edge. Then by a simple application of the pigeonhole principle, either $|V_R| \geq R(m-1,n)$ or $|V_B| \geq R(m,n-1)$ and, without loss of generality, we may assume that the former of these two inequalities is true.

Now by the very definition of $R(m-1,n)$, either V_R contains a blue K_n or a red K_{m-1}. If the former is true, then we are done. And if the latter is true, then the complete graph formed by the red K_{m-1} and the edges connecting it to v will give a red K_m. □

There are various directions in which we can hope to generalize Ramsey's theorem. Rather than restricting ourselves to red and blue edges, we might extend our palette to include an arbitrary, though finite, number of colours. Furthermore, rather than looking for monochromatic complete graphs, we might wish to look for other particular graphs of a given colour. For example, one can ask whether there exists a number N such that if the edges of K_N are coloured red, orange, yellow, green, blue, indigo or violet, then there must exist one of the following graphs:

- a red cycle of 2006 edges;
- an orange path of 11 edges;
- a yellow complete graph on 163 vertices;
- a green graph consisting of 42 edges sharing a common vertex;
- a blue graph consisting of 239 triangles sharing a common vertex;
- an indigo complete graph on 1729 vertices; or
- a violet subgraph whose vertices are in correspondence with all of the words in the English language and whose edges join two vertices if and only if they represent two distinct words with at least one letter in common.

To simplify the situation, note that when looking for a red cycle of 2006 edges, it is entirely sufficient, though far from necessary, to guarantee the existence of a red complete graph on 2006 vertices. Similarly, when looking for a graph on V vertices of a particular colour, it suffices to show the existence of a complete graph on V vertices of that colour. Therefore, we can restrict our attention to complete graphs. So we are now left with the problem of whether Ramsey numbers exist when more than two colours are allowed. This problem is answered by the following extension of Ramsey's theorem.

Ramsey's theorem (colourful version):
For every tuple of positive integers $(m_1, m_2, \ldots, m_C)$, there is a positive integer $R = R(m_1, m_2, \ldots, m_C)$ such that if the edges of K_R are coloured in one of the "colours" $1, 2, \ldots, C$, then there exists a complete subgraph on m_k vertices, all of whose edges have the colour k, for some value of k.

Proof. The $C = 2$ case is precisely the statement of Ramsey's theorem given earlier. We will now prove that

$$R(m_1, m_2, \ldots, m_C) \leq R(R(m_1, m_2), m_3, \ldots, m_C),$$

which shows by induction that the Ramsey numbers exist for any number of colours.

Consider a complete graph on $R(R(m_1, m_2), m_3, \ldots, m_C)$ vertices whose edges have been coloured in one of the colours $1, 2, \ldots, C$. Now suppose that the colours 1 and 2 correspond to red and green, respectively. Then a person who is red-green colour blind would only see $C-1$ colours. By definition, the graph must have a K_{m_k} whose edges are coloured k for

$k = 3, 4, \ldots, C$ or a complete graph on $R(m_1, m_2)$ vertices whose edges are all red-green. If the former is true, then we are done. And if the latter is true, then someone with perfect vision will be able to see a K_{m_1} whose edges are red or a K_{m_2} whose edges are green. □

Problem: Prove that at any party with nine people, there are either three people, all of whom know each other, or four people, all of whom do not know each other.

3 Calculating the Ramsey Numbers

The two colour Ramsey numbers are of central importance in Ramsey theory and we will concentrate on them for the remainder of the article. Given that the numbers $R(m, n)$ are guaranteed to be finite by Ramsey's theorem, the problem of calculating what the numbers actually are is entirely natural. As mentioned earlier, it is a trivial matter to prove that $R(2, n) = R(n, 2) = n$ for all n. The next case to consider is $R(3, 3) = 6$, which corresponds with the first party problem discussed in the article. And the problem posed at the end of the previous section asks to verify that $R(3, 4) = R(4, 3) = 9$. Unfortunately, despite the best efforts of mathematicians, there is no known formula for the Ramsey numbers in general. In fact, only seven other Ramsey numbers $R(m, n)$ are known for $m \leq n$. The following table shows these numbers, as well as the known upper and lower bounds for many of the other Ramsey numbers [4].

$n \backslash m$	3	4	5	6	7	8	9	10	11	12	13	14	15
3	6	9	14	18	23	28	36	40 43	46 51	52 59	59 69	66 78	72 88
4		18	25	35 41	49 61	56 84	73 115	92 149	97 191	128 238	133 291	141 349	153 417
5			43 49	58 87	80 143	101 216	125 316	143 442	159	185 848	209	235 1461	265
6				102 165	113 298	127 495	169 780	179 1171	253	262 2566	317	 5033	401
7					205 540	216 1031	233 1713	289 2826	405 4553	416 6954	511 10581	 15263	 22116
8						282 1870	317 3583	 6090	 10630	 16944	817 27490	 41525	861 63620
9							565 6588	580 12677	 22325	 39025	 64871	 89203	
10								798 23556		 81200			1265

So why are the Ramsey numbers so difficult to calculate? Well, suppose we decide to use a brute force approach to calculate the value of $R(5, 5)$. As witnessed from the above table, we have the quite reasonable bounds $43 \leq R(5, 5) \leq 49$. If we actually believed the answer to be 43, we might consider simply drawing all of the possible two-colourings of the complete graph on 43 vertices. Then, it would be a simple matter to examine each one to determine whether or not it contained a monochromatic copy of K_5. However, the number of edges in a graph on 43 vertices is precisely $\binom{43}{2} = 903$. Therefore, the number of distinct ways to colour the edges of the graph red or blue is $2^{903} \approx 6.76 \times 10^{271}$. So even if it were possible to analyze 10^{20} cases per second, the time required would still be of the order of 2×10^{244} years! Of course, it is possible to narrow down the number of cases by many orders of magnitude, but the computation is still far beyond our current technological capabilities.

To indicate the difficulty in calculating the Ramsey numbers, Paul Erdős, one of the most prolific mathematicians ever and a Ramsey theory enthusiast, used to tell the following story. He would claim that if a technologically superior race of aliens landed upon Earth and demanded the calculation of $R(5,5)$ within a year, then our best chance for survival would be to gather together all of the mathematicians and computing power in the world to work on the problem. On the other hand, if they demanded the calculation of $R(6,6)$, then Erdős claimed that our best chance for survival would be to gather together the world's military power in an attempt to destroy the aliens!

Given that the Ramsey numbers are considered to be so difficult to calculate, it seems reasonable to ask whether we can at least find bounds for them. Of most interest to mathematicians is the computation of $R(n,n)$, in which case we have the following bounds.

$$2^{\frac{n}{2}} \leq R(n,n) \leq 2^{2n-3}$$

The upper bound can be obtained quite easily from the inequality which was proved earlier,

$$R(m,n) \leq R(m-1,n) + R(m,n-1).$$

In conjunction with the initial values $R(2,n) = R(n,2) = n$, and the well-known recursion $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$, we obtain

$$R(m,n) \leq \binom{m+n-2}{m-1}.$$

Now it suffices to observe that

$$R(n,n) \leq \binom{2n-2}{n-1} = \binom{2n-3}{n-2} + \binom{2n-3}{n-1} \leq 2^{2n-3}.$$

The lower bound is more difficult to obtain, but its proof will allow us to showcase one of the mathematical legacies of Paul Erdős — namely, the probabilistic method. It is an extremely general principle which can be stated as follows.

> Suppose that the probability of an element chosen from a particular set having a certain property is less than 1. Then there must exist an element of the set with the desired property.

Simple as it seems, the addition of the probabilistic method to the combinatorialist's arsenal has yielded many new results as well as beautiful proofs of old results. The following was included in *Proofs from the Book* [1], an approximation to The Book, where Erdős believed that God stored the perfect proofs for mathematical theorems.

Theorem: $R(n,n) \geq 2^{\frac{n}{2}}$

Proof. Since $R(2,2) = 2$ and $R(3,3) = 6$, the theorem certainly holds true for $n = 2$ and $n = 3$. We will now prove that for $n \geq 4$ and $N < 2^{\frac{n}{2}}$, there exists a colouring of the edges of K_N red or blue which does not contain a monochromatic K_n. Observe that there are $2^{\binom{N}{2}}$ colourings in total. If we consider declaring each edge red or blue independently with probability $\frac{1}{2}$, then it is clear that any particular colouring occurs with probability $2^{-\binom{N}{2}}$.

Now given a set of vertices V, let V_R denote the event that the edges between the vertices of V are all red. Then the probability of the event V_R is simply $2^{-\binom{|V|}{2}}$. Let X_R denote the event that there is some complete graph on n vertices which is coloured red.

$$
\begin{aligned}
\Pr(X_R) &= \Pr\Big(\bigcup_{|V|=n} V_R\Big) \leq \sum_{|V|=n} \Pr(V_R) = \binom{N}{n} 2^{-\binom{n}{2}} \\
&= \frac{N(N-1)(N-2)\cdots(N-n+1)}{n(n-1)(n-2)\cdots 1} 2^{-\binom{n}{2}} \\
&\leq \frac{N^n}{n(n-1)(n-2)\cdots 1} 2^{-\binom{n}{2}} \leq \frac{N^n}{2^{n-1}} 2^{-\binom{n}{2}} \\
&< 2^{\frac{n^2}{2}-\binom{n}{2}-n+1} = 2^{1-\frac{n}{2}} \\
&\leq \frac{1}{2}
\end{aligned}
$$

So we have shown that for $n \geq 4$ and $N < 2^{\frac{n}{2}}$, the probability of a red K_n is less than $\frac{1}{2}$. By the symmetry of the problem, the probability of a blue K_n is also less than $\frac{1}{2}$. Therefore, the probability of a monochromatic K_n is less than 1, so there must exist a colouring which contains neither a red K_n nor a blue K_n. □

4 Complete Disorder is Impossible

Ramsey theory is concerned with more than simply determining the Ramsey numbers. In fact, there is a myriad of interesting results which possess the same flavour as the party problems considered above. The overarching theme behind Ramsey theory is the fact that within sufficiently large mathematical systems, there must exist subsystems containing a certain degree of order. This is often succinctly described by the Ramsey theorists' catch phrase, "Complete disorder is impossible!" We will conclude the article with a brief look at three Ramsey-type results.

Ramsey's Theorem — Infinite Version. The first result is Ramsey's theorem, as stated by Ramsey himself in his 1930 paper entitled, *On a problem of formal logic.* As witnessed by the title, he did not consider the result to be of great combinatorial importance, and it appeared only as a lemma towards what he considered a more substantial problem of formal logic. This presumably was a result of the fact that the foundations of mathematics were Ramsey's great passion, combined with the fact that combinatorics was a far less fashionable subject at that time than it is today. Ramsey's original result can be considered an infinite version of the theorem which is now attributed to him.

> For every pair of positive integers C and N, if the subsets of $\mathbb{N}$ with N elements are coloured in C colours, then there exists an infinite subset X of $\mathbb{N}$ such that all subsets of X with N elements are of the same colour.

Problem: Prove the infinite version of Ramsey's theorem and show that the finite version of Ramsey's theorem follows from the infinite version.

Van der Waerden's theorem. In 2004, Ben Green and Terence Tao announced their celebrated result that the primes contain arbitrarily long arithmetic progressions. One of the main ingredients in their proof was Szemerédi's theorem which states that if we take any positive fraction of the set of positive integers, a notion which can be made mathematically precise, then the resulting set must contain arbitrarily long arithmetic progressions. A precursor to Szemerédi's theorem is the following result, proved by Van Der Waerden in 1927.

> For every pair of positive integers C and P, there is a positive integer N such that if the numbers from 1 up to N are coloured in C colours, then there exists at least P numbers in arithmetic progression, all of the same colour.

The Hales-Jewett theorem. The final result we will consider is motivated by the game tic-tac-toe, in which players take turns to mark the squares of a 3×3 grid with the aim of occupying three squares along a column, row or diagonal. It is a well-known fact that in this traditional form of tic-tac-toe, both players can force a draw with optimal play. In stark contrast is the game of three-dimensional tic-tac-toe, played in a similar manner on a $3\times3\times3$ grid of cubes, where the first player has an easy win by occupying the central position on the first move. Actually, it is impossible to play out a draw in three-dimensional tic-tac-toe since any partition of the 27 cubes into two colours will always include a monochromatic column, row or diagonal. If we consider instead N-in-a-row tic-tac-toe played between C players, then the Hales-Jewett theorem guarantees similar behaviour.

> For every pair of positive integers C and N, there is a positive integer D such that if the unit hypercubes in a D-dimensional $N \times N \times \cdots \times N$ hypercube are coloured in C colours, then there exists at least one row, column or diagonal of N squares, all of the same colour.

Problem: Show that Van Der Waerden's theorem follows from the Hales-Jewett theorem.

Anyone wanting to find out more is strongly encouraged to consult the monograph entitled *Ramsey Theory* by Ronald Graham, Bruce Rothschild and Joel Spencer [2].

References

[1] M. Aigner and G.M. Ziegler, *Proofs from the Book*, 3rd ed. (Springer 2004).
[2] R.L. Graham, B.L. Rothschild and J. H. Spencer, *Ramsey Theory* (Wiley New York 1980).
[3] D.H. Mellor, *The eponymous F. P. Ramsey*, J. Graph Theory **7** (1983), 9–13.
[4] S.P. Radziszowski, *Small Ramsey numbers*, Electron. J. Comb. **DS1.11** (2006).

Department of Mathematics and Statistics, The University of Melbourne, VIC 3010
E-mail: N.Do@ms.unimelb.edu.au