7. Ramsey Theory

A major theme of graph theory has been the emergence of the understanding that, in some sense, complete chaos is impossible. Probably the first result in graph theory of this type is Ramsey's Theorem. It was proved in 1928 by **Frank Plumpton Ramsey**. It has turned out to be another example of a truly seminal theorem. Incidentally, Ramsey proved it, almost as an aside, in a paper on logic.

We have already seen an example in an assignment. If I take K_6 and colour the edges either red or blue, then the graph must either have a red triangle or a blue triangle. Ramsey's Theorem is a generalisation of this.

Let G be a graph, and let $U \subseteq V(G)$. We say that U is a *clique* if every pair of distinct vertices in U are adjacent. As you would expect, when U is a clique with $|U| = t$, we say that it is a clique of *size* t. In this section, we are primarily interested in the case that G is a simple graph; then, $U \subseteq V(G)$ is a clique of size t if $G[U]$ is isomorphic to K_t.

Let n be a positive integer, and let K be a complete graph on n vertices with the edges coloured red or blue. For $t < n$, we say that a subset U of t vertices of K is a *red* clique if all the edges in $K[U]$ are red, and is a *blue* clique if all the edges in $K[U]$ are blue. We say that U is a *monochromatic* clique if it is either a red or a blue clique.

Theorem 7.1 (Ramsey's Theorem, 1928)**.** *For every positive integer t, there is a number $r(t)$ such that the following holds: if G is a red-blue edge-coloured complete graph with at least $r(t)$ vertices, then G has a monochromatic clique of size t.*

In other words, given a very large 2-edge-coloured clique, we must have a large set of vertices where all edges between those vertices are the same colour. These "monochromatic cliques" are a structure that we cannot avoid.

It turns out that the best way to prove Theorem 7.1 is to prove a stronger theorem. We'll ask a slightly more sophisticated question. For positive integers s and t is it the case that, given a sufficiently large red-blue edge-coloured complete graph, we can always find a red clique of size s or a blue clique of size t?

This may seem like a more difficult thing to prove, but, in fact, it is easier.

Theorem 7.2. *For all positive integers s and t, there is a number $r(s, t)$ such that if G is a red-blue edge-coloured complete graph with at least $r(s, t)$ vertices,* *then G has either a red clique of size s, or a blue clique of size t. Moreover, when $s, t \geq 2$, we have $r(s,t) \leq r(s-1,t) + r(s,t-1)$.*

The natural way have a go at proving this theorem is by some sort of induction, but we have a problem. Induction works when we have a linear order, which is what happens when we are counting something like vertices or edges. But here we have ordered pairs (s, t).

The trick is to put a linear order on ordered pairs of positive integers. The ordering we will use is very standard, and it's worth looking at in more detail.

Lexicographic Ordering. Let S be a set with a relation $\prec$ on S. Then $\prec$ is a *linear order on S* if the following hold:

(i) for all $a, b \in S$, if $a \prec b$, then $b \not\prec a$;
(ii) for all $a, b, c \in S$, if $a \prec b$ and $b \prec c$, then $a \prec c$;
(iii) for all $a \in S$, we have $a \not\prec a$; and
(iv) for all $a, b \in S$ with $a \neq b$ we have either $a \prec b$ or $b \prec a$.

Strictly speaking, we didn't need condition (iii). It's really a question of whether we are defining $\preceq$ or $\prec$. In any case, the point is that a linear order is a partial order with the property that every pair of distinct elements is comparable one way or another.

Of course the natural numbers form a linear order with the normal ordering. So do the letters of the alphabet, with

$$a \prec b \prec c \prec \cdots \prec y \prec z.$$

Recall that a *word* from a set S is just an ordered string $s_1 s_2 \cdots s_n$ where $s_i \in S$ for all $i \in \{1, 2, \ldots, n\}$.

→ There is no fundamental difference between words and vectors; the only real difference between the word $s_1 s_2 \cdots s_n$ and the vector $(s_1, s_2, \ldots, s_n)$ is the notation.

→ The crucial point is that in words, and vectors, the order matters. The word *rat* is different from the word *art*, just as the vector (r, a, t) is different from the vector (a, r, t).

Let S be a linearly ordered set and let S^* be a set of words from S. The *lexicographic ordering* on the members of S^* is obtained as follows. Say $u = u_1u_2\cdots u_m$ and $v = v_1v_2\cdots v_n$ are distinct words in S^*, where $m \le n$.

- If $v_1v_2\cdots v_m = u_1u_2\cdots u_m$, then $u \prec v$.
- Otherwise, let i be the first index such that $u_i \neq v_i$. If $u_i \prec v_i$, then $u \prec v$; whereas if $v_i \prec u_i$, then $v \prec u$.

It's not hard to see that lexicographic ordering is a linear order. Of course, if you can use a dictionary, then you already understand lexicographic ordering. The word *limb* appears before *limbic* as an example of the first rule. The word *wilderness* appears before *wildfire* as an example of the second, since $e \prec f$.

Our original motivation was to obtain a linear order for pairs (a, b) of positive integers. In the lexicographic order, when (a, b) and (c, d) are distinct, we have $(a, b) \prec (c, d)$ if either $a < c$, or $a = c$ and $b < d$; otherwise $(c, d) \prec (a, b)$.

→ Perhaps this seems like a lot of effort to go to, just to define the above order. But lexicographic orderings are a useful theme in mathematics, not just for ordering pairs of integers.

Proof of Theorem 7.2. Let $\prec$ be the lexicographic order on pairs of positive integers. Trivially, any clique of size one is both red and blue. Hence we may set

$$r(1, t) = r(s, 1) = 1$$

for all s and t. This establishes the base case of our induction.

Now consider a pair (s, t) for $s > 1$ and $t > 1$. We may assume that for any (s', t') such that $(1, 1) \prec (s', t') \prec (s, t)$, the result holds. That is, for such an (s', t'), there exists a number $r(s', t')$ such that any red-blue edge-coloured complete graph with at least $r(s', t')$ vertices has either a red clique of size s', or a blue clique of size t'. When such an $r(s', t')$ exists, we say that $r(s', t')$ is *well defined.*

In the lexicographic order, $(1, 1) \prec (s-1, t) \prec (s, t)$ and $(1, 1) \prec (s, t-1) \prec (s, t)$. Thus, by the induction assumption, $r(s-1, t)$ and $r(s, t-1)$ are well defined. Now let G be a red-blue edge-coloured complete graph with

$$|V(G)| = r(s-1, t) + r(s, t-1).$$

Let v be a vertex of G. The degree of v is

$$d(v) = |V(G)| - 1 = r(s-1, t) + r(s, t-1) - 1.$$

Let R_v and B_v be the set of red and blue edges incident with v, respectively.

7.2.1. *Either* $|R_v| \geq r(s-1,t)$, *or* $|B_v| \geq r(s,t-1)$.

Subproof. If $|R_v| < r(s-1,t)$ and $|B_v| < r(s,t-1)$, then

$$d(v) \leq r(s-1,t) - 1 + r(s,t-1) - 1 < r(s-1,t) + r(s,t-1) - 1 = d(v).$$

But $d(v) < d(v)$ is contradictory, so we deduce that 7.2.1 holds. $\lhd$

Say that $|R_v| \geq r(s-1,t)$. Let S be the set of vertices joined to v by red edges. Then $|S| \geq r(s-1,t)$. Hence, as $r(s-1,t)$ is well defined, $G[S]$ has either a red clique of size $s-1$ or a blue clique of size t. If the latter case holds, then G has a blue clique of size t. In the former case $S \cup \{v\}$ is a red clique of size s.

On the other hand, suppose $|B_v| \geq r(s,t-1)$. Then we let S be the set of vertices joined to v by blue edges, so $|S| \geq r(s,t-1)$. As $r(s,t-1)$ is well defined, either $G[S]$ has either a red clique of size s, or a blue clique of size $t-1$, in which case $S \cup \{v\}$ is a blue clique of size t.

We have shown that if G is a red-blue edge-coloured complete graph with at least $r(s-1,t) + r(s,t-1)$ vertices, then G has either a red clique of size s, or a blue clique of size t. Therefore, $r(s,t)$ is well defined and, in particular, $r(s,t) \leq r(s-1,t) + r(s,t-1)$. The theorem now follows by induction. $\square$

We now get Theorem 7.1 as a straightforward corollary.

Proof of Theorem 7.1. For any positive integer t, we let $r(t)$ be the number $r(t,t)$ given by Theorem 7.2. Let t be a positive integer and let G be a red-blue edge-coloured complete graph with at least $r(t)$ vertices. Then, as $r(t) = r(t,t)$ where $r(t,t)$ is as given by Theorem 7.2, G has either a red clique of size t, or a blue clique of size t. In other words, G has a monochromatic clique of size t, as required. $\square$

→ Note that we never found the exact value of $r(s,t)$. All we said was that such a number exists. That is, the proof of Ramsey's theorem was *non-constructive.*

Once upon a time, in the not too distant past, it could be said that the dominant culture in mathematics was anti constructive techniques. But the computer revolution has changed that. The theory of algorithms and computability theory are now highly developed branches of mathematics. A non-constructive proof, like the proof of Ramsey's theorem that we just saw, may be elegant, but a constructive proof that leads to an efficient algorithm or an exact answer could be more useful.

Ramsey Numbers. Let s and t be positive integers. Now, we let $r(s,t)$ be the *minimum* possible number that guarantees that a red-blue edge-coloured complete graph with at least $r(s,t)$ vertices has either a red clique of size s or a blue clique of size t.

→ Theorem 7.2 showed that, for any (s,t), *there exists* a number N that guarantees that a red-blue edge-coloured complete graph with at least N vertices has either a red clique of size s or a blue clique of size t. But if N is such a number, so is $N+1$, and so is any number greater than N. Now, we are interested in finding the smallest possible value for N, and henceforth we take $r(s,t)$ to be this number.

Numbers that give this (minimum) value of $r(s,t)$, for different values of s and t, are called *Ramsey numbers*. Graph theorists have expended considerable effort in calculating Ramsey numbers.

The way to try to do this is to calculate *upper and lower bounds* for a Ramsey number. Consider a simple example.

Take K_5 and colour a 5-cycle red and the rest blue. Then it is easily checked that there is no red triangle and no blue triangle. Hence $r(3,3) \geq 6$. We have established that 6 is a *lower bound* for $r(3,3)$.

But we proved in an assignment that any red-blue edge-coloured K_6 has either a red triangle or a blue triangle. Hence $r(3,3) \leq 6$. In other words, 6 is an *upper bound* for $r(3,3)$.

Hey! The number 6 is both a lower and an upper bound. So we have proved that $r(3,3) = 6$.

How easy was that. Surely calculating the others cannot be too hard. It turns out that it is not that easy. It's not hard to see that $r(s,t) = r(t,s)$ so it's usual to focus on Ramsey numbers where $s < t$.

Also, we've already seen that $r(1,t) = 1$ for all t, and the next lemma is easy.

Lemma 7.3. *If $t \geq 2$, then $r(2,t) = t$.*

Proof. [redacted] □

The interesting cases are where $s \geq 3$ and $t \geq 3$.

The following table shows all known Ramsey numbers $r(s,t)$ where $3 \leq s \leq t$.

s	3	3	3	3	3	3	3	4	4
t	3	4	5	6	7	8	9	4	5
$r(s,t)$	6	9	14	18	23	28	36	18	25

That seems pretty pathetic. If we think of the cases where $s = t$, then all we know is $r(3,3)$ and $r(4,4)$. Surely $r(5,5)$ cannot be too hard to calculate. **Paul Erdős**, one of the most interesting mathematicians of all time, said the following in about 1990.

> Suppose aliens invade the earth and threaten to obliterate it in a year's time unless human beings can find $r(5,5)$. We could marshall the world's best minds and fastest computers, and within a year we could probably find the value. If the aliens demanded $r(6,6)$ we would have no choice but to launch a preemptive attack.

And it's not that Paul was naive about increasing computer power, as one can see from this interesting **Scientific American article**.

The problem is *combinatorial explosion* which is a way of describing just how incredibly quickly the number of possibilities can grow in discrete mathematics.

→ Combinatorial explosion can be a good thing, however. If we didn't have it, then we wouldn't be able to encode enough information for human beings in the DNA of an embryo. Also, cryptographic systems, which are now used ubiquitously (for example, in internet banking), rely on it for security.

While exact values of Ramsey numbers are beyond us, we can still try to find some upper and lower bounds. For example, it is known that $r(3,10) \geq 40$ and $r(3,10) \leq 41$ (and the latter was just proved in the last year, as you can see here).

The next theorem gives an upper bound for all Ramsey numbers.

Theorem 7.4. *For all positive integers s and t,*

$$r(s,t) \leq \binom{s+t-2}{s-1}.$$

Proof. We use induction, with the help of the lexicographic ordering on pairs of integers. Consider (s,t), for positive integers s and t. If either $s \leq 2$ or $t \leq 2$, then it is easily seen (with the help of Lemma 7.3 and the statement preceding it) that the result holds. So assume that $s \geq 3$ and $t \geq 3$, and that the result holds for all (m,n) such that $(1,1) \prec (m,n) \prec (s,t)$ in the lexicographic order.

By Theorem 7.2, and the induction hypothesis,

$$\begin{aligned} r(s,t) &\leq r(s,t-1) + r(s-1,t) \\ &\leq \binom{s+t-3}{s-1} + \binom{s+t-3}{s-2} \\ &= \frac{(s+t-3)!}{(s-2)!(t-2)!} \cdot \frac{(t-1)+(s-1)}{(s-1)(t-1)} \\ &= \binom{s+t-2}{s-1}. \end{aligned}$$

Thus, by induction, the theorem holds for all pairs of positive integers (s,t). □

Note that $\binom{s+t-2}{s-1}$ is the number of $(s-1)$-subsets of a $(s+t-2)$-element set and that 2^{s+t-2} is the total number of subsets. So, we see that

$$r(s,t) \leq 2^{s+t-2}.$$

In particular

$$r(s,s) \leq 2^{2s-2}.$$

This shows that Ramsey numbers grow at most exponentially. The next theorem of Erdős shows that Ramsey numbers of the form $r(k,k)$ grow at least exponentially. In other words, it gives an exponential *lower bound* for $r(k,k)$.

Theorem 7.5 (Erdős 1947)**.** *For all positive integers k,*

$$r(k,k) \geq 2^{k/2}.$$

Proof. Because $r(1,1) = 1$ and $r(2,2) = 2$, we may assume that $k \geq 3$. We denote by $\mathcal{G}_n$ the set of red-blue edge-coloured complete graphs with vertex set $\{v_1, v_2, \ldots, v_n\}$.

7.5.1. $|\mathcal{G}_n| = 2^{\binom{n}{2}}$.

Subproof. We have $\binom{n}{2}$ edges in the complete graph on n vertices. For each edge, there are two possibilities: we either colour it red or blue. In this way, we can obtain each possible red-blue edge-colouring. So there are $2^{\binom{n}{2}}$ graphs in $\mathcal{G}_n$. ◁

7.5.2. *The number of graphs in $\mathcal{G}_n$ having a particular set of k vertices as a red clique is $2^{\binom{n}{2}-\binom{k}{2}}$.*

Subproof. Consider our fixed set of k vertices that form a red clique. There are $\binom{k}{2}$ edges between these vertices, and we are guaranteed these are red. There are $\binom{n}{2} - \binom{k}{2}$ edges remaining. Arguing as before, we see that there are

$$2^{\binom{n}{2}-\binom{k}{2}}$$

ways of extending our colouring. $\lhd$

Let $\mathcal{R}_n^k$ be the set of those graphs which have a red clique of size k. There are $\binom{n}{k}$ distinct k-element subsets of $\{v_1, v_2, \ldots, v_n\}$. By this fact and 7.5.2, we have

$$|\mathcal{R}_n^k| \leq \binom{n}{k} 2^{\binom{n}{2}-\binom{k}{2}}$$

(A graph with more than one red clique of size k will get counted on the right more than once, so we can only guarantee an inequality here.)

Using the above inequality and 7.5.1, we get

$$\frac{|\mathcal{R}_n^k|}{|\mathcal{G}_n|} \leq \binom{n}{k} 2^{-\binom{n}{k}} \leq \frac{n^k 2^{-\binom{n}{k}}}{k!}$$

Now suppose that $n < 2^{\frac{k}{2}}$. Then

$$\frac{|\mathcal{R}_n^k|}{|\mathcal{G}_n|} < \frac{2^{k^2/2} 2^{-\binom{k}{2}}}{k!} = \frac{2^{k/2}}{k!} < \frac{1}{2}.$$

In other words, if $n < 2^{k/2}$, then fewer than half of the graphs in $\mathcal{G}_n$ contain a red k-clique. An identical argument shows that fewer than half contain a blue k-clique. Hence some graph contains neither a red k-clique nor a blue k-clique. Because this holds for any $n < 2^{k/2}$, we have $r(k,k) \geq 2^{k/2}$. $\square$

Let's take $k = 12$. By Theorem 7.5, there exists a red-blue edge-coloured complete graph with $2^{12/2} - 1 = 2^6 - 1 = 63$ vertices that does not contain a monochromatic clique of size 12. But it does not tell you how to find it!

All up, by Theorems 7.4 and 7.5 we have

$$2^{k/2} < r(k,k) < 2^{2k-2}.$$

But there is a lot of space in between those two numbers!

Stable sets and cliques. Let $G = (V, E)$ be a simple graph. Recall that a *clique* in G is a subset W of vertices of G such that $G[W]$ is a complete graph. A *stable set* in G is a subset W of vertices such that $G[W]$ has no edges. In other words, there is some $W \subseteq V(G)$ such that for every pair of distinct vertices u, v in W, there is no edge between u and v.

Recall that the *complement* of G, denoted $\overline{G}$, is the simple graph such that there is an edge between u and v in $\overline{G}$ if and only if there is no edge between u and v in G. The next lemma follows almost immediately from these definitions.

Lemma 7.6. *Let G be a simple graph. Then W is a stable set in G if and only if W is a clique in $\overline{G}$.*

While stable sets and cliques are very different structures, they are related by Lemma 7.6.

Many natural problems amount to finding a maximum-sized clique or stable set in a graph. Here is an example.

Attack of the Aliens. The aliens have given the world one year to find $r(5, 5)$. Victoria University of Wellington is doing its bit, and the mathematicians and computer scientists at VUW are forming a team to collaborate on this project. It turns out that with respect to such collaborations, pairs of people are either compatible or incompatible. The goal is to find the largest possible team such that every pair of people in the team is compatible.

One way to do this is to construct a graph whose vertices are the mathematicians and computer scientists with edges joining people if they are compatible. We seek a maximum sized clique in this graph.

On the other hand we could use the same set of vertices with edges joining people if they are incompatible. Then we are looking for a maximum sized stable set in this graph.

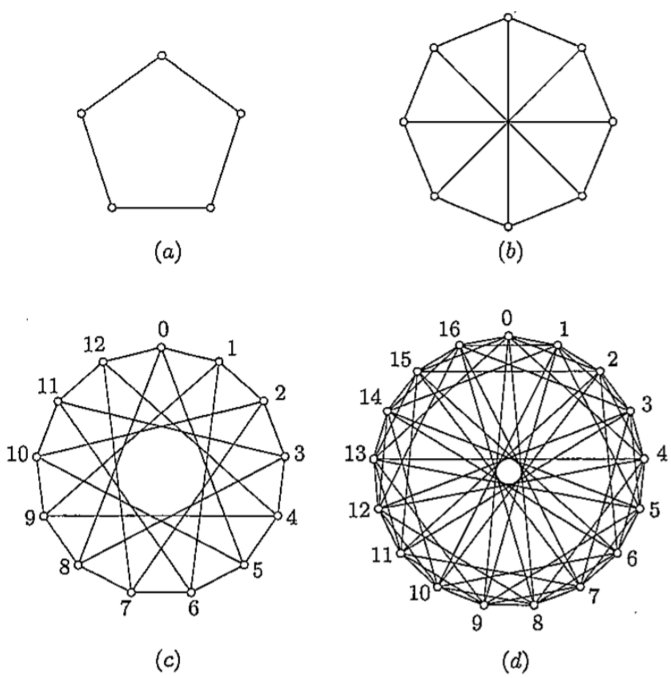
Maybe there are many incompatibilities, and we need to be a bit more creative with this problem. What if we took advantage of people's competitive instinct and were also happy to make a "team" of incompatible people?

Hopefully, this situation is starting to seem familiar. Take a complete graph whose vertices are the staff members. For an edge between two staff members, colour it blue if they are compatible, and red if they are not. A large blue clique gives us our compatible team, and a large red clique gives us our ruthless competitors.

Theorem 7.7. *Let G be a simple graph. If G has at least $r(s,t)$ vertices, then G has either a clique of size s or a stable set of size t.*

Proof. Say $G = (V, E)$. We use G to construct a complete graph K_n with vertex set V as follows. For each edge uv of this K_n, we colour it red if there is an edge between u and v in G, otherwise we colour it blue. By Theorem 7.2, the K_n has either a red clique of size s, giving a clique of size s in G; or K_n has a blue clique of size t, giving a stable set of size t in G. □

In fact, Ramsey theory is often presented in terms of cliques and stable sets. In the end, there is no real difference; the underlying theory is identical. One advantage of the clique/stable set perspective is that it makes it easier to draw pictures. An (s,t)-Ramsey graph is a graph G that has no clique of size s and no stable set of size t and is such that $|V(G)| = r(s,t) - 1$. In other words, for any graph with more vertices, there must be either a clique of size s or a stable set of size t. Given that we know very few Ramsey numbers exactly, we know very few (s,t)-Ramsey graphs. Here are a few:



(a) (b) (c) (d)

The 5-cycle in (a) is a $(3,3)$-Ramsey graph. The graph in (b) is our old friend the Möbius graph and is a $(3,4)$-Ramsey graph. The graph in (c) is a $(3,5)$-Ramsey graph, and the graph in (d) is a $(4,4)$-Ramsey graph.

Given the symmetries exhibited by these graphs, it is tempting to think that constructive techniques could exist to find higher-order Ramsey graphs. But it seems that such techniques, if they exist, are not easy to discover.

The Ramsey theory results we have seen so far can be generalised in many, many ways. Indeed Ramsey theory is a subject in its own right and has implications way beyond graph theory.

More Colours. One obvious way to extend Ramsey theory is to have more colours in our palette than just red and blue. Assume we have k "colours" $\{1, 2, \ldots, k\}$ for some $k \geq 2$, and that we k-edge-colour a complete graph G with these colours. We say that a clique S in G is *i-monochromatic* if all edges of $G[S]$ are coloured i.

For positive integers $t_1, t_2, \ldots, t_k$, we define $r(t_1, t_2, \ldots, t_k)$ to be the least integer n such that, for any k-edge-coloured complete graph K on at least n vertices, there will exist i such that K has an i-monochromatic clique S of size t_i.

Theorem 7.8. *For all positive integers $t_1, t_2, \ldots, t_k$, the number $r(t_1, t_2, \ldots, t_k)$ is well defined and satisfies*

$$\begin{aligned} r(t_1, t_2, \ldots, t_k) \leq r(t_1 - 1, t_2, \ldots, t_k) + r(t_1, t_2 - 1, \ldots, t_k) + \cdots \\ + r(t_1, t_2, \ldots, t_k - 1) - k + 2. \end{aligned}$$

when $t_i \geq 2$ for all i.

Proof. [redacted] □

The proof of Theorem 7.8 generalises the technique of the proof of Theorem 7.2 and is an excellent exercise.

An application to number theory. Consider the partition

$$(\{1, 4, 10, 13\}, \{2, 3, 11, 12\}, \{5, 6, 7, 8, 9\})$$

of the set $\{1, 2, 3, \ldots, 13\}$. You can check that in each part of the partition, you cannot find a, b, c such that

$$a + b = c.$$

But you can also check that, no matter how you partition $\{1, 2, \ldots, 14\}$ into three parts, you can always find a solution to $a+b = c$ in one of the parts. Schur proved that, in general, given any positive integer n, there exists an integer r_n such that whenever $\{1, 2, \ldots, r_n\}$ is partitioned into n parts, there will be a solution to $a + b = c$ in one of the parts.

Let $r_n = r(t_1, t_2, \ldots, t_n)$, where $t_1 = t_2 = \cdots = t_n = 3$; note that this is well defined by Theorem 7.8.

Theorem 7.9 (Schur 1916)**.** *Let* $\{A_1, A_2, \ldots, A_n\}$ *be a partition of* $\{1, 2, \ldots, r_n\}$ *into* n *subsets. Then some* A_i *contains three integers* x*,* y*, and* z *satisfying* $x + y = z$*.*

In other words, for any n, if we choose a sequence of integers of size at least r_n and partition its members into n subsets, we will always find a subset that contains integers x, y, and z satisfying $x + y = z$.

Proof. Consider the complete graph whose vertex set is $\{1, 2, \ldots, r_n\}$. Colour the edges of this graph by the rule that the edge uv gets colour i if $|u - v| \in A_i$. By Theorem 7.8, there exists a monochromatic triangle in this graph.

This means that there are three vertices a, b and c such that the edges ab, bc and ac all have colour j for some j. Assume without loss of generality that $a > b > c$. Write $x = a - b$, $y = b - c$ and $z = a - c$. Then x, y and z are all in A_j. Moreover $x + y = z$ as required. □

Theorem 7.9 is nothing more than a taster. The impact of Ramsey theory on number theory and other branches of mathematics has been quite profound. There is a copy of the excellent book *Ramsey Theory* by Graham, Rothschild and Spencer in the VUW library.