Friends and Strangers on a Party: Ramsey numbers

Jan Willem Klop

August 5, 2023

Abstract

On a party attended by 6 people, there is always a trio of mutual friends, or a trio of mutual strangers, possibly both. This popular math example is the starting point of a beautiful theory called Ramsey theory, invented by Frank Ramsey when he was in his early twenties.

There is a finite version, encountered in this chapter, and also an infinite version of Ramsey's main theorem, encountered later on. Both are powerful tools, the finite version in mathematical graph theory and combinatorics, the infinite version in logic and informatics. Most Ramsey numbers are still too difficult to be computed with the present state of the art in informatics. Possibly future quantum computers would come a bit farther, but even they would soon have to give up.

Some sixty years ago a Hungarian sociologist, studying patterns of friendship in groups of children, noticed that in a group of about twenty children, a certain regularity always seemed to manifest itself. Either the group included a quartet of mutual *friends*, or a quartet of children that were mutual *strangers*, i.e. not friends. Consulting about his observation some leading mathematicians in his country, he learned that there was a mathematical explanation for his sharp observation.
[1]

In fact, this story has its origin some ninety years ago, when the British mathematician and logician Frank Ramsey in the course of his logical research proved a combinatorial theorem that he needed to establish the decidability of a certain fragment of FOL, first-order predicate logic. [2]

[1] This story including names can be found in the Chapter by Alon and Krivelevich (p.562) in [2], the Princeton Companion to Mathematics (eds. T. Gowers et al., 2008).

[2] This concerns EPL, Effectively Propositional Logic, consisting of first-order sentences of the form $\exists^*\forall^*\phi$ where ϕ is quantifierfree and has no function symbols. Satisfiability for this class of so-called Bernays-Schönfinkel-Ramsey formulas is decidable.

Figure 1: *Frank Ramsey (1903-1930)*

0.1 Who was Frank Ramsey?

Frank Plumpton Ramsey (1903-1930) was a mathematician, logician and philosopher who put his mark on seveal areas, in spite of his short life; at the age of 26 he did not survive complications after surgery due to liver problems. By then he had founded a branch of combinatorics, which is now known as Ramsey theory. His main theorem in that area originated as a by-product of his logical investigations. See Figure 2.

Ramsey was the promotor of Ludwig Wittgenstein.

Ramsey's theorem comes in two versions, the finite version and the infinite. We will encounter the latter in another chapter. In fact, the infinite version implies the finite one. The infinite version is a powerful tool in logic. The entrance to the finite theorem of Ramsey is a simple version of the observation of the Hungarian sociologist. It states that in any group of six people there exists either a trio of mutual friends, or a trio of mutual strangers (i.e. not friends).

In the examples and figures of graphs that follow, a red connection will denote friendship, a blue one will denote that this pair consists of strangers to each other. Ramsey's main theorem is as follows.

Theorem 0.1. *For all positive natural numbers* n, m *there is a 'Ramsey number'*

264 F. P. RAMSEY [Dec. 13,

ON A PROBLEM OF FORMAL LOGIC

By F. P. RAMSEY.

[Received 28 November, 1928.—Read 13 December, 1928.]

This paper is primarily concerned with a special case of one of the leading problems of mathematical logic, the problem of finding a regular procedure to determine the truth or falsity of any given logical formula*. But in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest and are most conveniently set out by themselves beforehand.

I.

The theorems which we actually require concern finite classes only, but we shall begin with a similar theorem about infinite classes which is easier to prove and gives a simple example of the method of argument.

THEOREM A. *Let* Γ *be an infinite class, and* μ *and* r *positive integers; and let all those sub-classes of* Γ *which have exactly* r *members, or, as we may say, let all* r*-combinations of the members of* Γ *be divided in any manner into* μ *mutually exclusive classes* C_i $(i = 1, 2, \ldots, \mu)$, *so that every* r*-combination is a member of one and only one* C_i; *then, assuming the axiom of selections,* Γ *must contain an infinite sub-class* Δ *such that all the* r*-combinations of the members of* Δ *belong to the same* C_i.

Consider first the case $\mu = 2$. (If $\mu = 1$ there is nothing to prove.) The theorem is trivial when r is 1, and we prove it for all values of r by induction. Let us assume it, therefore, when $r = \rho - 1$ and deduce it for $r = \rho$, there being, since $\mu = 2$, only two classes C_i, namely C_1 and C_2.

* Called in German the *Entscheidungsproblem*; see Hilbert und Ackermann, *Grundzüge der Theoretischen Logik*, 72–81.

Figure 2: *Ramsey's paper [3] from 1928, containing the infinitary version of his theorem.*

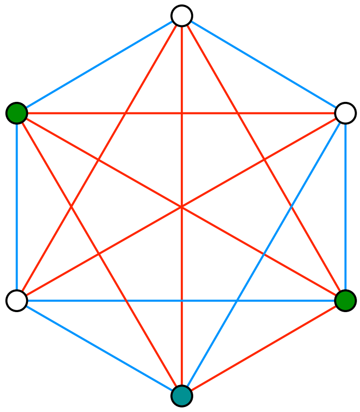


Figure 3: *A party of 6 always contains a trio of mutual friends, or a trio of mutual strangers. Red edges indicate pairs of friends, blue lines connect strangers. The three green nodes indicate the (only) trio of mutual friends.*

$R(n, m)$ *such that*

1. *In a finite gathering of* $R(n, m)$ *people there is a group of* n *mutual friends, or a group of* m *mutual strangers (not friends).*

   Or, in the terminology of graphs: for every bicolored finite graph, where every pair of nodes is connected by a red edge or a blue edge, there is a red 'clique' of size n*, or a blue clique of size* m*. (A clique in a graph is a subset of which all nodes are connected with each other.)*

2. $R(n, m)$ *is the least number with this property.*

Remarkably, we can prove that such numbers $R(n, m)$ exist for all n, m, while on the other hand the actual numbers $R(n, m)$ are unknown very soon when n, m are increased.

We will aim to prove in this chapter that $R(4, 4) = 18$, thus confirming the sociological observation above. [3] We will climb up to this theorem in two steps:

$$R(3,3) = 6 \implies R(3,4) = R(4,3) = 9 \implies R(4,4) = 18$$

$\boxed{R(3, 3 \leq 6}$

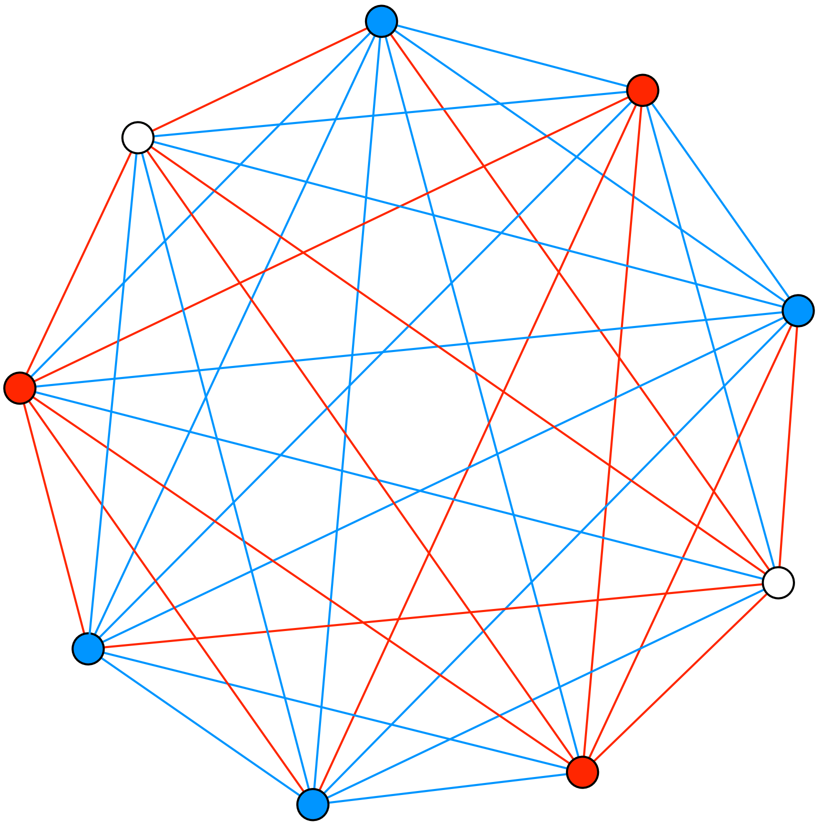
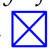


Figure 4: *A party of 9 people will always contain a trio*  *of mutual friends (red), or a quartet*  *of mutual strangers (blue).*

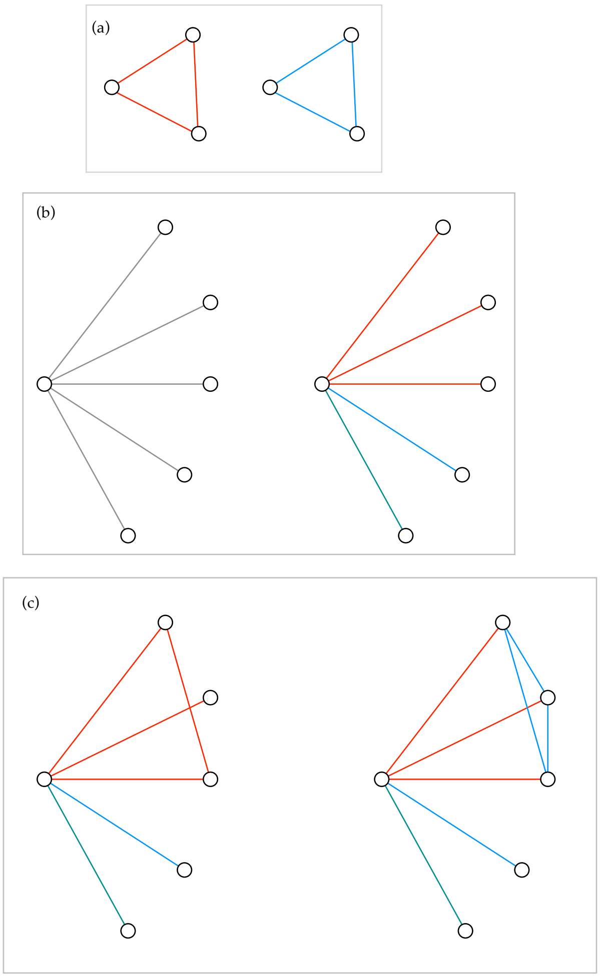


Figure 5: *Proof of* $\mathsf{R}(3,3) \leq 6$

The first step: $R(3,3) = 6$. See Figure 3.

Consider a graph $\mathcal{G}$ with 6 nodes, with its edges colored blue or red in some arbitrary way. We might choose the color by throwing a coin. (It does not need to be a fair coin.) We will consider the colors as being exclusive: an edge has one color, not both. *(Can this restriction be omitted?)*

*Claim.*The graph $\mathcal{G}$ contains a red triangle, that is, three nodes mutually connected by red edges , or a blue triangle .

How can we prove the claim? Trying out all possibilities for $\mathcal{G}$? This is somewhat cumbersome: there are 6 nodes, of which every pair is connected by an edge; so there are $5 + 4 + 3 + 2 + 1 = 15$ edges. These can be colored red or blue in $2^{15} = 32768$ ways. If we would be fast and check one case a minute, it would cost three weeks of exhausting labour.

Luckily there is a far better approach, using the power of logic. Each point p in $\mathcal{G}$ is connected to 5 companion nodes, by a 'fan' of red or blue edges as in Figure 5(b).

This means that there at at least 3 blue edges, *or* at least 3 red edges. For, the negation of this statement is that there are at most two blue, *and* at most two red in these 5. This cannot be the case: $2 + 2 = 4 < 5$.

In the first case p has 3 outgoing red edges. Now the end points of these three edges must form a blue triangle, otherwise there was a red triangle. $R(3,3) \neq 5$ The other case is similar, and also leads to the presence of a red or blue triangle. We are not yet done in establishing that $R(3,3) = 6$; we have to prove the minimality requirement. That is, we have to prove that a bicolored graph on 5 nodes does not necessarily have the property of containing a red triangle or a blue triangle. Such a counterexample indeed exists, see Figure 6. This is by the way the only counterexample for 5 points, but for swapping the colors. We now have proved that $R(3,3) = 6$.

The next step to prove that $R(4,4) = 18$, is to prove the intermediate result that $R(3,4) = R(4,3) = 9$. First we will prove that $R(3,4) \leq 10$, and $R(3,4) \leq 10$ then improve this to $R(3,4) = 9$. So consider a graph on 10 points. We aim to prove that it contains a blue triangle or a red quartet, the blue quartet also such that all 4 points are connected by a blue edge. Again we consider a point p fanning out to its 9 companion nodes. Of these 9 edges there must be at least 6 red, case 2, or at least 4 blue, case 1.This is so because $3 + 5 = 8 < 9$. See Figure 8.

Case 1. At least 4 blue. Figure 8(a). Then either there is a blue triangle, if we connect two end points of this fan with a blue line as in Figure 8(c); or

[3]In this chapter we have profited from the excellent classroom exposition by Imre Leader on Youtube, *'Order in Disorder'*, 2014. https://www.youtube.com/watch?v=AZnvP86N20I

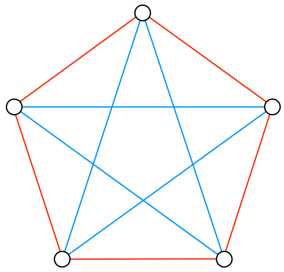


Figure 6: No △ *and no* △

a qed quartet arises as in Figure 8(c).

Case 2. At least 6 red edges in the fan out of p as in Figure 8(b). Now consider the green oval, containing 6 points. This makes it possible to apply the earlier result that $R(3,3) = 6$.

The counterexample showing that $R(3,4) \neq 8$ is shown in Figure 7, where the red edges are in the upper figure and the blue ones are below. Now it is clear that there is no red triangle. But also no blue quartet. This is easily seen by looking at one of the eight blue triangles in the lower figure, e.g. the one with heavy lines. Any additional point on the 'circle' extending those three would have distance (the number of chords on the circle separating them) 1 or 4 to one of the nodes of the heavy triangle. But that conflicts with the red coloring above, where the connection distances are just that, 1 or 4. So we have proved $R(3,4) = 9$. And of course, also $R(4,3) = 9$, by changing the colors.

$R(3,4) \neq 8$

Surprisingly, we can sharpen our result that $R(3,4) \leq 10$ to $R(3,4) \leq 9$. Consider a bicolored graph $\mathcal{G}$ with 9 points. Now if there is at least one point whose fan-out to the other 8 points splits into ≥ 4 blue or ≥ 6 red edges, we can apply the same reasoning as above. But what if *all* points of $\mathcal{G}$ split into < 4 blue and < 6 edges? Let us call such a point a 'bad' point.

Fortunately, a simple but amusing counting argument shows that this unfortunate situation cannot happen:

Claim. In a bicolored graph on 9 points it is not possible that all points are bad.

Proof of the Claim. Suppose all points in $\mathcal{G}$ are bad, i.e. they fan out to ≤ 3 blue and ≤ 5 red edges. Then they fan out to 3 blue and 5 red edges, because the sum must be 8. Neither of these conjuncts can be true. For, if from each of the 9 points there emanate 5 red edges, then there are 45 red 'half edges', or 'edge endings'. Each edge has two ends, so these 45 red edge endings arise from $45/2 = 22.5$ red edges. But half edges do not exist, so we have proved the claim by contradiction.

We will now embark on the proof that $R(4,4) = 18$. First we will show that $R(4,4) \leq 18$. So let $\mathcal{G}$ be a graph with 18 points. Let p be one of its points. Then p fans out to its 17 companion points. Of these 17 edges there are 9 red, or 9 blue. Less in both cases is not possible, as $8 + 8 = 16 < 17$. Say that p fans out to 9 points via red edges, as in Figure 9.

$\boxed{R(3,4) \leq 9}$

$\boxed{R(4,4) \leq 18}$

Then, by the earlier result that $R(3,4) = 9$, we find either a blue quartet, or a red triangle spanned by the endpoints of the 9 edges. In the first case we are done, see Figure 9(a), in the other case also, see Figure 9(b), then we have together with point p a red quartet. So we have proved $R(4,4) \leq 18$.

Finally we have to prove that $R(4,4) = 18$. Analogous to the situation above in three instances, we search for a counterexample graph on 17 points containing no red quartet and no blue quartet. While for the case of $R(3,3) = 6$ the search space was a mere 32.000 something, the search space is now much more prohibitive: a graph on 17 points has 136 edges, so that the number of possible red-blue colorings is 2^{136}. Yet, several of such counterexamples have been found. An example occurs as Exercise 12.3 (5), p.520 of [1]; it is included in Figure 10. Figure 11 explains why there is no monochrome quartet in this bicolored graph: a quartet contains two triangles sharing one side; it is easy to see that no combination of two triangles, that have the shape of the two heavy lined triangles, yield a clique of 4 points, neither for red and for blue.

Note that the red edges connect nodes at chord distance $1, 2, 5, 8$, while the blue edges connect points at chord distance $3, 4, 6, 7$. [4]

$\boxed{R(4,4) \neq 17}$

Theorem 0.2. *1.* $R(n,m) \leq R(n-1,m) + R(n,m-1)$

2. $R(n,m) \leq R(n-1,m) + R(n,m-1) - 1$ *if* $R(n-1,m)$ *and* $R(n,m-1)$ *are even.*

3. $R(2,m) = m$; $R(n,2) = n$

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

5. $R(n,n) < 4^n$

So far, we could reason without writing down any formula. But to give the formal proof that the Ramsey numbers $R(n,m)$ always exist, we obtain

[4]A similar but slightly different counterxample can be found at https://www.cut-the-knot.org/arithmetic/combinatorics/Ramsey44.shtml, there with chord distances $1, 2, 4, 8$ (red) and $3, 4, 7, 11$ (blue).

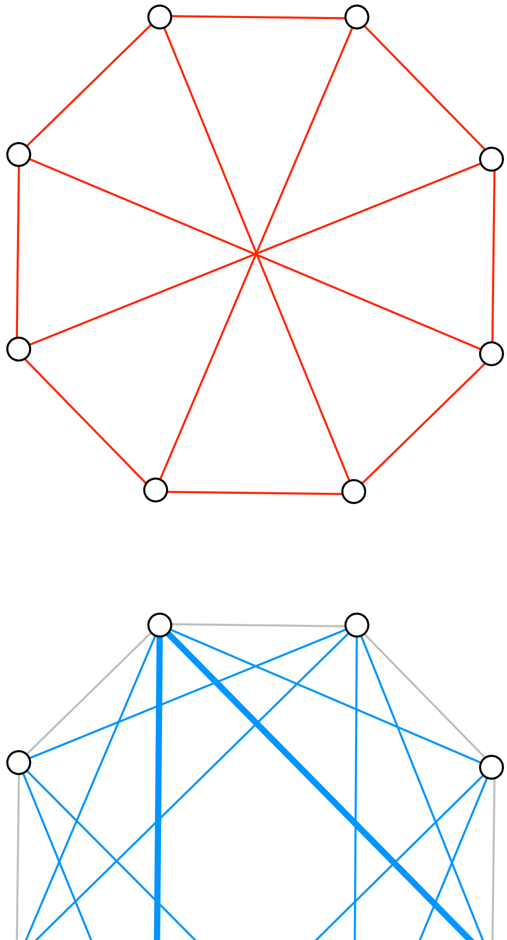


Figure 7: *Bi-colored graph on 8 points without red triangle of nodes and without blue quartet of nodes.*

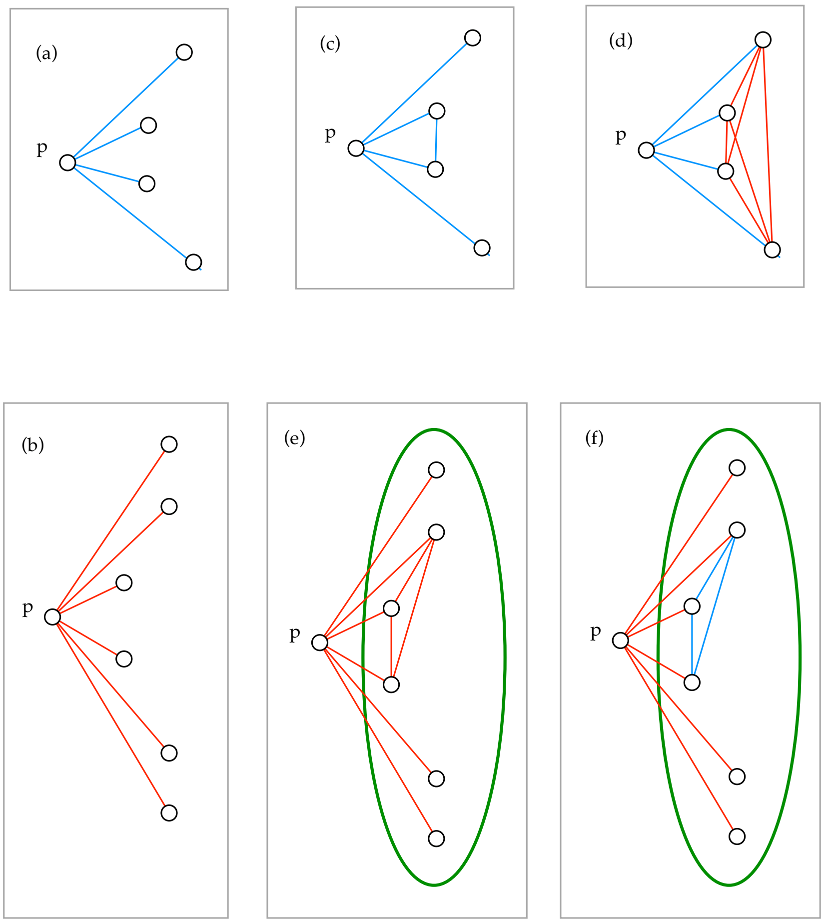


Figure 8: *Reasoning to prove that* $R(3, 4) \leq 10$

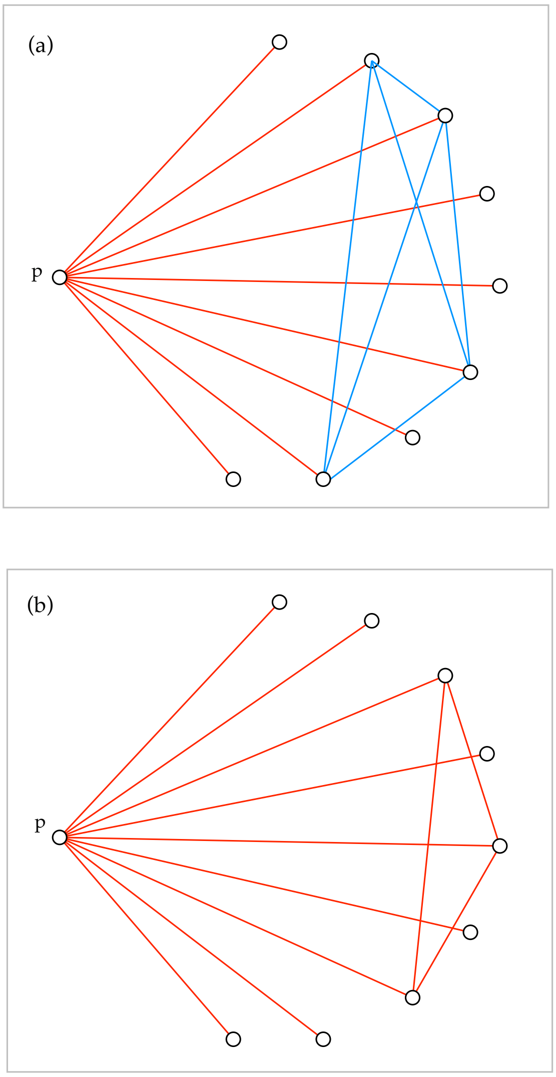


Figure 9: *Proving* $\mathrm{R}(4,4) \leq 18$*. The end points of a fan-out to 9 points span up a blue quartet or a red triangle.*

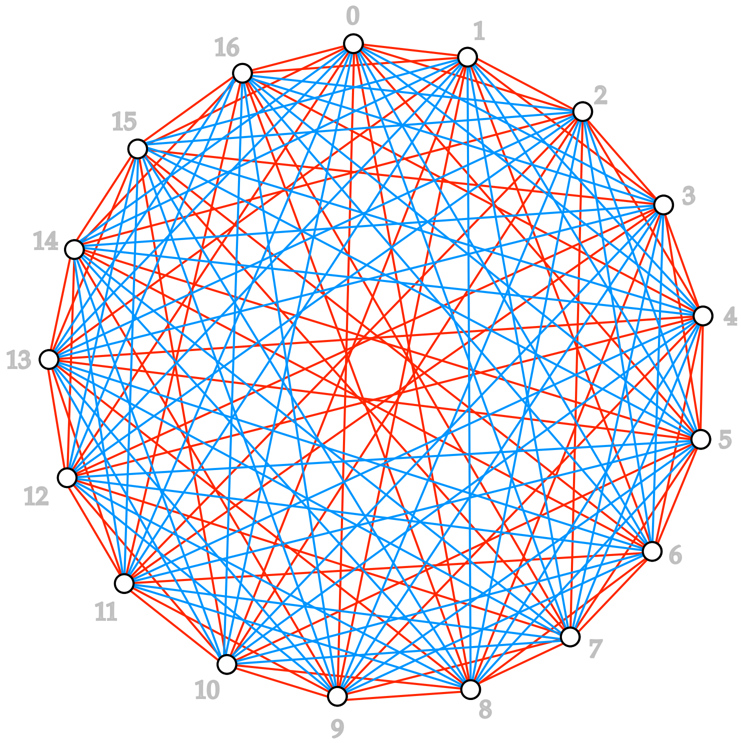


Figure 10: *No red quartet*  *of friends, no blue quartet*  *of strangers.*

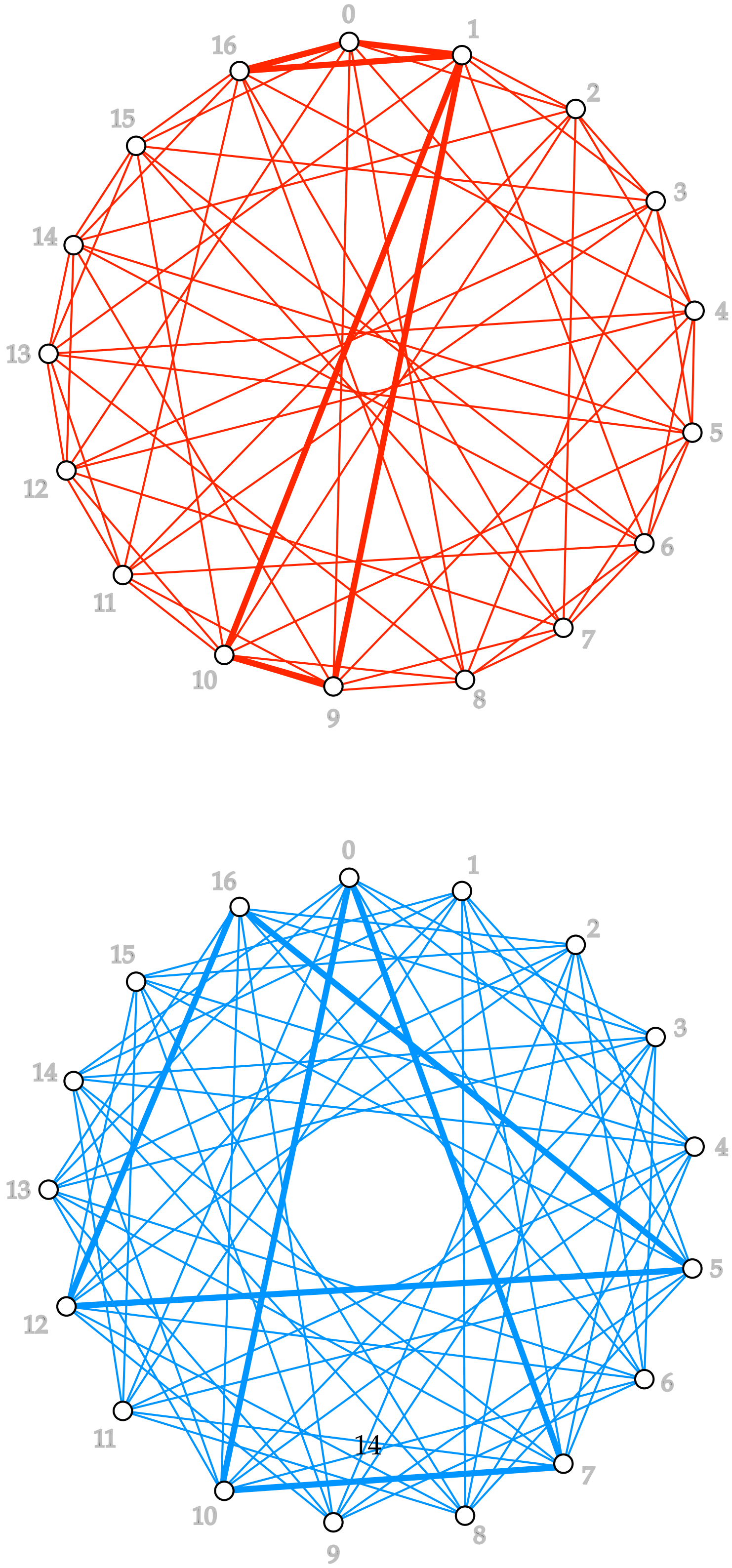


Figure 11: *No red quartet* ⊠ *of friends, no blue quartet* ⊠ *of strangers; sketch of proof*

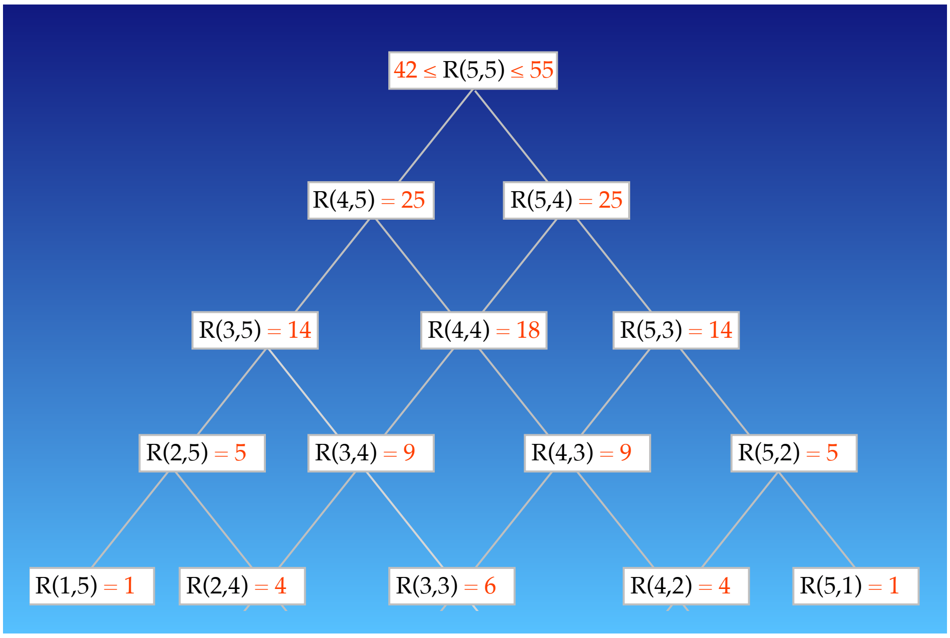


Figure 12: *Some initial Ramsey numbers. The precise value of* R(5.5)*, known to be between* 43 *and* 48 *is an open question. Determination of* R(6, 6)*, which is between* 102 *and* 165*, is considered to be totally impossible, like distant stars that never can be reached.*

by the same reasoning as used above for some actual n, m, and employing induction, the following facts, of which item 3 is trivial.

So what makes the actual determination of $R(n, m)$ so difficult, even though this theorem yields relatively small intervals for the $R(n, m)$ to be found? The problem lies in finding the various counterexamples as we did above in Figures 6, 7, 10.

E.g. checking all bicolored graphs on 50 points, we would have to deal with roughly 10^{400} cases, which dwarfs even the number of atoms in the universe, which is in the order of 10^{80}.

References

[1] Gary Chartrand, Linda Lesniak, and Ping Zhang. *Graphs & digraphs*. Chapman and Hall/CRC, 2010.

[2] Timothy Gowers, June Barrow-Green, and Imre Leader. *The Princeton companion to mathematics*. Princeton University Press, 2010.

[3] Frank P Ramsey. On a problem of formal logic. In *Classic Papers in Combinatorics*, pages 1–24. Springer, 2009.