Report on Progress at Clemson REU

Jay Heumann and Kelty Allen

Abstract

We divide our attention between two open problems. One of them is to find better lower bounds on Ramsey numbers. Our goal for this problem is to find graphs on n vertices that prove that $R(k, l) > n$. The other problem is to find a 4-saturated triangle-free graph. This is a very strong condition for a graph to fulfill, and no such graph has yet been found. We are still in the process of researching both these questions. In this paper we will provide some of the context for our work.

1 Lower Bounds for Unknown Ramsey Numbers

The *complete graph* K_n is the graph on n vertices with every edge between every pair of vertices drawn. Suppose that we take the edges in the complete graph and color them red or blue. This is called a red-blue coloring of K_n. Then the *Ramsey number* $R(k, l)$ is the smallest value of n so that every red-blue coloring of K_n contains a red K_k or a blue K_l. A trivial example of this is that $R(2, k) = k$, since any red edge is a red K_2.

This problem can be restated in terms of random graphs as opposed to colorings of the complete graph. A *random graph* is a graph on n vertices such that the probability that an edge exists between any two vertices is p, with p ranging between 0 and 1. Usually we take $p = 1/2$. Unless we state otherwise, we will be assuming throughout this paper that $p = 1/2$. We can thus restate the definition of the Ramsey number $R(k, l)$ as the smallest n such that every graph on n vertices contains either a K_k or an E_l, where E_l is the *independent set* of size l, or a set of l vertices such that no edge between any two exists.

The vast majority of Ramsey numbers is still unknown. In fact, only three values of $R(k, k)$ are known—$R(2, 2)$ is 2 (but this is trivial), $R(3, 3) = 6$, and $R(4, 4) = 18$. There are some other Ramsey numbers that are known, and there are many others that, while not known, are known to be bounded above and below. Current research on Ramsey numbers focuses partly on improving these bounds.

The simplest way to improve a bound for $R(k, l)$ (perhaps deceptively simple) is to find a graph on n vertices without a K_k or an E_l, where n is between the known bounds. This gives a new lower bound for $R(k, l)$. Unfortunately, this does not help to improve any upper bounds, and if such a graph cannot be found, the result is usually inconclusive.

A *cyclic graph* is a graph with vertex set $\{0, 1, 2, \ldots, n-1\}$ which is invariant under the transformation $i \to (i+1) \pmod{n}$. This is equivalent to saying that every edge chosen of a given distance forces every other edge of that distance to exist. For example, any complete graph is cyclic. In a cyclic graph, there are only $N/2$ independent edges that can be chosen; the existence of all other edges is dependent on them. Many of the best known examples of lower bounds are on cyclic graphs.

There is probabilistic and experimental evidence [2] that cyclic graphs on prime numbers of vertices are more likely to yield a Ramsey Graph than cyclic graphs on composite numbers of vertices. Thus, in our research, we focused more on cyclic graphs on prime numbers of vertices than composite.

Another approach to improving bounds is by using probabilistic methods. Here is one simple method: if we have a random graph on n vertices, and we know the probability that there exists a K_k and the probability that there exists an E_l, and we know that their sum is less than 1, then the probability that a graph has no K_k or E_l is greater than 0, and hence there exists such a graph on n vertices. This condition gives rise to another condition: instead of adding the probabilities, we can add the expected values. If that sum is less than 1, then we arrive at the same conclusion, since the expected values are always greater than or equal to the probabilities. These methods are sometimes worth examining, but we tended not to use them in our research after discovering what was involved in expected value calculations.

1.1 Expected Value of k-tuples

We set out to find the expected number of k-tuples in a random cyclic graph on a prime number of vertices (still call it n). We call a k-tuple *independent* if no edges in the k-tuple are forced to exist due to the existence of another edge and the cyclic nature of the graph. If a k-tuple is not independent, we say it is *t-dependent*, where t is the number of edges that are not independently chosen but are forced by other edges in the k-tuple.

We can apply these terms to vertices as well: suppose that we have chosen some subset of vertices less than k that we want to appear in our k-tuple. Then a vertex Q is *independent* if none of the edges between the vertices we have already chosen forces an edge from Q to one of these vertices. If this condition is not met, then Q is called *t-dependent*, where t is the number of forced edges.

A simple example of this is the expected value calculation for 3-tuples. The general idea is to determine the expected value by counting the number of independent k-tuples and t-dependent k-tuples for all possible values of t, multiplying each by the probability that the k-tuple exists on the vertices we've chosen.

On a graph with n vertices, we can label the vertices 0, 1, ..., $n-1$. Because we are looking at cyclic graphs, we can always start with vertex 0—were we to choose a different one, we could apply a cyclic automorphism until our starting point was in fact vertex 0. Next, there are always $n-1$ choices for the second vertex, but since we are looking on prime numbers of vertices, we can assume without loss of generality that this vertex is 1, by an isomorphism of graphs. (We will still multiply our final answer by $n-1$, though.) Now, with vertices 0 and 1 chosen, we claim that there are three 1-dependent vertices. The edge between 0 and 1 forces an edge between 1 and 2; hence vertex 2 is 1-dependent. For the same reason, vertex $n-1$ is also 1-dependent. Furthermore, the vertex $\frac{n+1}{2}$ is 1-dependent because if we were to choose this vertex, the edge from 0 to $\frac{n+1}{2}$ would force the edge from $\frac{n+1}{2}$ to 1. Therefore, if we were looking for an independent 3-tuple, there would be $n-5$ choices for the third vertex. The probability of this 3-tuple actually existing would be $(1/2)^3$.

Should we have wanted a 1-dependent 3-tuple, there would be 3 choices for the third vertex, and the probability that this exists is $(1/2)^2$. After we combine these terms, we must remember to multiply by $n-1$, and also to divide by 3!. The reason for this division is that we could have chosen the two non-zero vertices in either order, and we could move any vertex to 0 by cyclic automorphism, leading us to divide by $(2!)(3) = 3!$. Putting it all together, the formula is

$$\left(\frac{1}{3!}\right)(n-1)\left[\left(\frac{1}{2}\right)^3(n-5)+\left(\frac{1}{2}\right)^2(3)\right]$$

which simplfies to

$$\frac{(n-1)(n+1)}{48}$$

This method gets much more complicated for 4-tuples. The fewest number of distinct edges that need to be drawn is 3, so there will be 4 terms in the corresponding sum: one for independent 4-tuples, one for 1-dependent 4-tuples, one for 2-dependent 4-tuples, and one for 3-dependent 4-tuples.

Throughout this derivation we will assume $n \geq 17$.

Suppose we pick 0 as the first vertex and 1 as the second vertex in our 4-tuple (as above, we can do this without loss of generality). Already, we know that the vertices 2, $n-1$, and $(n+1)/2$ are 1-dependent vertices. However, unlike in the 3-tuple case, there are now also different cases for how we can pick the third vertex.

First, let's look for all the ways to choose an independent 4-tuple. Our third vertex cannot be a 1-dependent vertex. It turns out that are actually three different cases of independent vertices that produce different results; we will have to look at them all and add them up when we are finished. The first case (call this Case 1 for future reference) produces a total of one 3-dependent vertex, three 2-dependent vertices, and five 1-dependent vertices, leaving $n - 12$ independent vertices remaining to be chosen. It turns out that there are always six vertices that fall into Case 1: four vertices adjacent to the 1-dependent vertices and two vertices that have a forced edge connecting to a halfway point[1]. The second case (call this Case 2) gives a total of no 3-dependent vertices, four 2-dependent vertices, and seven 1-dependent vertices, leaving $n - 14$ independent vertices. There are also six vertices in Case 2. The final case (Case 3) leaves a total of no 3-dependent vertices, three 2-dependent vertices, nine 1-dependent vertices, and hence $n - 15$ independent vertices. Case 3 encompasses all vertices that are not 1-dependent or fall into the other two cases, so there are $n - 17$ of them. Because the probability of 6 distinct edges existing is $(1/2)^6$, the independent term of the formula is

$$(\frac{1}{2})^6((n-1)*(6)*(n-12)+(n-1)*6*(n-14)+(n-1)*(n-17)*(n-15))$$

For the t-dependent terms, we need only count the t-dependent vertices of each case. The only case we have not yet enumerated is the case in which we select a 1-dependent vertex as our third vertex. We already know there are 3 of these, and they all yield a total of two 2-dependent vertices, four 1-dependent vertices, and $n-9$ independent vertices (but remember that choosing a t-dependent vertex in this case means the k-tuple is $(t+1)$-dependent). So this means, because the probability that 5 distinct edges exist is $(1/2)^5$, that the 1-dependent term is

$$(\frac{1}{2})^5((n-1)*3*(n-9)+(n-1)*6*5+(n-1)*6*7+(n-1)*(n-17)*9)$$

Similarly, the probability that 4 distinct edges exist is $(1/2)^4$, so the 2-dependent term is

$$(\frac{1}{2})^4((n-1)*3*4+(n-1)*6*3+(n-1)*6*4+(n-1)*(n-17)*3)$$

Finally, the probability that 3 distinct edges exist is $(1/2)^3$, so the 3-dependent term is

$$(\frac{1}{2}^3((N-1)*3*2+(N-1)*6)$$

[1] A "halfway point" is our name for a vertex that, given two vertices a and b, is equal to $(a+b)/2$ (mod n). The point $(n+1)/2$, for example, is the halfway point between 0 and 1. Halfway points are usually 1-dependent, though in these special cases they can be 2-dependent or 3-dependent.

Now, as above, the whole sum must be divided by 4! to compensate for the number of ways we could pick the same 4-tuple. This means that our final result is

$$(\frac{1}{24})((\frac{1}{2})^6((n-1)*(6)*(n-12)+(n-1)*6*(n-14)+(n-1)*(n-17)*(n-15))$$
$$+(\frac{1}{2})^5((n-1)*3*(n-9)+(n-1)*6*5+(n-1)*6*7+(n-1)*(n-17)*9)$$
$$+(\frac{1}{2})^4((n-1)*3*4+(n-1)*6*3+(n-1)*6*4+(n-1)*(n-17)*3)$$
$$+(\frac{1}{2})^3((n-1)*3*2+(n-1)*6))$$

which simplifies to

$$(\frac{1}{1536})(n-1)(n^2+16n-9)$$

These calculations are so complicated for 5-tuples that it is not really worth doing. It is, however, worth noting that the formulas are always in the form of a constant multiplied by $n-1$ multiplied by a monic polynomial of degree $k-2$. Perhaps these facts can be used to expedite the process of finding these values.

However, more important than this is the fact that these values are almost always greater than 1, and do not improve on any lower bounds that have been found using the "brute force" method (searching for actual graphs). Therefore, we quickly[2] concluded that this avenue was probably a dead end. We have since been focusing our efforts on finding explicit Ramsey graphs.

2 Saturated Graphs

Let G be a graph on n vertices, with vertex set $V = \{0, 1, \ldots, n-1\}$. Then for any natural number k, $k > 2$, we say that G is *k-saturated* iff for any subset S .of V with size k, and any subset of S, called T, there exists some vertex in $G \backslash S$ which is connected to every point in T but not to any point in $S \backslash T$. This vertex is called a *witness* of (S,T).

[2]When we say "quickly," we mean "quickly after discovering the formula," not quickly in real time.

2.1 Infinitely Many Vertices

In this section we will explain the following:

Theorem 1 *The random graph on countably many vertices is k-saturated for all values of k. Furthermore, every graph on countably many vertices that is k-saturated for all k is isomorphic to this graph.*

First we will show that the random graph on countably many vertices is k-saturated for all k. Take an arbitrary subset of vertices and call it S, then take an arbitrary subset of S and call it T. Now for any given vertex outside of S, the probability that it is a witness for the pair (S, T) is $2^{-|S|}$. This means that the probability that it is not a witness is $(1 - 2^{-|S|})$. Now, the probability that n vertices in this graph are all not witnesses is $(1 - 2^{-|S|})^n$. However, since this graph is on countably many vertices, the probability that there is no witness for (S, T) is the limit of this expression as $n \to \infty$, which is of course 0. Therefore, every pair (S, T)—of any size—has a witness, making this graph k-saturated for all k. It is interesting to note that the same result holds if we change the probability of an edge existing, so long as it is not 0 or 1. This is because any number between 0 and 1, raised to a large power, will approach 0 as the exponent approaches infinity.

Next we must show that any two graphs G_1 and G_2 on countably many vertices that are k-saturated for all k are isomorphic. To do this, we start by giving an ordering to the vertices of each graph with the goal of constructing an isomorphism between the two graphs. Suppose we begin at vertex 1 of G_1. Now start by looking at whether vertex 2 is connected to vertex 1. Whatever the relationship is, we know, by the k-saturation, that on G_2, there is a vertex that has the same relationship to vertex 1 of G_2. Once we find this vertex, we pick the next vertex in the ordering of G_2. Whatever relationship this has to the first two vertices, we know that there is a vertex in G_1 with this same relationship. Then we pick the next vertex in the ordering of G_1 and repeat this process. Because the graphs are k-saturated for all k, this process will never fail to produce a desired vertex at any point, and this leads to an isomorphism of graphs.

Hence, there is one unique graph on countably many vertices—which can be given by a random graph—that is k-saturated for all k.

2.2 Finitely Many Vertices

Suppose that we fix k and ask how many vertices we need before we find a k-saturated graph G. We can use probabilistic methods to get an estimate on how big G should be. Start by fixing S, a subset of vertices of G with size k, and T, a subset of vertices of S. If G is a random graph, then the probability that any vertex v will be adjacent to the vertices in T but not in $S\backslash T$ is 2^{-k}. Since there are $n-k$ vertices in $G\backslash S$, the probability that there is *no* witness for a given (S,T) is $(1-2^{-k})^{n-k}$.

Once we know this, we want to bound the probability that there exists at least one pair (S,T) with no witness. Since there are $\binom{n}{k}$ ways to choose S, and 2^k ways of choosing T from within S, there must be $\binom{n}{k} 2^k$ ways to choose (S,T), and this means that

$$\Pr(\text{There exists an (S,T) pair with no witness}) \leq \binom{n}{k} 2^k(1-2^{-k})^{n-k}$$

This means that if $\binom{n}{k} 2^k(1-2^{-k})^{n-k} < 1$, then the probability that G is k-saturated is positive, and hence a k-saturated graph exists on n vertices. So to find n so that a k-saturated graph exists on n vertices, we need only choose n so that $\binom{n}{k} 2^k(1-2^{-k})^{n-k} < 1$. The only question that remains is, does this always happen for large enough n?

It turns out that it does. Since k is fixed, $1-2^{-k}$ is constant, as is 2^k. Because $n^k \geq \binom{n}{k}$, and $(1-2^{-k})^n \geq (1-2^{-k})^{n-k}$, we need only show that $n^k(1-2^{-k})^n \to 0$ as $n \to \infty$. But in fact this holds true for any expression $n^a b^n$ so long as $0 \leq b < 1$. Therefore, k-saturated graphs exist for large enough numbers of vertices.

A graph G with vertex set V is *k-saturated and triangle-free* iff the following two conditions are met:

1. No three points in V are all connected to each other; and

2. For any subset S of V with size k, and any independent set T contained in S, there exists a vertex in $G\backslash S$ that is connected to every point in T and no points in $S\backslash T$.

The difference here is that before, T could be any subset of S. Now T has to be an independent set. This seems like a weaker condition, but the triangle-free condition makes things more complicated. How much more complicated? It is not yet known whether any k-saturated graphs exist for the triangle-free case when $k > 3$. It is also unknown if there are any 3-saturated graphs of degree two or greater. Here, by *degree* we mean the minimum number of witnesses for every pair (S,T) on the entire graph. Our current goal in this area is to find a 4-saturated triangle-free graph, or a triangle free graph which is 3-saturated with degree 2, searching on cyclic graphs.

We have searched all cyclic graphs with 134 or fewer vertices, and there are no 4-saturated triangle-free cyclic graphs. Research is still being done in this area.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons Inc. New York, 1992.

[2] N. Calkin, P. Erdos and C. Tovey, *Improved Bounds on Ramsey Numbers from Cycle Graphs of Prime Order*, SIAM Discrete Math. Vol. 10, #3, 1997.

[3] G. Cherlin, A Problem on Triangle-Free Graphs,

`http://www.math.rutgers.edu/~cherlin/Problems/tfree.html`

Current as of July 12, 2004.