Symbolic Moment Calculus II.: Why is Ramsey Theory Sooooo Eeeenormously Hard?

Doron ZEILBERGER [1]

Dedicated to Ron Graham, on his $(106)_8$*-th Birthday*

Abstract. The short answer to the question in the title is that Ron Graham, one of the leaders of Ramsey theory, co-author of the definitive book (with Rothschild and Spencer) on the subject and co-prover of one of its Super-Six theorems (with Leeb and Rothschild), would not choose to work on an easy subject. A longer answer, from my enumerator's perspective, is that Ramsey theory, that according to Motzkin, proves that complete disorder is impossible, is equivalent to proving that for sufficiently large universes we are guaranteed islands of order. More precisely, if X is the random variable, "number of orderly islands", we have to find (or bound) the number of universes with X=0. If we knew all the moments of X, we would be done. Already the first moment, the expectation E[X], gives us some information (as was first observed by Erdős). The second moment is harder, but still tractable, even for humans. But for the third and fourth moments we need computers. Beyond that, even computers seem to get stumped.

Grade Inflation

Nothing is like it used to be. In the old days, an A was an A, and a C was a C. Nowadays an A is a C and a C is an F.

But not all inflation is bad. In the old days, seventy was *really old.* In fact Rabbi Judah son of Tema, in *pirke avot*, stated that *sixty* was already old, and *seventy* was *ancient* ("hoary head"). But if you look at Ron Graham, you see the epitome of youth. So, with all due respect to Rabbi Juda ben Tema, let me make a new, time-dependent, definition of **old**.

Definition: A person is old if his or her age is at least ten years older than the current age of Ron Graham.

Ramsey Theory

Speaking of Ron Graham, Ron is one of the leaders of this fascinating but extremely difficult subject (see abstract). Let's take the *original* Ramsey theorem, Ramsey's theorem.

For every k there exists an n such that any party with n guests is guaranteed either to have k people who all mutually love each other or k people who all mutually hate each other.

---

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg at math dot rutgers dot edu` , `http://www.math.rutgers.edu/~zeilberg/`.
First version: Nov. 24, 2005. Edited version of the transcript of a plenary talk given (Carrollton, GA, USA, 10:30-11:20, Oct. 28, 2005) at the INTEGERS 2005 conference in honor of Ron Graham's 70th Birthday, Oct. 27-30, organized by Bruce Landman. Accompanied by Maple package `SMCramsey` available from `http://www.math.rutgers.edu/~zeilberg/tokhniot/SMCramsey` . Supported in part by the NSF.

As we all know, the classical proof (that is not that hard) tells us that a party with $\binom{2k}{k}$ people would do the job, hence $R(k,k)$, the *smallest* party size that would guarantee a clique or anti-clique of size k is $\leq \binom{2k}{k}$. But what is it actually equal to? Well, $R(3,3)=6$, $R(4,4)=18$, but all we know about $R(5,5)$ is that it is between 43 and 49. As for $R(6,6)$, forget it! Erdős's proverbial aliens don't themselves know the answer, and they just used this question as an excuse to destroy us. Even the humble fact that $R(4,5)=25$ (due to Brendan McKay and Stanislaw Radziszkowski and their 1995 computer) took a year or so to figure out.

A pioneering result in explicit Ramsey numbers is due to the Birthday wife, Fan Chung (Graham), who in her doctoral thesis proved that $R(4,4,4,4)>50$ (*Discrete Math* **5** (1973), 317-321). The exact value is still unknown.

From a traditional mathematician's point of view a result such as $R(4,5)=25$ is *trivial*, since it is *routinely provable in finite time.* All you have to do is check all the 2^{300} possible parties, and check that each of them either has four people who mutually love each other or five people who mutually hate each other. This takes *only*

$$2^{300}\left(\binom{5}{2}\cdot\binom{25}{5}+\binom{4}{2}\cdot\binom{25}{4}\right)=(1.08\ldots)\cdot 10^{96}$$

operations. So we would have to wait many big bangs to do it by brute force.

It is also notoriously difficult to even find the *asymptotics.* As we all know, and Ron will remind us in *his* talk, it is not even known whether $\lim_{k\to\infty} R(k,k)^{1/k}$ exists, and if it does, all we know is that its value is $\geq\sqrt{2}$ and ≤ 4.

I am willing to bet that the Riemann Hypothesis will be proved by the time of Ron Graham's 100th birthday, and I am also willing to bet (a much larger amount!), that the exact value of $\lim_{k\to\infty} R(k,k)^{1/k}$, and the exact value of $R(6,6)$, would be still unknown.

An Enumerator's Perspective

Consider all $2^{\binom{n}{2}}$ parties of n people. Let $a(n,k)$ be the number of *good* parties, i.e., parties where you don't have any k-clique and any k-anticlique. Then $R(k,k)$ is *nothing but* the smallest n for which $a(n,k)=0$. So let's find a 'formula' for $a(n,k)$ using that venerable workhorse of enumeration called the *Principle of Inclusion-Exclusion* (PIE).

One way to formulate it is as follows. Let $\mathcal{U}$ be a set of elements and $\mathcal{P}$ be a set of properties. For each $u\in\mathcal{U}$ and $p\in\mathcal{P}$, u either does or does not enjoy property p. Let $X:\mathcal{U}\to\mathcal{Z}_{\geq 0}$ be the random variable defined by $X(u):=$ the number of properties in $\mathcal{P}$ enjoyed by u. Then PIE can be phrased as follows:

$$P(X=0)=\sum_{s\geq 0}(-1)^s E[\binom{X}{s}] \quad .$$

So if we can compute $E[\binom{X}{s}]$, for *all* s, or equivalently, $E[X^s]$ (for all s), we would be done!

Applying this to the set of all r-edge-colorings of K_n, with $X(C)$ the number of monochromatic K_k, we should be able to find the number of colorings with no monochromatic K_k. The beginning is encouraging. $E[X^0] = 1$, of course, and $E[X] = \binom{n}{k}/r^{\binom{k}{2}-1}$. Even $E[X^2]$ is not too hard (try it!). But for $E[X^3]$, we need a computer!.

So computing $P(X = 0)$ seems out of reach. But what about lower bounds? If we can show that $P(X = 0) > 0$, we know that they are good guys, and we immediately get a *lower bound* for the Ramsey number. Indeed Bonferroni's theorem tells us that, for any *odd* t,

$$P(X = 0) \geq \sum_{s=0}^{t} (-1)^s E[\binom{X}{s}] \quad ,$$

and for any even t,

$$P(X = 0) \leq \sum_{s=0}^{t} (-1)^s E[\binom{X}{s}] \quad .$$

The original *probabilistic method* (introduced by Erdős in 1947) consists of using $t = 1$. It already gives us some useful information: the lower bound $R(k, k) > 2^{k/2}$. I was hoping that using $t = 3$ would considerably improve this (asymptotic) lower bound, but disappointingly, it does not seem to improve it at all. It seems that we need more sophisticated *sieves*, in the style of the Brun and linear sieves in number theory, or a refinement of the Lovász Local Lemma.

So let's forget about the initial motivation. It is still interesting to have an explicit *closed form* formula, for *symbolic* n and r and *numeric* m and k for the m-th moment of the random variable "number of monochromatic K_k's in an r-coloring of the edges of K_n".

The Maple package SMCramsey

This article is accompanied by the Maple package `SMCramsey` available from

`http://www.math.rutgers.edu/~zeilberg/tokhniot/SMCramsey` .

Alas, already for the fourth moment it takes a very long time. I am sure that more clever programming would yield an explict expression for the fourth moment, but the fifth moment really seems out of reach.

More interesting than the output is the *method* of completely automatic **symbol-crunching**. Let's describe it briefly.

Write

$$X = \sum_{S} X_S \quad ,$$

where the sum is over all k-subsets of $\{1, \ldots, n\}$, and X_S is the *indicator* random variable that is 1 if the subgraph of K_n induced by S is monochromatic. Of course $E[X_S] = 1/r^{\binom{k}{2}-1}$. It follows by *linearity of expectation* and *symmetry* that

$$E[X] = \frac{\binom{n}{k}}{r^{\binom{k}{2}-1}} \quad .$$

Now

$$E[X^2] = E[\left(\sum_{S_1} X_{S_1}\right)\left(\sum_{S_2} X_{S_2}\right)] = \sum_{[S_1,S_2]} E[X_{S_1}X_{S_2}] \quad ,$$

where the sum is over all *ordered pairs* of k-subsets of $\{1,\dots,n\}$. Writing $S = S_1 \cup S_2$, we get that $k \leq |S| \leq 2k$. Of course for a given size of S, say, K, all are 'isomorphic', and there are $\binom{n}{K}$ of them. Without loss of generality we can label the vertices of $S_1 \cup S_2$ by $1, 2, \dots, K$. We have to see how to pick S_1 and S_2 such that $S = S_1 \cup S_2$ and for each of these, figure-out $E[X_{S_1}X_{S_2}]$. If S_1 and S_2 are disjoint or only share one vertex ($K = 2k$, $K = 2k-1$ respectively) then $E[X_{S_1}X_{S_2}] = (1/r^{\binom{k}{2}-1})^2$. If $e := |S_1 \cap S_2| \geq 2$ then all the edges must have the same color and then $E[X_{S_1}X_{S_2}] = 1/r^{2\binom{k}{2}-\binom{e}{2}-1}$. Adding up all the cases would give us the second moment.

Let's now consider the general case, the m-th moment. Now we get m-tuples $[S_1,\dots,S_m]$ of k-subsets of $\{1,\dots,n\}$ and, denoting $S := S_1 \cup \ldots \cup S_m$, we have $k \leq |S| \leq mk$. What is the contribution of a single $E[X_{S_1}X_{S_2}\cdots X_{S_m}]$ to the sum? Let $v := |S|$ and e the total number of edges in the union of the sets of edges of the complete graphs on S_i ($i = 1.\dots\dots m$). Also form a 'meta-graph' whose vertices are the S_i's and there is a meta-edge between S_i and S_j if $|S_i \cap S_j| > 1$. Let c be the number of connected components of this meta-graph. We have

$$E[X_{S_1}X_{S_2}\cdots X_{S_m}] = r^{c-e} \quad .$$

But each and every m-tuple $[S_1,\dots,S_m]$ is isomorphic to such an m-tuple in which the smallest label in $S := S_1 \cup \ldots \cup S_m$ is replaced by 1, the second smallest by 2, etc. For such a reduced $[S_1,\dots,S_m]$, define the *weight* to be

$$\binom{n}{v} r^{c-e} \quad .$$

Note that for *numeric* m and k, there are only *finitely many* such labelled m-tuples of k-sets $[S_1,\dots,S_m]$, whose union is $\{1,\dots,v\}$, for some v between k and mk. So a *naive* approach to finding the expression, in n and r, is to have the computer construct this finite set of objects, compute the weight for each of them, and then add up these *finitely many* expressions, getting an explicit (symbolic) algebraric expression in n and r.

Unfortunately, very soon, the 'finitely many' gets to be too many even for the fastest and largest computers, and we have yet another indication op why Ramsey Theory is **sooo** hard.

The above 'algorithm' is really *naive* weighted counting. One literally constucts the set itself, and adds-up the weights of its elements.

It is analagous (in ordinary counting) to computing 20! in `Maple` by doing `nops(permute(20));`. Of course, a much better way, due to Rabbi Levi Ben Gerson, is to establish, and then use, the recurrence $n! = n(n-1)!$.

What we need in our case is to partition our set, of m-tupes of k-sets, lets's call is $A(m,k)$, into one- or two- or whatever- parameter subsets $B(m,k;a_1,a_2,\ldots)$ and find a structre theorem that expresses these in terms of B's with smaller parameters. This structre theorem should enable one to take *weight* and get a recurrence scheme for the weight-enumerators of the B's, that can be added up to yield the weight-enumerator of $A(m,k)$, our desired quantity.

I believe that this approach should work with *numeric* (specific) k, but I doubt whether it would work with *symbolic* (i.e. general) k.

Two Computer-Generated Theorems

Theorem K_3. For any r-coloring of the edges of K_n, C, let $X(C)$ be the number of monochromatic triangles. Then, of course $E[X] = \binom{n}{3}/r^2$. The variance is:

$$\frac{\binom{n}{3}(r-1)(r+1)}{r^4} \quad ,$$

and the third moment (about the mean) is

$$\frac{\binom{n}{3}(r-1)(r^3+r^2-2r+6n-20)}{r^6} \quad .$$

Theorem K_4. For any r-coloring of the edges of K_n, C, let $X(C)$ be the number of monochromatic K_4. Then, of course $E[X] = \binom{n}{4}/r^5$. The variance is:

$$\frac{n(n-1)(n-2)(n-3)(r-1)(r^4+r^3+r^2+4rn-15r+4n-15)}{24r^{10}} \quad ,$$

and the third moment (about the mean) is

$$\frac{n(n-1)(n-2)(n-3)(r-1)}{24r^{15}} \cdot$$

$$(r^9+r^8+r^7+12nr^6-47r^6+24nr^5-95r^5+12nr^4-50r^4+40n^2r^3-348nr^3+750r^3+76n^2r^2$$
$$-672nr^2+1470r^2-44n^2r+348rn-690r+24n^3-401n^2+2097n-3510) \quad .$$