Big Numbers

hexy

October 30, 2019

Normal Large Numbers

Some numbers to keep in mind when we are talking about these very large numbers.

- 7.5×10^9: Number of people on Earth.
- 7.5×10^{18}: Number of grains of sand on Earth.
- 10^{24}: Number of stars in the universe.
- 10^{80}: Number of atoms in the observable universe.
- 10^{100}: A googol.
- $10^{10^{100}}$: A googolplex.

Very Large Counterexamples

We begin with some very large upper bounds that arose from the proofs of various facts, including in particular Graham's number.

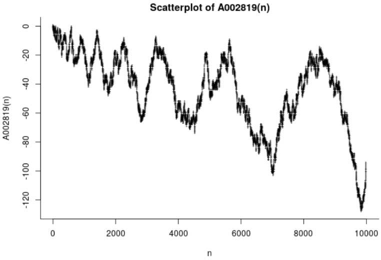
Pólya Conjecture

Consider a number, n, say 60. First, we factorize n into its prime factors:

$$60 = 2^2 \cdot 3 \cdot 5$$

Then we can see that 60 has 4 prime factors: two 2's, one 3, and one 5. In this case, 60 has an even number of prime factors. What we'd like to know is how many numbers have an even number of prime factors versus an odd number of prime factors?

Pólya Conjecture



Source: http://oeis.org/A002819/graph

This is a graph of the balance of the summatory Liouville function $L(n)$, which is the number of numbers $\leq n$ with an even number of prime factors minus the number of numbers with an odd number of prime factors.

Pólya Conjecture

The Pólya conjecture is the statement that there are always more numbers with an odd number of prime factors than there are with an even number of prime factors: alternatively, that $L(n) \leq 0$ for all $n > 1$. It was posed in 1919 by George Pólya.

It turns out that this conjecture is false: it was proven in 1958 using the assistance of the EDSAC I that $L(n) \geq 0$ when $n \approx e^{831.847}$, or 1.847×10^{361}!

Later in 1980, the smallest counterexample was shown to be $n = 906, 150, 257$, which is much smaller.

Mertens Conjecture

The Mertens conjecture is very similar to the Pólya conjecture, but with a different function. We define the Möbius function $\mu(x)$ to be the following:

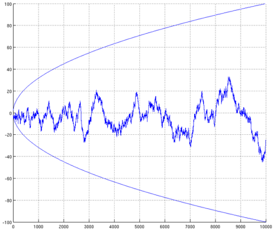
- $\mu(x) = 0$ if x is divisible by a square number.
- $\mu(x) = 1$ if x isn't divisible by a square number, and has an even number of prime factors.
- $\mu(x) = -1$ if x isn't divisible by a square number, and has an odd number of prime factors.

Mertens Conjecture

We then define the Mertens function to be

$$M(x) = \sum_{k=1}^{x} \mu(k).$$

Mertens conjecture is then that this function is bounded by $\sqrt{x}$, which can be seen in the graph below:



Source: https://commons.wikimedia.org/wiki/File:Congettura_Mertens.png

Mertens Conjecture

In 1987, Pintz was able to show that this conjecture must fail before $x = e^{3.21\times 10^{64}}$, improving the result that this conjecture had to fail, but couldn't cite a specific counterexample.

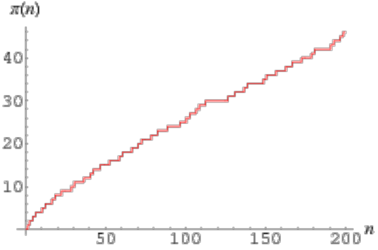
This result was later improved to $e^{1.59\times 10^{40}}$, but no improvement has been made since then.

Skewes's Numbers

In order to talk about Skewes's numbers (yes, there are multiple!), we first have to talk about the prime counting function. We define the prime counting function as

$$\pi(x) = \text{the number of primes} \leq n$$

Here is a graph of the function:



Source: http://mathworld.wolfram.com/PrimeCountingFunction.html

Skewes's Numbers

One of the most famous estimates of the prime counting function is the Prime Number Theorem, which states that

$$\pi(x) \approx \frac{x}{\ln(x)}$$

or formally,

$$\lim_{x \to \infty} \frac{\pi(x)}{\frac{x}{\ln(x)}} = 1.$$

This was proven in 1896 independently by Hadamard and de la Vallée Poussin.

Skewes's Numbers

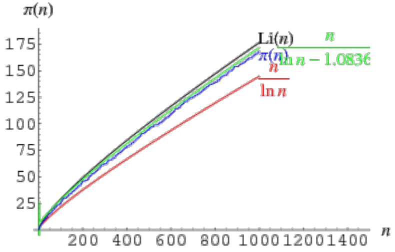
A later, better estimate was shown to be

$$\pi(x) \approx \mathrm{Li}(x)$$

where $\mathrm{Li}(x)$ is the logarithmic integral function, defined as

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\ln(t)}$$

A diagram of the two estimates can be seen below:



Source: http://mathworld.wolfram.com/PrimeNumberTheorem.html

Skewes's Numbers

One of the things you notice about the logarithmic integral funciton is that it is always larger than the prime counting function: this fact has been confirmed in the modern day up to 10^{25}. So it is natural to propose that $\pi(x) < \mathrm{Li}(x)$ for all $x > 2$.

Surprisingly, it was proved in 1914 by Littlewood that $\pi(x) > \mathrm{Li}(x)$ infinitely many times! Skewes was able to prove in 1933 assuming the Riemann Hypothesis that $\pi(x) > \mathrm{Li}(x)$ before

$$e^{e^{e^{79}}} \approx 10^{10^{10^{34}}} \text{ (First Skewes Number)}$$

and in 1955 without assuming the Riemann Hypothesis that $\pi(x) > \mathrm{Li}(x)$ before

$$e^{e^{e^{e^{7.705}}}} \approx 10^{10^{10^{10^{3}}}} \text{ (Second Skewes Number)}$$

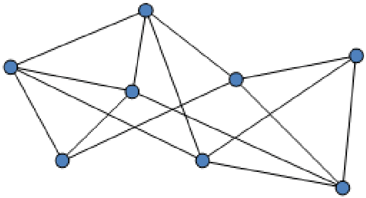
Graham's Number

Everyone has heard of Graham's number, but not many know that the Graham's number we usually talk about was not the original Graham's number!

In order to introduce the problem that Graham's number arises from, we need to introduce a couple of concepts: namely, Ramsey theory.

Ramsey Theory

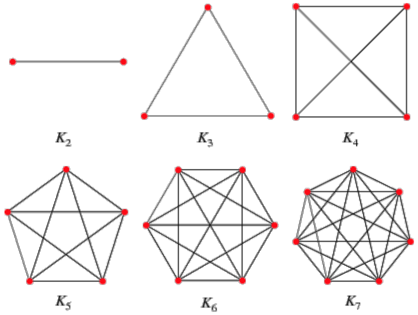
A graph is a collection of vertices (points) connected by edges between those vertices.



Source: https://www.uh.edu/engines/epi2467.htm

Ramsey Theory

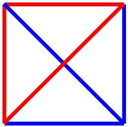
The complete graph on n vertices, K_n is the graph with n vertices where every vertex is connected to every other vertex. Below are the graphs for K_n for $2 \leq n \leq 7$.



Source: http://mathworld.wolfram.com/CompleteGraph.html

Ramsey Theory

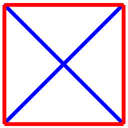
Now, suppose we take a complete graph and color each edge either red or blue, like so:



The question we want to ask is when can we find an all red or all blue triangle when we color a complete graph? In this case with K_4, we can: there is a red triangle.

Ramsey Theory

However, there is a way to color K_4 so there are no red or blue triangles whatsoever:

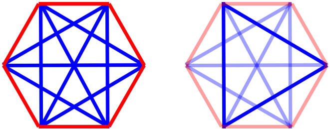


Any triangle in K_4 has to contain one of the diagonals, which are blue, and one of the sides, which are red.

Ramsey Theory

Is there a size of complete graph where we are actually guaranteed a triangle that is all red or all blue?

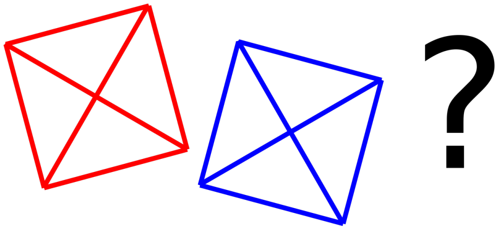
As it turns out, K_6 is guaranteed to have an all red or all blue triangle. If we try the same thing we did before, where the inside edges were blue and the outside edges were red, we can see that we have actually forced a blue triangle:



One can actually prove that there is no way to color the edges of K_6 such that there is no red or blue triangle.

Ramsey Theory

We can ask this questions for bigger subgraphs: is there a complete graph that is guaranteed to contain a K_4 that is all red or all blue?



Ramsey Theory

As it turns out, there is a complete graph that guarantees this: the first one is K_{18}.

Source: https://bl.ocks.org/bryik/a3d0d7a0d9d69e6afe0fd8b8b3becec1

This was proven in 1955 by Greenwood and Gleason, along with the previous result about having a red/blue triangle in a complete graph.

Ramsey Theory

In general, what we are talking about here is something called Ramsey numbers, denoted $R(p, q)$, which asks the question

What is the smallest n such that K_n must contain either an all red K_p or an all blue K_q?

The results we've just gone over are equivalent to the facts that $K(3, 3) = 6$ and $K(4, 4) = 18$.

Ramsey Theory

The general question we can ask is does $R(p, q)$ always exist? That is, is there always a complete graph big enough so that a graph of all one color is forced?

The answer is yes, and the result is known as Ramsey's theorem, named after F. P. Ramsey, who proved this fact in 1930 as the equivalent formulation in set theory:

"Let r, n, and μ be any integers greater than 0. Suppose we classify the r-element subsets of an m-element set S into μ mutually exclusive classes $C_1, \ldots, C_\mu$. Then there is some m_0 such that for all $m \geq m_0$, that there is an n-element subset T of S such that all r-element subsets of T belong to the same class."

Ramsey Theory

The following Ramsey numbers are the only ones that are known to date:

$$R(3,3) = 6$$
$$R(3,4) = 9$$
$$R(3,5) = 14$$
$$R(3,6) = 18$$
$$R(3,7) = 23$$
$$R(3,8) = 28$$
$$R(3,9) = 36$$
$$R(4,4) = 18$$
$$R(4,5) = 25$$

We only know rough bounds on all the others.

Graham's Number

Finally, we can talk about the origins of Graham's number.

Graham was interested in generalizing extensions of Ramsey's theorem to finite fields and affine spaces instead of sets. To do this, he invented something he calls 'n-parameter sets', which are too complicated to really describe here (also I don't understand them).

Graham's Number

The exact definition is given here, for fun.

Let $A = \{a_1, a_2, \ldots, a_t\}$ be a finite set with $t \geqq 2$. Let $H: A \to A$ be a permutation group acting on A. For $a \in A$, $\sigma \in H$, the action is denoted by $a \to a^\sigma$. Also, for $\sigma_1, \sigma_2 \in H$, $\sigma_1\sigma_2 \in H$ is defined by $a^{\sigma_1\sigma_2} = (a^{\sigma_1})^{\sigma_2}$ for all $a \in A$. For a nonempty

subset $B \subseteq A$, let $\bar{B} = \{\bar{b} : b \in B\}$ be the set of constant maps of A into A given by $x^{\bar{b}} = b$ for $x \in A$, $\bar{b} \in \bar{B}$. A^t denotes the cartesian product $A \times A \times \cdots \times A$ (t factors), which is just $\{(x_1, \ldots, x_t) : x_i \in A, 1 \leqq i \leqq t\}$.

For $x = (x_1, \ldots, x_t) \in A^t$, $\sigma \in H$, we define an action of $H: A^t \to A^t$ by

$$x^\sigma = (x_1, \ldots, x_t)^\sigma = (x_1^\sigma, \ldots, x_t^\sigma) \in A^t.$$

Similarly $\bar{B}$ acts on A^t by

$$x^{\bar{b}} = (x_1, \ldots, x_t)^{\bar{b}} = (x_1^{\bar{b}}, \ldots, x_t^{\bar{b}}) = (b, \ldots, b) \in A^t$$

for $x \in A^t$, $\bar{b} \in \bar{B}$.

For fixed integers $n > 0$ and $0 \leqq k \leqq n$, let $\Pi = \{S_0, S_1, \ldots, S_k\}$ be a partition of the set $I_n = \{1, 2, \ldots, n\}$ with $S_i \neq \varnothing$ for $1 \leqq i \leqq k$. $S_0 = \varnothing$ is possible. Let $f: I_n \to H \cup \bar{B}$ be a mapping with the property:

$$f(i) \in \bar{B} \quad \text{if } i \in S_0,$$
$$f(i) \in H \quad \text{if } i \in I_n - S_0.$$

The set $P(A, \bar{B}, H, \Pi, f, n, k) = P$ is defined by

$$P = \bigcup_{1 \leqq i_0, \ldots, i_k \leqq t} \{(x_1, \ldots, x_n) : x_j = a_{i_v}^{f(j)} \text{ if } j \in S_v\} \subseteq A^n.$$

Definition 1. A subset $P \subseteq A^n$ is said to be a *k-parameter set in* A^n if $P = P(A, \bar{B}, H, \Pi, f, n, k)$ for some meaningful choice of these variables.

Source:

https://www.ams.org/journals/tran/1971-159-00/S0002-9947-1971-0284352-8/S0002-9947-1971-0284352-8.pdf

Graham's Number

Graham then proves Ramsey's theorem for l-parameter subsets of n-parameter sets, whose statement is given here:

7. **The main result.** Before proceeding with the main result of the paper we make a remark on terminology.

DEFINITION 3. By an *r-coloring of a set X* we just mean a partition of X into r disjoint (possibly empty) classes.

Of course, the "r colors" correspond to the r classes into which X is partitioned. In general, we shall use this "chromatic" terminology in preference to that of partitions and classes.

THEOREM. *Given* A, B, H *and integers* $k, r, t_1, \ldots, t_r$, *there exists an* $N = N(A, \bar{B}, H, k, r, t_1, \ldots, t_r)$ *such that if* $n \geqq N$ *and* $P_n = P(A, \bar{B}, H, \Pi, f, w, n)$ *is any fixed n-parameter set in* A^w, *then for any r-coloring of the k-parameter subsets of* P_n *there exists an* i, $1 \leqq i \leqq r$, *such that there is some* t_i*-parameter subset of* P_n *with all its k-parameter subsets having color i.*

Graham's Number

The original statement of Graham's number is an aside on the last page of a 40-page paper about the bounds his theorem gives for various results, as stated here:

We point out that even though the techniques of the proof of the theorem are constructive so that upper bounds on the various N's of the corollaries can be given, these bounds are usually enormous, to say the least. To illustrate this, we consider the first nontrivial case of Corollary 12, the determination of an upper bound on $N(1, 2, 2)$. We recall that by definition $N(1, 2, 2)$ is an integer such that if $n \geqq N(1, 2, 2)$ and the $\binom{2^n}{2}$ straight line segments joining all possible pairs of vertices of a unit n-cube are arbitrarily 2-colored, then there always exists a set of four *coplanar* vertices which determines six line segments of the same color. Let N^* denote the least possible value $N(1, 2, 2)$ can assume. We introduce a calibration function $F(m, n)$ with which we may compare our estimate of N^*. This is defined recursively as follows:

$$F(1, n) = 2^n, \qquad F(m, 2) = 4, \qquad m \geqq 1,\ n \geqq 2,$$
$$F(m, n) = F(m-1, F(m, n-1)), \qquad m \geqq 2,\ n \geqq 3.$$

It is recommended that the reader calculate a few small values of F to get a feeling for its rate of growth, e.g., $F(5, 5)$ or $F(10, 3)$.

If the bounds generated by the recursive constructions needed for the proof of Corollary 12 are explicitly tabulated, the best estimate for N^* we obtain this way is roughly

$$N^* \leqq F(F(F(F(F(F(F(12, 3), 3), 3), 3), 3), 3), 3).$$

On the other hand, it is known only that $N^* \geqq 6$. Clearly, there is some room for improvement here.

The text here is quite dense, so we will break it down.

Graham's Number

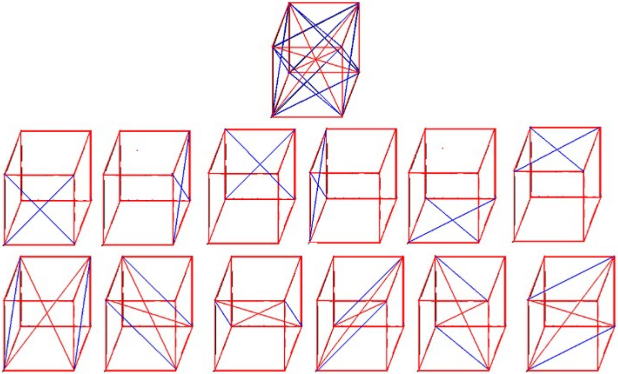
The problem being considered here is the following problem, which is like the Ramsey problems before, but with a twist:

Consider an n-dimensional hypercube, and connect all vertices with all edges to form a complete graph on 2^n vertices. What is the smallest n such that there is guaranteed to be a K_4 that is all red or all blue, and all its vertices are in the same plane?

We see that there is a geometric requirement here that is a twist on the usual problem, and makes the problem a lot more difficult to parse.

Graham's Number

Here is an example for a 3-dimensional cube that doesn't have a K_4 that is all red or all blue.



Source: Adapted from https://sites.google.com/site/largenumbers/home/3-2/Graham

Graham's Number

What Graham gives is an upper bound on what n can be, proving that the problem does have a solution, defined by the following function:

$$\begin{aligned} F(1,n) &= 2^n \\ F(m,2) &= 4 \\ F(m,n) &= F(m-1, F(m,n-1)) \end{aligned}$$

where the upper bound, Graham's number, is

$$F(F(F(F(F(F(F(12,3),3),3),3),3),3),3)$$

It is really unclear just how quickly this function grows, so we will compute some examples.

Graham's Number

What is $F(2,5)$?

$$
\begin{aligned}
F(2,5) &= F(1, F(2,4)) \\
&= F(1, F(1, F(2,3))) \\
&= F(1, F(1, F(1, F(2,2)))) \\
&= F(1, F(1, F(1,4))) \\
&= F(1, F(1, 2^4)) \\
&= F(1, 2^{2^4}) \\
&= 2^{2^{2^4}} \\
&= 2^{2^{2^{2^2}}}
\end{aligned}
$$

Graham's Number

What is $F(3,5)$?

$$\begin{aligned} F(3,5) &= F(2, F(3,4)) \\ &= F(2, F(2, F(3,3))) \\ &= F(2, F(2, F(2, F(3,2)))) \\ &= F(2, F(2, F(2,4))) \\ &= F(2, F(2, 2^{2^{2^{2}}})) \\ &= F(2, F(2, 65536)) \\ &\quad \text{uhhhh....} \end{aligned}$$

This quickly gets out of hand, and we're nowhere near $F(12,3)$, let alone Graham's number. There is a nice notation we can use to describe F, called Knuth's up arrow notation.

Knuth's Up Arrow Notation

We first define what single and double up arrows mean:

$$a \uparrow b = a^b$$
$$a \uparrow\uparrow b = \underbrace{a \uparrow a \uparrow \cdots \uparrow a}_{b \text{ times}}$$

So $2 \uparrow\uparrow 4 = 2 \uparrow 2 \uparrow 2 \uparrow 2 = 2^{2^{2^2}}$. We define the triple up arrow as follows:

$$a \uparrow\uparrow\uparrow b = \underbrace{a \uparrow\uparrow a \uparrow\uparrow \cdots \uparrow\uparrow a}_{b \text{ times}}$$

So $2 \uparrow\uparrow\uparrow 4 = 2 \uparrow\uparrow 2 \uparrow\uparrow 2 \uparrow\uparrow 2 = 2^{2^{\cdot^{\cdot^{2}}}} \Big\} 2^{2^{\cdot^{\cdot^{2}}}} \Big\} 2^{2^{2^2}}$ times, which is a complete mess.

Knuth's Up Arrow Notation

In general, if we call $a \uparrow^n b = a \underbrace{\uparrow\uparrow \cdots \uparrow}_{n \text{ times}} b$, then

$$a \uparrow^n b = \underbrace{a \uparrow^{n-1} a \uparrow^{n-1} \dots \uparrow^{n-1} a}_{b \text{ times}}.$$

Graham's Number

Going back to the function F, you can show that

$$F(m, n) = 2 \uparrow^m n,$$

and therefore Graham's number is

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots\cdots\cdots\cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots\cdots\cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots\cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow}\ 3$$

$$2 \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow 3$$

Graham's Number

There is a note we have to make here: This isn't actually Graham's number as we know it. "Graham's number" is usually defined as follows:

$$g_1 = 3 \uparrow\uparrow\uparrow\uparrow 3$$
$$g_n = 3 \uparrow^{g_{n-1}} 3$$

Then "Graham's number" is g_{64}, which looks like

$$\left.\begin{matrix} 3 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots\cdots\cdots\cdots\cdots\cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow} 3 \\ \vdots \\ \underbrace{\qquad\qquad\qquad\qquad\qquad\qquad} \\ 3 \underbrace{\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow \cdots \uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow} 3 \\ 3 \uparrow\uparrow\uparrow\uparrow 3 \end{matrix}\right\} \text{64 layers}$$

Graham's Number

So where does this definition come from? It actually comes from Martin Gardner's Mathematical Games column in Scientific American, as depicted below:

Graham's solution to a problem by Erdös

tion of Ramsey's theorem that they found in 1970. Finding the actual number, however, is something else. In an unpublished proof Graham has recently established an upper bound, but it is a bound so vast that it holds the record for the largest number ever used in a serious mathematical proof.

To convey at least a vague notion of the size of Graham's number we must first attempt to explain Knuth's arrow notation. The number written $3\uparrow 3$ is $3 \times 3 \times 3 = 3^3 = 27$. The number $3\uparrow\uparrow 3$ denotes the expression $3\uparrow(3\uparrow 3)$. Since $3\uparrow 3$ equals 27, we can write $3\uparrow\uparrow 3$ as $3\uparrow 27$ or 3^{27}. As a slanting tower of exponents it is

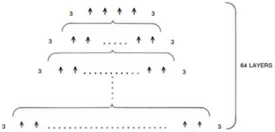
$$3^{3^3}.$$

The tower is only three levels high, but written as an ordinary number it is finite numbers are very much larger.

We are now in a position to indicate Graham's number. It is represented in the illustration below. At the top is $3\uparrow\uparrow\uparrow\uparrow 3$. This gives the number of arrows in the number just below it. That number in turn gives the number of arrows below it. This continues for 2^6, or 64, layers. It is the bottom number that Graham has proved to be an upper bound for the hypercube problem.

Now hold on to your hat. Ramsey-theory experts believe the actual Ramsey number for this problem is probably 6. As Stanislaw M. Ulam has said many times in his lectures, "The infinite we shall do right away. The finite may take a little longer."

Last month's problem was to guess the secret rule that determined the final layout for a round in the card game Eleusis. The rule was: "If the last card is lower than the preceding legally played card, play a card higher than the last card, otherwise play a lower one. The first card played is correct unless it is equal to the starter card."



Graham's upper bound for the solution to a Euclidean Ramsey problem

Source: https://www.scientificamerican.com/article/mathematical-games-1977-11/

Graham's Number

Why the different number? John Baez, a physicist and mathematician, asked Graham this question. He answered that it was easier to explain the problem with this larger upper bound than the one used in the paper. Unfortunately, the exact Google+ post which described this has been lost to time, and all that remains is this StackOverflow post from John Baez:

> 21
>
> I talked to Ronald Graham last night at the Joint Mathematics Meeting in San Diego and asked him the question here. He said he'd made up Graham's number when talking to Martin Gardner because 1) it was simpler to explain than his actual upper bound, the one that appears in his paper with Rothschild, and 2) it's bigger, so it's still an upper bound!
>
> So, apparently the comment on Wikipedia:
>
> > This weaker upper bound for the problem, attributed to an unpublished work of Graham [....]
>
> is a bit misleading, though still technically true. I'll try to fix it in a while.
>
> Nice question!
>
> share cite improve this answer
>
> edited Jan 11 '13 at 18:06 Andrés E. Caicedo 28.6k •5 •100 •176
>
> answered Jan 11 '13 at 17:26 John Baez 13.7k •1 •50 •103

Source: https://mathoverflow.net/questions/117006/reconstructing-the-argument-that-yields-grahams-number/118650#118650

Graham's Number

The original upper bound by Graham has since been improved to $2 \uparrow\uparrow\uparrow 5$ by Eryk Lipka in May of 2019, very recently! The lower bound of the problem at the time Graham gave his upper bound was 6, but this has since been improved to 13, giving us the ridiculous bound of

$$13 \leq n^* \leq 2 \uparrow\uparrow\uparrow 5$$

on this problem.

Fast Growing Functions

Aside from very large upper bounds to problems, we can talk about very fast growing functions that have been used to prove various things, or come about from the proofs of some theorems.

Ackermann Function

The Ackermann function is a function invented by Wilhelm Ackermann, which was originally defined with three arguments as follows:

$$\begin{aligned}
\varphi(a, b, 0) &= a + b \\
\varphi(a, 0, n) &= \alpha(a, n) \\
\varphi(a, b, n) &= \varphi(a, \varphi(a, b-1, n), n-1)
\end{aligned}$$

where

$$\alpha(a, n) = \begin{cases} 0, & n = 0 \\ 1, & n = 1 \\ a, & n > 1 \end{cases}$$

Ackermann Function

While the definition is unclear, this actually translates to something much simpler:

$$
\begin{aligned}
\varphi(a,b,0) &= a+b \\
\varphi(a,b,1) &= a\times b \\
\varphi(a,b,2) &= a^{b} \\
\varphi(a,b,3) &= a\uparrow\uparrow b \\
&\vdots
\end{aligned}
$$

What makes this function special is that it is one of the first functions that was shown to be computable but not primitive recursive. Think of computable as there being an algorithm to compute the number, and primitive recursive as there being a program with fixed length loops that is able to compute the number.

Ackermann Function

This original definition is not the one that is usually used; the modern definition of this function only has 2 arguments, and doesn't use the $\alpha(a, n)$ function:

$$\begin{aligned} A(0, n) &= n + 1 \\ A(m, 0) &= A(m-1, 1) \\ A(m, n) &= A(m-1, A(m, n-1)) \end{aligned}$$

One can show that $A(m, n) = (2 \uparrow^{m-2} (n+3)) - 3$.

Friedman's n(4)

This number has to do with strings of characters, like the following:

$$x = abcacbac$$

We call the first character of the string $x_1 = a$, the second character $x_2 = b$, and so on. Suppose the length of the string is n. What we are interested in is the following property of a string:

Property $*$: There is no $i < j \leq \frac{n}{2}$ such that $x_i, x_{i+1}, \ldots, x_{2i}$ is a subsequence of $x_j, x_{j+1}, \ldots, x_{2j}$.

Friedman's n(4)

This property is really confusing at first glance, so we are going to see if it holds for the string we had before, $x = abcacbac$.

1-2: ab

2-4: bca

3-6: $cacb$

4-8: $acbac$

Here we can see that Property $*$ doesn't hold since x_1, x_2 is a subsequence of x_3, x_4, x_5, x_6.

Friedman's n(4)

Here is an example of a string where Property $*$ does hold, $x = abbaaaaa$.

1-2: ab

2-4: bba

3-6: $baaa$

4-8: $aaaaa$

Friedman's n(4)

Now that we understand Property $*$ slightly better, we can finally define the $n(k)$:

$n(k) =$ the length of the longest string using k different characters that has Property $*$.

We can see that $n(1) = 3$. What is the longest string that has Property $*$ when you are allowed to only use a single character, a? This string is aaa, because if we added one more character (the string $aaaa$), then we would have that x_1, x_2 would be a substring of x_2, x_3, x_4.

Friedman's n(4)

What is $n(2)$? It can be proven that $n(2) = 11$, with the string being *abbbaaaaaaa*.

How about $n(3)$? By a result due to Harvey Friedman with the assistance of Ronald Dougherty, we have that

$$n(3) > 2 \uparrow^{7198} 158386$$

There is an explicit construction of the string that leads to this lower bound, but it is incredibly intricate and complicated.

Friedman's n(4)

The string looks like this:

$$\underbrace{u13^{n-1}}_{\text{special string } \alpha} 2 \overbrace{B_1 C_1 B_2 C_2 \cdots B_d C_d}^{x\text{, a strong } 2n, 3n, 3k+1, k, d \text{ sequence}}$$

The special string has the property that each $x_i, \ldots, x_{2i}$ has at least one a for $i \leq n-1$, and also has Property $*$.

x is a special type of sequence where each B_i consists of $3k+1$ repetitions of c, and each C_i consists of only b's and c's, where there are exactly k repetitions of c in C_i. The length of each C_i increases as i gets larger, but the length is hard to express simply.

Friedman's n(4)

$$3^{n-1}2\overbrace{B_1 C_1 B_2 C_2 \cdots B_d C_d}^{x,\text{ a strong } 2n, 3n, 3k+1, k, d \text{ sequence}}$$

The above string has the incredibly important property that if there is some i and j such that $x_i, \ldots, x_{2i}$ is a subsequence of $x_j, \ldots, x_{2j}$, then we have that $C_{i'}$ is a subsequence of $C_{j'}$ for some $i' < j' \leq d$.

As it turns out, there is a way to construct an extremely long sequence of C_i's (read: $2 \uparrow^{k-2} (2n - 4k - 1)$) that aren't subsequences of each other by creating a very long sequence of $(k-1)$-tuples and mapping those to C_i's in a special way that we will not describe here.

Friedman's n(4)

The above construction, while very convoluted, reduces to the fact that if you can find a special string of length $26k + 8$, then you can show that $n(3) > 2 \uparrow^{k-1} (22k + 8)$. Ronald Dougherty was able to find a special string of length $187196 \geq 26(7199) + 8$, hence the previous result.

Friedman's n(4)

All this, and we haven't even gotten to $n(4)$! As it turns out, Firedman was able to prove that if we define the function $f(n) = 2 \uparrow^n n$, then

$$n(4) > f^{f(187196)}(1)$$

where f^n means f composed with itself n times. As it happens, this shows that $n(4)$ is larger than Graham's number, from the original paper and the one invented with Martin Gardner!

Goodstein Function

Now, we move on to another, completely unrelated topic. We are going to go over Goodstein sequences. There are 3 steps to a Goodstein sequence:

1. Start with a number n, and write it in hereditary base 2 notation, like below:

$$5 = 2^2 + 1$$
$$9 = 2^3 + 1 = 2^{2^1+1} + 1$$
$$23 = 2^4 + 2^2 + 2 + 1 = 2^{2^2} + 2^2 + 2 + 1$$

   This is the starting number.

Goodstein Function

2. To get the next number in the sequence, increment the base by 1, then subtract 1 from the number. For instance, if we had started with $5 = 2^2 + 1$, we would do the following:

$$2^2 + 1 \xrightarrow[\text{add 1 to base}]{} 3^3 + 1 \xrightarrow[\text{subtract 1}]{} 3^3$$

3. Repeat until you hit 0. Here would be the next iteration from the previous example:

$$3^3 \xrightarrow[\text{add 1 to base}]{} 4^4 \xrightarrow[\text{subtract 1}]{} 3 \cdot 4^3 + 3 \cdot 4^2 + 3 \cdot 4 + 3$$

Goodstein Function

We will let the Goodstein function $G(n)$ be either the number of steps it takes for the Goodstein sequence starting from n to reach 0, or ∞ if it never does, since we don't even know if it does every time.

We can see that $G(1) = 1$ and $G(2) = 3$. What is $G(3)$?

As it so happens, $G(3) = 5$, as can be seen below:

$$2 + 1 \to 3 \to 3 \to 2 \to 1 \to 0$$

Goodstein Function

What is $G(4)$? Well...

$$\begin{aligned}
2^2 &\to 2 \cdot 3^2 + 2 \cdot 3 + 2 \\
&\to 2 \cdot 4^2 + 2 \cdot 4 + 1 \\
&\to 2 \cdot 5^2 + 2 \cdot 5 \\
&\to 2 \cdot 6^2 + 6 + 5 \\
&\quad \vdots \\
&\to 2 \cdot 11^2 + 11 \\
&\to 2 \cdot 12^2 + 11 \\
&\quad \vdots \\
&\to 2 \cdot 23^2 \\
&\to 24^2 + 23 \cdot 24 + 23 \\
&\quad \vdots
\end{aligned}$$

Goodstein Function

The sequence just keeps getting bigger. Does it stop? As it so happens, $G(4) = 3 \cdot 2^{402653211} - 1$.

The surprising result is that actually, no matter what number you start with, the Goodstein sequence always terminates! So $G(n)$ is always finite, but very large.

Even more surprising, it doesn't matter how fast the base grows at each step of the sequence, the Goodstein sequence will always terminate.

This function actually grows incredibly quickly, outpacing the $n(k)$ function we mentioned earlier by quite a bit. The proof of this fact involves going into ordinals, which are really complicated, and we will not cover them here.

Goodstein Function

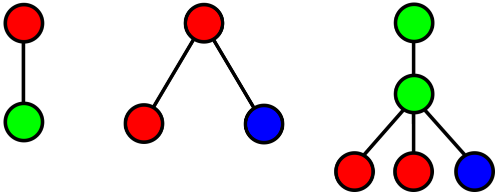
The reason why this function is interesting is because of the Kirby-Paris theorem, which roughly states that $G(n)$ grows so fast that Peano Arithmetic cannot prove that every Goodstein sequence will eventually terminate!

Peano Arithmetic is a lesser logical system to the one we use normally (ZF+Axiom of Choice), and its axioms are as follows:

1. 0 is a natural number.
2. If a is a natural number, then $S(a)$, the successor of a (think of successor as $a+1$), is a natural number.
3. There is no natural number such that $S(a) = 0$.
4. If $S(a) = S(b)$ then $a = b$.
5. If there is a set such that it contains 0 and if a is in the set, then $S(a)$ is in the set, then this set is the set of all natural numbers.

TREE(3)

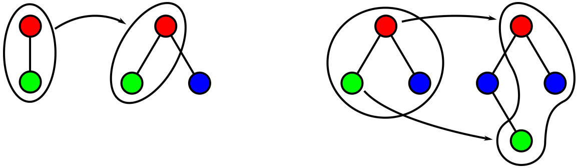
Now we move on to yet another topic, labeled trees. A labelled tree is a graph (think of this as points connected by lines) that has no cycles, and where each vertex is colored some color, like below:



For the purposes of the upcoming function, we will not need more than 3 colors.

TREE(3)

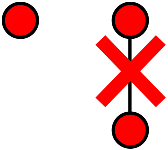
We say one tree is inf preserving and label preserving embeddable into another tree if we can map the vertices of one tree to the same color vertices of another tree, and descendants and common ancestors in one tree stay the same in the other tree, like the two examples below:



TREE(3)

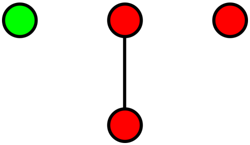
We then define TREE(n) to be the length of the longest sequence of trees $T_1, \ldots, T_k$ with up to n colors such that each T_i has a maximum of i vertices, and no T_i is inf preserving and label preserving embeddable into T_j for any $i < j$.

What is TREE(1)? TREE(1) $= 1$. The first tree can only have one vertex with a single color, and any tree after that will have a vertex with the same color as only 1 color is allowed.



TREE(3)

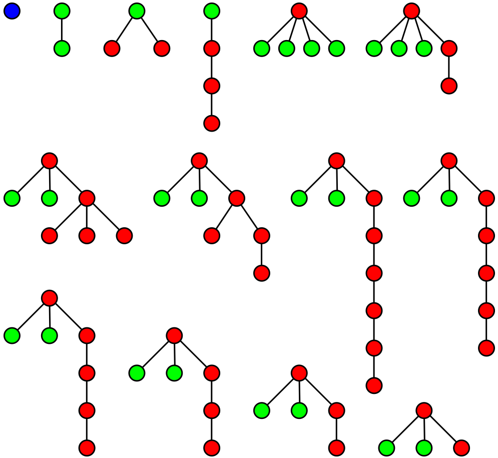
What is TREE(2)? $\mathrm{TREE}(2) = 3$, as shown below.



What is TREE(3)? We have that at the very minimum, $\mathrm{TREE}(3) > n(4)$ via an argument by Harvey Friedman, illustrated on the next slide.

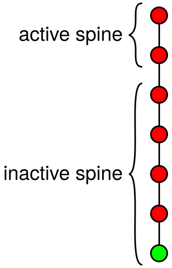
TREE(3)

We first start the sequence with 14 trees $T_1, \ldots, T_{14}$ as shown below:



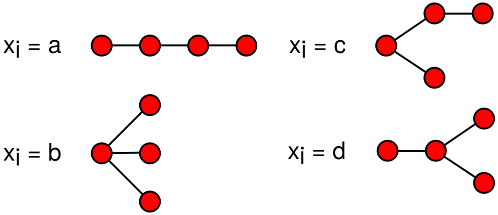
TREE(3)

We first construct T_{15} as a straight line of 7 vertices, the last one being green instead of red. The last 5 vertices are the *inactive spine*, which is necessary to make the construction possible, but does nothing. All the other vertices are the *active spine*, which we will be attaching 4-vertex trees to.



TREE(3)

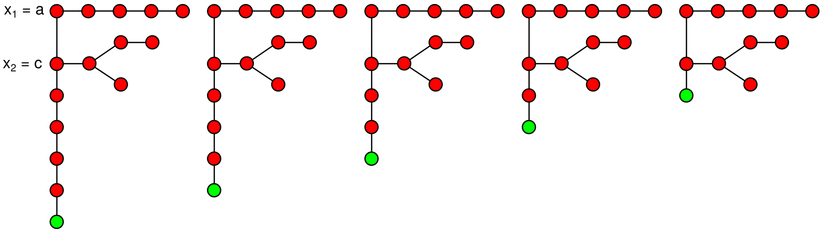
Now, we bring in $n(4)$. Let $x_1, \ldots$ be the sequence corresponding to $n(4)$. What we do next is we consider the subsequences $x_i, \ldots, x_{2i}$, and for each letter in the string we attach a specific 4-vertex tree to each vertex of the active spine, shown by the diagram below:



For T_{15}, the subsequence we use is x_1, x_2.

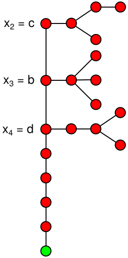
TREE(3)

Below, we show an example of what $T_{15}, \ldots, T_{19}$ look like. We progressively remove 1 vertex from T_{15} so that we have enough vertices to have a 3-vertex active spine for the next tree.



TREE(3)

Here is the example of the next tree, T_{20}. It can't be inf preserving and label preserving embeddable into T_{15} since for that to work, x_1, x_2 has to be a subsequence of x_2, x_3, x_4, which can't happen.



We can keep going for $5 \cdot n(4)$ steps, and even then at the end of the process we can still keep going with all-red trees.

TREE(3)

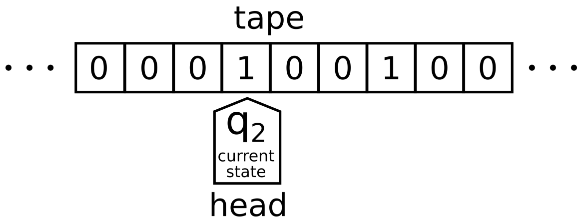
It should be noted that $n(4)$ is a very small upper bound on TREE(3): Friedman has remarked that TREE(3) $> n^{n(5)}(5)$, but provides no explicit construction for the tree in this case.

It has been shown that TREE(n) grows significantly faster than either $n(k)$ or $G(n)$, but the exact statement of this fact is somewhat complex.

Busy Beaver Function

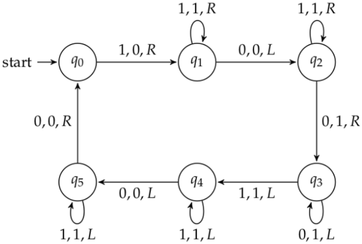
The very last function we will go over comes from computability theory. We will have to define a Turing machine, however.

A Turing machine is a very abstract computer, which consists of a tape to write 0's and 1's on, the machine head that moves around and actually writes the 0's and 1's, a state register that stores the current state of the Turing machine, and instructions which tell the Turing machine what to do.



Busy Beaver Function

At each step, we check the current state, and what is currently on the tape (0 or 1) at the head. Once these are determined, we write a 0 or 1 on the tape, move the head left or right by one cell, and change to another state.



Source: https://i1.wp.com/openlogicproject.org/wp-content/uploads/2016/01/tm.png?fit=720%2C466&ssl=1

Busy Beaver Function

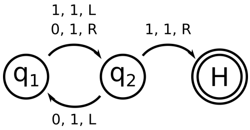
We define the Busy Beaver function $\Sigma(n)$ as follows:

Start with a Turing machine with at most n states, plus an additional Halt state that tells the Turing machine to stop. We run this machine on a tape that initially has all 0's, and if it halts, then we record the number of 1's the machine prints on the tape. Then $\Sigma(n)$ is the maximum number of 1's an n state Turing machine can print.

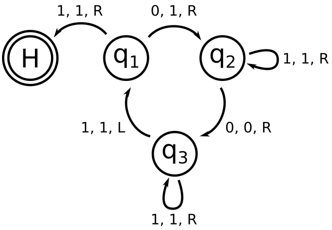
Busy Beaver Function

Here is a list of the values of $\Sigma(n)$ for $n \geq 2$:

- $\Sigma(2) = 4$

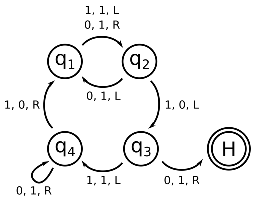


- $\Sigma(3) = 6$

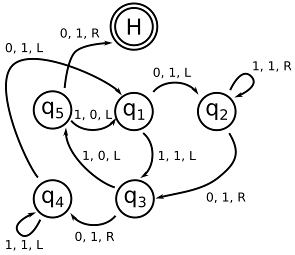


Busy Beaver Function

- $\Sigma(4) = 13$

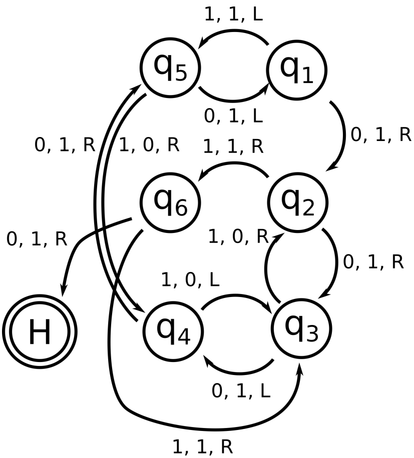


- $\Sigma(5) \geq 4098$



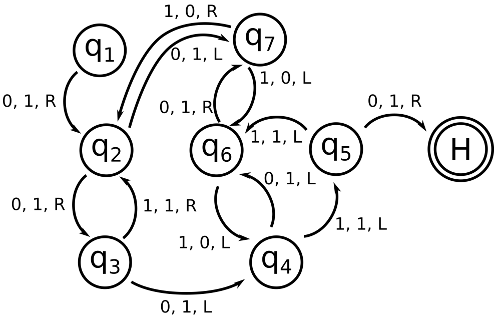
Busy Beaver Function

- $\Sigma(6) \geq 3.5 \times 10^{18267}$



Busy Beaver Function

- $\Sigma(7) \geq 10^{10^{10^{10^{18705352}}}}$



Busy Beaver Function

As it happens, the Busy Beaver function is uncomputable, and it grows faster than any computable function (calculable by a Turing machine). Every function we have talked about so far is computable, so this necessarily grows faster than pretty much any function you can define.

Conclusion

Thanks for listening!

Any questions?