An Enquiry Concerning Human (and Computer!) [Mathematical] Understanding

Doron ZEILBERGER[1]

Dedicated to my two favorite skeptics: David Hume and Gregory Chaitin

The Arrogance of Science and Mathematics

Science and mathematics seem to be huge success stories. Hence it is not surprising that most scientists and mathematicians think that science and mathematics are the most secure ways of acquiring *knowledge*, and that all knowledge could, at least in principle, be derived using either the *scientific method*, using *inductive* reasoning, and in the case of *mathematical* knowledge, using *deductive* reasoning.

In the 19th century, people were so impressed with science and mathematics that, starting with Comte, a movement called *Positivism*, that tried to apply the so-called scientific method to all domains of inquiry, gained prominence. But then the pendulum swung back, and many objected to what they called the *imperialism of science*, and Comte's *Empiric Positivism* gave way to Bergson's and others' *metaphysico-spiritual movement*, that emphasized the *heart* rather than the *brain*, and *intuition* rather than *deduction*. A century earlier, German *Romanticism* and *Idealism* were reactions against the *Rationalism* of the Enlightenment.

More recently, science came under attack by *post-modern* philosophers, and that got some scientists, most notably Alan Sokal, to fight back by making fun of them. Little did Alan know that the *joke is on him*, since while some of the details of the philosophical critiques of science were indeed erroneous and sometimes pure gibberish, the *spirit* of the critiques were very well-founded, since all that they were trying to say was that old standby, that goes back at least to Socrates: *We know that we don't know.*

Skeptics

I have always admired skeptics, from Pyrrho of Elis all the way to Jacques Derrida. But my two *favorite* skeptics are **David Hume** and **Gregory Chaitin**, who so beautifully and eloquently described the **limits of science** and the **limits of mathematics**, respectively.

David Hume's Critique of the Scientific Method

According to Bertrand Russell, there is a place in hell for philosophers who *believed* that they solved Hume's *problem of induction*. Of course, no one has yet solved it, and Hume's famous assertion that (physical) *induction*, i.e. *generalizing from* **finitely** *many cases*, has **no (logical) justification whatsoever**, has not yet been rebutted successfully.

---

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg@math dot rutgers dot edu`, `http://www.math.rutgers.edu/~zeilberg`.
First version: April 13, 2007. This version: April 18, 2007. Supported in part by the NSF.

Let's cite his doubts about the sun rising tomorrow:

That the sun will not rise tomorrow is no less intelligible a proposition, and implies no more contradiction, than the affirmation, *that it will rise.*

Another, more recent, attack on (physical) induction was launched by Nelson Goodman, who coined the term *grue* for an object that is green before Jan. 1, 2050, and is blue after it. So far all examined emeralds turned out to be green, hence, by (physical, incomplete) induction it is reasonable to state that "all emeralds are green". But, by the same token, so far all emeralds turned out to be grue, so stand by for Jan. 2, 2050, and dear old Goodman predicts that all emeralds will be blue then, since then grue would be blue, and we have such good empirical evidence that they are always grue.

Greg Chaitin and the Limits of Mathematics

Standing on the shoulders of Gödel, Turing (and Post, Church, Markov and others), Greg Chaitin gave the most *succinct*, *elegant*, and *witty* expression to the **limits** of our mathematical knowledge. It is his immortal **Chaitin's Constant**, Ω:

$$\Omega := \sum_{p \; halts} 2^{-|p|} \quad ,$$

where the sum ranges over all *self-delimiting* programs run on some Universal Turing Machine. As Greg puts it so eloquently, Ω is the epitome of *mathematical randomness*, and its digits are beautiful examples of *random mathematical facts*, true for "no reason". It also has the charming property of being *normal to all bases.*

How Real is Ω?

There is only one problem with Ω, it is a *real* number! As we all know, but most of us refuse to admit, "real" numbers are **not** real, but purely fictional, since they have **infinitely many** digits, and there is no such thing as infinity. Worse, Ω is **uncomputable**, since we know, thanks to Turing, that there is no way of knowing, *a priori*, whether p halts or not. It is true that many "real" numbers, for example $\sqrt{2}, \phi, e, \pi$ etc., can be *deconstructed* in finite terms, by renaming them 'algorithms', and we do indeed know that these are genuine algorithms since in each specific case, we can prove that any particular digit can be computed in a finite, pre-determined, number of steps. But if you believe in Ω, then you believe in God. God *does* know whether *any* program p will eventually halt or not, because God lives for ever and ever (Amen), and also can predict the future, so for God, Ω is as real as $\sqrt{2}$ or even 2 is for us mere mortals. So indeed, *if* God exists, then Ω exists as well, and God knows all its digits. Just because *we*, lowly mortals, will never know the digits of Ω, is just a reflection on *our* own limitations.

But what if you *don't* believe in God? Or, like myself, does not know for sure, one way or the other?

Do I believe in Ω?

Regardless of whether or not God exists, God has no place in mathematics, at least in *my* book. *My* God does not know (or care) whether a program p eventually halts or not. So Ω does **not** exist in my, ultra-finitistic, world-view. But, it does indeed exist as a *symbol*, and as a lovely *metaphor*, so like enlightened 'non-fundamentalist' religious folks, we can still enjoy and believe in the bible, even without taking it literally. I can still love and cherish and adore Chaitin's constant, Ω, the same way as I enjoy Adam and Eve, or Harry Potter, and who cares whether they are 'real' or 'fictional'.

Greg Chaitin's Advice About Experimental Mathematics

One interesting moral Greg Chaitin draws from his brainchild, *Algorithmic Information Theory*, and its crown jewel, Ω, is the advice to pursue Experimental Mathematics. Since so much of mathematical truth is inaccessible, it is stupid to insist on finding a proof for every statement, since for one, the proof may not exist (it may well be undecidable), or it may be too long and complicated for us mere humans, and even for our computers. So Greg suggests to take truths that we 'feel' are right (on heuristic or experimental grounds) and adopt them as new 'axioms', very much like physicists use Conservation of Energy and the Uncertainty Principle as "axioms". Two of his favorites are $P \neq NP$ and the Riemann Hypothesis. Of course, by taking these as new 'axioms' we give up on one of the original meanings of the word 'axiom', that it should be 'self-evident', but Hilbert already gave this up by making mathematical deduction into a formal game.

Stephen Wolfram's Vision

Another, even more extreme, advocate of Experimental Mathematics, is guru Stephen Wolfram, whose *New Kind of Science* and *New Kind of Mathematics* are *completely computer-simulation-*centric. Let's dump traditional equation-centric science and deduction-centric mathematics in favor of doing computer experiments, and watching the output.

Tweaking Chaitin's and Wolfram's Messages: The Many Shades of Rigor

I admire both Chaitin and Wolfram, but like true visionary prophets, they see the world as *black* and *white*. Since all truths that we humans can know with old-time certainty are doomed to be *trivial* (or else we wouldn't have been able to prove them completely), and conversely, *all* the *deep* results will never be able to be proved by us completely, with traditional standards, they advise us to abandon the old ways, and just learn how to ask our computers good questions, and watch its *numerical* output, and gain *insight* from it.

Things do not have to be so polarized. First, computers can help us find *completely rigorous* proofs, that we humans can never find by ourselves, for example the Four Color Theorem, or the many computer-generated proofs of WZ theory. Second, as I first suggested in my Oct. **1993** *Notices* manifesto, "*Theorems for a Price: Tomorrow's Semi-Rigorous Mathematical Culture*", one can try and prove things semi-rigorously.

So the great insight of Greg Chaitin and Stephen Wolfram can be *fine-tuned* and instead of the *"all or nothing"* mentality regarding rigor, we can introduce a whole *spectrum* of **rigor** and **certainty**.

The Greek Model for Mathematics and Meta-Mathematics

Meta-mathematics, starting with Frege, continuing through Russell, Whitehead and Hilbert, and culminating in Chaitin and others, has been using the *Euclidean* model of mathematics, trying to *emulate* and *formalize* Euclid's paradigmatic *Elements.* Start with a set of *axioms* (originally required to be *self-evident* but later considered *arbitrary*) and *rules of deduction*, and a notion of *formal proof* and try to derive *all* theorems from the axioms.

Alas, Hilbert's naive dream was shattered by Gödel (and later by Turing, and beautifully explicated by Chaitin) who (allegedly) proved that:

"*There exist true yet unprovable statements* ".

Of course, you can *meta-prove* them, but then there would be new statements that you could only meta-meta-prove *ad infinitum.*

Did Gödel Really Prove That There Exist True yet Unprovable Statements?

Of course not! All his "statements" were *meaningless*!

Every statement that starts : "for every integer $n \ldots$" or "there exists an integer $n \ldots$", is completely meaningless, since it tacitly assumes that there are *infinitely* many integers. Of course, there are only finitely many of them, since our *worlds*, both the *physical* and the *mathematical*, are *finite.*

More specifically, the meta-statement:

"P has a proof of length ≤ 1000000 *characters"* does make sense,

and even the meta-statement

"P has a proof of length $\leq$ *googolplex characters"* does make sense,

but the "statement":

"P is unprovable"

is the same as the following "statement":

"There does not exist an integer t such that P has a proof of length t characters",

and this "statement" is completely meaningless.

Ditto for the Gödel sentence that is "equivalent" to it, that contains lots of quantifiers.

So, all that Gödel meta-proved was the *conditional* statement:

If "P is unprovable" makes sense and *if* the Gödel sentence makes sense, *then* there exist true yet unprovable statements.

Gödel, being a devout infinitarian platonist, believed in the premises, but I, being a finitistic platonist, see Gödel's proof as a beautiful *reductio* proof that all statements that contain quantifiers are *a priori* meaningless, and only sometimes can be given an *a posteriori* meaning, when interpreted symbolically.

Very often one can *deconstruct* a seemingly 'infinitarian' statement by restating it *symbolically.*

The statement "$n + n = 2n$ for every integer n" is meaningless. It is only true for *every* **finite** integer. It is also true for *symbolic* n.

The statement "every integer has a successor" is meaningless, but one can say that $n + 1$ is the *symbolic* successor of n. Gödel's 'true' yet unprovable statements are simply statements that may not be resurrected for symbolic n. *A priori*, the statement "*there are infinitely many twin primes*" makes no sense, and neither does "*there are infinitely many primes*". *A posteriori* the latter can be made to make sense, by showing the validity of the algorithm implicit in Euclid's 2300-year-old proof, that manufactures 'yet another prime' (but symbolically!). I am sure that the twin-prime conjecture is also true, since it would turn out to be true for symbolic n. If $A(n)$ is the number of twin-prime pairs $\leq n$, then, some future sieving inequality (that will be found by computer!), will imply that

$$A(n) \geq C_1 \frac{n}{(\log n)^2} \quad ,$$

for *symbolic* n, and specific C_1, that would contradict the symbolic inequality $A(n) \leq C_2$, C_2 being a (symbolic!) constant.

The Chinese-Indian-Sumerian-Egyptian-Babylonian Model for Doing Mathematics

Euclid ruined mathematics by introducing that pernicious *axiomatic method* and making mathematics *deduction-centric.* But for thousands of years before Euclid, mathematics has been pursued *empirically* and *experimentally* and was *induction-centric.* It was what Richard Feynman called *Babylonian-style* mathematics. The reason Feynman liked it so much is that not only was it empirical, but it was also **algorithmic**.

Formalizing Algorithms: Turing Machines

Algorithms existed for at least five thousand years, but people did not know that they were algorithmizing. Then came Turing (and Post and Church and Markov and others) and *formalized* the notion. In the case of Turing, he introduced *Turing machines.* Of course, given an algorithm, it is nice to know that it is indeed an algorithm, and not just a Turing machine, in other words, that it *halts.* But the question "does T halt" is also meaningless. On the other hand: "does T halt in ≤ 1000 years" does make sense. So, by hindsight, just like in Gödel's case, it is not at all surprising that there in no decision algorithm for the halting problem. It was a stupid (in fact, worse, meaningless) question to begin with, and Turing just meta-proved that it was indeed very stupid to expect such an algorithm, and there is no way to make sense of it even *a posteriori*.

The Problem with the Chaitin-Kolmogorov Definition of Program-Size Complexity and Randomness

Greg Chaitin, and independently Andrey Kolmogorov and Ray Solomonoff, famously defined *program-size complexity* of a (finite or infinite) string as the *length of its shortest description* in some **fixed description language**. Now, that *description language* could be taken to be English, French, Hebrew, Spanish, or Chinese. But *natural* languages are notoriously fuzzy, and may be good media for love songs, but not for mathematics and computer science. The *lingua-franca* of theoretical computer science is the Turing machine. There are also numerous equivalent models, that are sometimes easier to work with. But even this is too vague, since we can't tell, thanks to Turing, whether our TM would halt or not, in other words whether it is a *genuine* algorithm or just an *algorithm wannabe*. Furthermore, even if it *does* halt, if my super-short computer program would take *googolplex* to the power *googolplex* years to generate my sequence, it can't do me much good. It is true that for *aesthetic* reasons, Greg Chaitin refused to enter time into his marvelous theory, and he preempted the criticism by the disclaimer that his theory is 'useless for applications'. But, I, for one, being, in part, a *naturalist*, find it hard to buy this nonchalance. Life is finite (alas, way too finite), and it would be nice to reconcile time-complexity with program-size complexity. Anyway, using *Turing machines* or any of the other computational models, for which the halting program is undecidable, makes this notion *meaningless*. Of course it has a great *metaphoric* and *connotative* meaning!.

So the notion of *Turing machine*-computable is way too general. Besides the Greek model, adopted by mathematicians and meta-mathematicians alike does not represent how most of mathematics is done in **practice**.

Most of mathematics, even logic, is done within narrow *computational frameworks*, sometimes explicit, but more often implicit. And what mathematicians do is *symbol-crunching* rather than *logical deduction*. Of course, formal logic is just yet another such symbolic-computational framework, and *in principle* all proofs can be phrased in that language, but this is *unnatural*, *inefficient*, and worse, sooo **boring**.

Let's call these computational frameworks **ansatzes**. In my humble opinion, mathematics should abandon the Greek model, and should **consciously** try to **explicate** more and more new ansatzes that formerly were only implicit. Once they are made explicit, one can teach them to our computers and do much more than any human.

The Ansatz Ansatz

Indeed, lots of mathematics, as it is *actually* practiced today, can be placed within well-defined *computational frameworks*, that are provably *algorithmic* and, of course, *decidable*. Sometimes the practitioners are aware of this, and in that case 'new' results are considered routine. For example, the theorem

$$198765487 \cdot 198873987 \;=\; 39529284877686669 \quad ,$$

is not very exciting today, since it belongs to the well-known class of **explicit arithmetical identities**.

On the other hand, the American Mathematical Monthly still publishes papers today in Euclidean Geometry, that, thanks to René Descartes, is reducible to *high-school algebra*, that is also routinely provable, of course in principle, but today also in practice, thanks to our powerful computer algebra systems.

The fact that multiplication identities are routinely provable is at least 5000-years old, and the fact that theorems in Plane Geometry are routinely-provable is at least 250-years old (and 40-year old in practice), but the fact that an identity like

$$\sum_{k=-n}^{n} (-1)^k \binom{2n}{n+k}^3 = \frac{(3n)!}{n!^3} \quad ,$$

discovered, and first proved in 1904 by Dixon, is also routinely provable, is only about 16-years old, and is part of so-called *Wilf-Zeilberger* Theory.

In each of these cases it is *nowadays* routine to prove an identity of the form $A = B$, since there is *a canonical form* algorithm $A \to c(A)$, and all we have to do is check that $c(A) = c(B)$. In fact, to prove that $A = B$, it suffices to have a *normal-form* algorithm, checking that $A - B$ is 'equivalent' to 0.

But before we can prove a statement of the form $A = B$, we have to find an *appropriate* **ansatz** to which they both belong.

At this time of writing, there are only a few explicitly known ansatzes. Let's first review one of my favorites.

The Polynomial Ansatz

David Hume is right that there is no formal, watertight, proof that the sun will rise tomorrow, since the Boolean-valued function

$$f(t) := evalb(The \quad Sun \quad Will \quad Rise \quad At \quad Day \quad t) \quad ,$$

has not yet been proved to belong to any known ansatz. Indeed, we now know, that for $t >> 0$, $f(t)$ is false, because the Sun will swallow Planet Earth, so all we can prove are vague probabilistic statements for small t (e.g. for $t =$ tomorrow).

The Clay Foundation is also right that there is not yet a formal, watertight, proof, of the Riemann Hypothesis, even though Andrew Odlyzko and Herman te Riele proved that the first ten billion, or whatever, complex zeros of $\zeta(s)$ lie on the critical line. This is because the sequence

$$f(n) := Re(z_n) \quad ,$$

where z_n is the n^{th} complex root of $\zeta(s) = 0$, has not yet been proved to belong to any known ansatz.

However, the following proof of the lovely identity

$$\sum_{i=1}^{n} i^3 = \left(\sum_{i=1}^{n} i\right)^2 \quad ,$$

is perfectly rigorous.

Proof: True for $n = 0, 1, 2, 3, 4$ (check!), hence true for all n. QED.

In order to turn this into a full-fledged proof, all you have to do is mumble the following incantation:

Both sides are polynomials of degree ≤ 4, hence it is enough to check the identity at five distinct values.

An Ansatz-based Chaitin-Kolmogorov Complexity

So let's define the *complexity* of an infinite (or finite) sequence always *relative* to a given *ansatz*, assuming that it indeed belongs to it. So our *descriptive language* is much more modest, but we can always determine its complexity, and everything is decidable. It does not have the *transcendental beauty* and *universal insight* of Chaitin's Algorithmic Information Theory, but on the other hand, we can always decide things, and nothing is unknowable (at least in principle).

It all depends on the data structure

Even within a specific ansatz, there are many ways of representing our objects. For example, since a polynomial P, of degree d is determined by its values at any $d+1$ values, we can represent it in terms of a finite sequence $[P(0), \dots, P(d)]$ that requires $d+1$ "bits" (units of information). Of course, we can also express it in the usual way, as a linear combination of the powers $\{1, n, n^2, \dots, n^d\}$, or in terms of any other natural base, for example $\{\binom{n}{k}, k = 0 \dots d\}$. Each of these data structures require $d+1$ "bits", in general, but in specific cases we can sometimes *compress* in order to get lower complexity. For example it is much shorter to write n^{1000} then to write $[0, 1, 2^{1000}, \dots, 1000^{1000}]$ (without the "…", and spelled-out).

The Strong N_0 property

An ansatz has the Strong N_0 property, if given any two sequences, A, B, within that ansatz, there exists an *easily computable* (say polynomial-time in the maximal size of A and B) number $N_0 = N_0(A, B)$ such that in order to prove that $A(n) = B(n)$ for all n, it suffices to prove it for any N_0 *distinct* values of n.

The *iconic example* of an ansatz having the strong N_0 property, already mentioned above, is the set of *polynomials*. For polynomials $P(x)$ of a single variable, $N_0(P(x))$ is $deg\, P + 1$. For a polynomial $P(x_1, \ldots, x_n)$ of degree d, $N_0(P)$ is $\binom{d+n}{n}$.

The Weak N_0 property

An ansatz has the Weak N_0 property, if in order to prove that $A(n) = B(n)$ (for all n), there exists an *easily computable* (say polynomial-time in the maximal size of A and B) number $N_0 = N_0(A, B)$ such that in order to prove that $A(n) = B(n)$ for all n, it suffices to prove it for the *first* N_0 values of n: $n = 1, n = 2, \ldots, n = N_0$.

A simple example of an ansatz that has the weak, but not the strong, N_0 property, are periodic sequences. If two sequences are known *a priori* to have periods d_1 and d_2, then if they are equal for the **first** $max(d_1, d_2)$ values, then they are identically equal. But the two sequences $f(n) := 1$ and $g(n) := (-1)^n$ coincide at infinitely many places (all the even integers), yet the two sequences are not identically equal.

Back to Science: The PEL Model

In Hugh G. Gauch's excellent book on the Scientific method "*Scientific Method in Practice*", he proposes the *PEL model*, PEL standing for "Presupposition, Evidence, Logic". So Hume's objection disappears if we are willing to concede that science is *theory laden*, and we have lots of presuppositions, both explicit and implicit.

Now the analog of presupposition in mathematics is ansatz. If we make the reasonable presupposition that the function

$$f(t) := evalb(The \quad Sun \quad Will \quad Rise \quad At \quad Day \quad t) \quad ,$$

belongs to the *constant* ansatz (at least for the next 100000 years), then checking it in *just one point*, say $t = today$, proves that the sun will indeed rise tomorrow.

On the other end, to prove that all emeralds are grue, presupposes that the color of emeralds belong to the *piece-wise constant* ansatz, since the notion of 'grue' belongs to it. In that case, $N_0 > 2050$, so indeed checking it for many cases but before 2050, does not suffice, even non-rigorously, to prove that all emeralds are grue.

The Probabilistic N_0 property

Sometimes N_0 is way too big, in other words, to get *complete certainty* will take too long. Then you might want to consider settling for $N_0(p)$.

An ansatz has the probabilistic N_0-property, if in order to prove that $A(n) = B(n)$ (for all n) with probability p, there exists an *easily computable* (say polynomial-time in the sizes of A and B) number $N_0(p) = N_0(A, B, p)$ such that it suffices to prove it for *any* $N_0(p)$ randomly chosen values of n.

The celebrated Schwartz-Zippel theorem establishes that multi-variable polynomials satisfy the $N_0(p)$ property (in addition to having the N_0 property, of course), and that $N_0(.9999999)$ is much smaller than $N_0(1)$, so it is stupid to pay for full certainty.

An Embarrassing Paper of Mine

Can you envision a professional mathematician publishing a paper entitled "A bijective proof of $10 \times 5 = 2 \times 25$", by concocting a nice bijection? Of course not! Today, all explicit arithmetical identities are known to be routinely provable.

Yet something analogous happened to me. In my web-journal, I published a paper that found an 'elegant' combinatorial proof of the identity

$$\sum_{i=0}^{2n} \binom{2n}{i} F_{2i} = 5^n F_{2n} \quad ,$$

where F_n are the Fibonacci numbers defined by $F_0 = 0, F_1 = 1, F_n = F_{n-1} + F_{n-2}$ $(n \geq 2)$. It was in response to a challenge by Arthur Benjamin and Jennifer Quinn, posed in their delightful book "*Proofs that really count*".

As "elegant" and "insightful" as my proof may have been, in Occam's and Chaitin's sense, the following proof is much more elegant.

Proof: Both sides are sequences that are solutions of second-order linear recurrence equations with constant coefficients. Hence, to prove that they coincide for all $n \geq 0$, it suffices to check that they coincide for $n = 0, 1, 2, 3$. Now just check that indeed

$n = 0: \quad 1 \cdot 0 = 0 \quad ,$

$n = 1: \quad 1 \cdot 0 + 2 \cdot 1 + 1 \cdot 3 = 5 \cdot 1$

$n = 2: \quad 1 \cdot 0 + 4 \cdot 1 + 6 \cdot 3 + 4 \cdot 8 + 1 \cdot 21 = 5^2 \cdot 3,$

$n = 3: \quad 1 \cdot 0 + 6 \cdot 1 + 15 \cdot 3 + 20 \cdot 8 + 15 \cdot 21 + 6 \cdot 55 + 1 \cdot 144 = 5^3 \cdot 8 \quad . \quad QED!.$

Of course, we have to *justify* the claims that both sides are solutions of linear recurrence equations with constant coefficients (by the way, such sequences are called C-finite), of second order. But these follow from the following easy claims, that can be proved *once and for all*, using elementary linear algebra (you do it!).

Claim 1: If a_n is a solution of a linear recurrence equation with constant coefficients of order d, then for any positive integer L

$$b_n := a_{nL} \quad ,$$

is likewise a solution of a (different) linear recurrence equation with constant coefficients of order d.

Claim 2: If a_n is a solution of a linear recurrence equation with constant coefficients of order d, then its **binomial transform**,

$$b_n := \sum_{i=0}^{n} \binom{n}{i} a_i \quad ,$$

is likewise a solution of a (different) linear recurrence equation with constant coefficients of order d.

Claim 3: If a_n is a solution of a linear recurrence equation with constant coefficients of order d, then $k^n a_n$ is likewise a solution of a (different) linear recurrence equation with constant coefficients of order d.

Claim 4: The algebra of C-finite sequences has the weak N_0-property, and two C-finite sequences of order $\leq d$ are identical if they are identical for $0 \leq n \leq 2d-1$.

For more complicated identities involving C-finite sequences, the following is needed.

Claim 5: If a_n and b_n are C-finite sequences of orders d_1 and d_2, then $a_n + b_n$ and $a_n b_n$ are C-finite sequences of orders $\leq d_1 + d_2$ and $\leq d_1 \cdot d_2$ respectively.

Since any polynomial sequence is C-finite (a polynomial of degree d satisfies the recurrence $(N-1)^{d+1} f(n) = 0$, where N is the forward-shift operator), it follows that the ansatz of C-finite sequences is a superset of the polynomial ansatz. The next ansatz is even bigger, and contains that of C-finite sequences.

The Schützenberger Ansatz

If the generating function of a sequence $\{a(n)\}_{n=0}^{\infty}$,

$$\phi(x) = \sum_{n=0}^{\infty} a(n) x^n \quad ,$$

satisfies a polynomial equation:

$$P(\phi(x), x) = 0 \quad ,$$

then it is called an *algebraic* formal power series. I call it the Schützenberger Ansatz, since it was Marco Schützenberger's favorite ansatz, and has gotten lots of attention by his illustrious disciple Xavier Viennot and Viennot's disciple the brilliant Mireille Bousquet-Mélou, and numerous others at the *école bordelaise.*

This is also an algebra, and every identity is decidable, and it, too, has the weak N_0 property.

Solving Functional Equations Empirically (Yet Rigorously!)

In many combinatorial problems, one is interested in a formal power series $F(x, y; t)$ that satisfies a functional equation of the form

$$A(x,y,t)F(x,y;t) + B(x,y,t)F(0,y;t) + C(x,y,t)F(x,0;t) + D(x,y,t)F(0,0;t) = E(x,y,t) \quad , \qquad (FunEq)$$

where A, B, C, D, E are polynomials in (x, y, t). Such an equation can be used to crank out the Maclaurin expansion of $F(x, y; t)$ to any desired order.

Often, we are really only interested in $\phi(t) := F(0, 0; t)$, that sometimes, surprisingly, happens to satisfy some nice algebraic equation $P(t, \phi(t)) = 0$, for no apparent reason, and the challenge is to prove that fact. $(FunEq)$ can't be used directly, since it involves (x, y, t) not just t, and plugging-in $x = 0, y = 0$ in $(FunEq)$ usually yields the fact $0 = 0$, that while true, is far from new, and does not help us with the conjecture at hand.

While we are **not** guaranteed, *a priori*, that this is the case, it is still worthwhile to try to conjecture that not only $\phi(t)$ is algebraic, but so is the full $F(x, y; t)$, i.e. there exists a polynomial $Q(F, x, y, t)$ such that

$$Q(F(x, y; t), x, y, t) \equiv 0 \quad . \qquad (AlgEq)$$

This Q can be found, empirically for now, by the method of *undetermined coefficients*. Use $(FunEq)$ to crank out the first 1000 or whatever terms of F, call the truncated version $\widetilde{F}$; let Q be a generic polynomial of four variables of a guessed degree d, with undetermined coefficients; ask the computer to compute $Q(\widetilde{F}(x, y; t), x, y, t)$, set the first 1000 terms to 0, get a huge system of equations for the undetermined coefficients, and solve them. If there is a non-zero solution, then it is great news! Otherwise, make d bigger, or give up.

Once we (or rather our computer) conjectured such a general algebraic equation, how do we prove it rigorously?

We have to prove that $(FunEq)$ implies $(AlgEq)$. By uniqueness, we can prove that $(AlgEq)$ implies $(FunEq)$. Defining $G(x, y; t)$ to be the unique solution of

$$Q(G(x, y; t), x, y, t) \equiv 0 \quad , \qquad (AlgEq)$$

it follows that $G(x, 0; t)$, $G(0, y; t)$, $G(0, 0; t)$ are all algebraic:

$$Q(G(x, 0; t), x, 0, t) \equiv 0 \quad , \qquad (AlgEq')$$

$$Q(G(0, y; t), 0, y, t) \equiv 0 \quad , \qquad (AlgEq'')$$

$$Q(G(0, 0; t), 0, 0, t) \equiv 0 \quad . \qquad (AlgEq''')$$

Now

$$H(x, y, t) := A(x, y, t)G(x, y; t) + B(x, y, t)G(0, y; t) + C(x, y, t)G(x, 0; t) + D(x, y, t)G(0, 0; t) - E(x, y, t)$$

is also algebraic and using the "Schützenberger calculator" one can find an equation satisfied by it, and prove that H is identically 0, and by uniqueness, $F = G$.

Now plugging-in $x = 0, y = 0$, into the *now-proved* algebraic equation $Q(F(x, y, t), x, y, t) = 0$, would yield a rigorous proof of the conjectured algebraic equation for $\phi(t) = F(0, 0, t)$, namely $Q(\phi(t), 0, 0, t) \equiv 0$.

The downside in the above empirical (yet *a posteriori* rigorous!) approach, is that the computations required to conjecture Q are very heavy, and for all but the simplest problems, the above method is beyond today's computers. Also, in practice it is more efficient to first conjecture algebraic equations for $F(x, 0; t)$ and $F(0, y; t)$ and use the "calculator" to derive what the algebraic equation for the $F(x, y; t)$ should be.

A yet more powerful ansatz, that contains all those considered so far is the *Holonomic Ansatz*, that is my absolute personal favorite.

The Holonomic Ansatz

A sequence $\{a(n)\}$ is holonomic if it satisfies a *linear recurrence equation* with **polynomial** coefficients. The sum and product of holonomic sequences is again holonomic, and one has a 'holonomic calculator' (The Salvy-Zimmerman Maple package `Gfun`).

Introducing the shift operator $Nf(n) := f(n+1)$, one can define a holonomic sequence in terms of its *annihilating operator* $P(N, n)$ and the initial conditions.

A discrete function of several variables $a(n_1, \ldots, n_k)$ is holonomic if for each variable n_i there is an annihilating operator $P_i(n_1, \ldots, n_k\,;\, N_i)$. This is the basis for so-called Wilf-Zeilberger theory and it is not only closed with respect to addition and multiplication, but also with respect to sums. For example, if $F(n, k)$ is holonomic, then $a(n) := \sum_k F(n, k)$ is holonomic as well.

Functional Equations and Holonomic Functions

Analogous remarks about the interface between functional equations and algebraic formal power series apply for finding a possible holonomic representation for a formal power series given as a solution of a functional equation.

In Search of New Ansatzes

The above ansatzes are just some of those known today. I am sure that the future will bring lots of new ansatzes that will trivialize and routinize large parts of mathematics.

Pólya's Heuristics Applied to Computer Generated Mathematics

One principle George Pólya was very fond of was "finding the right generalization". Suppose that you conjecture that $A(n) = B(n)$ but you can only prove it for $1 \le n \le 7$, because it takes too much time and space to verify it for $n = 8$ and beyond. Of course you can't generalize from seven cases! But if you can find *two-parameter* objects $C(m, n)$ and $D(m, n)$ such that $A(n) = C(n, n)$ and $B(n) = D(n, n)$, and you can prove that $C(n, m) = D(n, m)$ for $1 \le n \le 7$, for *all* $m \ge 0$, then the conjecture $C = D$ is true for infinitely many cases, so $C = D$ is very plausible, and hence $A = B$.

A Very Simple Toy Example

Let $A(n)$ be the number of words in the alphabet $\{1,2\}$ with exactly n 1's and exactly n 2's.

By direct enumeration you find that

$$A(0) = 1 \quad , \quad A(1) = 2 \quad , \quad A(2) = 6 \quad , \quad A(3) = 20 \quad , \quad A(5) = 252 \quad , \quad A(6) = 924 \quad ,$$

and this leads you to conjecture that $A(n) = B(n)$ where $B(n) = (2n)!/n!^2$.

How would you go about proving this conjecture?

Let's consider the *more general* problem of finding $C(m,n)$, the number of words in the alphabet $\{1,2\}$ with exactly m 1's and exactly n 2's. Then $C(m,0) = 1$, and you have the following recurrence, easily derived by looking at the number of 2's to the left of the rightmost 1:

$$C(m,n) = \sum_{i=0}^{n} C(m-1,i) \quad , \tag{1}$$

from which you can easily deduce the following special cases:

$$C(m,1) = \binom{m+1}{1} \quad , \quad C(m,2) = \binom{m+2}{2} \quad , \quad C(m,3) = \binom{m+3}{3} \quad ,$$

that naturally leads to the conjecture $C(m,n) = D(m,n)$, where $D(m,n) = \binom{m+n}{n}$. It can be verified for $n \leq 10$ easily by using (1) with specific n but general m, by *only* using polynomial summation. **Now** the more general statement, $C = D$, is much more plausible. Besides, this more general conjecture is much easier to prove, since you have more elbow room, and it is easy to prove that both $X = C$ and $X = D$ are solutions of the linear *partial recurrence boundary-value problem*:

$$X(m,n) = X(m-1,n) + X(m,n-1) \quad , \quad X(m,0) = 1 \quad , \quad X(0,n) = 1 \quad .$$

So in this case finding the right generalization first made our conjecture much more plausible, and then also made it easy to prove.

How to do it the hard way

In order for you to appreciate how much trouble could be saved by introducing a more general conjecture, let's do it, the **hard way**, sticking to the original one-parameter conjecture.

Let $b(n)$ be the number of words in $\{1,2\}$ with exactly n 1's and with exactly n 2's such that in addition , for any proper prefix, the number of 1's always exceeds the number of 2's. Analogously, Let $b'(n)$ be the number of words in $\{1,2\}$ with exactly n 1's and with exactly n 2's such that in addition the number of 2's always exceeds the number of 1's except at the beginning and end. By symmetry $b(n) = b'(n)$.

Then we have the non-linear recurrence

$$a(n) = \sum_{m=0}^{n-1} a(m)(b(n-m) + b'(n-m)) = 2\sum_{m=0}^{n-1} a(m)b(n-m) \quad . \tag{2}$$

obtained by looking at the longest (proper) prefix with the same number of 1's and 2's. Also, using a standard combinatorial argument, $b(n)$ can be shown to satisfy a non-linear recurrence

$$b(n) = \sum_{m=1}^{n-1} b(m)b(n-m) \quad ,$$

from which you can crank out many values of $b(n)$, that in turn, enable you to crank out many values of $a(n)$, and make your conjecture much more plausible. Using the above non-linear recurrence, you can generate the first few terms of the sequence $\{b(n)\}_{n=1}^{\infty}$: $1, 1, 2, 5, 14, 42, 132, \ldots$, and easily guess that $b(n) = \frac{(2n-2)!}{(n-1)!n!}$, and to prove it rigorously, all you need is to verify the binomial coefficient identity

$$\frac{(2n-2)!}{(n-1)!n!} = \sum_{m=1}^{n-1} \frac{(2m-2)!}{(m-1)!m!} \cdot \frac{(2n-2m-2)!}{(n-m-1)!(n-m)!} \quad ,$$

that can be done automatically with the WZ method, and then prove the identity

$$\binom{2n}{n} = 2\sum_{m=0}^{n-1} \binom{2m}{m} \cdot \frac{(2n-2m-2)!}{(n-m-1)!(n-m)!} \quad ,$$

that is likewise WZable.

Note: One can also do it, of course, with generating functions, staying within the Schützenberger ansatz rather than the holonomic anstaz. But it is still much harder than doing it via the 2-parameter generalization discussed above.

Pólya's Ode to Incomplete Induction

In Polya's masterpiece on the art of mathematical *discovery*, "**Induction and Analogy in Mathematics**" he lauded the use of incomplete induction as a powerful *heuristic* for discovering mathematical conjectures, and as a tool for discovering possible proofs. In particular he cites approvingly the great Euler who conjectured, long before he had a formal proof, many interesting results. For example:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad ,$$

that he verified numerically to six decimal places, noting that this implies that the probability that the left side and right side coincide by accident is less than one in a million. Many years later he found a complete proof, but he first had a "cheating proof" that proceeded by pretending that infinite products are like polynomials. Another notable example was the pentagonal number theorem, that he conjectured, and deduced important consequences from, based on expanding it to eighty terms. Only 25 years later did he find a formal proof.

Undoubtedly, the greatest conjecturer of all time was Srinivasa Ramanujan, who only needed very few special cases to formulate a conjecture, and was very seldom wrong.

The Law of Small Numbers

The conventional wisdom *against* the use of incomplete induction is called the *law of small numbers*, and there are many cases, many of them collected by Richard Guy in his two *Monthly* papers about that "Law", that should be cautionary tales against having insufficient data and "jumping to conclusions".

We all know the joke about the mathematician, physicist, and engineer, the mathematician saying "1 is a prime, 3 is a prime, 5 is a prime, 7 is a prime", hence all odd numbers are primes.

Deeper victims of the law of small numbers were Margie Readdy and Richard Ehrenborg who conjectured that the number of up-down involutions of length $2k$ is $k!$, based on data for $k = 1, 2, 3, 4, 5$. This was disproved, by Shalosh B. Ekhad, for $k = 6$ (and beyond). Later Richard Stanley explained why the sequence starts out like that.

Sometimes even 9 terms do not suffice. Neil Sloane, the great master-sequencer, pointed my attention to sequences **A0060041** and **A076912**, in his legendary **database**, that are known to be equal up to $n = 9$, but are believed to disagree for $n = 10$.

But the greatest source of such horror stories is *number theory.*

We all know how the great Fermat goofed when he conjectured that $2^{2^n} + 1$ is always prime, based on the five cases $n = 0, 1, 2, 3, 4$.

Another scary story involves a stronger version of the Riemann Hypothesis, due to Mertens. Recall that the Riemann Hypothesis is equivalent to the statement that the partial sums of the Möbius function:

$$M(n) := \sum_{i=1}^{n} \mu(i) \quad ,$$

satisfy

$$|M(n)| \leq C(\epsilon) n^{1/2+\epsilon} \quad .$$

Mertens, in 1897, conjectured the stronger conjecture that $|M(n)| \leq n^{1/2}$, and it was verified for n up to a very large number. Yet in 1985, Andrew Odlyzko and Herman te Riele disproved it.

Another notorious example concerns the Skewes Number, that is the smallest n for which $\pi(n)$, the number of prime numbers $\leq n$, is larger than $li(n)$, the logarithmic integral. No one knows its exact value, but it seems to be very large.

Inequalities vs. Equalities

By hindsight, it is not surprising that both $\pi(n) < li(n)$ and $|M(n)| < \sqrt{n}$ turned out to be false, even though they are true for so many values of n. First, prime numbers are very hazardous, and since often we have $\log\log$ and $\log\log\log$ showing up, it is reasonable to suspect that what seems large for us is really peanuts. But a better reason to distrust the ample empirical evidence is that *inequalities* need much more evidence than *equalities*.

A trivial example is the following. To prove that $P(x) = 0$ for a polynomial P of degree $\leq d$ (say given in some complicated way that is not obviously 0, for example $(x^4+1)(x+1) - x^5 - x^4 - x - 1$) it suffices to check $d+1$ special cases, but consider the "conjecture"

$$\frac{x}{10000000000000} - 1 < 0 \quad .$$

The left side is a polynomial of degree 1 in x, and the "conjecture" is true for the first 10000000000000 integer values of x, yet, of course, it is false in general.

The Art of Plausible Reasoning

Given a conjecture $P(n)$, depending on an integer parameter n, that has been verified for $1 \leq n \leq M$, how plausible is it?

If it has the form $A(n) \leq B(n)$, then no matter how big M, it would be very stupid to jump to conclusions, as shown in the above examples.

From now we will assume that it can *naturally* be phrased in the form $A(n) = B(n)$. Granted, every assertion $P(n)$, even an inequality like "$|M(n)| < \sqrt{n}$", is logically equivalent to an *equality* :

$$evalb(P(n)) \equiv true \quad ,$$

where $evalb(p)$ is true or false according to whether p is true or false. But of course this is contrived.

The most secure scenario is when *both* A and B are known to belong to a decidable *ansatz* with the strong or weak N_0 property, and it is easy to compute N_0, and it so happened that $M \geq N_0$. Then we immediately have a *rigorous* proof.

Next in line is when A and B both belong to an ansatz with the $N_0(p)$ property and $M \geq N_0(.999999)$ or whatever.

Next in line, as far as plausibility goes, is when there is a strong *heuristic* evidence, inspired by analogy and past experience, that both A and B belong to a *known* ansatz with an N_0 property, M is fairly large, and both A and B are not too complicated.

After that, in the certainty pecking-order, are cases where you have no ansatz in mind to which A and B may possibly belong, but you can *feel it in your bones* that there is a *yet to be discovered ansatz* that would have the N_0 property, and M is fairly large and A and B are not too complicated.

Finally, if the conjecture is so far-out or artificial, or A and B are so different, so that you have no reason to hope that there is a yet-to-be-discovered ansatz that would 'trivialize' $A = B$, and M is not that big, then I wouldn't even make a conjecture.

Also keep in mind the above remarks of finding the right generalization from a one-parameter identity to a multiple-parameter one, that not only can add plausibility to our conjecture, by verifying it for infinitely many cases, but sometimes also enables a formal proof.

Don't Get Hung-Up on the N_0-approach

In my eyes, an N_0 proof is the *most elegant.* It is also the **most fun**, since it defies that old and corny platitude, we mathematicians grew up with, that

"**checking finitely many cases, no matter how many, does not constitute a proof**".

But often the N_0 is way too big, and it may not be the most efficient way to prove identities. For example, for the identity

$$(n^{10000000} - 1)(n^{10000000} + 1) = n^{20000000} - 1 \quad ,$$

it would be stupid to verify it for $1 \leq n \leq 20000001$. Just use the "usual" algorithm for multiplying polynomials.

After all the N_0 approach is just *one* algorithm for proving identities within a given ansatz, and not necessarily always the most efficient one.

The Wilf-Zeilberger Algorithmic Proof Theory

A less trivial example of an ansatz that has the N_0 property, but using it is usually not feasible, is the WZ algorithmic proof theory, that can prove any conjectured identity of the form

$$\sum_{k=0}^{n} F(n,k) = \sum_{k=0}^{n} G(n,k) \quad ,$$

whenever $F(n,k)$ and $G(n,k)$ are products of binomial coefficients. By general nonsense we know that the sequence

$$a(n) := \sum_{k=0}^{n} F(n,k) - \sum_{k=0}^{n} G(n,k) \quad ,$$

is holonomic, i.e. satisfies a homogeneous linear recurrence equation with polynomial coefficients:

$$\sum_{i=0}^{L} p_i(n) a(n+i) \equiv 0 \quad ,$$

for some non-negative integer L and some polynomials $p_0(n), \dots, p_L(n)$. It is fairly easy to find relatively small *a priori* upper bounds for L, without actually finding the recurrence. If we knew beforehand that the leading coefficient, $p_L(n)$, has no positive integer zeros, then we could immediately deduce that $a(n)$ is identically 0 once it vanishes for $0 \leq n < L$. Lily Yen, in a 1993 Ph.D. thesis, written under the direction of Herb Wilf, found *a priori* bounds for the largest positive integer root of $p_L(n) = 0$, but they were enormous. It is possible that another approach could bring it down, but why bother? Yen's thesis was interesting *theoretically*, since it showed that WZ theory has the (weak) N_0 property, but as far as actually proving specific identities, it is much more efficient to use the Zeilberger algorithm to actually manufacture the recurrence, and then just look at $p_L(n)$ and convince ourselves that it has no positive integer roots, and if it does find them.

What is Mathematical Knowledge?; Reliablism

The standard definition of *knowledge* (see, e.g., Kwame Anthony Appiah's excellent introduction to contemporary philosophy, "**Thinking it Through**") is:

"**justified true belief**".

The problem is then "how justified is justified". In science one is willing to take ample empirical evidence as sufficient justification, but in mathematics, traditionally, one insisted on a formal rigorous proof, proved by human means, since a "proof by computer is only a physical experiment", and "you can't trust a computer", since programs have so many bugs.

Appiah talks about a movement in contemporary epistemology, co-pioneered by my Rutgers colleague Alvin Goldman, called **reliablism** (see also the Wiki entry), that modifies the definition of knowledge to be **true belief justified reliably**. The problem then is to introduce reliability standards.

I strongly believe that very soon most of serious mathematics will be computer-generated, and all the rest would be computer-assisted, so we do need to develop quality-control to maximize the chances that the computer-generated proofs are indeed valid.

One way to maximize reliability is to adopt what I call the **method of overlapping steps**.

Suppose that you have to devise an algorithm to do $S(n), n = 0, 1, 2, \ldots$, and you want to do it for as large n as possible. You should first write the **most naive program**, easy to write, and easy to check. Then let the computer output $S(0), S(1), \ldots, S(L_0)$, with L_0 rather small.

Unfortunately, the naive approach can't go very far. So you write a more sophisticated program, good for $n \leq L_1$, and compare its output with the output of the previous program for $n \leq L_0$. Then you write yet another, even more sophisticated program, valid for $n \leq L_2$ and check it against the previous ones, and so on and so forth.

It can also help if you have two entirely different approaches to tackle the same problem, and if the outputs match, then it is a great indication that they are *both* indeed correct.

There are hardly any isolated facts. As already noticed by Quine, all knowledge, in particular science and mathematics, consist of intricate *webs*. In the case of computer-generated mathematics, if all your programs are working together without contradiction, this fact simultaneously testifies that they are *all* OK. It is a little like the way computer scientists generate random bits at a fraction of the normal cost by using **expanders**.

How Necessary is Necessary and How Contingent is Contingent

Many of the traditional philosophical dichotomies like analytic/synthetic, *a priori*/a posteriori, induction/deduction, and especially necessary/contingent collapse once we realize that the **mathematical** universe is the **same** as the **physical** universe, and that our **unique** universe is **finite**. Also that everything is **computation**.

I will only dwell on the necessary/contingent dichotomy.

According to traditional thinking, the fact that the speed of light is constant is *contingent*, while the fact that the 100^{th} (decimal) digit of π is 9 is **necessary**. Nonsense. They are both necessary and both contingent. As Greg Chaitin said so beautifully about the digits of Ω, "they are true for no particular reason". But, even if you don't believe in Ω, lots of mathematical facts are, in some sense, contingent, and lots of efforts goes into explaining identities of the form $A = B$ by trying to *explain* them.

Alas, paraphrasing Greg, if the "explanation" is longer than the explanandum, then it is not much of an explanation.

So, the statement

"Amongst any eleven consecutive digits of π, two must be the same"

is much more necessary than the statement

"the 100^{th} digit of π is 9" ,

since the former is a special case of a **universal** result called the **pigeon-hole principle**, with **two** parameters m and n:

$P(m, n)$: If $m > n$ and m pigeons must be placed in n pigeon-holes, then at least two pigeons must be pigeon-hole-mates.

So a numeric result $a = b$, with a and b both numbers (in other words, they depend on zero parameters), is contingent, even if you have a formal proof. It is less contingent if it is a special case of $A(n) = B(n)$ with $n = n_0$. It is even less contingent if it is a special case of $A(m, n) = B(m, n)$ with $m = m_0, n = n_0$, and so on.

Depth vs. Elegance

The **depth** of a mathematical result is the smallest amount of computer-time it takes to prove it (within our ansatz). The **elegance** of a statement is how short is its statements.

Paul Erdös believed that God has a book with elegant (i.e. short) proofs of all theorems. I hope that he is wrong. Theorems with short proofs are **shallow**, and my favorite results are short statements that require long proofs.

Towards a New Kind of Mathematical Aesthetics

Truth is Beauty and Beauty is Truth, or so goes the Keatsian cliche. If Beauty is *elegance*, *symmetry*, and *shortness*, then Beauty is just *trivial* Truth. But if you care about *deep truth*, then you have to give up on the traditional standards of beauty.

Why is the Computer-Generated Proof of the Four Color Theorem so Beautiful in my Eyes?

Because the *idea* of the proof can be encapsulated in one short phrase:

There exists an unavoidable set of reducible configurations.

The rest are just details on how to teach the computer to construct such a set, and verify that it is indeed what we want.

In this case, a major open problem was reduced to finding **one** object, that can, and indeed was, searched for, and found, by computer. That **one**, specific, object **certified** that the statement of the Four Color Theorem was indeed true.

A proof of an identity in WZ theory also consists of displaying **one**, finite, object, the **WZ certificate**, that certifies its correctness. Thereby, apparently, proving "infinitely many cases". Of course, these 'infinitely' many numerical facts are just trivial consequences of just **one** symbolic fact.

Towards an Ansatz Based Mathematics and Meta-Mathematics

All thinking requires *logic*, but the *informal* logic of normal mathematical discourse is good enough. The reductionist attempts of logicism and formalism to reduce mathematics, at least in principle, to formal logic was unfortunate, even for human-generated mathematics, but especially for computer-generated mathematics. I believe that the logic-based approach that predominates **Automatic Theorem Proving** is not entirely satisfactory.

Also the "abstract nonsense", structuralist, approach, preached by Bourbaki, was not quite the right approach for humans, and is definitely not suited for computers. Hopefully, the future will bring some synthesis, but, at present, one should try to base mathematical research on *ansatzes.* Major breakthroughs will come not by solving specific open problems, and not even devising new human theories, but by finding new and powerful ansatzes where the open problems can be embedded. It is much more efficient to solve geometry problems using **algebra**, by using analytic geometry, rather than by **logic**, using synthetic geometry.

The traditional dichotomy between **numerical**, empirical, facts, and general, **theoretical** results, is only illusionary. In the eyes of God, $2+2=4$ is just as interesting as Fermat's Last Theorem. It is true that "$2+2=4$" has zero free parameters, while FLT has four $[P(a,b,c,n) := a^n+b^n-c^n \neq 0$ (if $n>2, abc \neq 0)]$, but this is a *quantitative* difference not a *qualitative* one.

Traditionally $(n+1)(n-1) = n^2 - 1$ is a "theorem", true for infinitely many n, while $3 \cdot 5 = 4^2 - 1$ is just one fact. However, viewed symbolically, they are both facts, the former with one parameter, and the latter with zero parameters.

To prove that 15 is not prime, all you have to do is come up with a factorization: $3 \cdot 5 = 15$. For large numbers, this is considered a difficult computational problem, but "conceptually" it is trivial, or so the conventional wisdom says.

To prove that

$$\sum_{k=-n}^{n} (-1)^k \binom{2n}{n+k}^3 = \frac{(3n)!}{n!^3} \quad ,$$

requires a proof, since this is a general statement, valid for all n, but thanks to WZ theory, there is just **one** object, a certain rational function $R(n,k)$, that certifies it. That certificate can be obtained empirically and algorithmically. So the 'proof' is just one object, like the pair $(3,5)$ in the case of the 'theorem' that 15 is composite.

If desired, it is always possible to convert a 'certificate proof' to a formal logic proof, but this is very artificial, and unnecessary.

Let's conclude this manifesto with:

Mathematicians and meta-mathematicians of the world unite, you have nothing to lose but your logic chains! Let's work together to develop an ansatz-based mathematics and meta-mathematics.