THEOREMS FOR A PRICE: Tomorrow's Semi-Rigorous Mathematical Culture

Doron Zeilberger[1]

Today

The most fundamental precept of the mathematical faith is *thou shalt prove everything rigorously.* While the practitioners of *mathematics* differ in their view of what constitutes a rigorous proof, and there are fundamentalists who insist on even a more rigorous rigor than the one practiced by the mainstream, the belief in this principle could be taken as the *defining property* of *mathematician.*

The Day After Tomorrow

There are writings on the wall that, now that the silicon savior has arrived, a new testament is going to be written. Although there will always be a small group of "rigorous" old-style mathematicians(e.g. [JQ]) who will insist that the true religion is theirs, and that the computer is a false Messiah, they may be viewed by future mainstream mathematicians as a fringe sect of harmless eccentrics, like mathematical physicists are viewed by regular physicists today.

The computer has already started doing to mathematics what the telescope and microscope did to astronomy and biology. In the future, not all mathematicians will care about absolute certainty, since there will be so many exciting new facts to discover: mathematical pulsars and quasars that will make the Mandelbrot set seem like a mere Jovian moon. We will have (both human and machine[2]) professional *theoretical* mathematicians, who will develop conceptual paradigms to make sense out of the empirical data, and who will reap Fields medals along with (human and machine) *experimental* mathematicians. Will there still be a place for *mathematical* mathematicians?

This will happen after a transitory age of *semi-rigorous mathematics*, in which identities (and perhaps other kinds of theorems) will carry price-tags.

A Taste Of Things To Come

To get a glimpse of how mathematics will be practiced in the not too far future, I will describe the case of algorithmic proof theory for *hypergeometric identities*[WZ*][Z*][Ca]. In this theory, it is possible to rigorously prove, or refute, any conjectured identity belonging to a wide class of identities, that includes most of the identities between the classical special functions of mathematical physics.

Any such identity is proved by exhibiting a *proof certificate*, that reduces the proof of the given

---

[1] Department of Mathematics, Temple University, Philadelphia, PA 19122, USA. `zeilberg@math.temple.edu` . Supported in part by the NSF. Based in part on a Colloquium talk given at Rutgers University.

[2] For example, my computer Shalosh B. Ekhad, and its friend Sol Tre, already have a non-trivial publication list, e.g. [E], [ET]

identity to that of a finite identity among rational functions, and hence, by clearing denominators, to that between specific polynomials.

This algorithm can be performed successfully on all "natural identities" we are now aware of. It is easy, however, to concoct artificial examples for which the running time, and memory, are prohibitive. Undoubtedly, in the future "natural" identities will be encountered whose complete proof will turn out to be not worth the money. We will see later, how, in such cases, one can get "almost certainty" with a tiny fraction of the price, along with the assurance that if we robbed a bank, we would be able to know for sure.

This is vaguely reminiscent of *transparent proofs* introduced recently in theoretical computer science[Ci][ALMSS][AS]. The result that there exist short theorems having arbitrarily long proofs, a consequence of Gödel's incompleteness theorem, also comes to mind [S].[3] I speculate that similar developments will occur elsewhere in mathematics, and will "trivialize" large parts of mathematics, by reducing mathematical truths to routine, albeit possibly very long, and exorbitantly expensive to check, *"proof certificates"*. These proof certificates would also enable us, by plugging in random values, to assert "probable truth" very cheaply.

Identities

Many mathematical theorems are *identities*: statements of type "=", which take the form $A = B$. Here is a sampler, in roughly an increasing order of sophistication.

1. $2 + 2 = 4$.

2. $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

3. $sin(x + y) = sin(x)cos(y) + cos(x)sin(y)$.

4. $F_{n+1}F_{n-1} - F_n^2 = (-1)^{n+1}$.

5. $(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$.

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

7. Let $(q)_r := (1 - q)(1 - q^2)\dots(1 - q^r)$, then

$$\sum_{r=0}^{n} \frac{q^{r^2}}{(q)_r (q)_{n-r}} = \sum_{r=-n}^{n} \frac{(-1)^r q^{(5r^2-r)/2}}{(q)_{n-r}(q)_{n+r}} \quad .$$

7'. Let $(q)_r$ be as in 7.

[3] Namely, the ratio (proof length)/(theorem length) grows fast enough to be non-recursive. Adding an axiom can shorten proofs by recursive amounts [G], [D].

$$\sum_{r=0}^{\infty} \frac{q^{r^2}}{(q)_r} = \prod_{i=0}^{\infty} (1-q^{5i+1})^{-1}(1-q^{5i+4})^{-1} \quad .$$

8. Let H_n be given by,

$$H_n = H_n(q) = \frac{(1+q)}{(1-q)} \frac{(1+q^2)}{(1-q^2)} \cdots \frac{(1+q^n)}{(1-q^n)} \quad ,$$

then

$$\left(\sum_{k=0}^{n} \frac{2(-q^{n+1})^k}{1+q^k} H_k \right)^4 \sum_{k=-n}^{n} \frac{4(-q)^k}{(1+q^k)^2} \frac{H_{n+k}}{H_n} \frac{H_{n-k}}{H_n} = (\sum_{k=-n}^{n} (-q)^{k^2})^4 \ .$$

8'.

$$(\sum_{k=-\infty}^{\infty} q^{k^2})^4 = 1 + 8 \sum_{k=1}^{\infty} \frac{q^k}{(1+(-q)^k)^2} \ .$$

9. Analytic Index=Topological Index.

10. $Re(s) = \frac{1}{2}$ for every non-real s such that $\zeta(s) = 0$.

All the above identities are trivial, except possibly the last two, which I think quite likely will be considered trivial in two hundred years. I will now explain.

Why are the first 8 identities trivial?

The first identity, while *trivial* nowadays, was very deep when it was first discovered, independently, by several anonymous cave-dwellers. It is a general, abstract theorem, that contains, as special cases, many apparently unrelated theorems: Two bears and Two bears make Four bears, Two apples and Two apples make Four apples, etc. It was also realized that in order to prove it rigorously, it suffices to prove it for any one special case, say, marks on the cave's wall.

The second identity: $(a+b)^3 = a^3+3a^2b+3ab^2+b^3$, is of one level of generality higher. Taken literally (in the *semantic* sense of the word *literally*), it is a fact about *numbers*. For any specialization of a and b we get yet another correct numerical fact, and as such it requires a "proof", invoking the commutative, distributive and associative "laws". However, it is completely routine when viewed *literally*, in the syntactic sense, i.e. in which a and b are no longer symbols *denoting* numbers, but rather represent themselves, qua (commuting) literals. This shift in emphasis roughly corresponds to the transition from Fortran to Maple, i.e. from *numeric* computation to *symbolic* computation.

Identities **3** and **4** can be easily embedded in classes of routinely verifiable identities in several ways. One way is by *defining* $cos(x)$ and $sin(x)$ by $(e^{ix}+e^{-ix})/2$ and $(e^{ix}-e^{-ix})/(2i)$, and the Fibonacci numbers F_n by Binet's formula.

Identities **5-8** were, until recently, considered genuine non-trivial identities, requiring a human demonstration. One particularly nice human proof of **6**, was given by Cartier and Foata[CF]. A one-line computer-generated proof of identity **6** is given in [E]. Identities **7** and **8** are examples of so-called *q-binomial coefficient identities*, (a.k.a. *terminating q-hypergeometric series.*) All such identities are now routinely provable[WZ2](see below.) The machine-generated proofs of **7** and **8** appear in [ET] and [AEZ] respectively. Identities **7** and **8** immediately imply, by taking the limit $n \to \infty$, identities **7'** and **8'**, which, in turn, are equivalent to two famous number-theoretic statements: The first Rogers-Ramanujan identity, that asserts that the number of partitions of an integer into parts that leave remainder 1 or 4 when divided by 5 equals the number of partitions of that integer into parts that differ from each other by at least 2, and Jacobi's theorem that asserts that the number of representations of an integer as a sum of 4 squares, equals 8 times the sum of its divisors that are not multiples of 4.

The WZ Proof Theory

Identities **5-8** involve sums of the form

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

where the summand, $F(n,k)$, is a *hypergeometric term*, (in **5**, and **6**), or a *q-hypergeometric term*, (in **7**, and **8**), in both n and k, which means that both quotients $F(n+1,k)/F(n,k)$ and $F(n,k+1)/F(n,k)$ are *rational functions* of (n,k) $((q^n,q^k,q)$ respectively).

For such sums, and multi-sums, we have ([WZ2]) the following result.

The Fundamental Theorem of Algorithmic Hypergeometric Proof Theory:

Let $F(n;k_1,\ldots,k_r)$ be a *proper* (see [WZ2]) hypergeometric term in all of $(n;k_1,\ldots,k_r)$. Then there exist polynomials $p_0(n),\ldots,p_L(n)$ and rational functions $R_j(n;k_1,\ldots,k_r)$ such that $G_j := R_jF$ satisfy

$$\sum_{i=0}^{L} p_i(n)F(n+i;k_1,\ldots,k_r) =$$

$$\sum_{j=1}^{r}[G_j(n;k_1,\ldots,k_j+1,\ldots,k_r) - G_j(n;k_1,\ldots,k_j,\ldots,k_r)] \qquad (multiWZ)$$

and hence, if for every specific n, $F(n;-)$ has compact support in $(k_1,\ldots,k_r)$, the definite sum,

$g(n)$, given by

$$g(n) := \sum_{k_1,\dots,k_r} F(n; k_1, \dots, k_r) \qquad (multiSum)$$

satisfies the linear recurrence equation with polynomial coefficients:

$$\sum_{i=0}^{L} p_i(n) g(n+i) = 0 \quad . \qquad (P - recursive)$$

$(P - recursive)$ follows from $(multiWZ)$ by summing over $\{k_1, \dots, k_r\}$, and observing that all the sums on the right telescope to zero.

If the recurrence happens to be first order, i.e. $L = 1$ above, then it can be written in *closed form*: for example the solution of the recurrence $(n+1)g(n) - g(n+1) = 0$, $g(0) = 1$, is $g(n) = n!$.

This "existence" theorem also implies an algorithm for finding the recurrence (i.e. the p_i) and the accompanying certificates R_j, see below.

An analogous theorem holds for q-hypergeometric series[WZ2][K].

Since we know how to find, and prove, the recurrence satisfied by any given hypergeometric sum or multi-sum, we have an effective way of proving any equality of two such sums, or the equality of a sum with a conjectured sequence. All we have to do is check whether both sides are solutions of the same recurrence, and match the appropriate number of initial values. Furthermore, we can also use the algorithm to find new identities. If a given sum yields a first-order recurrence, it can be solved, as mentioned above, and the sum in question turns out to be explicitly evaluable. If the recurrence obtained is of higher order, then most likely the sum is not explicitly-evaluable (in closed form), and Petkovsek's algorithm[P], that decides whether a given linear recurrence (with polynomial coefficients) has *closed form* solutions, can be used to find out for sure.

Almost Certainty For An ε Of The Cost

Consider identity $(multiSum)$ once again, where $g(n)$ is "nice". Dividing through by $g(n)$ and letting $F \to F/g$ (Herb Wilf's wonderful trick), we can assume that we have to prove an identity of the form

$$\sum_{k_1,\dots,k_r} F(n; k_1, \dots, k_r) \; = \; 1 \quad . \qquad (Nice)$$

The WZ theory promises you that the left side satisfies some linear recurrence, and if the identity is indeed true, then the sequence $g(n) = 1$ should be a solution (in other words $p_0(n) + \ldots + p_L(n) \equiv 1$).

For the sake of simplicity, let's assume that the recurrence is minimal, i.e. is $g(n+1) - g(n) = 0$, (this is true anyway in the vast majority of the cases.) To prove the identity, by this method, we have to find *rational functions* $R_j(n; k_1, \ldots, k_r)$ such that the $G_j := R_j F$ satisfy:

$$F(n+1; k_1, \ldots, k_r) - F(n; k_1, \ldots, k_r) =$$

$$\sum_{j=1}^{r} [G_j(n; k_1, \ldots, k_j + 1, \ldots, k_r) - G_j(n; k_1, \ldots, k_j, \ldots, k_r)] \,. \qquad (multiWZ')$$

By dividing $(multiWZ')$ through by F, and clearing denominators, we get a certain functional equation for the $R_1, \ldots, R_r$, from which it is possible to determine their denominators $Q_1, \ldots, Q_r$. Writing $R_j = P_j/Q_j$, the proof boils down to finding *polynomials* $P_j(k_1, \ldots, k_r)$, with coefficients that are rational functions in n and possibly other (auxiliary) parameters. It is easy to predict upper bounds for the degrees of the P_j in $(k_1, \ldots, k_r)$. We then express each P_j symbolically, with "undetermined" coefficients, and substitute into the above-mentioned functional equation. We then expand, and equate coefficients of all monomials $k_1^{a_1} \ldots k_r^{a_r}$, and get an (often huge) system of inhomogeneous linear equations with *symbolic* coefficients. The proof boils down to proving that this inhomogeneous system of linear equations has a solution. It is very time-consuming to solve a system of linear equations with *symbolic* coefficients. By plugging in specific values for n and the other parameters, if present, one gets a system with *numerical* coefficients, which is much faster to handle. Since it is unlikely that a random system of inhomogeneous linear equations with more equations than unknowns can be solved, the solvability of the system for a number of special values of n and the other parameters is a very good indication that the identity is indeed true. It is a waste of money to get absolute certainty, unless the conjectured identity in question is known to imply the Riemann Hypothesis.

Semi-Rigorous Mathematics

As wider classes of identities, and perhaps even other kinds of classes of theorems, become routinely-provable, we might witness many results for which we would know how to find a proof (or refutation), but we would be unable, or unwilling, to pay for finding such proofs, since "almost certainty" can be bought so much cheaper. I can envision an abstract of a paper, c. 2100, that reads : "We show, in a certain precise sense, that the Goldbach conjecture is true with probability larger than 0.99999, and that its complete truth could be determined with a budget of \$10B."

It would be then OK to rely on such a priced theorem, provided that the price is stated explicitly. Whenever statement A, whose price is p, and statement B, whose price is q, are used to deduce statement C, the latter becomes a priced theorem priced at $p + q$.

If a whole chain of boring identities would turn out to imply an interesting one, we might be tempted to redeem all these intermediate identities, but we would not be able to buy out the whole store, and most identities would have to stay unclaimed.

As absolute truth becomes more and more expensive, we would sooner or later come to grips with the fact that few non-trivial results could be known with old-fashioned certainty. Most likely we will wind up abandoning the task of keeping track of price altogether, and complete the metamorphosis to non-rigorous mathematics.

Note: Maple programs for proving hypergeometric identities are available by anonymous `ftp` to `math.temple.edu`, in directory `pub/zeilberger/programs`. a Mathematica implementation of the single-summation program, can be obtained from Peter Paule at `paule@risc.uni-linz.ac.at` .

References

[AZ] G. Almkvist and D. Zeilberger, *The method of differentiating under the integral sign*, J. Symbolic Computation **10**(1990), 571-591.

[AEZ] G. E. Andrews, S. B. Ekhad, and D. Zeilberger *A short proof of Jacobi's formula for the number of representations of an integer as a sum of four squares*, Amer. Math. Monthly, **100**(1993), 274-276.

[ALMSS] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy, *Proof verification and intractability of approximation problems*, Proc. 33rd Symp. on Foundations of Computer Science (FOCS), IEEE Computer Science Press, Los Alamitos, 1992, pp. 14-23.

[AS] S. Arora, and M. Safra, *Probabilstic checking of proofs*, ibid, pp. 2-13.

[Ca] P. Cartier, *Démonstration "automatique" d'identités et fonctions hypergéometriques [d'apres D. Zeilberger]*, Séminaire Bourbaki, exposé n^o 746, Astérisque **206**, $41-91$, SMF, 1992.

[CF] P. Cartier and D. Foata, *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes Math. v. **85**, Springer 1969.

[Ci] B. Cipra, *Theoretical computer scientists develop transparent proof techniques*, SIAM News, **25/3**(May 1992).

[D] J. Dawson, *The Gödel incompleteness theorem from a length of proof perspective*, Amer. Math. Monthly **86**(1979), 740-747.

[E] S. B. Ekhad, *A very short proof of Dixon's theorem*, J. Comb. Theo. Ser. A **54**(1990), 141-142.

[ET] S. B. Ekhad and S. Tre, *A purely verification proof of the first Rogers-Ramanujan identity*, J. Comb. The. Ser. A **54**(1990), 309-311.

[G] K. Gödel, *On length of proofs*, (in German) *Ergeb. Math. Colloq.* **7**(1936), 23-24. Translated in "The Undecidable" (M. Davis, Ed.), Raven Press, Hewitt, NY, 1965, 82-83.

[JQ] A. Jaffe and F. Quinn, *"Theoretical mathematics": Toward a cultural synthesis of mathematics* *and theoretical physics*, Bulletin(N.S.) of the Amer. Math. Soc. **29**(1993), 1-13.

[K] T. H. Koornwinder, *Zeilberger's algorithm and its q-analogue*, University of Amsterdam Mathematics department preprint.

[P] M. Petkovsek, *Hypergeometric solutions of linear recurrence equations with polynomial coefficients*, J. Symbolic Computation **14**(1992), 243-264.

[S] J. Spencer, *Short theorems with long proofs*, Amer. Math. Monthly **90**(1983), 365-366.

[WZ1] H. S. Wilf and D. Zeilberger, *Rational functions certify combinatorial identities*, J. Amer. Math. Soc. **3**(1990), 147-158.

[WZ1a] ——,*Towards computerized proofs of identities*, Bulletin(N.S.) of the Amer. Math. Soc. **23**(1990), 77-83.

[WZ2] ——-, *An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities*, Invent. Math. **108**(1992), 575-633.

[WZ2a] ——, *Rational function certification of hypergeometric multi-integral/sum/"q" identities*, Bulletin(N.S.) of the Amer. Math. Soc. **27**(1992) 148-153.

[Z1] D. Zeilberger, *A Holonomic systems approach to special functions identities*, J. of Computational and Applied Math. **32**(1990), 321-368.

[Z2]——, *A Fast Algorithm for proving terminating hypergeometric identities*, Discrete Math **80**(1990), 207-211.

[Z3]——, *The method of creative telescoping*, J. Symbolic Computation **11**(1991), 195-204.

[Z4]——, *Closed Form (pun intended!)*, "Special volume in memory of Emil Grosswald", M. Knopp and M. Sheingorn, eds., Contemporary Mathematics **143**, 579-607, AMS, Providence, 1993.

July 1993 ; Revised: August 1993.

Added Dec. 7, 2011: I thank John Murray for correcting a typo in the statement of Dixon's identity **6**.