The HOLONOMIC ANSATZ
II. Automatic DISCOVERY(!) And PROOF(!!) of Holonomic Determinant Evaluations

Doron ZEILBERGER [1]

Prerequisites

I assume that readers have read [Z1].

Experimental Mathematics: The diminishing role of humans

In a wonderful essay[W] on Experimental Mathematics, Herb Wilf outlines the four steps of doing Experimental Mathematics, in the way it is usually practiced today.

1. Wondering, by a human, what a "particular situation looks like in detail".

2. Some computer experimentation to show the structure of that situation for a selection of small values of the parameters.

3. The [human] mathematician *gazes* at the computer output, attempting to see or to codify some pattern, that hopefully leads him or her to formulate a *conjecture.*

4. Human-made proof of the human-made conjecture (that was computer-inspired).

Under this scheme, only step **2** employs the computer. In the present series of articles, I illustrate, **by example**, how computers can be used, **without any human intervention**, to also do steps **3** and **4**. As for step **1**, the wondering, this can also be done by machinekind -it is not too hard to teach the computer how to wonder. All that we, humans, ultimately would have to do is *meta-wonder.* In other words, make up new **ansatzes** and write **once and for all** computer programs teaching the computer how to *wonder* in these ansatzes, **then** *gaze* at the pattern, **then** *formulate a conjecture* (within the given ansatz) and **then**, finally, *prove* the conjecture, **all by itself**, without any human intervention! No longer just **computer-assisted** but fully **computer-generated**.

Shalosh B. Ekhad vs. Some Great Human Mathematicians

Consider the determinant evaluation

$$\det\left(\binom{\mu+i+j}{2i-j}\right)_{0\le i,j\le n-1} = \prod_{i=1}^{n-1} Nice(\mu,n,i) \quad , \qquad (MRR)$$

where $Nice$ is some explicit expression whose exact form I omit right now in order not to detract form the essence. This determinant-evaluation, discovered and first proved by William Mills, David

[1] Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA. `zeilberg at math dot rutgers dot edu` , `http://www.math.rutgers.edu/~zeilberg/`.
First version: Feb. 14, 2006. Accompanied by the Maple package DET available from `http://www.math.rutgers.edu/~zeilberg/tokhniot/DET` . Supported in part by the NSF.

Robbins, and Howard Rumsey [MRR1], was so attractive that other great mathematicians, notably George Andrews and Dennis Stanton[AS], Marko Petkovsek and Herb Wilf [PW], and Christian Krathenthaler [K] took the trouble to find other proofs.

So I estimate that at least several months of human-mathematician time was spent in discovering, and proving (MRR) in various ways.

But if you have Maple, and downloaded the Maple package `DET` into your computer, and gotten into Maple by typing `maple`, and typed `read DET:`, then typing:

`RproofP(binomial(m+n+p,2*m-n+1),m,n,N,30,R,p,40,60):` ,

and waiting 256 seconds of CPU time will, completely *ab initio*, conjecture, a *symbolic* explicit evaluation, $Nice(\mu, n, i)$ of the determinant and immediately proceed to prove it *fully rigorously.* It is entirely seamless, and all the rôle of the human was to type the above line. And of course, at present, write the Maple package `DET`, and create Maple, and invent the computer (and the transistor, and the chip etc.).

Once written, `DET` can *discover(!)* and *prove(!!)* countless other determinant evaluations, provided they belong to the right *ansatz*, the holonomic ansatz in our case. You are welcome to look at the webpage of this article

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzII.html`

for numerous examples of input and output, in addition to the Mills-Robbins-Rumsey evaluation mentioned above. If you have Maple, you can generate many more examples on your own.

Caveat

Unlike WZ theory, where the computer is guaranteed to give an output (time and space permitting), this is not the case here. It is not known, and is probably false, that it is always the case that if $a(i,j)$ is holonomic, and setting $A(n) := det(a(i,j))_{0 \le i,j \le n-1}$, then $B(n) := A(n+1)/A(n)$ is holonomic in n. But it happens often enough to justify asking the computer to give it a try.

The General Idea

Of course, the seed of the method *did* originate from humans. It can be formulated in two equivalent ways. We have:

Problem: Given some explicit expression $m(i,j)$, define the $n \times n$ matrix M_n to be

$$M_n := (m(i,j))_{0 \le i,j \le n}$$

Find an *explicit* (in some sense) evaluation (in n) for

$$A(n) := \det M_n \quad .$$

George Andrews's Approach

George Andrews pulls out of a hat an *upper-triangular* matrix U_n whose entries are "nice", and whose diagonal entries are all 1's, and such that $L_n := M_n U_n$ is lower-triangular and has a "nice" diagonal (but the other entries are possibly ugly). Then since $\det(M_n) = \det(L_n)/\det(U_n)$, and $\det(U_n) = 1$, we can express $\det(M_n)$ as a product of nice things, and hence it is nice itself.

Of course, the reader is never told how U_n was conjectured, it is just pulled out of the blue. To prove the assertion one has to prove that L_n is indeed lower-triangular, i.e. the entries of $M_n U_n$ above the diagonal are all 0, which boils down to (usually) proving a hypergeometric identity. Next, one has to prove that the diagonal entries of L_n are as claimed, which involves another hypergeometric identity. These are sometimes proved by computer, using the Zeilberger algorithm, but still in a piecemeal, **human-centric**, way.

Dave Robbins's Approach

Although mathematically equivalent to George Andrews's LU approach, I find Dave Robbins's approach, described in [MRR2], more conducive for teaching a computer.

Consider the $n+1$ by $n+1$ matrix $(a(i,j))$, for $0 \le i,j \le n$. Let $A(n,i)$ $(0 \le i \le n)$, be the cofactor of the (n,i) entry. Then of course

$$\det(a(i,j))_{0 \le i,j \le n} = \sum_{i=0}^{n} A(n,i)a(n,i)$$

Now the $n+1$ unknowns $A(n,j), j = 0 \ldots n$ are uniquely determined, up to a normalization factor (that only depends on n), by the n linear homogeneous equations

$$\sum_{j=0}^{n} A(n,j)a(i,j) = 0 \quad , \quad (i = 0 \ldots n-1) \quad .$$

The normalized cofactors $A'(n,j)$ defined by $A'(n,j) = A(n,j)/A(n,n)$, are then determined uniquely, for each specific n, by the system of linear equations

$$\sum_{j=0}^{n} A'(n,j)a(i,j) = 0 \quad , (i = 0 \ldots n-1) \quad . \qquad (Dave1)$$

subject to

$$A'(n,n) = 1 \quad . \qquad (Dave2)$$

Now the human can ask the computer to crank out $A'(n,i)$ for $0 \le i,n \le N$ for some N (say 100), and ask the **computer** (rather than doing it himself) to *gaze* at the output, and to conjecture some 'explicit' form for $A'(n,i)$ and then prove that they indeed satisfy $(Dave1)$ and $(Dave2)$.

Finally, if in luck,

$$B(n) := \sum_{j=0}^{n} A'(n,j)a(n,j) \quad , \qquad (Dave3)$$

turns out to be "explicit" , and that one is clever enough to prove it. Since $A(n,n) = \det(a(i,j))_{0 \le i,j \le n-1}$, tt follows that

$$\frac{\det(a(i,j))_{0 \le i,j \le n}}{\det(a(i,j))_{0 \le i,j \le n-1}} = B(n) \quad ,$$

that finally entails that

$$\det(a(i,j))_{0 \le i,j \le n} = \prod_{j=0}^{n} B(j) \quad ,$$

which should be considered 'nice' if $B(j)$ is.

The Limit of Humans

If $A'(n,i)$ turns out to be *closed-form* then some clever humans, like George Andrews, can gaze at it and conjecture an expression for it. Other humans, like Christian Krattenthaler will 'cheat' and use a computer program (like Krattenthaler's `rate`) to do do the guessing, but *interactively*.

If nothing nice emerges, then these humans use some dirty tricks of the trade, that only they know, and are unwilling (and usually also unable) to divulge, to express the co-factors as a single-sum or double-sum or whatever. Then they go over excruciating pain, to verify $(Dave1)$ and evaluate $B(n)$, using $(Dave3)$. The most challenging case is when the general normalized cofactor, $A'(n,i)$, does not happen to be a hypergeometric term, but is rather a hypergeometric sum (or multisum).

Enter Computers

But computers do not play favorites. When instructed to operate within the *Holonomic Ansatz* (see [Z1]), they have no particular fondness for hypergeometric terms. A discrete function of two variables being a hypergeometric term is just the *special case* of the defining recurrences being *first-order*. So, staying within the *holonomic ansatz*, once the computer generated enough numerical data $A'(n,j)$ for $0 \le j,n \le N$ for a big enough N, instead of *gazing*, it keeps **guessing** linear recurrence equations with polynomial coefficients (see [Z1]) in the j direction and in the m direction, not necessarily of the first-order. Once conjectured, the rest, i.e., proving $(Dave1)$, $(Dave2)$ and 'evaluating' $B(n)$ is algorithmically decidable thanks to [Z2], at least in principle, but thanks to Frederic Chyzak's beautiful work[C], probably also in practice.

What If It Takes Too Long? Let's Settle for a Semi-Rigorous Proof

Since we *know* that if we had a big enough computer, and good enough software, we can prove $(Dave1)$ and $(Dave3)$ (once we conjectured recurrences for $A'(n,j)$ and a recurrence for $B(n)$), is it really worth the trouble? We can use the conjectured linear recurrences for $A'(n,j)$ to compile a much larger table of conjectured values, much faster than by direct computation using the system of equations $(Dave1)$.

So let's temporarily call these new values $A''(n,j)$, and once found, plug them into $(Dave1)$ and $(Dave3)$ for say $n \leq 10000$. If $(Dave1)$ and $(Dave3)$ hold for all these 10000 values of n, it shows that our conjecture is right at least up to 100000. This is almost a rigorous proof! It is like proving that two polynomials are the same by plugging in enough special values. In the case of polynomials of one variable, there is an easy parameter, the degree plus 1, that tells you how many special values you have to plug-in in order to rigorously prove a conjectured identity between two given polynomials. The problem now that it is not so easy to find the analog of the *degree* for a holonomic function, but if we plug it in for $n \leq 10000$, this is most likely more than enough. The notion of *semi-rigorous* proof was introduced in [Z3] and critiqued in [A].

A Note on Guessing

There now exist powerful guessing programs, for example `superseeker` in Neil Sloane's famous site, Bruno Salvy and Paul Zimmermann's Maple packages `Gfun` and `Mgfun` and Christian Krattenthaler's Mathematica program `rate`, that could have been used, as subroutines in the present endeavor. But I found it easier to program everything *ab initio*, borrowing freely, of course, from these pioneering programs.

The Maple Package DET

Everything here is implemented in the Maple package `DET` available from

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

The webpage of this article

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzII.html`

contains some input and output.

The main functions are

`Rproof, RproofP, SRproof, SRproofI`, to find out about them, in `DET`, type `ezra(FunctionName);`. For example, with help for `Rproof`, type `ezra(Rproof);`.

The pillars of the above functions are `DaveH` and `DaveV`, that guess horizontal and vertical pure recurrences respectively for the normalized cofactors $A'(n,m)$. Full details can be gotten by reading the source code of `DET`. Let me just mention that `DaveH` works by conjecturing recurrences for many rows (i.e. for fixed n), and then using the procedure `GR1` that guesses rational functions in order to guess a unified recurrence, featuring n symbolically.

Apology and Future Directions

This work is a little half-baked, and I hope that other people can continue to perfect `DET` and make it more efficient.

It should be relatively easy to write the q-analog of `DET`, but the running times would be slower because of the extra symbol q. Also translating from Maple to C or Matlab should speed things up.

REFERENCES

[AS] G.E. Andrews and D.W. Stanton, *Determinants in Plane Partition Enumeration*, Europ. J. Combin. **19** (1998), 273-282.

[K] C. Krattenthaler, *Advanced Determinant Calculus*, Sém. Lothar. Comb. **42** (1999), B42q. ("The Andrews Festschrift", D. Foata and G.-N. Han (eds.))

[MRR1] W.H. Mills, D.H. Robbins, and H. Rumsey, *Enumeration of a Symmetry Class of Plane Partitions*, Discrete Mathematics **67** (1987), 43-55.

[MRR2] W.H. Mills, D.H. Robbins, and H. Rumsey, *Proof of the Macdonald conjecture*, Invent. Math. **66** (1982), 73-87.

[PW] M. Petkovv̂sek and H. Wilf, *A high-tech proof of the Mills-Robbins-Rumsey determinant evaluation*, Elec. J. of Combinatorics **3** (no.2 "The Foata Festschrift") (1996), #R19.

[W] Herbert S. Wilf, *Mathematics: An Experimental Science*, to appear in "Princeton Companion of Mathematics" (W.T. Gowers, ed.), Princeton University Press. Available from
`http://www.math.upenn.edu/~wilf/reprints.html`.

[Z1] Doron Zeilberger, *The HOLONOMIC ANSATZ I. FOUNDATIONS and Applications to Lattice Walk Counting*, submitted. Available from

`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzI.html`.

[Z2] Doron Zeilberger, *A Holonomic systems approach to special functions identities*, J. of Computational and Applied Math. **32**, 321-368 (1990).

[Z3] Doron Zeilberger, *Theorems for a price: Tomorrow's semi-rigorous mathematical culture*, Notices of the Amer. Math. Soc. **40 # 8**, 978-981 (Oct. 1993). Reprinted: Math. Intell. **16**, no. 4, 11-14 (Fall 1994).