Automated Enumerative Combinatorics

Doron Zeilberger

Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019. `zeilberg@math.rutgers.edu`
`http://www.math.rutgers.edu/~zeilberg/` .

Summary of Results from Previous NSF Support: DMS 0901226

1. The current NSF award number is DMS-0901226

(`http://www.nsf.gov/awardsearch/showAward?AWD_ID=0901226`)

for the period 2009-2013 (five summers), totaling $344,191 (including the supplements for graduate student support).

2. Its title is: "*Rigorous Experimental Combinatorics*".

3. Summary of the results of the completed work

It is often the case, in scientific research (and elsewhere), that *the whole is larger than the sum of its parts*. I believe that in the case of my own research, and of my students, the whole is *much* larger than the sum of its parts, not that the parts have anything to be ashamed of.

The Whole

[This preamble is the same as in my last grant proposal]

My research students and I continued to practice a new research methodology, that can be loosely called *rigorous experimental mathematics*. It has something in common with both "mainstream" experimental mathematics (as preached by the Borwein brothers, David Bailey, Victor Moll, and their collaborators, see e.g. the masterpiece [BB], and the recent collection [BBCGLM]), and *automated theorem proving* (as practiced in computational logic), but is definitely distinct from them. It is based on what I call *the ansatz ansatz* ([Z1][Z2], see [Z3] for a philosophical discussion), described briefly in my interview with (at the time) MAA president Joe Gallian[Gal]. In this methodology, one "teaches" the computer how to "conjecture an answer" to a problem, and then "teaches" that very same computer to prove its own conjectures rigorously. The novelty is that *both* the conjecturing and the proving are automatically done by the computer. This does not mean that human mathematicians are superfluous. Quite the contrary! Someone has to "teach" the computer, i.e. design algorithms and meta-algorithms for both proving and conjecturing. In my experience, this act of "teaching" the computer how to do mathematics is at least as challenging as doing mathematics "by hand", and in my humble opinion, time much better spent, since the vast potential of the computer is still very underutilized and underrated, and it is important to have mathematicians, like myself and my students, who are dedicated to that activity, that I believe will soon revolutionize mathematics.

The Parts

In the last five years, many in collaboration with my students and other researchers, I published 31 research articles in peer-reviewed journals and refereed conference proceedings. These can be viewed (with links to the actual papers) in the webpage

`http://www.math.rutgers.edu/~zeilberg/papers1.html` .

In addition there are about fifty new items (since 2009) in the "Personal Journal of Shalosh B. Ekhad and Doron Zeilberger", a freely available on-line journal, see http://www.math.rutgers.edu/~zeilberg/pj.html .

While some of the entries there are "silly" (or links to videotaped lectures), most of them are substantial, usually heavily computational, research articles (in collaboration with my beloved silicon servant, Shalosh B. Ekhad). All my papers that are authored only by me and/or Shalosh B. Ekhad (who is not interested in promotion or tenure) no longer get submitted to official peer-reviewed journals, that in my humble opinion will soon be obsolete.

But, the **crux** of my current activity, that I propose to continue in this proposal, in a more systematic and 'professional' way, is the development-and implementation of-algorithms for computer-generated and computer-assisted (usually rigorous, but sometimes semi-rigorous, and even non-rigorous) discovery of new, interesting, and *deep* results in enumerative combinatorics, number theory, and other areas where symbolic computation promises to be crucial. Most of my papers are accompanied by often sophisticated Maple packages, all *freely available* from

http://www.math.rutgers.edu/~zeilberg/programs.html .

These Maple packages (53 new ones since 2009) are not only interesting, and (at least potentially) useful, for their own sake, but also constitute many *case studies* for the efficient use of symbolic computation in combinatorics, and elsewhere. Unfortunately, they are written in a programming language that requires a *commercial* software, Maple^{TM}. One of my dreams is to learn SAGE well enough to be able to start coding in it, but more realistically, interest my students, and other people, to write SAGE versions. Also, since I am not a professional software developer, the organization is amateurish, and could use a more professional handling.

Some highlights

Perhaps the most striking accomplishment([KKZ]), in collaboration with Manuel Kauers and Christoph Koutschan, is the computer-assisted proof of the long-standing *q-TSPP conjecture*, published in the Proceedings of the National Academy of Science :

http://www.pnas.org/content/early/2011/01/20/1019186108.abstract .

This conjecture, made independently, more than thirty years ago, by George Andrews and David Robbins, was the only remaining 'survivor' in Richard Stanley's influential 'baker's dozen' list of conjectures([St]), and resisted the efforts of the greatest combinatorial enumerators. Our proof showed the symbiotic relationship between human conceptual insight, human computational and algorithm-development insight, and the sheer power of computers, as it relied on a beautiful theorem of Soichi Okada[0] that expressed the desired generating function in terms of a certain explicit determinant, that nevertheless resisted the efforts of the greatest human determinant-evaluators like Christian Krattenthaler([K1][K2]). It used my non-trivial extension ([Z2]) of the Wilf-Zeilberger algorithmic proof theory ([PWZ] [WZ]) to the context of determinant evaluations.

This discovery was written up in the Austrian and Swiss general media, as well as in *Pour La Science* (the French analog of Scientific American)

(http://www.pourlascience.fr/ewb_pages/a/actualite-un-theoreme-pour-des-empilements-de-cubes-26504.php).

This method was recently extended to the evaluation of Pfaffians in a beautiful paper [KT] by my former student Thotsaporn "Aek" Thanatipanonda and Christoph Koutschan, proving yet more intriguing conjectures of Krattenthaler and George Andrews.

In a different direction, Andrew V. Sills and I ([SZ]) refuted a long-standing conjecture of the great 20th-century number theorist, Hans Rademacher,

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

While our disproof is not fully rigorous, it is more convincing than many humanly-provably allegedly rigorously- proved theorems, that rely on (often flawed) human peer-reviews.

Another article that I am fond of, is [EGZ] (joint with Shalosh B. Ekhad and Evangelos Georgiadis), that, I believe, has potential far-reaching applications to mathematical finance.

In yet another direction([Z4]) is the computerized discovery (and proof!) of 144 new Collatz-like theorems that may one day, lead, who knows?, to a proof (or disproof!) of the notorious $3x + 1$ conjecture.

Going completely out of combinatorics and number theory is the article([HZ]) that my former student, Emilie Hogan, and I wrote, about computerized proofs of theorems about Global Asymptotic Stability of Discrete Dynamical Systems.

4. List of Publications resulting from the previous NSF award 2009-2013

See: `http://www.math.rutgers.edu/~zeilberg/papers1.html` (for publications in peer-reviewed journals) and also
`http://www.math.rutgers.edu/~zeilberg/pj.html` , mentioned above.

5. As already mentioned above, most of my papers are accompanied by Maple packages that are available, free of charge, from the webpage

`http://www.math.rutgers.edu/~zeilberg/programs.html` .

Some of them are of a rather general scope, and should be useful to researchers in combinatorics, number theory, analysis, statistics, statistical physics, and possibly other areas. They are all **freely available source-code**.

6. A large part of the proposed research is a direct continuation of the previous research, but there are also new directions, in which the connection is less obvious.

7. Education and Human Resources Statement

Specific Graduate Education: Ph.D. students

During the discussed period (2009-2013), five students received their Ph.D. degree under my guidance: Eric Rowland (2009, currently at University of Québec at Montréal), Paul Raff (2009, Applied Researcher, Big Data Mining at Microsoft), Emilie Hogan (2011, currently a permanent computational mathematical scientist at Pacific Northwest Laboratory), Andrew Baxter (2011, currently at Penn State), and Brian Nakamura (2013, currently a postdoc at Dimacs-CCICADA).

All their defenses have been videotaped and are available on-line, see the links in the webpage:

`http://www.math.rutgers.edu/~zeilberg/banim.html` , where one can also find links to the actual theses.

In addition, I have a very talented *disciple*, Edinah Gnang, `http://paul.rutgers.edu/~gnang/`, who is not my official student, since his PhD is in computer science, but with whom I collaborated([GZ]) (and still do), and who is currently (AY 2013-2014) at the Institute for Advanced Study in Princeton.

Currently I have two official graduate students, Matthew Russell (joint with Vladimir Retakh), and Nathaniel Shar, who are making very good progress, and a prospective student, Nathan Fox.

General Graduate Education

Eleven years ago, I started teaching a graduate course called *Experimental Mathematics* that in addition to making the students 'computer-algebra whizes' and skilled and sophisticated Maple programmers, also implicitly introduces them to the methodology of doing computer experiments to rigorously solve open problems. To see detailed logs of what has been done in the various classes (each time dedicated to a different topic!), see the links in the webpage

`http://www.math.rutgers.edu/~zeilberg/teaching.html` .

Quite a few of the students, of other advisors, who took this class told me how useful they found the computer-algebra skills that they learned in my class, and indirectly, this also benefits their advisors, who are mostly computer illiterate. Speaking of 'Experimental Mathematics', my students and I are organizing a weekly seminar by that name that is very well-attended, both by faculty and graduate students, and that has a great diversity of speakers, from George Andrews (past president of the AMS), Jon Borwein (of AGM and Experimental Math fame), Freeman Dyson (of QED fame), Tom Hales (of Kepler fame), Neeraj Kayal(the 'K' of AKS, of PRIME in P fame), John Nash (of Beautiful Mind fame), Neil Sloane (of OEIS fame), Noga Alon (of the Probabilistic Method fame), Gil Kalai (of convexity fame), Avi Wigderson (of NP vs. P (and Zero-Knowledge proof) fame), Robert Trivers (of Altruism fame) to just drop a few names, all the way to first-year graduate students. In fact, on Nov. 7, 2013, we are going to have a talk by a *high school* student, William Kang. Starting Winter 2010, most of the lectures are videotaped and uploaded to the internet, where the whole world can benefit from them. See the links to the archives from the seminar page `http://www.math.rutgers.edu/~russell2/expmath/`.

Undergraduate Education

I take my undergraduate teaching very seriously, and post on-line lecture notes that I believe are also used by students at other universities. See my homepage `http://www.math.rutgers.edu/~zeilberg/` for links.

Implicitly, my former PhD student, Lara Pudwell (`http://faculty.valpo.edu/lpudwell/`), has designed a pioneering *undergraduate* course, called Experimental Mathematics, inspired by my Rutgers graduate course, that I hope would be emulated at other colleges and universities.

PROPOSED RESEARCH: Automated Enumerative Combinatorics

I. Towards a General Theory of Automated Enumeration based on Enumeration Schemes

I.0 The Painful Dichotomy between the Doable and Undoable

We humans pride ourselves on the great things that we *did* do, and all the seemingly hard problems that we have solved. But, alas, the number of problems that we can **not** do, and the number of numbers that we can **not** compute (and never will!) far outnumber those problems that we **can** solve and those numbers that we **can** compute.

While writing this, I opened another window on my computer, running Maple^{TM}, and typed

```
with(linalg):ra:=rand(0..1):  A:=[seq([seq(ra(),i=1..15)],j=1..15)]:  time(det(A)), time(permanent(A));
```

and got `0.004, 0.844` . Then I typed

```
with(linalg):ra:=rand(0..1):  A:=[seq([seq(ra(),i=1..17)],j=1..17)]:  time(det(A)), time(permanent(A));
```

and got `0.008, 5.544`.

Then I typed

```
with(linalg):ra:=rand(0..1): A:=[seq([seq(ra(),i=1..20)],j=1..20)]: time(det(A)), time(permanent(A));
```

and got `0.012, 102.478`.

For the readers of this proposal who are not familiar with Maple, let me explain. The command '`with(linalg)`', loads the linear algebra built-in Maple package, `linalg`, that has the commands `det` and `permanent`, for computing the *determinant* and *permanent* of a matrix, so once a matrix A has been defined, '`det(A);`' gives its determinant, and '`permanent(A);`' gives its permanent. The command '`ra:=rand(0..1)`'=, generates a (pseudo-) random number generator, so once defined, '`ra();`' outputs either 0 or 1, each with probability $\frac{1}{2}$. Hence typing

```
A:=[seq([seq(ra(),i=1..n)],j=1..n)]:
```

for any positive integer n, generates a random $0-1$ matrix with roughly as many 0's as 1's.

Finally for any command, '`time(command);`' returns the time it takes to execute `command` rather then the output of the command itself.

What Maple is trying to tell us is that if we take a random 15×15 $0-1$ matrix, with roughly as many 0's as 1's, it takes it 0.004 seconds to compute the determinant and 0.844 seconds to compute the permanent. The analogous computation for a random 17×17 matrix takes 0.008 and 5.544 seconds respectively, and for a 20×20 matrix it takes 0.012 and 102.478 seconds respectively. Maple can still do the *determinant* of a 300×300 matrix in reasonable time. For example

```
with(linalg):ra:=rand(0..1): A:=[seq([seq(ra(),i=1..300)],j=1..300)]: time(det(A));
```

returns `137.640`, indicating that it took it 137.640 seconds. On the other hand

```
with(linalg):ra:=rand(0..1): A:=[seq([seq(ra(),i=1..300)],j=1..300)]: time(permanent(A));
```

would take several zillion years!

The reason, as we all know (at least since Gauss, but implicitly much longer), is that computing *determinants*, of numerical matrices, is **fast**. On the other hand, the fastest known algorithms for computing the permanent of a numerical matrix take exponential time (the naive algorithm takes about $(n+1)!$ operations, of course, and that is worse than exponential, but an algorithm due to Ryser is 'only' exponential, see the wikipedia entry).

In a famous seminal paper, written in 1979, Leslie Valiant proved that computing permanents is a #P-complete task, meaning, in analogy with the notion of NP-completeness for *decision* problems, that it is 'too good to be true' that a fast (polynomial time) algorithm for computing permanents would ever be found. Of course, no one has (yet) any clue on how to prove this impossibility, but no one in his or her (or its) right mind doubts this impossibility.

While the problem of computing the permanent of a $0-1$ matrix can be, strictly speaking, considered a problem in *enumeration*, that inputs an arbitrary bipartite graph (with an edge between Mr i and Ms. j iff $A_{ij} = 1$), and outputs the total number of *perfect matchings* (note that the analogous *decision* problem, deciding whether there *exists* a matching, i.e. whether *permanent*(A) is strictly positive, is famously (thanks to Philip Hall) polynomial time), is too 'general' to qualify as a *genuine* problem in (traditional) enumerative combinatorics. A typical problem in enumerative combinatorics has usually one, or at most, few, *inputs*, usually non-negative integers, let's call it n (or $(n_1, n_2, \ldots,))$ and a description of a *natural* combinatorial set $A(n)$ (or $A(n_1, n_2, \ldots))$ describing combinatorial objects satisfying a certain set conditions that depend on n (or $(n_1, n_2, \ldots)$).

Of course, from a purely *logical* point of view, one can formulate the problem of computing the permanent of a general $n \times n$ $0-1$ matrix by only using *two* integer arguments, n and m, where

m is an integer between 0 and $2^{n^2}-1$ whose binary representation yield the n^2 entries of the matrix read by rows, but this is very *artificial.* Instead let's consider three *case studies* of *natural* enumeration problems.

I.1 First Case Study: Lattice Animals vs. Directed Lattice Animals

A (2D) (fixed) *lattice animal* (alias a *polyomino*) of size n is any *connected* set of lattice points on the square grid, up to translation-equivalence. For the sake of definiteness let the leftmost points have coordinate $y=0$, and amongst those, let the lowest point be the origin. It is sequence `http://oeis.org/A001168` in the celebrated OEIS, and at this time of writing only 56 terms are known, computed by F. Jensen, see `http://oeis.org/A001168/b001168.txt`. How many fixed lattice animals are there with 1000 points?, no one knows, and, **most probably**, no one will *ever* know, even though, once found, it would fit on half a screen.

Now define a *directed* animal to be a lattice animal such that every point is reachable from the origin, $(0,0)$ by a path that only uses eastbound unit steps and northbound unit steps. Superficially the problem of enumerating directed animals does not seem any easier than the former, but, there is a nice explicit expression (quoted in `http://oeis.org/A005773`), due to Dhar, for the *generating function*

$$\sum_{n=0}^{\infty} a(n)x^n \;=\; \frac{2x}{3x-1+\sqrt{1-2x+3x^2}} \quad ,$$

that easily enables Maple to compute $a(1000)$. In fact, typing into a Maple session

```
coeff(taylor(2*x/(3*x-1+sqrt(1-2*x-3*x**2)),x=0,1003),x,1000);
```

immediately yields a certain 476-digit integer (starting with $1361733190\dots$).

An even faster way would be to use the supplied three-term recurrence

$$n\,a(n) = 2n\,a(n-1) + 3(n-2)\,a(n-2) \quad , a(0)=1 \quad , \quad a(1)=1 \quad .$$

What makes directed animals so easy to count, while unrestricted animals are seemingly (and most probably) *intractable*?

A beautiful derivation of Dhar's formula was given by Bétréma and Penaud[BeP] and exposited in [Z5]. What made things work out was a subtle, but a posteriori almost-trivial, *decomposition* that mapped a directed animal into smaller directed animals, thereby getting a (non-linear) *recurrence.* So there was a *scheme* that expressed $a(n)$ in terms of $a(i)$ for $i<n$. No such analogous scheme is known for the original lattice animals, and I am (almost) sure, that none exists.

Xavier Viennot showed that counting directed lattice animals is equivalent to counting certain *domino towers.* Once the *idea* behind the [BeP] proof was made clear, I taught it to my computer, and it was able to derive *explicit* generating functions for many, far more complicated, families, see [Z8].

I.2 Second Case Study: Enumerating Pattern-Avoiding Permutations

The subject of *pattern-avoiding permutations*, pioneered by Herbert Wilf, Rodica Simion, and Richard Stanley, is currently a very active field, with an annual conference dedicated to it. For a beautiful, and very lucid introduction, see Chapter 4 of Miklós Bóna's book [Bo1], and the wikipedia entry `http://en.wikipedia.org/wiki/Permutation_pattern` , initiated, and largely written, by my former student, Vince Vatter.

Let's recall what it means for a permutation p, of length n, to *contain* a pattern q of length k (with $k \le n$), by citing Definition 4.1 of [Bo1].

Definition: Let $q = (q_1, q_2, \ldots, q_k) \in S_k$ be a permutation, and let $k \leq n$. We say that the permutation $p = (p_1, p_2, \ldots, p_n)$ contains q as a pattern if there are k entries $p_{i_1}, p_{i_2}, \ldots, p_{i_k}$ in p so that $i_1 < i_2 < \ldots < i_k$, and $p_{i_a} < p_{i_b}$ if and only of $q_a < q_b$. Otherwise we say that p avoids q.

For example, the permutation 3451267 avoids the pattern 321 as it does not contain a decreasing subsequence of length 3. It contains the pattern 2134 as shown by the entries 4267.

It is well-known (see there) that the number of 123-avoiding permutations of length n, as well as the number of 132-avoiding permutations, are both given by the ubiquitous *Catalan numbers* $(2n)!/(n!(n+1)!)$, `http://oeis.org/A000108`. When two patterns yield the same enumeration, they are said to be *Wilf-equivalent*, so the patterns 123 and 132 are Wilf-equivalent.

It is also well-known that there are only three Wilf classes for patterns of length 4: 1342, 1234, and 1324, all the remaining $4! - 3 = 21$ patterns are Wilf-equivalent to one of these, either trivially or non-trivially (thanks to theorems of Julian West and Zvezdelina Stankova).

The number of 1234-avoiding permutations of lenght n has been derived by Ira Gessel [G] and is given in `http://oeis.org/A005802`, with an explicit formula that enables one to compute as many terms as desired, and makes the computation of $a(10000)$ (for example) almost immediate (on a fast computer).

Seven years later, Miklós Bóna derived an explicit, and fast, formula for the number of 1342-avoiding permutations of lenght n, and it is given in `http://oeis.org/A022558`. Perhaps the fastest way to find the first 10000 terms (say) is via the recurrence that is implied by Bóna's formula

$$n\,a(n) = (7n-22)\,a(n-1) + 4(2n-1)\,a(n-2) \quad , \quad a(0) = 1 \quad , \quad a(1) = 1 \quad .$$

But, in an often-quoted claim that I once made

"Not even God knows the number of 1324*-avoiding permutations of length* 1000*"* .

According to Sloane's OEIS `http://oeis.org/A061552/b061552.txt`, there are currently only 25 terms known. Very recently (in preparation), my former student, Brian Nakamura, together with Fredrik Johansson, in a computational tour-de-force, computed seven more terms, using an extension of the functional equation approach developed in [NZ].

The reason that 1234-avoiding and 1342-avoiding permutation do have 'nice' formulas, that enables one to easily compute the first 10000, say, terms, is that they posses *enumeration schemes* and a *hidden recursive* structure. In the case of the former, this is made explicit in the notion of *enumeration scheme* that I introduced in [Z6], and that was vastly extended by Vince Vatter[V] (see also [Z7]). These lead to *linear recurrences.*

In the case of 1342-avoiding permutations, a careful reading of [Bo2] shows that, hidden behind the argument, is (another kind) of scheme that leads to a *non-linear* recurrence, and hence to an *algebraic* generating function. Since this is fairly complicated, let's illustrate it in terms of the far simpler enumeration of 132-avoiding permutation.

Given a permutation π, of $\{1, 2, \ldots, n\}$ that avoids the pattern 132, let n be at the i^{th} place, i.e. $\pi_i = n$. Then $\{\pi_1, \ldots, \pi_{i-1}\}$ must consist of the $i-1$-largest remaining entries, $\{n-i+1, \ldots, n-1\}$, and $\{\pi_{i+1}, \ldots, \pi_n\}$ must consist of the $n-i$ *smallest* entries $\{1, 2, \ldots, n-i\}$, or else there would emerge a 132 pattern. Of course both $\pi_1, \ldots, \pi_{i-1}$ and $\pi_{i+1}, \ldots, \pi_n$ must be 132-avoiding as well, so we have the non-linear recurrence

$$a(n) = \sum_{i=1}^{n} a(i-1)a(n-i) \quad ,$$

that implies that $a(n)$ are the Catalan numbers.

I am sure that this kind of argument can be 'taught' to a computer, who would be able to handle much more complicated scenarios. I believe that it should be possible to fuse both kinds of recurrences, and arrive at a formalized, yet-more general, notion of enumeration scheme, that of course should also work for *multiple patterns*.

Once there is such a scheme, already existing Maple packages, available from my website, can *automatically* handle the sequences, and derive, recurrence relations, asymptotics, and algorithms for random generation.

Once a general notion of enumeration scheme, that would include both linear and non-linear recurrences, would be defined and implemented, it would hopefully lead to insight that would explain why the enumeration of 1324-avoiding permutations is intractable.

I. 3 Third Case Study: Automatic Counting of generalized Ménages permutations

In chapters 7 and 8 of his classic book([R]), the great combinatorial enumerator, John Riordan, develops a beautiful theory for enumerating permutations with restricted position, using *rook polynomials*. The simplest instance is the enumeration of *derangements*, i.e. permutations π such that $\pi_i - i \neq 0$ (for $1 \leq i \leq n$) . The next case, of *Ménages permutation* is (essentially) counting permutations for which $\pi_i - i \neq 0, 1$. The first satisfactory formula was given by Touchard in 1934. In an unpublished 1975 Bell Labs internal publication (recently kindly uploaded by Neil Sloane), John Riordan spends lots of pages enumerating permutations for which (for $1 \leq i \leq n$) $\pi_i - i \neq 0, 1, 2$, and things start to get complicated for humans. In a still under-construction Maple package,

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

that would hopefully accompany a forthcoming paper, I 'taught' the computer how to derive such formulas for an *arbitrary* finite set S. In other words, the user inputs a set, S, of integers (possibly including negative, as well as positive entries) and would get a (possibly complicated) linear recurrence equation with polynomial coefficients satisfied by enumerating sequence for the number of permutations of length n such that, for all i between 1 and n, $\pi_i - i \notin S$.

A typical output file can be viewed here:

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oMENAGES3` .

I'd like to emphasize that the (still under-construction) Maple package `MENAGES` contains numerous subpackages, and everything is streamlined. First there is a complete automation of computing generating functions for rook-polynomials for the given board, let's call it $R(t, z)$. Then the quantity itself, using Riordan's general theory of rook polynomials (thanks to Euler's famous $\int_0^\infty t^n \, e^{-t} \, dt = n!$) can be written as

$$a(n) = \frac{1}{2\pi i} \int_0^\infty \int_{|z|=1} \frac{R(-t, \frac{z}{t})}{z^{n+1}} \, e^{-t} \, dz \, dt \quad .$$

Now one uses WZ theory (specifically the so-called *Multi-Almkvist-Zeilberger* algorithm implemented in the Maple package

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

accompanying [AZ], that has been 'embedded' in the present package, `MENAGES` in order to make it self-contained, to get a (rigorously proved!) linear recurrence equation with polynomial coefficients.

If someone is in a rush, one can 'cheat' and get a *semi-rigorously-proved* recurrence, by plain *guessing*, where one can either use the built-in Maple package `gfun`, or my own implementation

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

that is also embedded in `MENAGES`. Once you have the a linear recurrence, one uses my Maple package `AsyRec` available from

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

to compute an asymptotic expansion to *any* desired order (the leading asymptotics can be shown to be $n!/e^k$ where k is the 'width' of the strip) .

I.4: Conclusion of Section I

The above are but three specific *case studies*. In addition to pursuing them for their own sake, it is hoped that a general enough theory of *enumeration schemes* would be developed for which one may be able to prove, in a yet-to-be-found sense, *intractability*. Judging from the wide open holy grail of computer science, the P vs. NP problem (and its close cousin the notion of #P-completeness), this task is, most probably, (meta)-intractable, so one should settle on finding some *reduction procedures* that would establish that many seemingly intractable enumeration problems are 'equivalent' (modulo polynomial-time reduction), thereby developing an enumerative-combinatorics-analog of the notion of #P-completeness.

II. A Symbolic-Computational Approach to Limit Laws of Combinatorial Statistics

II.1 Probability Limit Laws

One of the central themes of modern probability theory are *limit laws*, the most celebrated one being the Central Limit Theorem, that roughly says that if you repeat the same experiment many times, and the "atomic" experiment can have an arbitrary probability distribution (with finite variance), then in the limit, after one "centralizes" and "normalizes" (divides by the so-called standard deviation) one gets the (continuous) *Standard Normal Distribution*:

$$Pr(a \leq X \leq b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\, dx \quad .$$

The *iconic* example of a discrete probability distribution is the random variable "number of Heads" upon tossing a (loaded) coin n times, whose probability distribution is given by the **Binomial Distribution**, usually denoted by $\mathbf{B}(n,p)$. It describes the experiment of tossing a coin n times with the probability of Heads being p. The "sample space" is the set of all 2^n outcomes $\{H,T\}^n$, and the probability of an "atomic" event is $p^{NumberOfHeads}(1-p)^{NumberOfTails}$, and hence the probability of the "compound event", *NumberOfHeads*$=k$, is $\binom{n}{k}p^k(1-p)^{n-k}$. If we call this random variable X_n, then its mean is np and its variance is $\sigma^2 := np(1-p)$. Introducing the *centralized and normalized* random variable

$$Z_n := \frac{X_n - np}{\sqrt{np(1-p)}} \quad ,$$

the "original" (De Moivre-Laplace) Central Limit Theorem asserts that

$$Z_n \to \mathcal{N} \quad ,$$

where $\mathcal{N}$ is the Standard Normal Distribution.

A *symbolic-computational* approach, based on the *method of moments*, to combinatorial statistics, taken *one at a time* has been initiated in [Z9], and implemented in the Maple packages

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

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

But often in combinatorics we have *several* interesting combinatorial statistics, for example, in the symmetric group, the *number of inversions* (*inv*) and the *major index* (*maj*), and it is of interest to prove *joint-limit laws*, and getting the *multi-variate* limiting distribution.

II.2 Multi-Variate Combinatorial Statistics

A first step towards this goal was done in collaboration with my former student, Andrew Baxter, in [BZ], where we proved that the above-mentioned combinatorial statistics, namely *inv* and *maj*, are asymptotically joint-independently-normal. Quite recently, a student of Christian Krattenthaler, Marko Thiel, beautifully generalized our approach to *words* (alias *multiset permutations*), see his paper [T]. But this is still but the *tip of the iceberg*. What about considering three (or more) combinatorial statistics simultaneously? A fascinating playground would be to consider the statistics of 'number of occurrences of permutation patterns', using, as a starting point, my recent paper with Brian Nakamura ([NZ]) (already mentioned above in a different context).

III. Wilf-Zeilberger Algorithmic Proof Theory

WZ-theory has been implemented in all the major computer-algebra systems. In Maple it is even part of the software itself, where one can find the following commands

`SumTools[Hypergeometric][Zeilberger]` and `DEtools[Zeilberger]`,

(that also use enhancements and generalizations due to Sergey Abramov and Wolfram Koepf).

These algorithms (and general theory!) are widely used in combinatorics, of course, but also in topology (knot theory) (e.g. [G],[GL]), Nuclear Physics(e.g. [ABKS]), Calabi-Yao differential equations (that are important in String theory) (e.g. [Al]), Number Theory (e.g. [Su], [GR]), the discovery (and proof!) of infinite series that converge very fast to famous constants like $\frac{1}{\pi^2}$ (e.g. [Gu]), and most recently in (mainstream) multivariate statistics (e.g. [HNTT]) (not to be confused with multivariate combinatorial statistics mentioned above). I hope to continue to enlarge the scope of WZ theory, make it more efficient, and find novel applications.

Conclusion: The Medium is (a large part of) the Message [1].

The **intellectual merit** of the present proposal should be assessed on (at least) three levels. On the 'lowest' level this research contributes to combinatorics, an important field of mathematics with many applications to almost every branch of science, technology and human endeavor (the World Wide Web and telephone communication, CD players and pictures from Mars would be impossible without it, to mention just a few things that come to mind). The fact that I am using computers extensively in order to do my research in combinatorics should not be held for or against it, it is just another (legitimate) tool.

On the 'middle level', by using symbolic computation on a day-to-day basis, and trying to develop new algorithms (like in WZ theory), this research contributes, both directly and indirectly, to computer algebra, which is emerging as an indispensable tool not only in mathematics, but in all of science and technology.

Finally, on the 'highest-level', this research contributes to a new outlook and *awareness* in mathematical research. Mathematical research, until now, was *paper and pencil* and *a priori*, and people like Appel and Haken and Thomas Hales have to be apologetic and defensive about using computers. The computer is a mighty tool, go forth and use it! But we *humans* must think of *creative* ways of utilizing its immense potential, over-and-above its obvious use as a 'numerical and symbolic calculator' and 'brute force number- (and symbol-) cruncher'. We urgently need to develop new *methodologies* to enable us to make full use of computers. The potential applications are unforeseeable, but I am sure that future computer mathematics will make all past and present mathematics look like Mickey-Mouse stuff. But these new advances will not come by themselves. The role of the human mathematician would have to change from that of 'athlete' to that of 'coach', and this would also necessitate a change in mentality. I hope that my preaching (in my papers and the opinion column of my website), courses and seminars on Experimental Mathematics, and especially research (both the research papers viewed as case studies, and the more philosophical and methodological

[1] This is the same conclusion as in my previous proposal. It is even truer today.

papers), will form a modest, yet strictly positive, beginning. **If we build it** (a new experimental methodology for mathematics), **they will come** (present and future mathematicians will practice it.)

Similarly, the **broader impact** should also be judged on more than one level. Combinatorics per se, and Computer Algebra, are both essential to science, technology, and even entertainment. But, more generally, *mathematics*, as a whole, is one of the greatest pillars of our civilization and culture, both spiritually and materially. Helping change the way we practice mathematics (for the better, I am sure), would have the broadest impact on mathematics itself. **And what's good for mathematics is good for humanity**.

References Cited

[ABKS] J Ablinger, J Blümlein, S Klein, and C Schneider *Modern Summation Methods and the Computation of 2- and 3-loop Feynman Diagrams*, Nuclear Physics B - Proceedings Supplements **205-206** (2010), 110-115.
`http://arxiv.org/pdf/1006.4797` .

[Al] Gert Almkvist, *Calabi-Yau differential equations of degree 2 and 3 and Yifan Yang's pullback*, preprint.
`http://arxiv.org/abs/math/0612215` .

[AZ] Moa Apagodu and Doron Zeilberger, *Multi-Variable Zeilberger and Almkvist-Zeilberger Algorithms and the Sharpening of Wilf-Zeilberger Theory*, Adv. Appl. Math. **37**(2006), (Special issue in honor of Amitai Regev), 139-152.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/multiZ.html` .

[BZ] Andrew Baxter and Doron Zeilberger, *The Number of Inversions and the Major Index of Permutations are Asymptotically Joint-Independently Normal (Second Edition!)*, Personal Journal of Shalosh B. Ekhad and Doron Zeilberger (refereed by a group of eight world-class experts!), Feb. 4, 2011.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/invmaj.html` .

[BeP] J. Bétréma and J.-G. Penaud, *Modèles avec particules dures, animaux dirigé et séries en variables partiellement commutatives* .
`http://fr.arxiv.org/abs/math/0106210` .

[Bo1] Miklós Bóna, *"Combinatorics of Permutations"*, Chapman and Hall/CRC, 2004.

[Bo2] Miklós Bóna, *Exact enumeration of 1342-avoiding permutations: a close link with labeled trees and planar maps*, Journal of Combinatorial Theory, Series A **80** (1997), 257-272.

[BB] J. Borwein and D. Bailey, *Mathematics by Experiment.* A.K. Peters, Natick, 2004.

[BBCGLM] D. Bailey, J. Borwein, N. Calkin, R. Gingensohn, D. Luke, and V. Moll, *"Experimental Mathematics in Action"*, A.K. Peters, 2007.

[EGZ] Shalosh B. Ekhad, Evangelos Georgiadis and Doron Zeilberger, *How to Gamble If You're In a Hurry*, Journal of Difference Equations and Application, **19** (2013) , 520-526.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/himurim.html` .

[Gal] J. Gallian, *Interview with Doron Zeilberger*, MAA Focus May/June 2007, 14-17. (Available on-line at `www.maa.org`).

[Gar] Stavros Garoufalidis, *On the characteristic and deformation varieties of a knot*, Geometry and Topology **7** (2004), 291-309.
`http://arxiv.org/pdf/math.GT/0306230` .

[GL] Stavros Garoufalidis and T.T.Q. Le, *The colored Jones function is q-holonomic*, Geometry and Topology **9** (2005), 1253-1293.
`http://arxiv.org/pdf/math/0309214` .

[Ge] Ira M. Gessel (1990), *Symmetric functions and P-recursiveness*, Journal of Combinatorial Theory, Series A **53** (1990), 257-285.
`http://people.brandeis.edu/~ssel/homepage/papers/dfin.pdf` .

[GZ] Edinah K. Gnang and Doron Zeilberger, *Zeroless Arithmetic: Representing Integers ONLY using*, to appear in J. Difference Equations and Applications.
`http://math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/arif.html` .

[Gu] Jesús Guillera, *A new Ramanujan-like series for* $1/\pi^2$, Ramanujan J. **26**(2011), 369 - 374.
`http://arxiv.org/abs/1003.1915` .

[GR] Jesús Guillera and Mat Rogers, *Mahler measure and the WZ-algorithm*, Proc. of the Amer. Math. Soc., to appear.
`http://arxiv.org/abs/1006.1654` .

[HNTT] Hiroki Hashiguchi, Yasuhide Numata, Nobuki Takayama, Akimichi Takemura, *Holonomic gradient method for the distribution function of the largest root of a Wishart matrix*, Journal of Multivariate Analysis **117**(2013), 296-312.
`http://arxiv.org/abs/1201.0472` .

[HZ] Emilie Hogan and Doron Zeilberger, *A New Algorithm for Proving Global Asymptotic Stability of Rational Difference Equations*, Journal of Difference Equations and Applications **18**(2012), 1853-1873.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/gas.html` .

[KKZ] Christoph Koutschan, Manuel Kauers, and Doron Zeilberger, *A Proof Of George Andrews' and David Robbins' q-TSPP Conjecture*, Proceedings of the National Academy of Science, **108 #6** (Feb. 8, 2011), 2196-2199.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/qtsppRig.html` .

[KT] Christoph Koutschan and Thotsaporn Thanatipanonda, *Advanced Computer Algebra for Determinants*, Annals of Combinatorics, to appear.
`http://arxiv.org/abs/1112.0647` .

[K1] C. Krattenthaler, *Advanced Determinant Calculus*, Sém. Lothar. Comb. **42** (1999), B42q. ("The Andrews Festschrift", D. Foata and G.-N. Han(eds.)).
`http://arxiv.org/abs/math/9902004`

[K2] C. Krattenthaler, *Advanced Determinant Calculus: A Complement*, Linear Algebra Appl. **411**(2005), 68-166.
`http://arXiv.org/abs/math.Co/0503507` .

[NZ] Brian Nakamura and Doron Zeilberger, *Using Noonan-Zeilberger Functional Equations to Enumerate (in Polynomial Time!) Generalized Wilf Classes*, Advances in Applied Mathematics **50** (2013), 356-366.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/Gwilf.html` .

[O] Soichi Okada, *On the generating functions for certain classes of plane partitions*, J. Comb. Theory Ser. A **53**(1989), 1-23.

[PWZ] Marko Petkovsek, Herbert Wilf, and Doron Zeilberger, *"A=B"*, AK Peters, Wellesley, (1996). `Web Page: http://central.cis.upenn.edu/~wilf/AeqB.html` .

[R] John Riordan, "*Introduction to Combinatorial Analysis*", John Wiley, 1958. Reprinted by

Dover, 2002.

[SZ] Andrew V. Sills and Doron Zeilberger, *Rademacher's Infinite Partial Fraction Conjecture is (almost certainly) False*, Journal of Difference Equations and Applications **19** (2013), 680-689.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/hans.html` .

[St] Richard Stanley, *A baker's dozen of conjectures concerning plane partitions.* In *"Combinatoire énumérative"*, ed. G. Labelle and P. Leroux. Lecture Notes in Mathematics **1234** (1986), 285-293. Springer Verlag, New York.
`http://www-math.mit.edu/~rstan/pubs/pubfiles/67.pdf` .

[Su] Zhi-Wei Sun, *Super congruences and Euler numbers*, China Mathematics **54** (2011), 2509-2535.
`http://arxiv.org/abs/1001.4453` .

[T] Marko Thiel, *The inversion number and the major index are asymptotically jointly normally distributed on words*, preprint.
`http://arxiv.org/abs/1302.6708` .

[WZ] Herbert S. Wilf and Doron Zeilberger, *An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities*, Invent. Math. **108**(1992), 575-633.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/multiwz.html` .

[V] Vince Vatter, *Enumeration schemes for restricted permutations*, Combinatorics, Probability and Computing **17**(2008), 137-159.
`http://www-groups.mcs.st-andrews.ac.uk/~vince/publications/wilfplus/` .

[Z1] Doron Zeilberger, *The HOLONOMIC ANSATZ I. Foundations and Applications to Lattice Path Counting*, Annals of Combinatorics **11**(2007), 227-239 .
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzI.html` .

[Z2] Doron Zeilberger, *The HOLONOMIC ANSATZ II. Automatic DISCOVERY(!) and PROOF(!!) of Holonomic Determinant Evaluations*, Annals of Combinatorics **11**(2007), 241-247 .
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/ansatzII.html` .

[Z3] Doron Zeilberger, *An Enquiry Concerning Human (and Computer!) [Mathematical] Understanding*, in: C.S. Calude, ed., "Randomness & Complexity, from Leibniz to Chaitin", World Scientific, Singapore, Oct. 2007.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/enquiry.html` .

[Z4] Doron Zeilberger, *Teaching the Computer how to Discover(!) and then Prove(!!) (all by Itself(!!!)) Analogs of Collatz's Notorious 3x+1 Conjecture*, J. of Difference Equations and Applications **17**(2011), 375-386.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/collatz.html` .

[Z5] Doron Zeilberger, *The Amazing 3^n Theorem and its even more Amazing Proof [Discovered by Xavier G. Viennot and his École Bordelaise gang]*, Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, Aug 10, 2012.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/bordelaise.html` .

[Z6] Doron Zeilberger, *Enumeration Schemes, and More Importantly, Their Automatic Generation*, Annals of Combinatorics **2**(1998), 185-195 .
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/wilf.html` .

[Z7] Doron Zeilberger, *On Vince Vatter's Brilliant Extension of Doron Zeilberger's Enumeration Schemes for Counting Herb Wilf's Classes*, Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, Dec. 29, 2006.
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/vatter.html` .

[Z8] Doron Zeilberger, *Automated Counting of Towers Á La Bordelaise [Or: Footnote to p. 81 of the Flajolet-Sedgewick Chef-d'oeuvre]*, Personal Journal of Shalosh B. Ekhad and Doron Zeilberger, Dec. 17, 2012 .
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/migdal.html` .

[Z9] Doron Zeilberger, *The Automatic Central Limit Theorems Generator (and Much More!)*, Advances in Combinatorial Mathematics: Proceedings of the Waterloo Workshop in Computer Algebra 2008 in honor of Georgy P. Egorychev", chapter 8, pp. 165-174, (I.Kotsireas, E.Zima, eds. Springer Verlag, 2009) .
`http://www.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/georgy.html` .

[Zu] Wadim Zudilin, *Ramanujan-type supercongruences*, Journal of Number Theory **129**(2009), 18481857.
`http://arxiv.org/pdf/0805.2788` .