How to do MONTHLY problems with your computer

István Nemes, Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger

Fortunately, on 20 April 1977, all of this kludgery was rendered obsolete when I found a decision procedure for this problem. (A discrete analog to the Risch algorithm for indefinite integration.)

— R. William Gosper, Jr., *Indefinite Hypergeometric Sums in MACSYMA (1977)*

1 Introduction

The problem of finding simple evaluations of major classes of sums that involve factorials, binomial coefficients, and their q-analogues, has been completely solved. Sums that have the rather general form specified in Section 3 can all be done algorithmically, that is to say, you can do them on your own PC. Your computer evaluates the sum as a simple formula, if that's possible, and gives you a proof that you can check, or gives you a proof that your sum cannot be "done" in simple closed form, if that is the case.

We first briefly describe the algorithms and the theory that have achieved this goal. Second, to illustrate both the scope of the method and the fact that in some interesting cases human intervention still helps, we show how these computer methods would have fared in attacking 27 problems that have appeared over the years in the Problems section of this MONTHLY.

It happens (coincidentally, of course) that three of the authors of this article (PWZ) have just written a book [8] that describes the theoretical foundations of the solution of this problem, and also gives the software by means of which everyone can perform these sums *sans peine* (almost).

2 The methods

The methods that have achieved the complete solution of this class of problems are the following:

- Sister Celine's method [1]
- Gosper's algorithm [3]
- Zeilberger's algorithm `ct` ("creative telescoping") [11]
- Wilf and Zeilberger's WZ method [9]
- Petkovšek's algorithm `Hyper` [6]

Here is a brief description of the scope of each of these algorithms (full descriptions are in [8]). Computer programs, in *Maple* or *Mathematica* versions, that carry out each of these algorithms are available free at `http://www.cis.upenn.edu/~wilf/AeqB.html`.

Sister Celine's algorithm has been superseded by faster ones, but her work contains the original ideas on which the later algorithms have built. What it does can be stated pretty simply: it finds recurrences for hypergeometric summands. The fundamental theorem of this subject, which we state precisely in Section 3, holds that every proper hypergeometric summand does indeed satisfy a recurrence relation. For instance, if, under your summation sign, there lurks

$$F(n,k) = \binom{n}{k}^2,$$

then her method informs you that

$$\begin{aligned} &nF(n,k) - (2n-1)(F(n-1,k) + F(n-1,k-1)) \\ &\quad + (n-1)(F(n-2,k) - 2F(n-2,k-1) + F(n-2,k-2)) = 0. \end{aligned}$$

Why did you want to know that? Well, if you sum this recurrence over all integer k, you'll find immediately (try it!) that the sum $f(n) = \sum_k \binom{n}{k}^2$ satisfies $f(n) = 2(2n-1)f(n-1)/n$, and so by induction, $f(n) = \binom{2n}{n}$, and you have evaluated your sum. But, you say, you already knew that the sum of the squares of the binomial coefficients is $\binom{2n}{n}$? Sure you did, but the same method works on any sum of factorials and binomial coefficients and powers in the world, provided it's of the form described in Section 3 below. So it wasn't finding that one particular sum that was the revolutionary event. It was the fact that Sister Celine's method can find recurrences satisfied by any one of a huge class of summands, and, as was realized much later, from the recurrence for the summand there comes the recurrence for the sum, and from that comes the closed form evaluation of the sum, if it has one. We now have algorithms that handle all of those pieces.

Gosper's algorithm completely solves the problem of *indefinite* hypergeometric summation. Given a summand $F(k)$ that is a *hypergeometric term* in k (i.e., $F(k+1)/F(k)$ is a rational function of k), Gosper's algorithm finds a hypergeometric term $G(k)$ such that $F(k) = G(k+1) - G(k)$, if one exists, or prove that none exists, if that be the case. Thus it solves the discrete analogue of the antidifferentiation problem: instead of exhibiting a given *integrand* as the *derivative* of something, thereby enabling *integration* in finite terms, it exhibits a given *summand* as the *difference* of something, thereby enabling *summation* in finite terms. Examples of the operation of this algorithm are in Section 5.

Zeilberger's algorithm `ct` finds a recurrence for a given hypergeometric summand $F(n,k)$. To that extent, it solves the same problem that Sister Celine's method solves. The form of the recurrence that it finds is different, however, and that allows an enormous speedup in its operation time. His algorithm finds a recurrence for $F(n,k)$ in the form

$$\sum_{j=0}^{d} a_j(n) F(n+j,k) = G(n,k+1) - G(n,k), \tag{1}$$

in which G/F is a rational function (which the output exhibits) and the $a_j(n)$'s are polynomials in n. The power of this result derives from the fact that if we sum both sides of this recurrence over a certain range of k, the sum on the right side telescopes, and so is easy to handle, and we obtain a recurrence for the sum, $\sum_k F(n,k)$, that we are trying to deal with. The fundamental theorem guarantees that such recurrences always exists if F is a proper hypergeometric summand (see Section 3).

Wilf and Zeilberger's WZ method is at once a special case and a generalization of Zeilberger's method. In order to prove an identity of the type $\sum_k F(n,k) = 1$, it finds a recurrence of the form

$$F(n+1,k) - F(n,k) = G(n,k+1) - G(n,k), \tag{2}$$

where G/F is a rational function called the *proof certificate* of the identity. This form is clearly a special case of (1) above. A recurrence of this form does not always exist. When it does, one gets two benefits: first a very short proof of one's summation identity, and second, because of the symmetry of (2) in F and G, one finds a new identity, involving G, from the original one, involving F.

Petkovšek's algorithm `Hyper` finds closed form solutions $f(n)$ to linear difference equations with polynomial coefficients,

$$\sum_{j=0}^{d} a_j(n)f(n+j) = 0,$$

when such solutions exist, or it proves that they do not exist, when they do not. We use the phrase "closed form" in the following precise sense: $f(n)$ is said to be of (hypergeometric) *closed form* if it is equal to a linear combination of a fixed number, r, say, of hypergeometric terms in n. Thus `Hyper` completes the job of doing the summation problem because the methods just described, while they are guaranteed to give you a recurrence for your unknown sum, are not guaranteed to give you one of minimum order! But `Hyper` knows how to solve such recurrences in closed form, or to prove the impossibility of solving the recurrence in closed form, if that be the case.

3 The theory

Are these algorithms just more tricks, that might or might not work, to try on sums? Quite the contrary. In fact, the algorithms are accompanied by theorems that precisely describe circumstances under which they are *guaranteed* to work. So these are definitely not of the let's-see-if-it-works genre. They *will* work if the hypotheses of the relevant theorems are satisfied.

We are talking about sums of the form $f(n) = \sum_{k=a(n)}^{b(n)} F(n,k)$. The whole method rests on the fact that if $F(n,k)$ is a suitable summand then it satisfies a recurrence relation of a certain form. A summand $F(n,k)$ is suitable (*proper hypergeometric*) if it is of the form

$$F(n,k) = P(n,k)\frac{\prod_{i=1}^{I}(a_i n + b_i k + c_i)!}{\prod_{j=1}^{J}(u_j n + v_j k + w_j)!}x^n y^k, \tag{3}$$

in which

- $P(n,k)$ is a polynomial in n and k, whose degree is a specific integer, and
- the limits I, J on the products are fixed specific nonnegative integers, and
- the quantities a_i, b_i, u_j, v_j are specific integers, and
- the quantities c_i, w_j, x, y may depend on parameters.

Suppose we have a summand of that kind. What can we expect?

Theorem 1 *If $F(n,k)$ is proper hypergeometric then there exist a nonnegative integer d, a rational function $R(n,k)$, and polynomials $\{p_j(n)\}_{j=0}^{d}$, independent of k, such that $F(n,k)$ satisfies*

$$\sum_{j=0}^{d} p_j(n)F(n+j,k) = G(n,k+1) - G(n,k)$$

where $G(n,k) = R(n,k)F(n,k)$.

This theorem of Zeilberger, and the `creative telescoping` algorithm that carries it out, are used to find recurrences for given sums. The existence part of the proof follows from an earlier algorithm of Sister Mary Celine Fasenmyer that was used by her to find recurrences for hypergeometric polynomials.

From the recurrence for the summand one gets a recurrence for the sum. From the recurrence for the sum one gets the evaluation of the sum in closed form, if possible, or a proof of impossibility. The latter follow from algorithm `Hyper`, if the recurrence obtained is of order greater than 1. Just as the problem of finding recurrences has a life of its own, aside from its uses in evaluating sums, so algorithm `Hyper` has a life aside from finding out if sums have closed forms. Combinatorics is full of enumeration problems that lead to recurrences. With `Hyper` we can now find solutions of these, or else prove that closed forms do not exist, for the first time. In this way a large number of combinatorial sequences have been proved not to be of closed form, such as those in the following theorem.

Theorem 2 *None of the following famous sequences can be expressed in hypergeometric closed form:*

- *the sum of the cubes (also the fourth and fifth powers) of the binomial coefficients of order n,*
- *the number of $3 \times n$ Latin rectangles,*
- *the number of involutions on n letters,*
- *the derangement numbers,*
- *the sum of the first n of the binomial coefficients of order pn $(p > 2)$ [7],*

For the whole story of this remarkable current of mathematical thought, see [8].

4 Three recipes for success

Given a sum $S(n) = \sum_{k=a(n)}^{b(n)} F(n,k)$ with a proper hypergeometric summand, Zeilberger's algorithm `ct` yields a linear recurrence relation $\mathcal{L}$ with polynomial coefficients, of order $d \geq 0$, satisfied by $S(n)$. This is very helpful in the following situations that interest us here (and in many other situations too):

1. To prove that $S(n) = t_n$ where t_n is given in closed form, simply verify that t_n also satisfies $\mathcal{L}$, and that it agrees with $S(n)$ for d sufficiently large consecutive values of n.

2. To prove equality of two such sums use algorithm `ct` on both, and find a common multiple, $\mathcal{M}$, of the two resulting recurrences. If the order of $\mathcal{M}$ is m, verify that the two sums agree for m sufficiently large consecutive values of n.

3. To find a closed form evaluation of $S(n)$, note first that if $d = 0$, or $d = 1$ and $\mathcal{L}$ is homogeneous, then such an evaluation is immediate from $\mathcal{L}$. Otherwise, for various special reasons we might be able to solve $\mathcal{L}$ by inspection; it might be homogeneous with constant coefficients, for instance. But if no solutions are immediately apparent, then

   (a) If $d = 1$ and $\mathcal{L}$ is inhomogeneous, then $S(n)$ can be expressed in terms of an "indefinite" sum which Gosper's algorithm will put into closed form provided such a form exists.

   (b) Otherwise, use `Hyper` to find all closed form solutions of $\mathcal{L}$. Homogenize first if $\mathcal{L}$ is inhomogeneous. If you are lucky and $\mathcal{L}$ is satisfied by $t > 0$ linearly independent hypergeometric terms, then:

      i. If $t = d$, any solution of $\mathcal{L}$ can be put into closed form by choosing an appropriate linear combination of hypergeometric solutions.

      ii. If $t < d$, try to find a linear combination of hypergeometric solutions that agrees with $S(n)$ for d sufficiently large consecutive values of n.

      iii. If this fails, use hypergeometric solutions to reduce the order of $\mathcal{L}$. Repeat these steps with the new recurrence.

This procedure is *guaranteed* to decide whether $S(n)$ has a closed form evaluation (and to find it when it exists) whenever $F(n,k)$ is proper hypergeometric, and the limits of summation are either infinite (recall that summands often have compact support) or linear in n. But sometimes, with a little help, it works even when $F(n,k)$ is not proper hypergeometric (cf. problems E 3258, 10206, 10223, 10388 in the next Section). We refer the interested reader to [8, Chapter 8] for more details.

During the reviewing of this paper, one of the readers asked us for an example of a problem that we had not been able to do by these methods, even though it may have appeared to be a candidate. Of course such a problem would have to violate the hypotheses of Theorem 1, while at the same time seeming, at first glance anyway, to satisfy them. A good example of such a problem is an identity whose truth was conjectured by Borwein and Bradley, and which has recently been proved by Almqvist and Granville. It states that

$$\frac{5}{2}\sum_{k=1}^{n}\binom{2k}{k}\frac{k^2}{4n^4+k^4}\prod_{j=1}^{k-1}\frac{n^4-j^4}{4n^4+j^4}=\frac{1}{n^2}, \qquad (n \geq 1).$$

After factoring the fourth degree polynomials that appear in the summand one discovers that it has exactly the form (3), except that, for instance, one of the numbers a_i is $\sqrt{-1}$, which is not a specific integer, so the conditions are not satisfied.

5 Problems and solutions

We looked through MONTHLY problems on sums and recurrences that have been published since 1978, and selected 27 of the kind we're considering here. We used algorithms `ct` and `Hyper`, following the recipes given in the previous section. Wherever possible we used Gosper's algorithm and the WZ method, which technically are special cases of Zeilberger's algorithm `ct` corresponding to $d = 0$, and to $d = 1$ with given closed form evaluation, respectively (d being the order of the resulting recurrence). Besides our own implementations, we used the outstanding implementation of Zeilberger's algorithm in *Mathematica* by P. Paule and M. Schorn [5], which excels especially when the resulting recurrence is not homogeneous.

Many of the problems were solved completely automatically, while others required a little human help. For example, in several sums that involve the floor function we humans carried out the replacement

$$\sum_k F(n,k,\lfloor k/2\rfloor) = \sum_k (F(n,2k,k)+F(n,2k+1,k)) \qquad (4)$$

in which, if $F(n,k,m)$ is hypergeometric in n, k, m, the summand on the left *is not* hypergeometric, but the one on the right *is*. Other examples of human intervention include the choice of the "best" recurrence variable when the summand depends on more than one parameter, etc.

A notable exception in the amount of necessary human aid is the double sum in Problem E 3376, which required the sharp eyes of P. Paule [4] to notice a special relationship among the coefficients of the recurrence. In principle, of course, multiple sums can be handled by the methods of [10], but it is nice to be able to "do" a double sum with single-sum methods.

And now, here are the problems and their solutions!

Problem 6407 (Proposed in 1982, p. 703; solution in 1984, p. 315)

Define $\begin{bmatrix} n \\ k \end{bmatrix}$ *by means of the relation*

$$\begin{bmatrix} n \\ k \end{bmatrix} = F_{n,k}/F_{k,k} \quad F_{n,k} = (q^n - 1)(q^{n-1} - 1)\ldots(q^{n-k+1} - 1),$$

so that $\begin{bmatrix} n \\ k \end{bmatrix}$ *is the so-called Gaussian polynomial. Prove the identity*

$$\sum_{k=1}^{n} \frac{q^k}{1-q^k} = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{1-q^k} q^{k(k+1)/2} \begin{bmatrix} n \\ k \end{bmatrix}. \tag{5}$$

Denote the right side of (5) by $S(n)$. The q-version of algorithm `ct` yields the recurrence $S(n) - S(n-1) = q^n/(1-q^n)$, which is also satisfied by the left side of (5). As they agree for $n = 0$, the identity is proved.

Problem E 3021 (Proposed in 1983, p. 645; solution in 1986, p. 652)

Let

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k}^2 (1+x)^k (1-x)^{n-k}. \tag{6}$$

Express $p_n(x)$ *as an explicit function of* $1 - x^2$.

We provide a partial solution as follows. By algorithm `ct`

$$4(n+1)x^2 p_n(x) - 2(2n+3)p_{n+1}(x) + (n+2)p_{n+2}(x) = 0. \tag{7}$$

With $p_0(x) = 1$ and $p_1(x) = 2$, we see from this recurrence that $p_n(x)$ is a polynomial in x^2, and hence in $1 - x^2$. By comparing (7) with the three-term recurrence

$$(n+1)P_n(x) - (2n+3)xP_{n+1}(x) + (n+2)P_{n+2}(x) = 0$$

satisfied by the Legendre polynomials $P_n(x)$, we find that $p_n(x) = (2x)^n P_n(1/x)$.

Problem E 3022 (Proposed in 1983, p. 645; solution in 1986, p. 736)

Show that, for any $\alpha > 0$ and any positive integer N,

$$\sum_{k=1}^{N}(-1)^{k-1}\binom{N}{k}\frac{k}{1+(k-1)\alpha} = \prod_{k=1}^{N-1}\left(\frac{k+1}{k+\alpha^{-1}}\right).$$

We are to show that $\sum_k F(N,k) = 1$, where

$$F = \frac{(-1)^{k-1}(\alpha^{-1})^{\overline{N}}}{(k-1)!(N-k)!(k-1+\alpha^{-1})},$$

and $(x)^{\overline{n}}$ is the rising factorial. The WZ method does this with the rational proof certificate $(k-1)(\alpha^{-1}+k-1)/(N(k-N-1))$, and a check of the case $N=1$.

Problem E 3065 (Proposed in 1984, p. 649; solution in 1987, p. 378)

Let $n \geq 0$ be any integer and let k be any integer such that $k \geq n+1$. Then find a closed formula for

$$\sum_{j=0}^{n}\frac{(-1)^j}{j+1}\binom{k}{j}\binom{k-1-j}{n-j}.$$

Let $S(n)$ be the sum in question. Algorithm `ct` finds that

$$S(n)+S(n+1) = \frac{\binom{k}{n+1}}{n+2},$$

whence

$$S(n) = (-1)^n\left(\sum_{j=1}^{n}\frac{(-1)^j}{j+1}\binom{k}{j}+S(0)\right) = (-1)^n\sum_{j=0}^{n}\frac{(-1)^j}{j+1}\binom{k}{j}.$$

Gosper's algorithm gives the final answer

$$S(n) = \frac{1}{k+1}\left(\binom{k}{n+1}+(-1)^n\right).$$

Problem E 3088 (Proposed in 1985, p. 359; solution in 1987, p. 685)

Show that, for every positive integer n,

$$\sum_{k=1}^{n}\frac{k\cdot k!}{n^k}\binom{n}{k} = n. \qquad (8)$$

Let t_k denote the summand in (8). Gosper's algorithm finds that $t_k = s_{k+1} - s_k$ where $s_k = -nt_k/k$. Summing this recurrence on k from 1 to n gives the sum as $s_{n+1} - s_1 = n$.

Problem 6519 (Proposed in 1986, p. 403; solution in 1988, p. 156)

Let

$$F(a,b,m,n) = \sum_{k=0}^{m} \binom{a+m+n-2k}{n-k} \binom{a+n}{k} \binom{b+m}{m-k}, \tag{9}$$

where m and n are nonnegative integers. Show that $F(a,b,m,n) = F(b,a,n,m)$.

For $F(a,b,m,n)$ and $F(b,a,n,m)$ we compute recurrence relations with respect to n using algorithm `ct`. As it turns out, both sums satisfy the same recurrence of order 4:

$$\begin{aligned}
&- 3(1+a+n)(2+a+n)(3+a+n)S(n) \\
&+ 2(2+a+n)(3+a+n)(10+a+b-2m+4n)S(n+1) \\
&- (3+a+n)(53+5a+5b-ab-26m-4am-4bm \\
&\qquad\qquad\qquad\qquad + 36n+2an+2bn-8mn+6n^2)S(n+2) \\
&- (7+a+b+2n)(1+4a+4b+ab+8m+am+bm+an+bn+2mn)S(n+3) \\
&+ (4+n)(4+b+n)(4+a+b+n)S(n+4) = 0.
\end{aligned}$$

Checking that $F(a,b,m,n) = F(b,a,n,m)$ for $0 \le n \le 3$ therefore completes the proof.

Problem E 3190 (Proposed in 1987, p. 181; solution in 1988, p. 877)

Show that

$$\sum_{r=0}^{j} \frac{(-1)^r (N-2r)\binom{j}{r}}{(N-r)\dots(N-r-j)} = 0 \tag{10}$$

for $j > 0$ and $N > 2j$.

Let t_r denote the summand in (10). Gosper's algorithm finds that $t_r = s_{r+1} - s_r$ where $s_r = r(r+j-N)t_r/(j(N-2r))$. Summing this recurrence on r from 0 to $j-1$ gives the sum as $s_j - s_0 + t_j = 0$.

Problem E 3207 (Proposed in 1987, p. 456; solution in 1990, p. 67)

If m is a positive integer, let

$$F_m(x) = \sum_{k=0}^{m} \left(1 - \frac{k}{m}\right) \binom{2m}{k} x^k (1-x)^{2m-k}.$$

Show that

$$F_{m-1}(x) - F_m(x) = \frac{1}{2(2m-1)}\binom{2m}{m}x^m(1-x)^m, \quad \text{for } m > 1. \tag{11}$$

Zeilberger's algorithm `ct` instantly yields (11).

Problem E 3258 (Proposed in 1988, p. 259; solution in 1989, p. 651)

Prove that

$$\sum_{j=0}^{n}\binom{n}{j}2^{n-j}\binom{j}{\lfloor j/2\rfloor} = \binom{2n+1}{n}.$$

If we use the transformation (4) here, the sum in question becomes

$$\sum_j \frac{n+2}{2j+1}\binom{n}{2j}2^{n-2j-1}\binom{2j+1}{j}.$$

If we divide the summand by the claimed right side, $\binom{2n+1}{n}$, the WZ method proves the identity with a proof certificate of $4j(j+1)/((2n+3)(2j-n-1))$.

Problem E 3335 (Proposed in 1989, p. 525; solution in 1990, p. 927)

Solve the recurrence

$$x_0 = a,\ x_1 = b,\ x_{n+2} = x_{n+1} + x_n/(n+1) \quad \textit{for } n = 0, 1, 2, \ldots$$

both exactly (in terms of familiar functions of n) and asymptotically.

Algorithm `Hyper` gives one solution, $n+1$. Then by reducing the order we find that $(n+1)\sum_{k=0}^{n}(-1)^k/(k+1)!$ is another. So

$$x_n = (n+1)\left(a + (b-2a)\sum_{k=0}^{n+1}\frac{(-1)^k}{k!}\right),$$

which is asymptotic to $(n+1)\,(a + (b-2a)/e)$.

Problem E 3352 (Proposed in 1989, p. 838; solution in 1991, p. 369)

Show that

$$\sum_{n=0}^{\infty}\frac{1}{n!(n^4+n^2+1)} = \frac{e}{2}.$$

This is equivalent to

$$\sum_{n=0}^{\infty}\left(\frac{1}{n!(n^4+n^2+1)}-\frac{1}{2n!}\right)=0. \tag{12}$$

Let t_n denote the summand in (12). Gosper's algorithm finds that $t_n = s_{n+1} - s_n$ where $s_n = n^2/(2n!(n^2-n+1))$. Summing this recurrence on n from 0 to ∞ gives the sum as $s_\infty - s_0 = 0$.

Problem E 3376 (Proposed in 1990, p. 240; solution in 1992, p. 63)

Prove that

$$\sum_{i=0}^{N}\sum_{j=0}^{N}\binom{i+j}{j}^2\binom{4N-2i-2j}{2N-2j}=(2N+1)\binom{2N}{N}^2,$$

for any positive integer N.

Let $S(N,i)$ denote the inner sum on the left. Algorithm `ct` finds the recurrence

$$\sum_{j=0}^{3} p_j(N,i)S(N,i+j)=0 \tag{13}$$

where

$$\begin{aligned}
p_0(N,i) &= (1+i)(-i+N)(-1-2i+2N),\\
p_1(N,i) &= -18-32i-22i^2-6i^3-11N+4iN+8i^2N-30N^2-20iN^2,\\
p_2(N,i) &= (2+i)(27+23i+6i^2+9N-4iN+18N^2),\\
p_3(N,i) &= -2(2+i)(3+i)^2.
\end{aligned}$$

Following Paule [4], we notice that $\sum_{j=0}^{3} p_j(N,i-j) = -2(2N+1)^2$ is independent of i. Summing recurrence (13) on i from -3 to N, and changing the order of summation to take advantage of this, we obtain

$$\sum_{i=1}^{3}\sum_{j=0}^{3-i}(S(N,-i)p_j(N,-i-j)+S(N,N+i)p_{j+i}(N,N-j))-2(2N+1)^2\sum_{i=0}^{N}S(N,i)=0. \tag{14}$$

Since for positive integer i the sums $S(N,-i)$ and $S(N,N+i)$ contain only i nonzero terms, the result

$$\sum_{i=0}^{N}S(N,i)=(2N+1)\binom{2N}{N}^2$$

can be readily computed from (14).

Problem E 3439 (Proposed in 1991, p. 437; solution in 1993, p. 188)

If M and N are nonnegative integers, prove that

$$\binom{M+N}{M} = \sum_{0\le a\le \frac{M-1}{2}} \binom{M-a-1}{a}\binom{N+a}{2a+1} + \sum_{0\le a\le \frac{M}{2}} \binom{M-a}{a}\binom{N+a}{2a}. \tag{15}$$

If $M = 0$ both sides are 1. When $M > 0$ the two sums on the right can be combined into a single hypergeometric sum

$$S(N) = \sum_{a=0}^{M-1} \frac{MN - 2aN + aM + M - a}{(2a+1)(M-a)} \binom{M-a}{a}\binom{N+a}{2a}$$

for which creative telescoping finds the recurrence

$$(N+1)S(N+1) - (M+N+1)S(N) = 0$$

satisfied by $\binom{M+N}{M}$. As the two sides of (15) agree at $N = 0$, the identity is proved.

Problem 10206 (Proposed in 1992, p. 266; solution in 1995, p. 657)

If m and k are positive integers, prove that

$$\sum_r \binom{r}{k-r}\binom{m}{r} = \sum_j \binom{\lfloor j/2\rfloor}{k-j}\binom{m-k+\lfloor 3j/2\rfloor}{j}. \tag{16}$$

If we apply the transformation (4) to the sum on the right of (16) it becomes

$$\sum_j \frac{km - k^2 + 3jk + 3j - 2jm + 1}{(2j+1)(j+1)} \binom{j+1}{k-2j}\binom{m-k+3j}{2j}. \tag{17}$$

The creative telescoping algorithm with respect to k shows that the same recurrence

$$(k-2m)S(k) + (1-m+k)S(k+1) + (k+2)S(k+2) = 0$$

is satisfied both by (17) and by the left side of (16). Since they agree for $k = 0, 1$, the identity is proved.

Problem 10223 (Proposed in 1992, p. 462)

For $p \in \mathbb{R}$, $q = 1 - p$, and positive integers n, prove

$$\sum_{k=n}^{2n-1} \binom{k-1}{n-1} (p^n\, q^{k-n} + p^{k-n}\, q^n) = 1.$$

Write the sum as $S_n(p)+S_n(1-p)$ where $S_n(p) = \sum_{k=n}^{2n-1} \binom{k-1}{n-1} p^n\, (1-p)^{k-n}$. By algorithm `ct`

$$S_{n+1}(p) - S_n(p) = \frac{2p-1}{2}(p\,(1-p))^n \binom{2n}{n},$$

therefore

$$S_{n+1}(1-p) - S_n(1-p) = \frac{1-2p}{2}(p\,(1-p))^n \binom{2n}{n},$$

which implies that $S_n(p) + S_n(1-p)$ is constant. Evaluating the sum for $n = 1$ completes the proof.

Problem 10229 (Proposed in 1992, p. 570; solution in 1994, p. 797)

Given that m and p are integers with $m \geq p \geq 1$, evaluate

$$\sum_{j=1}^{p} \binom{1/2}{m-j+1} \binom{1/2}{m+j}. \tag{18}$$

Let t_j denote the summand in (18). Gosper's algorithm finds that $t_j = s_{j+1} - s_j$ where $s_j = (j-1)(j+m)(2m-2j+1)t_j/(m(2m+1))$. Summing this recurrence on j from 1 to p gives the sum as $s_{p+1} - s_1$ which is

$$\binom{1/2}{m-p+1} \binom{1/2}{m+p} \frac{p(m-p+1)(2m+2p-1)}{m(2m+1)}.$$

Problem 10332 (Proposed in 1993, p. 796; solution in 1996, p. 702)

If n and k are integers with $0 \leq k \leq n$, prove that

$$\binom{2n}{n+k} = \sum_j 2^{n-k-2j} \binom{n}{j} \binom{n-j}{j+k}.$$

If we divide the summand on the right by the claimed left side, $\binom{2n}{n+k}$, the WZ method proves the identity with a proof certificate of $4j(j+k)/((2j+k-n-1)(2n+1))$.

Problem 10357 (Proposed in 1994, p. 75)

Define integers $a_{m,n}$ *by*

$$\frac{1}{1-u-v+2uv} = \sum_{m,n=0}^{\infty} a_{m,n} u^m v^n.$$

Show that $(-1)^j a_{2j,2j+2}$ *is the Catalan number* $\binom{2j}{j}/(j+1)$.

We expand the function on the left into a power series in u and v, using first the geometric series, then the Binomial Theorem, and finally the derivatives of the geometric series:

$$\begin{aligned}
\frac{1}{1-u-v+2uv} &= \frac{1}{(1-v)(1-u\frac{1-2v}{1-v})} = \frac{1}{1-v}\sum_{m=0}^{\infty}\left(\frac{1-2v}{1-v}\right)^m u^m \\
&= \frac{1}{1-v}\sum_{m=0}^{\infty}\left(1-\frac{v}{1-v}\right)^m u^m = \sum_{m,k=0}^{\infty}\binom{m}{k}\frac{(-v)^k}{(1-v)^{k+1}}u^m \\
&= \sum_{m,n,k=0}^{\infty}(-1)^k\binom{m}{k}\binom{n}{k}u^m v^n.
\end{aligned}$$

From this we see that $a_{m,n} = \sum_k (-1)^k \binom{m}{k}\binom{n}{k}$. Using Zeilberger's algorithm `ct` on $S(j) := (-1)^j a_{2j,2j+2} = \sum_k (-1)^{k+j}\binom{2j}{k}\binom{2j+2}{k}$ we obtain the recurrence

$$\begin{aligned}
&16(1+j)(1+2j)(9+4j)\,S(j) - 2(7+4j)(21+28j+8j^2)\,S(j+1) \\
&+ (3+j)(5+2j)(5+4j)\,S(j+2) = 0,
\end{aligned}$$

which is satisfied by $c_j = \binom{2j}{j}/(j+1)$. As $S(j) = c_j$ for $j = 0, 1$, the proof is complete.

Problem 10363 (Proposed in 1994, p. 175)

If m, n *are integers satisfying* $1 \le m \le n-1$, *prove that*

$$\binom{2n-m-1}{2n-2m-1} - \binom{n-1}{m} = \sum_k \sum_j \binom{k+j}{k}\binom{2n-m-2k-j-3}{2(n-m-k-1)}.$$

Let

$$S(k) = \sum_{j=0}^{m-1}\binom{k+j}{k}\binom{2n-m-2k-j-3}{2(n-m-k-1)}.$$

Algorithm `ct` finds the first-order recurrence

$$(2m-2n+k+1)\,S(k)+(2n-m-k-2)S(k+1)=0,$$

for $0 \le k \le m-n-2$. From this, $S(k)=C(n,m)\binom{2m-2n+k}{k}/\binom{m-2n+k+1}{k}$, where

$$C(n,m)=S(0)=\sum_{j=0}^{m-1}\binom{2n-m-j-3}{2(n-m-1)}=\binom{2n-m-2}{m-1},$$

by Gosper's algorithm. Finally,

$$\sum_{k=0}^{n-m-1} S(k)=\binom{2n-m-2}{m-1}\sum_{k=0}^{n-m-1}\frac{\binom{2m-2n+k}{k}}{\binom{m-2n+k+1}{k}}=\binom{2n-m-1}{m}-\binom{n-1}{m},$$

by Gosper's algorithm again.

Problem 10375 (Proposed in 1994, p. 362)

Find the complete solution of the recurrence

$$U_{n+2}=2(2n+3)^2U_{n+1}-4(n+1)^2(2n+1)(2n+3)U_n, \quad \textit{for} \quad n \ge 0.$$

Petkovšek's algorithm `Hyper` finds that $(2n)!$ satisfies the recurrence. Then by reducing the order we find that $(2n)!H_n$ also satisfies it, where $H_n = 1+1/2+\ldots+1/n$ is the n-th harmonic number. Hence the complete solution is $U_n=(2n)!\,(C_1+C_2H_n)$ where C_1 and C_2 are arbitrary constants.

Problem 10388 (Proposed in 1994, p. 474)

Find

$$\sum_{k=0}^{n}\binom{n}{k}\binom{\frac{n-3}{4}-\frac{k}{2}+p}{2p}$$

where n and p are positive integers.

The summand is not hypergeometric due to the non-integral coefficient of k. Denote the sum by $S(n,p)$, and write $S(n,p)=S_1(n,p)+S_2(n,p)$ where

$$S_1(n,p)=\sum_k\binom{n}{2k}\binom{\frac{n-3}{4}-k+p}{2p}, \quad S_2(n,p)=\sum_k\binom{n}{2k+1}\binom{\frac{n-3}{4}-k-\frac{1}{2}+p}{2p}.$$

Zeilberger's algorithm `ct` finds that both sums satisfy the same recurrence with respect to p, viz.,

$$\begin{aligned}&(n-4p-7)(n-4p-5)(n-4p-3)(n-4p-1)S_i(n,p)-\\&16(n-4p-7)(n-4p-5)(5n+4pn-8p^2-20p-14)S_i(n,p+1)+\\&512(n-2p-3)(n-2p-4)(2p+3)(p+2)S_i(n,p+2)=0,\quad \text{for } i\in\{1,2\}.\end{aligned}\tag{19}$$

Then $S(n,p)$ also satisfies (19). Algorithm `Hyper` finds the complete solution of this recurrence as

$$S(n,p)=C_1(n)\frac{\binom{\frac{n-3}{4}}{p}\binom{\frac{n-1}{4}}{p}}{(-4)^p\binom{-\frac{1}{2}}{p}\binom{\frac{n-2}{2}}{p}}+C_2(n)\frac{\binom{\frac{n-3}{4}}{p}\binom{\frac{n-1}{4}}{p}}{4^p\binom{\frac{n-1}{2}}{p}}.$$

The initial conditions $S(n,0)=2^n$ and

$$S(n,1)=\sum_{k=0}^{n}\binom{n}{k}\binom{\frac{n-3}{4}-\frac{k}{2}+1}{2}=\frac{n-3}{32}2^n,$$

can be found using algorithm `ct` again. From these $C_1(n)=0$ and $C_2(n)=2^n$, whence

$$S(n,p)=\frac{2^n}{4^p}\frac{\binom{\frac{n-3}{4}}{p}\binom{\frac{n-1}{4}}{p}}{\binom{\frac{n-1}{2}}{p}}.$$

Problem 10396 (Proposed in 1994, p. 681)

Let $\alpha>0$ and let $\langle b_n : n\ge 1\rangle$ be defined recursively by $b_1=\alpha$, $b_2=3\alpha$,

$$b_{n+1}=(2n+1)b_n-(n^2+\alpha^2)b_{n-1}\quad (n\ge 2).\tag{20}$$

Prove that $\langle b_n\rangle$ contains infinitely many positive and infinitely many negative terms.

Algorithm `Hyper` finds the complete solution of (20) as $C_1(1+\alpha i)(2+\alpha i)\cdots(n+\alpha i)+C_2(1-\alpha i)(2-\alpha i)\cdots(n-\alpha i)$ where $i^2=-1$ and C_1, C_2 are arbitrary constants. From the initial conditions, $C_1=1/(2i)$ and $C_2=-1/(2i)$, so that

$$b_n=\Im(1+\alpha i)(2+\alpha i)\cdots(n+\alpha i).$$

Write $b_n=\Im z_n$, where $z_0=1$ and $z_n=(n+\alpha i)z_{n-1}$ (for $n\ge 1$). As $\lim_{n\to\infty}\arctan\alpha/n=0$ but $\sum_{n=1}^{\infty}\arctan\alpha/n=\infty$, it is clear that the imaginary part of z_n is positive and negative infinitely often.

Problem 10403 (Proposed in 1994, p. 792)

Define a sequence $\langle y_n \rangle$ recursively by $y_0 = 1$, $y_1 = 3$ and

$$y_{n+1} = (2n+3)y_n - 2ny_{n-1} + 8n \tag{21}$$

for $n \geq 1$. Find an asymptotic formula for y_n.

Algorithm `Hyper` finds that $2^n n!$ satisfies the homogeneous part of (21). By reduction of order we obtain

$$y_n = 2^n n! \sum_{k=0}^{n} \frac{1 + 8\sum_{m=1}^{k-1} m}{2^k k!} = 2^n n! \sum_{k=0}^{n} \frac{1 + 4k(k-1)}{2^k k!} = 2^{n+1} n! \sum_{k=0}^{n} \frac{1}{2^k k!} - 2n - 1,$$

which is asymptotic to $2^{n+1} n! \sqrt{e}$.

Problem 10424 (Proposed in 1995, p. 70)

Evaluate the sum

$$\sum_{0 \leq k \leq \frac{n}{3}} 2^k \frac{n}{n-k} \binom{n-k}{2k}. \tag{22}$$

Denote the sum in (22) by $S(n)$. The creative telescoping algorithm yields the constant-coefficient recurrence

$$S(n+3) - 2S(n+2) + S(n+1) - 2S(n) = 0 \quad (n \geq 1). \tag{23}$$

The roots of the characteristic polynomial are $2, \pm i$, and the solution of (23) satisfying $S(1) = S(2) = 1$ is

$$S(n) = 2^{n-1} - \sin\frac{(n-1)\pi}{2} = 2^{n-1} + \left\{ \begin{array}{rll} 0, & n \equiv 1 & \pmod 2 \\ -1, & n \equiv 2 & \pmod 4 \\ 1, & n \equiv 0 & \pmod 4 \end{array} \right. .$$

Problem 10466 (Proposed in 1995, p. 654)

For $x \in \mathbb{C}$ and $n \in \mathbb{N}$, prove the following identities between polynomials:

$$(a) \qquad (-4)^n \sum_{j=0}^{n} \binom{x+\frac{1}{2}}{j}\binom{n-1-x}{2n-j} = \binom{2n}{n} \sum_{j=0}^{n} \binom{x+j}{2j}\binom{x-j}{2n-2j}$$

(b) *For all $m \in \mathbb{N}$, with $0 \leq m \leq 2n$, generalize (a) to*

$$(-4)^n \sum_{j=0}^{n} \binom{x+\frac{1}{2}}{j}\binom{n-1-x}{2n-j} = \binom{2n}{n} \sum_{j=-\lfloor\frac{m}{2}\rfloor}^{n-\lfloor\frac{m}{2}\rfloor} \binom{x+j}{2j+m}\binom{x-j}{2n-m-2j}$$

(a) Denote the sums on the left and right by $S(n)$ and $T(n)$, respectively. Algorithm `ct` finds recurrence relations

$$(1+2n)S(n) - 2(1+n)S(n+1) = \frac{(1+3n-2x)(1+2x)}{1+n}\binom{2x}{2n}\binom{2n}{n} \tag{24}$$

and

$$\begin{aligned}&2(1+n)(2+n)^2(1+3n-2x)T(n+2) - \\ &(1+n)(6+41n+67n^2+30n^3 - x(28+98n+68n^2-40x-56nx+16x^2))T(n+1) + \\ &2(1+2n)(4+3n-2x)(x-n)(2x-2n-1)T(n) = 0.\end{aligned}$$

The latter recurrence turns out to be the homogenization of the former, so $S(n)$ and $T(n)$ satisfy the same recurrence of order 2. As they agree for $n = 0, 1$, they are identical. – As a bonus, from (24) we can express $S(n)$ in terms of an "indefinite" sum in which the summand does not depend on n:

$$S(n) = \frac{\binom{2n}{n}}{4^n}\left(1 - (2x+1)\sum_{k=0}^{n-1} 4^k \frac{3k+1-2x}{(k+1)(2k+1)}\binom{2x}{2k}\right).$$

(b) Denote the sum on the right by $U(n,m)$. Algorithm `ct` with respect to m finds the recurrence

$$(2n-m-1)\,U(n,m+2) - 2n\,U(n,m+1) + (m+1)\,U(n,m) = 0. \tag{25}$$

For $U(n,1) = \binom{2n}{n}\sum_{j=0}^{n-1}\binom{x+j}{2j+1}\binom{x-j}{2n-1-2j}$, algorithm `ct` finds the same recurrence as for $T(n)$ (for $n \geq 1$). As they agree for $n = 1, 2$, they agree for all $n \geq 1$. So $U(n,1) = T(n) = U(n,0)$ (for $n \geq 1$). It follows from (25) that $U(n,m) = T(n) = S(n)$.

Problem 10473 (Proposed in 1995, p. 745)

Prove that there are infinitely many positive integers m such that

$$\frac{1}{5\cdot 2^m}\sum_{k=0}^{m}\binom{2m+1}{2k}3^k \tag{26}$$

is an odd integer.

Denote (26) by $S(m)$. Algorithm `ct` yields the constant-coefficient recurrence

$$S(m+2) - 4S(m+1) + S(m) = 0. \tag{27}$$

The sequence $T(m) = 5S(m)$ satisfies (27) as well and starts out as $\langle 1, 5, \ldots \rangle$, hence it is integral. Let $T_5(m) = T(m) \bmod 5$ and $T_2(m) = T(m) \bmod 2$. Using (27) mod 5 and mod 2, respectively, we see that $T_5 = \langle 1, 0, 4, 1, 0, 4, \ldots \rangle$ and $T_2 = \langle 1, 1, 1, \ldots \rangle$, so that $S(3k+1) = T(3k+1)/5$ is an odd integer for all $k \geq 0$.

Problem 10494 (Proposed in 1996, p. 74)

For each positive integer n, evaluate the sum

$$\sum_{k=0}^{2n} (-1)^k \binom{4n}{2k} \Big/ \binom{2n}{k}. \tag{28}$$

Let t_k denote the summand in (28). Gosper's algorithm finds that $t_k = s_{k+1} - s_k$ where $s_k = (2k-1)t_k/(2(1-2n))$. Summing on k from 0 to $2n-1$ gives the sum as $s_{2n} - s_0 + t_{2n} = 1/(1-2n)$.

6 Conclusion

Quite often, MONTHLY problems require evaluation of a single or double sum in closed form, or a proof of equality of two such sums. When the summand involves binomial coefficients, factorials, products of rational functions, and exponential functions with constant base, there are very good chances that such a problem can be solved automatically by Gosper's algorithm or by its generalizations: the WZ method, Zeilberger's algorithm `ct`, and algorithm `Hyper`. Although Gosper's algorithm is now over 19 years old (see the quotation on the title page!), it seems that it is not as widely known as it deserves to be.

To help spread the word, we surveyed MONTHLY problems that have appeared since the publication of Gosper's algorithm in 1978. We have presented here a selection of those on which these methods are successful. For a similar list of earlier problems, see the Web site `http://www.math.temple.edu/~zeilberg/Monthly.html`.

References

[1] Sister Mary Celine Fasenmyer, *Some Generalized Hypergeometric Polynomials*, Ph.D. dissertation, Univ. of Michigan, 1945.

[2] R. W. Gosper, Jr., Indefinite hypergeometric sums in MACSYMA, *Proc. MACSYMA Users Conference*, Berkeley CA, 1977, 237 – 252.

[3] R. W. Gosper, Jr., Decision procedure for indefinite hypergeometric summation, *Proc. Natl. Acad. Sci. USA* **75** (1978) 40 – 42.

[4] P. Paule, *Solution of a Séminaire homework example (28th SLC)*, RISC-Linz Report Series No. 92-59, Linz 1992.

[5] P. Paule and M. Schorn, A *Mathematica* version of Zeilberger's algorithm for proving binomial coefficient identities, *J. Symb. Comput.* **20** (1995) 673–698.

[6] M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients, *J. Symb. Comput.* **14** (1992) 243–264.

[7] M. Petkovšek and H. S. Wilf, When can the sum of $(1/p)$th of the binomial coefficients have closed form?, *Electronic J. Combinatorics*, to appear.

[8] M. Petkovšek, H. S. Wilf and D. Zeilberger, $A = B$, A K Peters, Ltd., Wellesley, MA, 1996.

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

[10] H. S. Wilf and D. Zeilberger, An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities, *Inv. Math.* **108** (1992) 575 – 633.

[11] D. Zeilberger, The method of creative telescoping, *J. Symb. Comput.* **11** (1991) 195 – 204.

Nemes:
RISC Linz
Johannes Kepler University
Linz, Austria
Istvan.Nemes@risc.uni-linz.ac.at

Petkovšek:
University of Ljubljana
Ljubljana, Slovenia
Marko.Petkovsek@mat.uni-lj.si

Wilf:
University of Pennsylvania
Philadelphia, PA, USA
wilf@math.upenn.edu

Zeilberger:
Temple University
Philadelphia, PA, USA
zeilberg@euclid.math.temple.edu