Discrete Mathematics 80 (1990) 207–211
North-Holland

COMMUNICATION

A FAST ALGORITHM FOR PROVING TERMINATING HYPERGEOMETRIC IDENTITIES

Doron ZEILBERGER*
Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA

Communicated by D. Foata
Received October 11 1989

An algorithm for proving terminating hypergeometric identities, and thus binomial coefficients identities, is presented. It is based upon Gosper's algorithm for indefinite hypergeometric summation. A MAPLE program implementing this algorithm succeeded in proving almost all known identities. Hitherto the proof of such identities was an exclusively human endeavor.

1. Introduction

I will describe an explicit algorithm for *proving* binomial coefficients identities (≡terminating hypergeometric identities) of the form

$$\sum_k F(n,k) = \text{rhs}(n), \tag{1}$$

where $F(n, k)$ has the form

$$F(n,k) = \binom{n}{k} \frac{\prod_{i=1}^{A} (a_i n + a_i' k + a_i'')!}{\prod_{i=1}^{B} (b_i n + b_i' k + b_i'')!} z^k, \tag{2a}$$

where the a_i, a_i', b_i, b_i' have to be *constant*, specific, (positive or negative) integers, but z, a_i'', and b_i'' can be any complex numbers or parameters, and $x!$ means $\Gamma(x+1)$.

The right side of (1), rhs(n), may be given explicitly, in the form

$$\text{rhs}(n) = C \frac{\prod_{i=1}^{\bar{A}} (\bar{a}_i n + \bar{a}_i')!}{\prod_{i=1}^{\bar{B}} (\bar{b}_i n + \bar{b}_i')!} \bar{x}^n, \tag{2b}$$

where $\bar{a}_i$ and $\bar{b}_i$ are specific (positive or negative) integers, and $\bar{a}_i'$, $\bar{b}_i'$, C, and $\bar{x}$, are complex numbers or parameters. Another possibility is that the right side is given implicitly in terms of a minimal (ordinary) linear recurrence operator with polynomial coefficients, conj(n, N), that annihilates rhs(n), together with the appropriate initial conditions.

* Research supported in part by NSF grant DMS88-00663.

The theory of hypergeometric series forms a venerable part of classical analysis [12, 3] and is nowadays making a fast comeback [1, 2]. The classical identities of Vandermonde–Chu, Pfaff–Saalschutz, Dixon, Whipple, Watson and Dougall [3] form an important part of this theory. Recently Gosper discovered many "strange" hypergeometric identities, most of which were proved by Gessel and Stanton [9]. My algorithm, implemented on MAPLE (running on an AT & T 3B1 UNIX PC) was able to prove all the above mentioned classical identities and almost all those in [9], including one ((6.2)) that they were unable to prove, and that was only proved recently by Gasper and Rahman [8].

A MAPLE program implementing the algorithm is available from the author upon request. A q-analog will appear in a forthcoming paper.

A sequence $a(k)$ is said to have *closed form* if $a(k+1)/a(k)$ is a rational function in k.

2. The slow algorithm

In [15] I show that I.N. Bernstein's theory of holonomic systems [4, 5] implies algorithms for proving a large class of special function identities, that includes all terminating hypergeometric identities. For the latter case I gave an explicit algorithm that succeeded in proving some non-trivial identities, but that ran out of memory space, at least on my PC, on most of the deeper identities. Although of limited practical use, this algorithm is important for theoretical reasons, since its validity implies the validity of my new, more efficient, algorithm. I will therefore first recall this slow algorithm [15].

For any discrete function of two variables $F(n, k)$ let N and K be the shift operators in the n and k directions respectively: $NF(n, k) := F(n+1, k)$ and $KF(n, k) := F(n, k+1)$. Now form $F(n+1, k)/F(n, k) = A(n, k)/B(n, k)$ and $F(n, k+1)/F(n, k) = A'(n, k)/B'(n, k)$. Due to the form (2a) of $F(n, k)$ it follows that A, B, A', B' are polynomials in n and k. Let $P(N, K, n, k) := B(n, k)N - A(n, k)$ and $Q(N, K, n, k) := B'(n, k)K - A'(n, k)$. It follows that $P(N, K, n, k)F \equiv 0$, and $Q(N, K, n, k)F \equiv 0$. By using an elimination algorithm in the (non-commutative) algebra of linear partial difference operators with polynomial coefficients, that is an adaptation of the classical elimination method from commutative algebra [13], one can find operators $C(N, K, n, k)$ and $D(N, K, n, k)$ such that $R(N, K, n) := C(N, K, n, k)P(N, K, n, k) + D(N, K, n, k)Q(N, K, n, k)$ is non-zero and does not involve k. It follows that $R(N, K, n)F \equiv 0$. It is then shown that $S(N, n) := R(N, 1, n)$ is non-zero and we write $R(N, K, n) = S(N, n) - (K-1)\bar{R}(N, K, N)$. It follows that $S(N, n)F(n, k) = (K-1)\bar{R}(N, K, n)F(n, k)$. Now let $G(n, k) := \bar{R}(N, K, n)F(n, k)$, which is easily seen to be a certain multiple of $F(n, k)$ by a rational function (and thus of *closed form*), and we get

$$S(N, n)F(n, k) = G(n, k+1) - G(n, k). \tag{3}$$

and by summing (3) with respect to k we get that $\text{lhs}(n) := \sum_k F(n, k)$ satisfies the linear ordinary recurrence equation with polynomial coefficients $S(N, n)\text{lhs}(n) \equiv 0$.

If the conjectured right side of (1), $\text{rhs}(n)$, is given explicitly, by (2b), we compute $\text{rhs}(n+1)/\text{ths}(n)$, a certain rational function, cross multiply, and get a first order recurrence operator annihilating $\text{rhs}(n)$, let us call it $\text{conj}(N, n)$. If $\text{rhs}(n)$ is not given explicitly, then $\text{conj}(N, n)$ is given from the outset, where it is assumed that it has minimal order. Now use the (Euclidean) division algorithm in the algebra of linear recurrence operators to write $S(N, n) = qu(N, n)\text{conj}(N, n) + \text{rem}(N, n)$, where $\text{rem}(N, N)$ has smaller order than $\text{conj}(N, n)$. If indeed $\text{lhs}(n) \equiv \text{rhs}(n)$, this forces $\text{rem}(N, n)\text{rhs}(n) \equiv 0$, and by the minimality of the order of $\text{conj}(N, n)$, $\text{rem}(N, n)$ has to be identically zero. We should thus have for some operator $qu(N, n)$:

$$qu(N, n)[\text{conj}(N, n)F(n, k)] = G(n, k+1) - G(n, k), \tag{4}$$

where $G(n, k)$ has *closed form*. It follows, upon summing with respect to k, that $qu(N, n)[\text{conj}(N, n)\text{lhs}(n)] = 0$, and we conclude that $\text{conj}(N, n)\text{lhs}(n) \equiv 0$ provided it is true in the appropriate initial values, that are easily checked, and thus $\text{lhs}(n) \equiv \text{rhs}(n)$, provided it is true at the appropriate initial value(s).

3. The fast algorithm

The elimination algorithm alluded to in the above section is very time- and space-consuming. However, note that once we have gone through the trouble of finding $G(n, k)$ and $qu(N, n)$ the proof that $\text{conj}(N, n)\text{lhs}(n) = 0$ is an immediate consequence, modulo purely routine verification. Perhaps there is *another* way of coming up with the winning team of $qu(N, n)$ and $G(n, k)$? Indeed there is. Considering n as an auxiliary parameter, and given $qu(N, n)$, we have to look for *closed form* solutions $G(n, k)$ of the "difference equation" (4). A general algorithm that given a *closed form sequence* $a(k)$, decides whether there is a closed form solution $S(k)$ of the difference equation $S(k) - S(k-1) = a(k)$, and finds such an $S(k)$ in the affirmative case, was given by Gosper [10]. The only trouble is that we do not know $qu(N, n)$ to begin with. So we start being optimistic and try $qu(N, n) = 1$, and this works for 99% of the cases in practice. If this fails we put $qu(N, n) = b_1(n)N + b_0(n)$, plug it in (4), and let b_0 and b_1 be additional unknowns to those in Gosper's algorithm. If there are such b_0 and b_1, Gosper's algorithm will tell us what they are. If this fails too we try $qu(N, n) =$

$b_2N^2 + b_1N + b_0$, until we get a closed form solution $G(n, k)$ to (4), together with the corresponding $qu(N, n)$.

Example. For Dixon's classical identity [3],

$$F(n, k) = (-1)^k \binom{n+b}{n+k}\binom{n+c}{c+k}\binom{b+c}{b+k}, \quad \mathrm{rhs}(n) = \frac{(n+b+c)!}{n!\,b!\,c!}.$$

It follows that $\mathrm{conj}(N, n) = (n+1)N - (n+b+c+1)$, and it took 30 seconds to find $qu(N, n) = 1$, and $G(n, k) := ((k+b)(k+c)/2(n-k+1))F(n, k)$.

4. Conclusion

The above method was dubbed *creative telescoping* in [11], where it was used to prove that the Apery numbers satisfy their recurrence. Only there it is described as a sequence of *mirabilia*. By the general considerations of Section 2 we now know that there is always a miracle (possibly modified by the $qu(N, n)$) and thanks to Gosper's algorithm, we can always perform the miracle, not only in principle, like in [15], but also in practice.

I hope that the present algorithm will relieve humans from the tedium of devising proofs that G.H. Hardy used to call "essentially verifications", and will encourage them to pursue elegant and insightful proofs. That there still is need for such proofs can be demonstrated by Foata's beautiful proof [7] of Dixon's identity. The ideas in Foata's proof lead to the proof of a deep multi-variate identity [16]. Neither the present algorithm, nor the more general algorithms of [15], hold for general multi-variate identities with an arbitrary number of variables, and it is possible that there can never be such an algorithm.

On the other hand we should not be too chauvinistic and assume that a computer-generated proof can give *no* insight. The team human-computer is a mighty one, and an open-minded human can draw inspiration from all sources, even from a machine.

5. Postscript (written a few months later in the revised version)

When I wrote the last sentence I did not realize how soon my prophecy would come true. On December 24, 1988, around 11:00 pm, I received a phone call from Herb Wilf telling me how to bring insight into the proofs generated by my computer. Thus were born the notions of "rational function certification" and "WZ pairs" [14] (see also [6]).

Note added in proof. The full writeup of my algorithm, with many applications and examples, will appear in my paper "The method of creative telescoping", that will appear in the Journal for Symbolic Computation.

References

[1] G.E. Andrews, q-Series: their development and applications in analysis, number theory, combinatorics, physics, and computer algebra, CBMS series 66 (Amer. Math. Soc., Providence, 1986).

[2] R. Askey, Ramanujan and hypergeometric and basic hypergeometric series, preprint.

[3] W.N. Bailey, Generalized hypergeometric series, Cambridge Math. Tracts 32 (Cambridge University Press, London, 1935). (Repreinted: Hafner, New York, 1964.)

[4] I.N. Bernstein, Modules over a ring of differential operators, study of the fundamental solutions of equations with constant coefficients, Functional Analysis and Its Applications 5 (2) (1971), Russian original:1–16, English translation:89–101.

[5] J.E. Björk, Rings of Differential Operators (North-Holland, Amsterdam, 1979).

[6] B.A. Cipra, How the Grinch stole mathematics, Science 245 (1989) 595.

[7] P. Cartier and D. Foata, Problèmes combinatoires de commutation et réarrangements, Lecture Notes in Mathematics 85 (Springer, Berlin, 1969).

[8] G. Gasper and M. Rahman, An idefinite bibasic summation formula and some quadratic, cubic, and quartic summation and transformation formulas, preprint.

[9] I. Gessel and D. Stanton, Strange evaluations of hypergeometric series, SIAM J. Math. Anal. 13 (1982) 295–308.

[10] R.W. Gosper, Decision procedures for indefinite hypergometric summation, Proc. Natl. Acad. Sci. USA 75 (1978) 40–42.

[11] A. van der Poorten, A proof that Euler missed . . . Apery's proof of the irrationality of $\zeta(3)$, Math. Intell. 1 (1979) 195–203.

[12] E.D. Rainville, Special Functions (Macmillan, New York, 1960). (Reprinted: Chelsea, Bronx, NY, 1971).

[13] B.L. van der Waerden, Modern Algebra, Vol. I and II (Frederik Ungar Publ. Co., New York).

[14] H.S. Wilf and D. Zeilberger, Rational functions certify combinatorial identities, J. Amer. Math. Soc., to appear.

[15] D. Zeilberger, A holonomic systems approach to special function identities, preprint.

[16] D. Zeilberger and D. Bressoud, A proof of Andrew's q-Dyson conjecture, Discrete Math. 54 (1985) 201–224.