4.3 The 'Snake Oil' method for easier combinatorial identities

Combinatorial mathematics is full of dazzling identities. Legions of them involving binomial coefficients alone fill text- and reference books (see below for some references). It is a fine skill for a working discrete mathematician to have if he/she is able to evaluate or simplify complicated looking sums that involve combinatorial numbers, because they have a way of turning up in connection with problems in graphs, algorithms, enumeration, etc. (they're fun, too!).

In the past, one had to have built up a certain arsenal of special devices, the more the better, in order to be able to trot out the correct one for the correct occasion. Recently, however, a good deal of quite dramatic systematization has taken place, and there are unified methods for handling vast sub-legions of the legions referred to above.

In this section we are going to do two things. First we will give a single method (the *Snake Oil** method) that uses generating functions to deal with the evaluation of combinatorial sums. That one method is capable of handling a great variety of sums involving binomial coefficients, but there's nothing special about binomial coefficients in this respect. The method also works beautifully, within its limitations, on sums involving other combinatorial numbers. The philosophy is roughly this: don't try to evaluate the sum that you're looking at. Instead, find the generating function for the whole parameterized family of them, then read off the coefficients.

Second, we will confess that Snake Oil doesn't cure them all. Some combinatorial sums are really hard. Many of the very hardest binomial coefficient sums can now be proved by computers using the method of rational functions, which we will discuss next. Not only that, but use of the computer has resulted in some new proofs of classical identities. The hallmarks of these proofs are that (a) they are very short compared to the previously known proofs, (b) they seem extremely unmotivated to the reader, but (c) nothing is left out, and they really are proofs. The computerized proof techniques rely on a very simple-looking observation, which we will describe and illustrate.

Therefore, in this section you can expect to see one unified method that works on a lot of relatively easy sums, and one other unified method that works on many more kinds of binomial coefficient sums, including some fiendishly difficult ones.

First let's talk about the Snake Oil method.

The basic idea is what I might call the *external* approach to identities

---

* The Random House Dictionary of the English Language defines 'snake oil' as a purported cure for everything, and gives the example *The governor promised to lower taxes, but it was the same old snake oil.* The date of the expression is given as '1925-30, Amer.'

rather than the usual *internal* method.

To explain the difference between these two points of view, suppose we want to prove some identity that involves binomial coefficients. Typically such a thing would assert that some fairly intimidating-looking sum is in fact equal to such-and-such a simple function of n.

One approach that is now customary, thanks to the skillful exposition and deft handling by Knuth in [Kn], and by Graham, Knuth and Patashnik in [GKP], consists primarily of looking inside the summation sign ('internally'), and using binomial coefficient identities or other manipulations of indices *inside* the summations to bring the sum to manageable form.

The method that we are about to discuss is complementary to the internal approach. In the *external*, or generatingfunctionological, approach that we are selling here, one begins by giving a quick glance at the expression that is inside the summation sign, just long enough to spot the 'free variables,' i.e., what it is that the sum depends on after the dummy variables have been summed over. Suppose that such a free variable is called n.

Then instead of trying to grapple with the sum, just sweep it all under the rug, as follows:

The Snake Oil Method for Doing Combinatorial Sums

(a) Identify the free variable, say n, that the sum depends on. Give a name to the sum that you are working on; call it $f(n)$.
(b) Let $F(x)$ be the opsgf whose $[x^n]$ is $f(n)$, the sum that you'd love to evaluate.
(c) Multiply the sum by x^n, and sum on n. Your generating function is now expressed as a double sum over n, and over whatever variable was first used as a dummy summation variable.
(d) Interchange the order of the two summations that you are now looking at, and perform the inner one in simple closed form. For this purpose it will be helpful to have a catalogue of series whose sums are known, such as the list in section 2.5 of this book.
(e) Try to identify the coefficients of the generating function of the answer, because those coefficients are what you want to find.

If that seems complicated, just wait till you see the next seven examples. By then it will seem quite routine.

The success of the method depends on favorable outcomes of steps (d) and (e). What is surprising is the high success rate. It also has the 'advantage' of requiring hardly any thought at all; when it works, you know it, and when it doesn't, that's obvious too.

We will adhere strictly to the customary conventions about binomial coefficients and the ranges of summation variables. These are: first that the binomial coefficient $\binom{x}{m}$ vanishes if $m < 0$ or if x is a nonnegative integer that is smaller than m. Second, a summation variable whose range is not otherwise explicitly restricted is understood to be summed from $-\infty$ to ∞. Thus we have, for integer $n \geq 0$,

$$\sum_k \binom{n}{k} = 2^n,$$

in the sense that the sum ranges over all positive and negative and 0 values of k, the summand vanishes unless $0 \leq k \leq n$, and the sum has the value advertised. These conventions will save endless fussing over changing limits of summation when the dummy variables of summation get changed. For example, we find that

$$\sum_k \binom{n}{r+k} x^k = x^{-r} \sum_k \binom{n}{r+k} x^{r+k} = x^{-r} \sum_s \binom{n}{s} x^s = x^{-r}(1+x)^n,$$

for nonnegative integer n and integer r, without ever even thinking about the ranges of the summation variables.

The series evaluations that are most helpful in the examples that follow are, first and foremost,

$$\sum_{r\geq 0} \binom{r}{k} x^r = \frac{x^k}{(1-x)^{k+1}} \qquad (k \geq 0), \tag{4.3.1}$$

which is basically a rewrite of (2.5.7). Also useful are the binomial theorem

$$\sum_r \binom{n}{r} x^r = (1+x)^n \tag{4.3.2}$$

and (2.5.11), which we repeat here for easy reference:

$$\sum_n \frac{1}{n+1} \binom{2n}{n} x^n = \frac{1}{2x}(1 - \sqrt{1-4x}). \tag{4.3.3}$$

Example 1. Openers

Consider the sum

$$\sum_{k\geq 0} \binom{k}{n-k} \qquad (n = 0, 1, 2, \ldots).$$

The free variable is n, so let's call the sum $f(n)$. Write it out like this:

$$f(n) = \sum_{k\geq 0} \binom{k}{n-k}.$$

OK, now multiply both sides by x^n and sum over n. You have now arrived at step (c) of the general method, and you are looking at

$$F(x) = \sum_n x^n \sum_{k \geq 0} \binom{k}{n-k}.$$

Ready for step (d)? Interchange the sums, to get

$$F(x) = \sum_{k \geq 0} \sum_n \binom{k}{n-k} x^n.$$

We would like to 'do' the inner sum, the one over n. The trick is to get the exponent of x to be exactly the same as the index that appears in the binomial coefficient. In this example the exponent of x is n, and n is involved in the downstairs part of the binomial coefficient in the form $n-k$. To make those the same, the correct medicine is to multiply inside the sum by x^{-k} and outside the inner sum by x^k, to compensate. The result is

$$F(x) = \sum_{k \geq 0} x^k \sum_n \binom{k}{n-k} x^{n-k}.$$

Now the exponent of x is the same as what appears downstairs in the binomial coefficient. Hence take $r = n - k$ as the new dummy variable of summation in the inner sum. We find then

$$F(x) = \sum_{k \geq 0} x^k \sum_r \binom{k}{r} x^r.$$

We recognize the inner sum immediately, as $(1 + x)^k$. Hence

$$F(x) = \sum_{k \geq 0} x^k (1 + x)^k = \sum_{k \geq 0} (x + x^2)^k = \frac{1}{1 - x - x^2}.$$

The generating function on the right is an old friend; it generates the Fibonacci numbers (see Example 1.3 of chapter 1). Hence $f(n) = F_{n+1}$, and we have discovered that

$$\sum_{k \geq 0} \binom{k}{n-k} = F_{n+1} \qquad (n = 0, 1, 2, \ldots).$$

■

Example 2. Another one

Consider the sum

$$\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1} \qquad (m,n \geq 0). \tag{4.3.4}$$

Can it be that the same method will do this sum, without any further infusion of ingenuity? Indeed; just pour enough Snake Oil on it and it will be cured. Let $f(n)$ denote the sum in question, and let $F(x)$ be its opsgf. Dive in immediately by multiplying by x^n and summing over $n \geq 0$, to get

$$\begin{aligned}
F(x) &= \sum_{n\geq 0} x^n \sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1} \\
&= \sum_k \binom{2k}{k}\frac{(-1)^k}{k+1} x^{-k} \sum_{n\geq 0} \binom{n+k}{m+2k} x^{n+k} \\
&= \sum_k \binom{2k}{k}\frac{(-1)^k}{k+1} x^{-k} \sum_{r\geq k} \binom{r}{m+2k} x^{r} \\
&= \sum_k \binom{2k}{k}\frac{(-1)^k}{k+1} x^{-k} \frac{x^{m+2k}}{(1-x)^{m+2k+1}} \qquad \text{(by (4.3.1))} \\
&= \frac{x^m}{(1-x)^{m+1}} \sum_k \binom{2k}{k}\frac{1}{k+1}\left\{\frac{-x}{(1-x)^2}\right\}^k \\
&= \frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\sqrt{1+\frac{4x}{(1-x)^2}}\right\} \\
&= \frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\frac{1+x}{1-x}\right\} \\
&= \frac{x^m}{(1-x)^m}.
\end{aligned}$$

The original sum is now unmasked: it is the coefficient of x^n in the last member above. But that is $\binom{n-1}{m-1}$, by (4.3.1) again, and we have our answer. See exercise 16 for a generalization of this sum.

If the train of manipulations seemed long, consider that at least it's always the *same* train of manipulations, whenever the method is used, and also that with some effort a computer could be trained to do it! ∎

Example 3. A discovery

Is it possible to write the sum

$$f_n = \sum_{k\leq \frac{n}{2}} (-1)^k \binom{n-k}{k} y^{n-2k} \qquad (n \geq 0) \tag{4.3.5}$$

in a simpler closed form?

This example shows the whole machine at work again, along with a few new wrinkles. The first step is to let $F \overset{ops}{\longleftrightarrow} \{f_n\}$, and try to find the generating function F instead of the sequence $\{f_n\}$.

To do that we multiply (4.3.5) on both sides by x^n and sum over $n \geq 0$ to obtain

$$F(x) = \sum_{n\geq 0} x^n \sum_{k \leq \frac{n}{2}} (-1)^k \binom{n-k}{k} y^{n-2k}.$$

The next step is invariably to interchange the summations and hope. To try to make the innermost summation as clean looking as possible, be sure to take to the outer sum any factors that depend only on k. This yields

$$F(x) = \sum_k (-1)^k y^{-2k} \sum_{n\geq 2k} \binom{n-k}{k} x^n y^n.$$

Now focus on (4.3.1), and try to make the inner sum look like that. If in our inner sum the powers of x and y were $x^{n-k}y^{n-k}$, then those exponents would match exactly the upper story of the binomial coefficient $\binom{n-k}{k}$, and so after a change of dummy variable of summation we would be looking exactly at the left side of (4.3.1).

Hence we next multiply inside the inner sum by $x^{-k}y^{-k}$, and outside the inner sum by $x^k y^k$. Now we have

$$\begin{aligned}
F(x) &= \sum_k (-1)^k y^{-2k} x^k y^k \sum_{n\geq 2k} \binom{n-k}{k} x^{n-k} y^{n-k} \\
&= \sum_k (-1)^k x^k y^{-k} \sum_{a\geq k} \binom{a}{k} (xy)^a \\
&= \sum_{k\geq 0} (-1)^k x^k y^{-k} \frac{(xy)^k}{(1-xy)^{k+1}} \qquad \text{(by (4.3.1))} \\
&= \frac{1}{1-xy} \sum_{k\geq 0} \left\{ \frac{-x^2}{1-xy} \right\}^k \\
&= \frac{1}{1-xy} \frac{1}{1+\frac{x^2}{1-xy}} \\
&= \frac{1}{1-xy+x^2}.
\end{aligned} \tag{4.3.6}$$

(Question: Why, after the third equals sign above, did the range of k get restricted to '$k \geq 0$'?)

We now expand (4.3.6) in partial fractions to obtain a closed form for the sum (4.3.5). This gives

$$\begin{aligned} F(x) &= \frac{1}{(1-xx_+)(1-xx_-)} \\ &= \frac{x_+}{(x_+-x_-)(1-xx_+)} - \frac{x_-}{(x_+-x_-)(1-xx_-)}, \end{aligned}$$

where

$$x_\pm = \frac{y \pm \sqrt{y^2-4}}{2}.$$

Hence, for $n \ge 0$ the coefficient of x^n is

$$f_n = \frac{1}{\sqrt{y^2-4}} \left\{ \left(\frac{y+\sqrt{y^2-4}}{2} \right)^{n+1} - \left(\frac{y-\sqrt{y^2-4}}{2} \right)^{n+1} \right\}.$$

We now have our answer, but just for a demonstration of the effectiveness of cleanup operations, let's invest a little more time in making the answer look as neat as possible. Because of the ubiquitous appearance of $\sqrt{y^2-4}$ in the answer, we replace y formally by $x+(1/x)$. Then

$$\sqrt{y^2-4} = x - \frac{1}{x},$$

and our formula becomes

$$\sum_{k \le \frac{n}{2}} (-1)^k \binom{n-k}{k} (x^2+1)^{n-2k} x^{2k} = \frac{x^{2n+2}-1}{x^2-1} \qquad (n \ge 0).$$

Finally we write $t = x^2$ to obtain the pretty evaluation

$$\sum_{k \le \frac{n}{2}} (-1)^k \binom{n-k}{k} (t+1)^{n-2k} t^k = \frac{1-t^{n+1}}{1-t} \qquad (n \ge 0). \tag{4.3.7}$$

For instance, the value $t = 1$ gives

$$\sum_{k \le \frac{n}{2}} (-1)^k \binom{n-k}{k} 2^{n-2k} = n+1 \qquad (n \ge 0). \tag{4.3.8}$$

As a final touch, we can read off the coefficient of t^m in (4.3.7) to discover the interesting fact that

$$\sum_{k \le \frac{n}{2}} (-1)^k \binom{n-k}{k} \binom{n-2k}{m-k} = \begin{cases} 1, & \text{if } 0 \le m \le n; \\ 0, & \text{otherwise.} \end{cases} \tag{4.3.9}$$

Try this identity with $n = 2$ and watch what happens. ■

Here is another example of the same technique.

Example 4.

Evaluate the sums

$$f_n = \sum_k \binom{n+k}{2k} 2^{n-k} \qquad (n \geq 0). \tag{4.3.10}$$

Without stopping to think, let F be the opsgf of the sequence, multiply both sides of (4.3.10) by x^n, sum over $n \geq 0$, and interchange the two sums on the right. This produces

$$\begin{aligned} F &= \sum_k 2^{-k} \sum_{n \geq 0} \binom{n+k}{2k} 2^n x^n \\ &= \sum_k 2^{-k} (2x)^{-k} \sum_{n \geq 0} \binom{n+k}{2k} (2x)^{n+k} \\ &= \sum_{k \geq 0} 2^{-k} (2x)^{-k} \frac{(2x)^{2k}}{(1-2x)^{2k+1}} \qquad \text{(by (4.3.1))} \\ &= \frac{1}{1-2x} \sum_{k \geq 0} \left\{ \frac{x}{(1-2x)^2} \right\}^k \\ &= \frac{1}{1-2x} \frac{1}{1 - \frac{x}{(1-2x)^2}} \\ &= \frac{1-2x}{(1-4x)(1-x)} \\ &= \frac{2}{3(1-4x)} + \frac{1}{3(1-x)}. \end{aligned}$$

It is now a triviality to read off the coefficient of x^n on both sides and discover the answer:

$$\sum_k \binom{n+k}{2k} 2^{n-k} = \frac{2^{2n+1}+1}{3} \qquad (n \geq 0). \tag{4.3.11}$$

■

Example 5.

Our next example will be of a sum that we won't succeed in evaluating in a neat, closed form. However, the generating function that we obtain will be rather tidy, and that is about the most that can be expected from this family of sums.

The sum is

$$f_n(y) = \sum_k \binom{n}{k}\binom{2k}{k} y^k \qquad (n \geq 0). \tag{4.3.12}$$

Follow the usual prescription. Define $F(x,y) = \sum_{n\geq 0} f_n(y)x^n$. To find F, multiply (4.3.12) by x^n, sum over $n \geq 0$ and interchange the inner and outer sums, to obtain

$$\begin{aligned} F(x,y) &= \sum_k \binom{2k}{k} y^k \sum_{n\geq 0}\binom{n}{k} x^n \\ &= \sum_k \binom{2k}{k} y^k \frac{x^k}{(1-x)^{k+1}} \\ &= \frac{1}{1-x}\sum_k \binom{2k}{k}\left(\frac{xy}{1-x}\right)^k. \end{aligned} \tag{4.3.13}$$

Now since

$$\sum_k \binom{2k}{k} z^k = \frac{1}{\sqrt{1-4z}}, \tag{4.3.14}$$

by (2.5.11), we obtain

$$\begin{aligned} F(x,y) &= \frac{1}{(1-x)\sqrt{1-\frac{4xy}{1-x}}} \\ &= \frac{1}{\sqrt{(1-x)(1-x(1+4y))}}. \end{aligned} \tag{4.3.15}$$

For general values of y, that's about all we can expect. There are two special values of y for which we can go further. If $y = -1/4$, we find that

$$\sum_k \binom{2k}{k}\binom{n}{k}(-\frac{1}{4})^k = 2^{-2n}\binom{2n}{n} \qquad (n \geq 0). \tag{4.3.16}$$

If $y = -1/2$, then

$$\begin{aligned} F(x,-1/2) &= 1/\sqrt{1-x^2} \\ &= \sum_m \binom{2m}{m}(x/2)^{2m} \qquad \text{(by (2.5.11))}. \end{aligned}$$

Hence we have Reed Dawson's identity

$$\sum_k \binom{2k}{k}\binom{n}{k}(-1)^k 2^{-k} = \begin{cases} \binom{n}{n/2}2^{-n} & \text{if } n \geq 0 \text{ is even,} \\ 0 & \text{if } n \geq 0 \text{ is odd,} \end{cases} \tag{4.3.17}$$

and Snake Oil triumphs again. ∎

Example 6.

Suppose we have two complicated sums and we want to show that they're the same. Then the generating function method, if it works, should be very easy to carry out. Indeed, one might just find the generating functions of each of the two sums independently and observe that they are the same.

Suppose we want to prove that

$$\sum_k \binom{m}{k}\binom{n+k}{m} = \sum_k \binom{m}{k}\binom{n}{k}2^k \qquad (m, n \geq 0)$$

without evaluating either of the two sums.

Multiply on the left by x^n, sum on $n \geq 0$ and interchange the summations, to arrive at

$$\begin{aligned}\sum_k \binom{m}{k} x^{-k} \sum_{n\geq 0} \binom{n+k}{m} x^{n+k} &= \sum_k \binom{m}{k} x^{-k} \frac{x^m}{(1-x)^{m+1}} \\ &= \frac{x^m}{(1-x)^{m+1}}\left(1+\frac{1}{x}\right)^m \\ &= \frac{(1+x)^m}{(1-x)^{m+1}}.\end{aligned}$$

If we multiply on the right by x^n, etc., we find

$$\begin{aligned}\sum_k \binom{m}{k} 2^k \sum_{n\geq 0} \binom{n}{k} x^n &= \frac{1}{(1-x)} \sum_k \binom{m}{k} \left(\frac{2x}{(1-x)}\right)^k \\ &= \frac{1}{(1-x)}\left(1+\frac{2x}{1-x}\right)^m \\ &= \frac{(1+x)^m}{(1-x)^{m+1}}.\end{aligned}$$

Hence the two sums are equal, even if we don't know what they are! ∎

Example 7.

There are, in combinatorics, a number of *inversion formulas*, and generating functions give an easy way to prove many of those. An inversion formula in general is a relationship that expresses one sequence in terms of another, along with the inverse relation, which recovers the original sequence from the constructed one.

We have already seen a couple of famous examples of these. One is the Möbius inversion formula, which is the pair (2.6.11), (2.6.12). Another is the pair (4.2.3), (4.2.7) that occurred in the sieve method. We repeat that pair here, for ready reference. It states that if we compute a sequence $\{N_r\}$ from a sequence $\{e_r\}$ by the relations

$$N_r = \sum_{t \geq 0} \binom{t}{r} e_t \qquad (r = 0, 1, 2, \ldots), \tag{4.3.18}$$

then we can recover the original sequence ('invert') by means of

$$e_t = \sum_j (-1)^{j-t} \binom{j}{t} N_t \qquad (t \geq 0).$$

To give just one more example of such a pair of formulas, consider the relation

$$a_r = \sum_s \binom{r}{s} b_s \qquad (r \geq 0), \tag{4.3.19}$$

which differs from the previous pair in that the summation is over the lower index in the binomial coefficient. How can we find the relations that are inverse to (4.3.19)? That is, how can we solve for the b's in terms of the a's?

The answer is that we convert the relation (4.3.18) between two sequences into a relation between their exponential generating functions, which we then invert. By (2.3.3) we have $A(x) = e^x B(x)$, where A and B are the egf's. Hence $B(x) = e^{-x} A(x)$, and therefore

$$b_n = \sum_m \binom{n}{m} (-1)^{n-m} a_m \qquad (n \geq 0). \tag{4.3.20}$$

■

An inversion formula of a somewhat deeper kind appears in (5.1.5), (5.1.6).

Example 8. Snake Oil vs. hypergeometric functions.

Many combinatorial identities are special cases of identities in the theory of hypergeometric series (we'll explain that remark, briefly, in a moment). However, the Snake Oil method can cheerfully deal with all sorts of identities that are not basically about hypergeometric functions. So the approaches are complementary.

A hypergeometric series is a series

$$\sum_k T_k$$

in which the *ratio* of every two consecutive terms is a rational function of the summation variable k. That means that

$$\frac{T_{k+1}}{T_k} = \frac{P(k)}{Q(k)},$$

where P and Q are polynomials, and it takes in a lot of territory. Many binomial coefficient identities, including all of the examples in this chapter so far, are of this type. There are some general tools for dealing with such sums, and these are very important considering how frequently they occur in practice. For a discussion of some of these tools, see, for example, the article by Roy [Ro].

In this example we want to emphasize that the scope of the Snake Oil method includes a lot of sums that are not hypergeometric. Consider, for instance, the following sum-

$$f(n) = \sum_k \begin{bmatrix} n \\ k \end{bmatrix} B_k,$$

where the $[\ ]$'s are the Stirling numbers of the first kind, and the B's are the Bernoulli numbers.

Now one thing, at least, is clear from looking at this sum: it is not hypergeometric. The ratio of two consecutive terms is certainly not a rational function of k. The Snake Oil method is, however, unfazed by this turn of events. If you follow the method exactly as before, you could define $F(x)$ to be the egf of the sequence $\{f(n)\}$, multiply by $x^n/n!$, sum on n, interchange the indices, etc., and obtain

$$\begin{aligned}
F(x) &= \sum_n \frac{f(n)x^n}{n!} \\
&= \sum_n \frac{x^n}{n!} \sum_k \begin{bmatrix} n \\ k \end{bmatrix} B_k \\
&= \sum_k B_k \sum_n \begin{bmatrix} n \\ k \end{bmatrix} \frac{x^n}{n!} \\
&= \sum_k B_k \left\{ \frac{1}{k!} \left(\log \frac{1}{1-x} \right)^k \right\} \qquad \text{(by (3.5.3))} \\
&= \sum_k \frac{B_k}{k!} u^k \qquad (u = \log \frac{1}{1-x}) \\
&= \frac{u}{e^u - 1} \qquad \text{(by (2.5.8))} \\
&= \frac{1-x}{x} \log \frac{1}{1-x}.
\end{aligned}$$

If we now read off the coefficient of $x^n/n!$ on both sides, we find that the unknown sum is

$$\sum_k \begin{bmatrix} n \\ k \end{bmatrix} B_k = -\frac{(n-1)!}{n+1} \qquad (n \geq 1). \tag{4.3.21}$$

Example 9. The scope of the Snake Oil method

The success of the Snake Oil method depends upon being given a sum to evaluate in which there is a free variable that appears in only one place. Then, after interchanging the order of the summations, one finds one of the basic power series (4.3.1) or (4.3.2) to sum.

At the risk of diminishing the charm of the method somewhat by adding gimmicks to it, one must remark that in many important cases this limitation on the scope is easy to overcome. This is because it frequently happens that when an identity is presented that has a free variable repeated several times, that identity turns out to be a special case of a more general identity in which each of the repeated appearances of the free variable is replaced by a *different* free variable. Before abandoning the method on some given problem, this possibility should be explored.

Consider the identity

$$\sum_i \binom{n}{i}\binom{2n}{n-i} = \binom{3n}{n}.$$

At first glance the possibilities for successful Snake Oil therapy seem dim because of the multiple appearances of n in the summand. However, if we generalize the identity by splitting the appearances of n into different free variables, we might be led to consider the sum

$$\sum_i \binom{n}{i}\binom{m}{r-i},$$

which is readily evaluated by the Snake Oil method. It is characteristic of the subject of identities that it is usually harder to prove special cases than general theorems. Multiple appearances of a free variable are often a hint that one should try to find a suitable generalization.

4.4 WZ pairs prove harder identities

Computers can now *find proofs of* combinatorial identities, including most of the identities that we did by the Snake Oil method in the previous section, as well as many, many more. In this section we will say how that is done. Although *finding* proofs this way requires more work than a human would care to do, the result, after the computer is finished, is a neat and compact proof that a human can often easily check, and can always check