Massachusetts Institute of Technology | Handout 45
6.042J/18.062J: Mathematics for Computer Science | November 7, 1995
Professors David Karger and Tom Leighton

Notes 17

1 Manipulating OGF's

Let $G(x) = \sum_n g_n x^n$. Then

$$x^k G(x) = \sum_{n=k}^{\infty} g_{n-k} x^n,$$

for all $k \geq 0$. In other words, multiplying the OGF by x^k prefixes the sequence with k zeroes. Similarly,

$$(1/x^k)(G(x) - g_0 - g_1 x - \cdots - g_{k-1} x^{k-1}) = \sum_{n=k}^{\infty} g_{n+k} x^n,$$

for all $k \geq 0$. The sequence is left-shifted k times.

Some other useful identities are:

$$\begin{aligned}
\alpha F(x) + \beta G(x) &= \sum_n (\alpha f_n + \beta g_n) x^n \\
G(cx) &= \sum_n c^n g_n x^n \\
1/2 \cdot (G(x) + G(-x)) &= \sum_n g_{2n} x^{2n} \qquad (1) \\
1/2 \cdot (G(x) - G(-x)) &= \sum_n g_{2n+1} x^{2n+1} \\
G'(x) &= \sum_n (n+1) x^n \\
1/x \int_0^x G(t)dt &= \sum_n \frac{g_n}{n+1} x^n \\
F(x)G(x) &= \sum_n (\sum_{k=0}^{n} f_k g_{n-k}) x^n \qquad (2) \\
G(x)/(1-x) &= \sum_n (\sum_{k=0}^{n} g_{n-k}) x^n \qquad (3)
\end{aligned}$$

Equation 1 gives the sequence $\langle g_0, 0, g_2, 0, g_4, \ldots\rangle$. If we want to get rid of those zeros, consider the function

$$H(y) = \frac{G(\sqrt{y}) + G(-\sqrt{y})}{2}$$

Equation 2 describes the *convolution* of sequences $\{f_i\}$ and $\{g_i\}$. To choose a size-n combination from $A \cup B$, you have to, for some k, pick a size k combination from A and a size $n-k$ combination from B. Equation 3 is a special case of the previous one with $F(x) = 1/(1-x)$ as the generating function for $\langle 1, 1, 1, 1, \ldots\rangle$.

Example 1 The Fibonacci numbers have the generating function

$$F(z) = \frac{z}{1 - z - z^2}$$

therefore the *even* Fibonacci numbers with odds zeroed out have OGF

$$\frac{1}{2}\left(\frac{z}{1 - z - z^2} + \frac{-z}{1 + z - z^2}\right) = \frac{z^2}{1 - 3z^2 + z^4}$$

To eliminate all the unnecessary 0's just use

$$\frac{z}{1 - 3z - z^2}$$

Example 2 What is $\sum_{i=1}^{n} k^2$? First find the OGF of $< 0^2, 1^2, 2^2, \ldots >$

- $\sum_n x^n = \frac{1}{1-x}$
- Taking derivatives and multiplying by x: $\sum_n nx^n = \frac{x}{(1-x)^2}$
- Doing the same again: $\sum_n n^2 x^n = \frac{x+x^2}{(1-x)^3}$

Thus, by Equation 3 we get that the OGF of $< 0^2, 0^2 + 1^2, \ldots >$ is $F(x) = \frac{x+x^2}{(1-x)^4}$. All we need to do is determine the coefficient of x^n in $F(x)$. We know that

$$(1-x)^{-k} = \sum_n \binom{k+n-1}{n} x^n.$$

Applying this with $k = 4$ we get that

$$(1-x)^{-4} = \sum_n \binom{n+3}{n} x^n = \sum_n \frac{(n+1)(n+2)(n+3)}{6} x^n.$$

This means that the coefficient of x^n in $F(x)$ is $n(n+1)(n+2)/6 + (n-1)n(n+1)/6 = n(n+1)(2n+1)/6$.

Example 3 How many ways can I tile a $2 \times n$ rectangle with 2×1 dominoes? Well, $t_n = t_{n-1} + t_{n-2}; t_0 = 1, t_1 = 1$, so the answer is Fibonacci numbers. What about a $3 \times n$ rectangle? Well, $t_0 = 1$ and $t_1 = 0$. But how can we build a recurrence? Suppose we let v_n be the number of ways to tile a $3 \times n$ rectangle with a missing corner square. Then $v_0 = 0$ and $v_1 = 1$. Furthermore,

$$\begin{aligned} t_n &= 2v_{n-1} + t_{n-2} \\ v_n &= t_{n-1} + v_{n-2} \end{aligned}$$

Therefore,

$$\begin{aligned} T(z) &= 2zV(z) + z^2T(z) + 1 \\ V(z) &= zT(z) + z^2V(z) \end{aligned}$$

The second equation implies $V(z) = zT(z)/(1 - z^2)$. Plugging into first equation,

$$\begin{aligned} T(z) &= 2z^2T(z)/(1 - z^2) + z^2T(z) + 1 \\ T(z) &= \frac{1 - z^2}{1 - 4z^2 + z^4} \end{aligned}$$

Expanding $(1 - 4z^2 + z^4)^{-1}$ as a function of z^2 yields the n^{th} coefficient

$$w_n = \frac{3 + 2\sqrt{3}}{6}(2 + \sqrt{3})^n + \frac{3 - 2\sqrt{3}}{6}(2 - \sqrt{3})^n.$$

Then just note that $t_{2n} = w_n - w_{n-1}$.

2 Multiple Convolutions

How many ways can we divide a set of n people into groups, each with a leader? We first need to pick a number m of groups. Then, for $i = 1, \ldots, m$, we need to pick a number k_i of people to go in the i^{th} group. Of course, we need $k_1 + k_2 + \cdot + k_m = n$. Then, in group i, we have k_i choices of a leader. Therefore, if s_n is the number of ways to do the choices,

$$s_n = \sum_{m>0} \sum_{k_1+k_2+\cdots+k_m=n} k_1k_2\cdots k_m$$

For fixed m, the inner sum is the coefficient of x^n in the convolution of $< 0, 1, 2, 3, \ldots >$ with itself m times. This makes sense, since $\langle 0, 1, 2, \cdots \rangle$ is the sequence of "ways to pick a leader from one group." Thus, the inner sum has generating function $G(z)^m$, where $G(z)$ is the OGF for $\langle 0, 1, 2 \ldots \rangle$, namely $z/(1-z)^2$. Thus, the OGF for s_n is

$$S(z) = G(z) + G(z)^2 + G(z)^3 + \cdots = \frac{G(z)}{1 - G(z)}.$$

Thus

$$S(z) = \frac{z}{(1-z)^2 - z} = \frac{z}{1 - 3z + z^2}$$

This is the OGF for the even Fibonacci numbers in the first example. So that's our answer: $s_n = f_{2n}$, the $2n$th Fibonacci number.

3 Binary Trees

We want to compute the number b_n of binary trees on n nodes. (In a binary tree every node has one or two children). Now it is clear that $b_0 = 1$, $b_1 = 1$, $b_2 = 2$, $b_3 = 5$. It is not hard to write up a recurrence for the b_n. Indeed, if the left son of the root is the root of a binary tree of k nodes, then the right son of the root must be the root of a binary tree of $n - k - 1$ nodes. This yields

$$b_n = \sum_{k=0}^{n-1} b_k b_{n-k-1}$$

if $n > 0$ and $b_0 = 1$. This is a messy, nonlinear recurrence which would be hard to solve directly. Therefore, we resort to the OGFs. Let $B(x)$ be the generating function of the sequence $\langle b_0, b_1, \ldots \rangle$. Then

$$\sum_{n=1}^{\infty} b_n x^n = \sum_{n=1}^{\infty} (\sum_{k=0}^{n-1} b_k b_{n-k-1}) x^n$$

that is (replacing n by $n+1$),

$$B(x) - b_0 = \sum_n (\sum_{k=0}^{n} b_k b_{n-k}) x^{n+1} = x \sum_n (\sum_{k=0}^{n} b_k b_{n-k}) x^n,$$

which implies by the convolution rule that the right hand side is $xB(x)B(x)$. Therefore,

$$x(B(x))^2 - B(x) + 1 = 0$$

Solving this equation for $B(x)$ we get that

$$B(x) = \frac{1 \pm \sqrt{1 - 4x}}{2x}$$

Which of the two possibilities is right? Well, if we rewrite

$$2xB(x) = 1 \pm \sqrt{1-4x}$$

and plug in $x = 0$, we see that $0 = 1 \pm \sqrt{1-4x}$ is satisfied only when we take the negative sign. Thus,

$$B(x) = \frac{1-\sqrt{1-4x}}{2x}.$$

To get the coefficients b_n, we need to expand this quantity in a Taylor series. We can do this using the binomial theorem, which applies to non-integer exponents. Some might object that the binomial theorem only works this way when we are inside the radius of convergence of the series, but remember that generating functions are only *formal* power series; we don't care if they converge or not. Anyway, the binomial theorem says that

$$(1-4x)^{1/2} = \sum \binom{1/2}{k}(-4x)^k$$

where $\binom{1/2}{0} = 1$ and, for $k > 0$,

$$\binom{1/2}{k} = \frac{(1/2)(-1/2)(-3/2)\cdots(3/2-k)}{k!}$$

Now b_n is the coefficient of x^n in $B(x)$ (note that the coefficient of x^{-1} is 0, so this is a well-defined OGF). Looking at $B(x)$, we see that this is $(-1/2)$ times the coefficient of x^{n+1} in the expansion of $(1-4x)^{1/2}$. In other words,

$$\begin{aligned}
b_n &= -\frac{1}{2}\binom{1/2}{n+1}(-4)^{n+1} \\
&= \frac{1}{-2}\cdot\frac{(1/2)(-1/2)(-3/2)\cdots(1/2-n)}{(n+1)!}(-4)^{n+1} \\
&= \frac{(-1/2)(-3/2)\cdots(1/2-n)}{(n+1)!}(-4)^n \\
&= \frac{(1)(3)(5)\cdots(2n-1)}{(n+1)!}2^n \\
&= \frac{1\cdot 3\cdot 5\cdots(2n-1)}{(n+1)!}\cdot\frac{2\cdot 4\cdot 6\cdots 2n}{2^n n!}\cdot 2^n \\
&= \frac{1\cdot 2\cdot 3\cdots 2n}{n!(n+1)!} \\
&= \frac{1}{n+1}\binom{2n}{n}
\end{aligned}$$

The number b_n is called the n^{th} *Catalan number.* You also encountered it in recitation, and showed how it could be determined by a reflection principle. But generating functions let you find the answer just by working through the rules, without having to have the insight about a slick way to solve the problem.

4 Exponential Generating Functions

Another way to represent the sequence $\{a_r\}$ is to use the *exponential generating function (EGF)*

$$\sum a_r \frac{x^r}{r!}$$

We will see that EGFs are useful for enumerating *permutations* while OGF's enumerate *combinations.*

Example 4 • Let $P(n,r)$ be the number of ways to pick an ordered sequence of r items from n. The EGF for this sequence is

$$\sum P(n,r)\frac{x^r}{r!} = (1+x)^n$$

(The OGF for this sequence is ugly).

- EGF for $\langle 1,1,1,\ldots\rangle$ is

$$\sum \frac{x^r}{r!} = e^x$$

- The EGF for an r-permutation of n objects with arbitrary repetitions is

$$\sum \frac{n^r}{r!}x^r = e^{nx}$$

The operations on OGF have analogues for EGFs. If $G(x)$ is the EGF for $\{g_n\}$, then

- $\{ng_{n-1}\}$ has EGF $xG(x)$
- $< g_1, g_2, g_3, \ldots >$ has EGF $G'(x)$
- Addition and subtraction work as for OGF's

The most interesting operation is *convolution.* Given EGF's $F(x), G(x)$ for $\{f_n\}$ and $\{g_n\}$, the function $H(x) = F(x)G(x)$ is the EGF for

$$h_n = \sum_k \binom{n}{k} f_k g_{n-k}.$$

Because
$$\frac{h_n}{n!} = \sum_k \frac{f_k}{k!} \frac{g_{n-k}}{(n-k)!}$$
Combinatorially speaking, if f_r describes the number of r-permutations obeying certain rules from a set A, and g_r describes the number of r-permutations from a disjoint set B, then h_r describes the number of r-permutations of A and B that "interleave" a valid permutation from A with a valid permutation from B. (Note that in the convolution of *OGFs* the analogous relation was
$$h_n = \sum_k f_k g_{n-k}.$$
This is because OGFs deal with combinations and so there is no "interleaving" involved, we just have to pick a valid combination from A and one from B.

Example 5 How many ways to permute 4 letters from PEPPER?

- Ways to permute P's: sequence $< 1,1,1,1 >$ has EGF
$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$$
This is because for each collection of k Ps, $k = 0, 1, 2, 3$, there is exactly one permutation of k Ps.
- Ways to permute E's: sequence $< 1,1,1 >$ has EGF
$$1 + x + \frac{x^2}{2!}!$$
- Ways to permute R has EGF $1 + x$

Thus, EGF for PEPPER is
$$(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!})(1 + x + \frac{x^2}{2!})(1 + x)$$
which expands to
$$1 + 3x + \frac{8}{2!}x^2 + \frac{19}{3!}x^3 + \frac{38}{4!}x^4 + \frac{60}{5!}x^5 + \frac{60}{6!}x^6$$
Hence the number of 4-permutations from PEPPER is 38.

What if we require an even number of P's and at least one E?
$$(1 + \frac{x^2}{2!})(x + \frac{x^2}{2!})(1 + x) = x + \frac{3}{2!}x^2 + \frac{18}{3!}x^3 + \frac{18}{4!}x^4 + \frac{30}{5!}x^5$$
Hence the number of 4-permutations from PEPPER containing an even number of P's and at least one E is 18.