CDM

Generating Functions

Klaus Sutner

Carnegie Mellon University

30-gener-func 2017/12/15 23:17

1. Algebra and Infinity

- Generating Functions
- Recurrences
- Regular Languages and Generating Functions

Finite state machines are a great example of a finite, algorithmically attractive representation of an infinite object (the acceptance language or, in case of infinite words, even just a single accepted word).

Register machines and Turing machines also describe infinite objects, but they are much less amenable to manipulation by algorithms.

Are there any other compelling examples along the lines of finite state machines? How about something more algebraic?

Suppose we have an unlimited supply of coins of denominations $1 \leq d_1 < d_2 < \ldots < d_n$.

How many ways are there to make change for $x \geq 0$?

Call this number $C(x)$.

In other words, we want to count the number of solutions of the equation

$$x = \sum_{i=1}^{n} x_i d_i \qquad x_i \in \mathbb{N}$$

The standard approach to computing $C(x)$ is dynamic programming.

More precisely, in the bottom-up approach we compute a $n \times (x+1)$ array $C(k,z)$ where $1 \leq k \leq n$ and $0 \leq z \leq x$ with the intent that

$$C(k,z) = \#\text{ of ways to make change for } z \text{ using } d_1, \ldots, d_k$$

Then

$$C(k,z) = C(k-1,z) + C(k, z-d_k)$$

where $C(k,0) = 1$ and $C(k,y) = 0$ whenever $y < 0$ or $k = 0$.

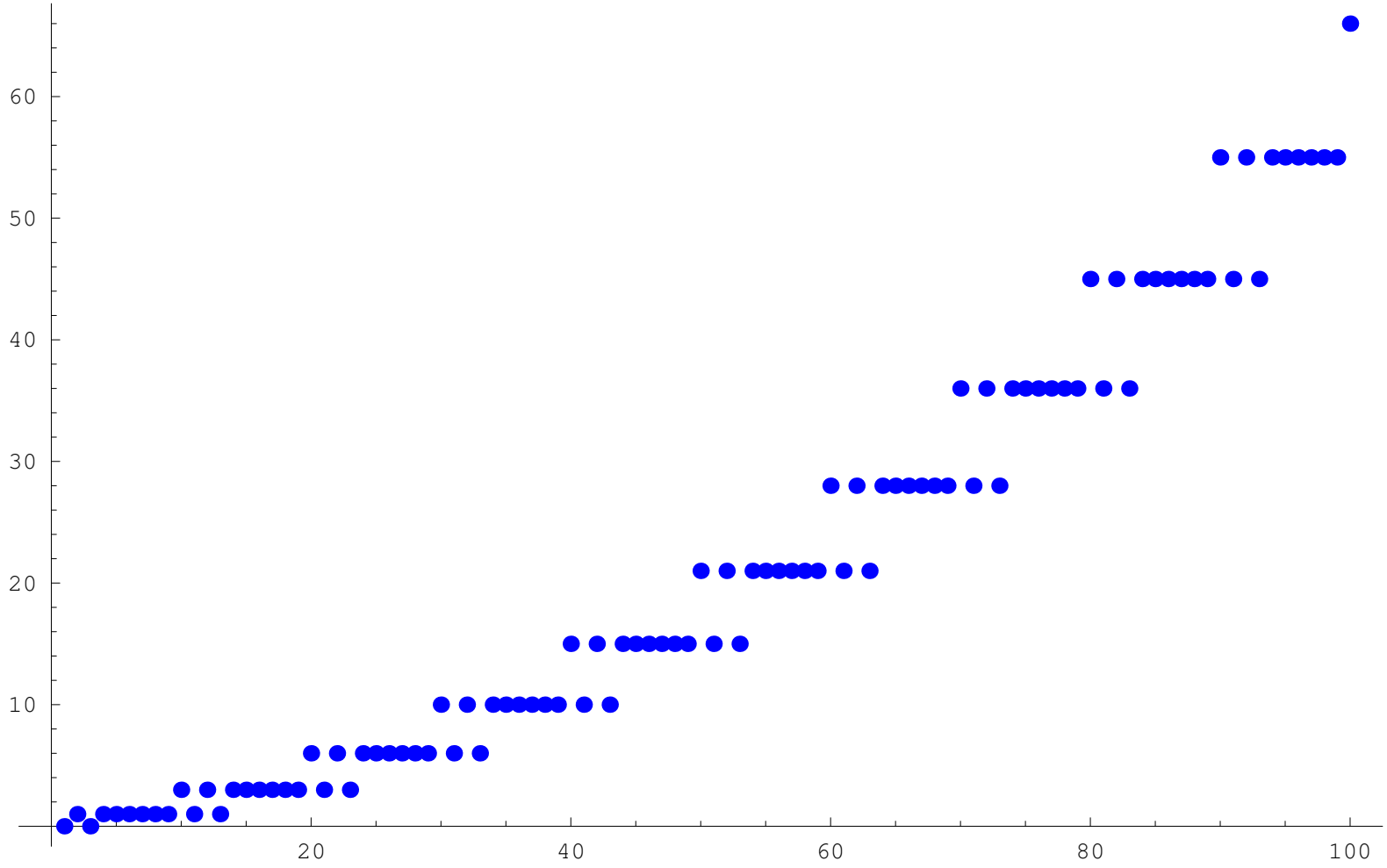
So we can get $C(z)$, $z \leq x$, in $\Theta(nx)$ steps.

The table for $\mathbf{d} = (2, 5, 10)$ and $x = 18$.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	1	1	1	1	1	2	1	2	1	2	2	2	2	2
1	0	1	0	1	1	1	1	1	1	3	1	3	1	3	3	3	3	3

Computing a few more values produces the following plot. Ponder deeply.

$C(x)$ for $x \leq 100$, $\mathbf{d} = (2, 5, 10)$.



Another way to get numerical answers is to multiply together polynomials

$$p_d(z) = \sum_{i \leq \lfloor x/d \rfloor} z^{i\,d}$$

for $d = d_1, \ldots, d_n$.

Claim

The coefficient of z^s in $\prod p_{d_i}(z)$ is $C(s)$, for all $s \leq x$.

This follows immediately from the definition of polynomial multiplication.

We write

$$[z^i]\, p(z)$$

for the coefficient of term z^i in the polynomial $p(z)$.

So

$$C(s) = [z^s]\, p_{d_1}(z) \ldots p_{d_n}(z)$$

for all $s \leq x$.

The bound $s \leq x$ is annoying. It would be nice to be able to deal with the whole infinite sequence $(C(s))_{s \geq 0}$.

We would like to set $x = \infty$ in the definition of $p_d(z)$.

That's actually OK, except that we are now dealing with a power series instead of a polynomial.

$$P_d(z) = 1 + z^d + z^{2d} + \ldots = \sum z^{id}$$

Then

$$C(s) = [z^s]\, P_{d_1}(z) \ldots P_{d_n}(z)$$

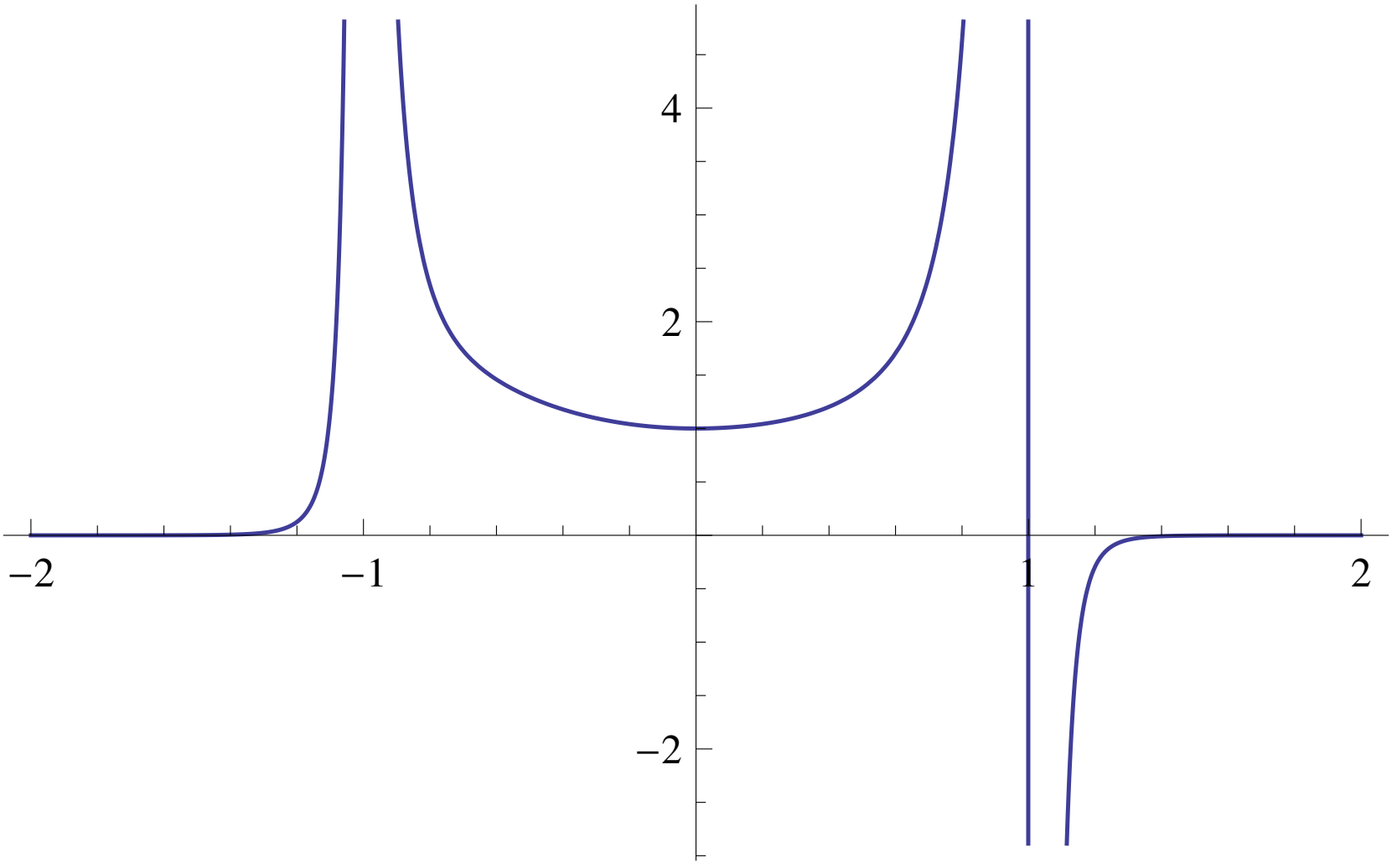
for all $s \geq 0$.

So far, we have not really gained much: to compute the coefficients of the product power series we have to multiply out polynomial of sufficiently high degrees, just as before.

But note that

$$P_d(z) = \frac{1}{1-z^d} = \frac{1}{1-z^d}$$

Instead of using infinitary power series we can use perfectly finitary rational functions!


4
2
−2
−2
−1
1
2

There is price we have to pay, though: given a rational function, it becomes a bit more difficult to extract the coefficients of the corresponding power series.

For example, expansion about $z = 0$ produces

$$f(z) = f(0) + f'(0)z + f''(0)/2\, z^2 + f^{(k)}(0)/k!\, z^k + \dots$$

Evaluating the derivatives requires quite a bit of work.

Getting a nice closed form can be very hard.

The rational function looks like $\frac{1}{(1-z^2)(1-z^5)(1-z^{10})}$ and $C(s)$ has the form $\gamma/(2000(-7+3\sqrt{5}))$ where γ is this:

$$-40(-1)^{9s/5}(7-4\sqrt[5]{-1}-4(-1)^{2/5}+7(-1)^{3/5})+40(-1)^{s/5}(7-7(-1)^{2/5}+4(-1)^{3/5}+4(-1)^{4/5})-$$
$$40(-1)^{7s/5}(-7+11\sqrt[5]{-1}-11(-1)^{3/5}+7(-1)^{4/5})+40(-1)^{3s/5}(7+7\sqrt[5]{-1}-11(-1)^{2/5}+11(-1)^{4/5})-25$$
$$(-1)^s(-7+3\sqrt{5})(17+2s)+8(-1)^{4s/5}(341-284\sqrt[5]{-1}+148(-1)^{2/5}-103(-1)^{3/5}+224(-1)^{4/5}+2(7-4\sqrt[5]{-1}$$
$$4(-1)^{2/5}+7(-1)^{3/5})s)+8(-1)^{6s/5}(-99+216\sqrt[5]{-1}-337(-1)^{2/5}+292(-1)^{3/5}-156(-1)^{4/5}+2(7-7(-1)^2$$
$$4(-1)^{3/5}+4(-1)^{4/5})s)-8(-1)^{2s/5}(395-363\sqrt[5]{-1}+572(-1)^{2/5}-737(-1)^{3/5}+633(-1)^{4/5}+2(-7+11\sqrt[5]{-}$$
$$11(-1)^{3/5}+7(-1)^{4/5})s)+8(-1)^{8s/5}(685-447\sqrt[5]{-1}+343(-1)^{2/5}-508(-1)^{3/5}+717(-1)^{4/5}+2(7+7\sqrt[5]{-1}$$
$$11(-1)^{2/5}+11(-1)^{4/5})s)-5(-7+3\sqrt{5})(123+2s(17+s))$$

It is not even clear that this simplifies to a integer value (it does, for all $s \geq 0$). The best way to check is by machine.

- Algebra and Infinity

2. Generating Functions

- Recurrences

- Regular Languages and Generating Functions

Suppose we have a sequence $\mathbf{a} = (a_i)_{i \geq 0}$ of numbers (naturals, integers, reals, whatever).

Definition

The generating function of $\mathbf{a}$ is the power series

$$A(z) = \sum_{i \geq 0} a_i z^i = a_0 + a_1 z + a_2 z^2 + \ldots + a_n z^n + \ldots$$

Note that z is just a variable and could be replaced by any other variable.

At first glance, $A(z)$ may look slightly more complicated than $\mathbf{a}$, so why bother?

We will mostly ignore issues of convergence here, we are not really interested in evaluating the series $A(z)$.

Still, note that for "reasonable" a_i and sufficiently small z we will actually get a function from, say, the reals to the reals.

Lemma

Let $r > 0$ such that $|a_n| \leq r^n$ for all $n > 0$. Then $A(z)$ converges for all z such that $|z| < 1/r$.

One can use the machinery of analysis to determine properties of this function and, hopefully, get a better understanding of $\mathbf{a}$.

In particular, one can use Taylor expansion to retrieve the coefficients given some suitable representation of the series, more later.

The real goal is to the algebra of power series–rather than manipulating the sequence directly, we manipulate the series. For example,

$$aF(z) + bG(z) = \sum (af_i + bg_i)z^i$$

$$z^m \cdot F(z) = \sum f_{i-m} z^i$$

$$F(az^k) = \sum a^i f_i z^{ik}$$

$$G'(z) = \sum (i+1) g_{i+1} z^i$$

$$F(z) \cdot G(z) = \sum_i \Big(\sum_{s+t=i} f_s g_t \Big) z^i$$

So we can perform certain operations on sequences:

- Add two sequences.
- Multiply a sequence by a scalar.
- Multiply a sequence by 1,2,3,...
- Shift a sequence to the right or left.
- Space out a sequence.
- Convolve two sequences.

Given a few basic sequences this allows us to construct a whole class of sequences.

$$1 = 1, 0, 0, 0, 0, \ldots, 0, \ldots$$

$$1/(1-z) = 1, 1, 1, 1, 1, \ldots, 1, \ldots$$

$$-\ln(1-z) = 0, 1, 1/2, 1/3, 1/4, \ldots, 1/n, \ldots$$

$$e^z = 1, 1, 1/2, 1/3!, 1/4!, \ldots, 1/n!, \ldots$$

$$1/(1-z)^m = 1, m, m(m+1)/2, m(m+1)(m+2)/6, \ldots, \binom{m+n-1}{m-1}, \ldots$$

How do we concoct the generating function for

$$1, 1, 2, 2, 4, 4, 8, 8, \ldots$$

No problem:

$$1/(1-z) = 1, 1, 1, 1, 1, 1, 1, \ldots$$

$$1/(1-2z^2) = 1, 0, 2, 0, 4, 0, 8, \ldots$$

$$z/(1-2z^2) = 0, 1, 0, 2, 0, 4, 0, \ldots$$

$$(1+z)/(1-2z^2) = 1, 1, 2, 2, 4, 4, 8, \ldots$$

You have an urn containing 30 red, 40 blue and 50 white balls. How many ways are there to extract 70 balls from the urn (order does not matter)?

Letting

$$p = \left(\sum_{i \le 30} z^i\right)\left(\sum_{i \le 10} z^i\right)\left(\sum_{i \le 50} z^i\right)$$

we need to calculate $[z^{70}]\, p$.

We could expand out the whole polynomial, but a better way is to rewrite the geometric sums as

$$p = \frac{1-z^{31}}{1-z}\,\frac{1-z^{41}}{1-z}\,\frac{1-z^{51}}{1-z} = \frac{(1-z^{31})(1-z^{41})(1-z^{51})}{(1-z)^3}$$

The denominator term $1/(1-z)^3$ can be expanded according to the binomial theorem as

$$1 + 3z + 6z^2 + 10z^3 + 15z^4 + \ldots + \binom{2+n}{2} z^n + \ldots$$

Since we are looking for z^{70} we only need the part

$$1 - z^{31} - z^{41} - z^{51}$$

from the numerator, all other terms have higher degree. So the answer is

$$[z^{70}]\, p = \binom{72}{2} - \binom{41}{2} - \binom{31}{2} - \binom{21}{2} = 1061.$$

A pair of standard dice produce the following relative frequencies for the 11 possible outcomes:

sum	2	3	4	5	6	7	8	9	10	11	12
#ways	1	2	3	4	5	6	5	4	3	2	1

Weird Question: Can we get the same frequencies using non-standard dice? Let's agree that the dice have 6 faces and the number of spots on each face is positive.

A priori there are 12 variables to deal with, one for each face. Of course, there are constraints that simplify matters, but it's still a bit tedious to try out "random" spot assignments.

George Sicherman seems to be the first to have solved this problem.

We can describe an ordinary die as

$$p = z^1 + z^2 + \ldots + z^6.$$

So we would like to find positive exponents a_i and b_i such that

$$p^2 = (z^{a_1} + z^{a_2} + \ldots + z^{a_6})(z^{b_1} + z^{b_2} + \ldots + z^{b_6})$$

Better is to factor p:

$$p^2 = z^2(1+z)^2(1+z+z^2)^2(1-z+z^2)^2$$

Then the two new polynomials must factor as

$$q_1 = z^{c_1}(1+z)^{c_2}(1+z+z^2)^{c_3}(1-z+z^2)^{c_4}$$
$$q_2 = z^{d_1}(1+z)^{d_2}(1+z+z^2)^{d_3}(1-z+z^2)^{d_4}$$

where $c_i + d_i = 2$, $c_i, d_i \geq 0$.

First note that we must have $c_1 = d_1 = 1$ since $q_i(0) = 0$: each face must have a positive number of spots.

We must have $c_2 = d_2 = c_3 = d_3 = 1$ since $q_i(1) = 6$: there are six faces.

This leaves c_4 and d_4.

One possibility is, of course, $c_4 = d_4 = 1$ and we have ordinary dice. The only other possibility is, say, $c_4 = 0$ and $d_4 = 2$ producing two dice

$$(1, 2, 2, 3, 3, 4) \qquad (1, 3, 4, 5, 6, 8)$$

That's it, there are no other solutions.

- Algebra and Infinity
- Generating Functions

3. Recurrences

- Regular Languages and Generating Functions

Generating functions suggest the following general approach towards solving recurrences:

- Write the solution (a_i) as $A(z) = \sum_{i \geq 0} a_i z^i$.
- Apply the recurrence to the terms of the power series and massage things into a nice closed form.
- Extract the terms of the sequence from the closed form of $A(z)$, e.g., by Taylor expansion.

The recurrence here has the form

$$f_n = \begin{cases} n & \text{if } n \leq 1, \\ f_{n-1} + f_{n-2} & \text{otherwise.} \end{cases}$$

Hence we have the generating function

$$\begin{aligned} F(z) &= \sum_{i \geq 0} f_i \cdot z^i \\ &= z + z^2 + 2z^3 + 3z^4 + 5z^5 + 8z^6 + 13z^7 \ldots \end{aligned}$$

By applying the recurrence to the series we get

$$\begin{aligned} F(z) &= \sum_{i \geq 0} f_i \cdot z^i \\ &= 0 + z + \sum_{i \geq 2} f_i \cdot z^i \\ &= z + \sum_{i \geq 2} (f_{i-1} + f_{i-2}) \cdot z^i \\ &= z + z \cdot \sum_{i \geq 0} f_i \cdot z^i + z^2 \cdot \sum_{i \geq 0} f_i \cdot z^i \\ &= z + zF(z) + z^2 F(z) \end{aligned}$$

and therefore

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

Since $F(z)$ is a simple rational function, we can determine the coefficients of the Taylor series easily using a partial fraction decomposition.

$$\begin{aligned} F(z) &= \frac{z}{1-z-z^2} \\ &= \frac{z}{(1-\varphi z)(1-\widehat{\varphi} z)} \\ &= \frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi z} - \frac{1}{1-\widehat{\varphi} z}\right) \end{aligned}$$

Here $\varphi = (1+\sqrt{5})/2$ and $\widehat{\varphi} = (1-\sqrt{5})/2$. Since

$$\frac{\partial^n}{\partial z^n}\frac{1}{1-\alpha z} = \frac{n!\,\alpha^n}{(1-\alpha z)^{n+1}}$$

we have

$$f_n = \frac{1}{\sqrt{5}}(\varphi^n - \widehat{\varphi}^n)$$

Let BT_n be the number of (ordered) binary trees on n nodes (which is the same as FT_{n+1}, the number of full binary trees on $n+1$ leaves).

Brute force table:

n	BT_n	FT_n
0	1	1
1	1	1
2	2	1
3	5	2
4	14	5
5	42	14
6	132	42
7	429	132
8	1430	429
9	4862	1430
10	16796	4862

The recurrence for binary trees takes the form

$$b_n = \begin{cases} 1 & \text{if } n = 0, \\ \sum_{i=0}^{n-1} b_i \cdot b_{n-i-1} & \text{otherwise.} \end{cases}$$

Let $B(z)$ be the generating function and note that

$$\begin{aligned} B(z)^2 &= 1 + 2z + 5z^2 + 14z^3 + 42z^4 + \dots \\ &= \sum \Big(\sum_{s+t=i} b_s b_t \Big) z^i \\ &= \sum b_{i-1} z^i \\ &= (B(z) - 1)/z \end{aligned}$$

Therefore

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

Performing Taylor expansion is somewhat difficult in this case, though of course not a problem for a computer algebra system:

$$1+z+2\,z^2+5\,z^3+14\,z^4+42\,z^5+132\,z^6+429\,z^7+1430\,z^8+4862\,z^9+16796\,z^{10}+O(z^{11})$$

It turns out to be easier to use the Binomial Theorem to get a better description.

$$\sqrt{1+\alpha} = \sum \binom{1/2}{i} \alpha^i$$

and thus

$$\begin{aligned} B(z) &= \frac{1}{2z}\left(1 - \sum \binom{1/2}{i} (-4z)^i\right) \\ &= 2\sum \binom{1/2}{i+1} (-4z)^i \end{aligned}$$

A simple calculation shows

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

Hence, $b_n = C_n$, the nth Catalan number.

- Algebra and Infinity
- Generating Functions
- Recurrences

4. Regular Languages and Generating Functions

Recall a definition from a few weeks back:

Definition

Given a language $L \subseteq \Sigma^\star$ its growth function is defined by $\gamma_L : \mathbb{N} \to \mathbb{N}$,

$$\gamma_L(n) = |L \cap \Sigma^n|$$

For regular languages growth functions are particularly simple: they have rational generating functions.

How does one compute γ_L?

Suppose we have a DFA M on states $[n]$ that accepts L. Define the full transition matrix of M to be $A \in \mathbb{N}^{n \times n}$ where

$$A(i,j) = \#\text{ transitions from } i \text{ to } j$$

Let $I^{\rightarrow}$ be the row unit vector with a 1 at the initial state.

Let $F^{\downarrow}$ be the column vector with 1's at the final states, 0's elsewhere.

Claim

$$\gamma_L(k) = I^{\rightarrow} A^k F^{\downarrow}$$

Note that this does not work for nondeterministic machines.

The full matrix for the even/even DFA looks like so:

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

In this case it is easy to calculate A^k:

$$A^{2k+1} = 2^{2k} A \qquad\qquad A^{2k} = 2^{2k-1}(\mathbf{1} - A)$$

Hence $\gamma_L(0) = 1$ and for $k > 0$ we have

$$\gamma_L(k) = 2^{k-2}(1 + (-1)^k)$$

It would seem natural to try to determine the generating function for γ_L.

How about something like

$$\sum_{i \geq 0} A^i z^i = \frac{I}{I - Az}$$

Looks great, but A is a matrix, not a real; the fraction on the right hand side requires a bit of explanation (and might just be meaningless).

I is the identity matrix, $I - Az$ is a matrix of linear polynomials in z. Such a matrix may well have an inverse, so we should interpret the fraction as a matrix inverse.

The inverse of $(I - Az)$ for the even/even DFA looks like so:

$$\frac{1}{1-4z^2}\begin{pmatrix} 1-2z^2 & z & z & 2z^2 \\ z & 1-2z^2 & 2z^2 & z \\ z & 2z^2 & 1-2z^2 & z \\ 2z^2 & z & z & 1-2z^2 \end{pmatrix}$$

Since $q_0 = 1$ and $F = \{1\}$ we only need the first term. Taylor expansion produces

$$\frac{1-2z^2}{1-4z^2} = 1 + 2z^2 + 8z^4 + 32z^6 + 128z^8 + 512z^{10} + \ldots$$

Works! But why?

To do this real justice one has to take a closer look at the algebraic structure of all formal power series whose coefficients are $n \times n$ integer matrices (which is often written $\mathbb{Z}^{n,n}[[z]]$ and has lots of interesting properties).

Less formally, we want to justify

$$(I - Az)^{-1} = \sum A^i z^i$$

Multiplying both side this simply means:

$$I = (I - Az) \cdot (I + Az + A^2 z^2 + A^3 z^3 \ldots)$$

The last equation is easily verified; just multiply out the RHS.

Incidentally, in this formal power series structure there is a Kleene star: $A^\star = \sum A^i z^i$, the unique solution of $X = AX + I$.

Write $g_i(z)$ for the generating function counting the words that lead from q_0 to i. Let $G(z) = (g_1(z), \ldots, g_n(z))^T$ be the column vector of all these functions.

Then we need to solve the linear system

$$(I - Az)\, G(z) = e$$

where e is the unit column vector indicating the initial state.

The generating function for γ_L is then $\sum_{i \in F} g_i(z)$.

Note that this type of computation can be handled easily on a computer algebra system; it's a bit unpleasant to do by hand.

Recall the problem of building a DFA for all words over $\{a, b\}$ that contain exactly three occurrences of the subword ab (subword, not factor).

The state complexity of this language is 7 and the matrix $I - Az$ for the minimal DFA has the form

$$\begin{pmatrix} 1-z & -z & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -z & -z & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -z & -z & 0 \\ 0 & 0 & 0 & 1 & -2z & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -z & -z \\ 0 & 0 & 0 & 0 & 0 & 1-2z & 0 \\ 0 & 0 & 0 & 0 & 0 & -z & 1-z \end{pmatrix}$$

1 is the initial state, 6 is a sink and 7 is final.

The generating function for γ_L is

$$\frac{3z^4}{(1-z)^2}$$

A little bit of pattern matching produces

$$\gamma_L(n) = \begin{cases} 0 & \text{if } n < 3, \\ 3(n-3) & \text{otherwise.} \end{cases}$$

Not exactly earth shattering, but very elegant.

- Generating functions are another example of a compact description of an infinite object.
- Combinatorial operations on sequences correspond to algebraic operations on the corresponding functions.
- Generating functions can be used to solve recurrence equations.
- There is a deep connection between rational generating functions and regular languages; in particular growth functions of regular languages are rational.