

1 Basics of generating functions

- Introduction [Wilf 1–3]:

- how to define a sequence: exact formula, recurrent relation (Fibonacci), algorithm (the sequence of primes); there are uncomputable sequences (programs that do not stop)
- a new way: power series (members of the sequence as coefficients in the series)
- advantages: many advanced tools from analytical theory of functions
- very powerful: works on many sequences where nothing else is known to work
- allows to get asymptotic formulas and statistical properties
- powerful way to prove combinatorial identities
- “Konečne vidím, že je tá matalýza aj na niečo dobrá. Keby mi to bol niekto predtým povedal. . .”

- Two examples [Wilf 3–7]:

- $a_{n+1} = 2a_n + 1$ for $n \geq 0$, $a_0 = 0$
- write few members, guess $a_n = 2^n - 1$, provable by induction
- multiply by x^n , sum over all n , assign gf: $\frac{A(x)}{x} = 2A(x) + \frac{1}{1-x}$
- partial fraction expansion: $A(x) = \frac{x}{(1-x)(1-2x)} = \frac{1}{1-2x} - \frac{1}{1-x}$
- the method stays basically the same for harder problems
- $a_{n+1} = 2a_n + n$ for $n \geq 0$, $a_0 = 1$
- exact formula not obvious; no unqualified variables in the recurrence
- obstacle: $\sum_{n \geq 0} nx^n = x/(1-x)^2$; solution: differentiation
- concern: is differentiation allowed? discussed later, but in principle yes: in formal power series (as an algebraic ring) or via convergence (if we care about analytical properties)
- $A(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{A}{(1-x)^2} + \frac{B}{1-x} + \frac{C}{1-2x} = \frac{-1}{(1-x)^2} + \frac{2}{1-2x}$
- $1/(1-x)^2$ is just $x/(1-x)^2$ (see above) shifted by 1
- $a_n = 2^{n+1} - n - 1$

- The method [Wilf 8]:

- 1. make sure variables in the recurrence are qualified (e.g. range for n)
- 2. name and define the gf
- 3. multiply by x^n , sum over all n in the range
- 4. express both sides in terms of the gf
- 5. solve the equation for gf
- 6. calculate coefficients of gf power series expansion
- useful notation: $[x^n]f(x)$; e.g.

$$[x^n]e^x = 1/n! \quad [t^r]\frac{1}{1-3t} = 3^r \quad [v^m](1+v)^s = \binom{s}{m}$$

- Solve $a_n = 5a_{n-1} - 6a_{n-2}$ for $n \geq 2$, $a_0 = 0$, $a_1 = 1$. [$G(x) = \frac{x}{(1-2x)(1-3x)}$; $a_n = 3^n - 2^n$]

- Fibonacci [Wilf 8–10]:

- three-term recurrence: $F_{n+1} = F_n + F_{n-1}$ for $n \geq 1$, $F_0 = 0$, $F_1 = 1$.

– apply the method ($r_{\pm} = (1 \pm \sqrt{5})/2$):

$$F(x) = \frac{x}{1-x-x^2} = \frac{x}{(1-xr_+)(1-xr_-)} = \frac{1}{r_+ - r_-} \left(\frac{1}{1-xr_+} - \frac{1}{1-xr_-} \right)$$

– $F_n = \frac{1}{\sqrt{5}}(r_+^n - r_-^n)$

– the second term is < 1 and goes to zero, so the first term $\frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})^n$ gives a good approximation

- Find ogf for the following sequences (always $n \geq 0$) [W1.1]:

- | | | |
|-----|--|--|
| (a) | $a_n = n$ | [introduce $x\mathbf{D}$; $(x\mathbf{D})\frac{1}{1-x} = \frac{x}{(1-x)^2}$] |
| (b) | $a_n = \alpha n + \beta$ | [$\alpha x/(1-x)^2 + \beta/(1-x)$] |
| (c) | $a_n = n^2$ | [$(x\mathbf{D})^2 1/(1-x) = \frac{1+x}{(1-x)^3}$] |
| (d) | $a_n = n^3$ | [$(x\mathbf{D})^3 1/(1-x)$] |
| (e) | $a_n = P(n)$; P is a polynomial of degree m | [$P(x\mathbf{D})\frac{1}{1-x}$] |
| (f) | $a_n = 3^n$ | [$1/(1-3x)$] |
| (g) | $a_n = 5 \cdot 7^n - 3 \cdot 4^n$ | [$\frac{5}{(1-7x)} - \frac{3}{1-4x}$] |
| (h) | $a_n = (-1)^n$ | [$1/(1+x)$] |

- Find the following coefficients [W1.5]:

- | | | |
|-----|---------------------------------------|--|
| (a) | $[x^n] e^{2x}$ | [$2^n/n!$] |
| (b) | $[x^n/n!] e^{\alpha x}$ | [α^n] |
| (c) | $[x^n/n!] \sin x$ | [$(-1)^m$ if $n = 2m + 1$ is odd, 0 otherwise] |
| (d) | $[x^n] 1/(1-ax)(1-bx)$ ($a \neq b$) | [$(a^{n+1} - b^{n+1})/(a-b)$] |
| (e) | $[x^n] (1+x^2)^m$ | [$[2 \mid n] \binom{m}{n/2}$] |

- Compute $\square_n = \sum_{k=1}^n k^2$.

– assign ogf to the sequence $1^2, 2^2, \dots, n^2$: $f(x) = \sum_{k=1}^n k^2 x^k$

– $(x\mathbf{D})^2[(x^{n+1} - 1)/(x - 1)] = x \frac{-2n^2 x^{n+1} + n^2 x^{n+2} + n^2 x^n - 2n x^{n+1} + x^{n+1} + 2n x^n + x^n - x - 1}{(x-1)^3}$

– note that $\square_n = f(1) = \lim_{x \rightarrow 1} (x\mathbf{D})^2[(x^{n+1} - 1)/(x - 1)] = n(n+1)(2n+1)/6$

- Find the sequence with gf $1/(1-x)^3$.

- Find a linear recurrence going back two sequence members that has a solution that contains $n \cdot 3^n$ (possibly plus some linear combination of other exponential or polynomial factors).

- Find explicit formulas for the following sequences [W1.6, R2, R3, R7]:

- | | | |
|-----|---|---|
| (a) | $a_{n+1} = 3a_n + 2$ for $n \geq 0$; $a_0 = 0$ | [$3x/(1-x)(1-3x)$; $3^n - 1$] |
| (b) | $a_{n+1} = \alpha a_n + \beta$ for $n \geq 0$; $a_0 = 0$ | [$\beta x/(1-x)(1-\alpha x)$; $\frac{\alpha^n - 1}{\alpha - 1} \beta$] |
| (c) | $a_{n+1} = a_n/3 + 1$ for $n \geq 0$; $a_0 = 1$ | [$\frac{3/2}{1-x} - \frac{1/2}{1-x/3}$; $\frac{3^{n+1}-1}{2 \cdot 3^n}$] |
| (d) | $a_{n+2} = 2a_{n+1} - a_n$ for $n \geq 0$, $a_0 = 0$, $a_1 = 1$ | [$x/(1-x)^2$; n] |
| (e) | $a_{n+2} = 3a_{n+1} - 2a_n + 3$ for $n \geq 0$; $a_0 = 1$, $a_1 = 2$ | [$\frac{4}{1-2x} - \frac{3}{(1-x)^2}$; $2^{n+2} - 3n - 3$] |
| (f) | $a_n = 2a_{n-1} - a_{n-2} + (-1)^n$ for $n > 1$; $a_0 = a_1 = 1$ | [$\frac{1/2}{(1-x)^2} - \frac{1/4}{1-x} + \frac{1/4}{1+x}$; $\frac{2n+3+(-1)^n}{4}$] |
| (g) | $a_n = 2a_{n-1} - n \cdot (-1)^n$ for $n \geq 1$; $a_0 = 0$ | [$\frac{x/9-2/9}{(1+x)^2} + \frac{2/9}{1-2x}$; $\frac{2^{n+1}-(3n+2)(-1)^n}{9}$] |
| (h) | $a_n = 3a_{n-1} + \binom{n}{2}$ for $n \geq 1$; $a_0 = 2$ | [$\frac{1}{8}(19 \cdot 3^n - 2n(n+2) - 3)$] |
| (i) | $a_n = 2a_{n-1} - a_{n-2} - 2$ for $n \geq 2$; $a_0 = a_{10} = 0$ | [$n(a_1 + 1 - n)$, so with a_{10} , $a_n = n(10 - n)$] |
| (j) | $a_n = 4(a_{n-1} - a_{n-2}) + (-1)^n$ for $n \geq 2$; $a_0 = 1$, $a_1 = 4$ | [$\frac{1+x+x^2}{(1+x)(1-2x)^2}$; $\frac{(-1)^n}{9} - \frac{5}{18} \cdot 2^n + \frac{7}{6}(n+1)2^n$] |
| (k) | $a_n = -3a_{n-1} + a_{n-2} + 3a_{n-3}$ for $n \geq 3$; $a_0 = 20$, $a_1 = -36$, $a_2 = 60$ | [$5(-3)^n + 18(-1)^n - 3$] |

2 Ordinary generating functions

- From the homework: solve $a_n = 2a_{n-1} - a_{n-2} - 2$ for $n \geq 1$; $a_0 = a_{10} = 0$.

Applying the standard method, while keeping a_1 as a parameter, we get

$$A(x) = \frac{a_1x - a_1x^2 - 2x^2}{(1-x)^3} = \frac{a_1x}{(1-x)^2} + \frac{x(1-x)}{(1-x)^3} - \frac{x^2+x}{(1-x)^3},$$

so $a_n = (a_1 + 1)n - n^2$. From $a_{10} = 0$ we get $a_1 = 9$, thus $a_n = n(10 - n)$.

- Another way for boundary problems (this particular example is motivated by splines, Wilf 10–11):

- consider $au_{n+1} + bu_n + cu_{n-1} = d_n$ for $1 \leq n \leq N-1$; $u_0 = u_N = 0$.
- similar to Fibonacci with two given non-consecutive terms (but more general)
- define $U(x) = \sum_{j=0}^N u_j x^j$ (unknown); $D(x) = \sum_{j=1}^{N-1} d_j x^j$ (known)
- derive $a \cdot \frac{U(x) - u_1 x}{x} + bU(x) + cx(U(x) - u_{N-1}x^{N-1}) = D(x)$
- $(a + bx + cx^2)U(x) = xD(x) + au_1x + cu_{N-1}x^N$ (*)
- plug in suitable values of x (roots r_+ and r_- of the quadratic polynomial on the LHS)
- solve the system of two linear equations and two unknowns u_1, u_{N-1}
- if the roots are equal, differentiate (*) to obtain the second equation

- Mutually recursive sequences [Knuth 343, Example 3]

- consider the number u_n of tilings of $3 \times n$ board with 2×1 dominoes
- define v_n as the number of tilings of $3 \times n$ board without a corner
- $u_n = 2v_{n-1} + u_{n-2}$; $u_0 = 1$; $u_1 = 0$
- $v_n = v_{n-2} + u_{n-1}$; $v_0 = 0$; $v_1 = 1$
- derive

$$U(x) = \frac{1-x^2}{1-4x^2+x^4}, \quad V(x) = \frac{x}{1-4x^2+x^4}$$
- consider $W(z) = 1/(1-4z+z^2)$; $U(x) = (1-x^2)W(x^2)$, so $u_{2n} = w_n - w_{n-1}$
- hence $u_{2n} = \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} + \frac{(2-\sqrt{3})^n}{3+\sqrt{3}} = \left\lceil \frac{(2+\sqrt{3})^n}{3-\sqrt{3}} \right\rceil$ (derivation as a homework)

- Given $f(x) \xleftrightarrow{\text{ogf}} (a_n)_{n \geq 0}$, express ogf for the following sequences in terms of f [W1.3]:

- | | | |
|-----|---|--|
| (a) | $(a_n + c)_{n \geq 0}$ | $\left[f(x) + c/(1-x) \right]$ |
| (b) | $(na_n)_{n \geq 0}$ | $\left[xDf(x) \right]$; napísať im $(P(n)a_n)_{n \geq 0} \longleftrightarrow P(xD)f(x)$ |
| (c) | $0, a_1, a_2, a_3, \dots$ | $\left[f(x) - a_0 \right]$ |
| (d) | $0, 0, 1, a_3, a_4, a_5, \dots$ | $\left[f(x) - a_0 - a_1x + (1-a_2)x^2 \right]$ |
| (e) | $(a_{n+2} + 3a_{n+1} + a_n)_{n \geq 0}$ | $\left[(f - a_0 - a_1x)/x^2 + 3(f - a_0)/x + f \right]$ |
| (f) | $a_0, 0, a_2, 0, a_4, 0, a_6, 0, \dots$ | $\left[(f(x) + f(-x))/2 \right]$ |
| (g) | $a_0, 0, a_1, 0, a_2, 0, a_3, 0, \dots$ | $\left[f(x^2) \right]$ |
| (h) | $a_1, a_2, a_3, a_4, \dots$ | $\left[(f(x) - a_0)/x \right]$ |
| (i) | a_0, a_2, a_4, \dots | $\left[(f(\sqrt{x}) + f(-\sqrt{x}))/2 \right]$ |

- introducing a new variable and changing the order of summation can help

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

- alternatively, one can use binomial theorem (Knuth 199/5.56 and 5.57):

$$\begin{aligned}\frac{1}{(1-z)^{n+1}} &= (1-z)^{-n-1} = \sum_{k \geq 0} \binom{-n-1}{k} (-z)^k \\ &= \sum_{k \geq 0} \frac{(-n-1)(-n-2)\dots(-n-k)}{k!} (-z)^k = \sum_{k \geq 0} \binom{n+k}{n} z^k\end{aligned}$$

Formal power series [Wilf chapter 2]

- a ring with addition and multiplication $\sum_n a_n x^n \sum_n b_n x^n = \sum_n \sum_k (a_k b_{n-k}) x^n$
- if $f(0) \neq 0$, then f has a unique reciprocal $1/f$ such that $f \cdot 1/f = 1$
- composition $f(g(x))$ defined iff $g(0) = 0$ or f is a polynomial (cf. e^{e^x-1} vs. e^{e^x})
- formal derivative D : $D \sum_n a_n x^n = \sum n a_n x^{n-1}$; usual rules for sum, product etc.
- exercise: find all f such that $Df = f$

Rules for manipulation [Wilf 2.1, Knuth 334]. Assume that $f \xleftrightarrow{\text{ogf}} (a_n)_{n=0}^\infty$.

- **Rule 1:** for a positive integer h , $(a_{n+h}) \xleftrightarrow{\text{ogf}} (f - a_0 - \dots - a_{h-1}x^{h-1})/x^h$
- **Rule 2:** if P is a polynomial, then $P(xD)f \xleftrightarrow{\text{ogf}} (P(n)a_n)_{n \geq 0}$
 - example: $(n+1)a_{n+1} = 3a_n + 1$ for $n \geq 0$, $a_0 = 1$; thus $f' = 3f + 1/(1-x)$
 - example: $\sum_{n \geq 0} \frac{n^2+4n+5}{n!}$; thus $f = \sum_{n \geq 0} (n^2+4n+5) \frac{x^n}{n!} = ((xD)^2+4xD+5)e^x = (x^2+5x+5)e^x$
we need $f(1) = 11e$; works because the resulting f is analytic in a disk containing 1 in the complex plane (that is, it converges to its Taylor series)

- **Rule 3:** if $g \xleftrightarrow{\text{ogf}} (b_n)$, then $fg \xleftrightarrow{\text{ogf}} (\sum_{k=0}^n a_k b_{n-k})_{n \geq 0}$

$$\sum_{k=0}^n (-1)^k k = (-1)^n \sum_{k=0}^n k \cdot (-1)^{n-k} = (-1)^n [x^n] \frac{x}{(1-x)^2} \cdot \frac{1}{1+x} = \frac{(-1)^n}{4} (2n+1 - (-1)^n)$$

- **Rule 4:** for a positive integer k , we have $f^k \xleftrightarrow{\text{ogf}} \left(\sum_{n_1+n_2+\dots+n_k=n} a_{n_1} a_{n_2} \dots a_{n_k} \right)_{n \geq 0}$
 - example: let $p(n, k)$ be the number of ways n can be written as an ordered sum of k nonnegative integers
 - according to R4, $(p(n, k))_{n \geq 0} \xleftrightarrow{\text{ogf}} 1/(1-x)^k$, so $p(n, k) = \binom{n+k-1}{n}$ thanks to (1)

- **Rule 5:** $\frac{f}{(1-x)} \xleftrightarrow{\text{ogf}} \left(\sum_{k=0}^n a_k \right)_{n \geq 0}$

$$\text{– example: } (\square_n)_{n \geq 0} \xleftrightarrow{\text{ogf}} \frac{1}{1-x} \cdot (xD)^2 \frac{1}{1-x} = \frac{x(1+x)}{(1-x)^4}, \text{ so by (1), } \square_n = \binom{n+2}{3} + \binom{n+1}{3}$$

1. Using Rule 5, prove that $F_0 + F_1 + \cdots + F_n = F_{n+2} - 1$ for $n \geq 0$ [Wilf 38, example 6].
[Compare gfs of both sides, left is $f/(1-x)$, where $f = x/(1-x-x^2)$, i.e. Fibonacci.]
2. Solve $g_n = g_{n-1} + g_{n-2}$ for $n \geq 2$, $g_0 = 0$, $g_{10} = 10$.
[$g_n = \frac{g_{10}}{F_{10}} F_n$, try the “boundary method” described above, computer necessary]
3. Solve $a_n = \sum_{k=0}^{n-1} a_k$ for $n > 0$; $a_0 = 1$. [R16]
[$a_n = 2^{n-1}$ for $n \geq 1$]
4. Solve $f_n = 2f_{n-1} + f_{n-2} + f_{n-3} + \cdots + f_1 + 1$ for $n \geq 1$, $f_0 = 0$ [Knuth 349/(7.41)]
[$F(x) = x/(1-3x+x^2)$; $f_n = F_{2n}$]
5. Solve $g_n = g_{n-1} + 2g_{n-2} + \cdots + ng_0$ for $n > 0$, $g_0 = 1$. [K7.7]
[$G(x) = 1 + x/(1-3x+x^2)$; $g_n = F_{2n} + [n=0]$]
6. Solve $g_n = \sum_{k=1}^{n-1} \frac{g_k + g_{n-k} + k}{2}$ for $n \geq 2$, $g_1 = 1$.
7. Solve $g_n = g_{n-1} + 2g_{n-2} + (-1)^n$ for $n \geq 2$, $g_0 = g_1 = 1$. [Knuth 341, example 2]
[$G(x) = \frac{1+x+x^2}{(1-2x)(1+x)^2}$; $g_n = \frac{7}{9}2^n + \frac{1}{9}(3n+2)(-1)^n$]
8. Solve $a_{n+2} = 3a_{n+1} - 2a_n + n + 1$ for $n \geq 0$; $a_0 = a_1 = 1$. [R24]
[$A(z) = \frac{2}{1-2z} - \frac{1}{(1-z)^3}$; $a_n = 2^{n+1} - \binom{n+2}{2}$]
9. Prove that $\ln \frac{1}{1-x} = \sum_{n \geq 1} \frac{1}{n} x^n$. [consider $\int \frac{1}{1-x}$]

3 Skipping sequence elements, Catalan numbers

Discovering combinatorial identities via gfs [Knuth 198, Vandermonde and 5.55]

- $(1+x)^r = \sum_{k \geq 0} \binom{r}{k} x^k$; consider $(1+x)^r(1+x)^s = (1+x)^{r+s}$
- comparison of coefficients yields $\sum_{k \geq 0} \binom{r}{k} \binom{s}{n-k} = \binom{r+s}{n}$ — Vandermonde
- by considering $(1-x)^r(1+x)^r = (1-x^2)^r$, we obtain

$$\sum_{k=0}^n \binom{r}{k} \binom{r}{n-k} (-1)^k = (-1)^{n/2} \binom{r}{n/2} [2 \mid n]$$

Every third binomial coefficient [Wilf 51, example 4]

- why $\frac{1}{2}(A(x) + A(-x)) \xrightarrow{\text{ogf}} a_0, 0, a_2, 0, a_4, \dots$ works: $\frac{1}{2}(1^n + (-1)^n) = [2 \mid n]$
- in general, for ω being r -th root of unity, $\frac{1}{r} \sum_{j=0}^{r-1} (\omega^j)^n = \frac{1}{r} \sum_{j=0}^{r-1} e^{2\pi i j n / r} = [r \mid n]$
— just a geometric progression, or a consequence of $t^r - 1 = (t-1)(t^{r-1} + \cdots + t + 1)$
- problem: find $S_n = \sum_k (-1)^k \binom{n}{3k}$
- if we knew $f(x) = \sum_k \binom{n}{3k} x^{3k}$, we would have $S_n = f(-1)$
- for $A(x) = (1+x)^n$, we have $f(x) = \frac{1}{3}(A(x) + A(x\omega) + A(x\omega^2))$ for $\omega = e^{2\pi i / 3}$
- and so $S_n = f(-1) = \frac{1}{3}[(1-\omega)^n + (1-\omega^2)^n] =$

$$= \frac{1}{3} \left[\left(\frac{3 - \sqrt{3}i}{2} \right)^n + \left(\frac{3 + \sqrt{3}i}{2} \right)^n \right] = 2 \cdot 3^{\frac{n}{2}-1} \cos\left(\frac{\pi n}{6}\right)$$

- consider the number of possibilities c_n of how to specify the multiplication order of $A_0 A_1 \dots A_n$ by parentheses; let $C(x) = \sum_{n \geq 0} c_n x^n$

- divide possibilities by the place of last multiplication; $c_n = \sum_{k=0}^{n-1} c_k c_{n-1-k}$ for $n > 0$; $c_0 = 1$

- many ways to deal with the recurrence:

(1) shift the recurrence to $c_{n+1} = \sum_{k=0}^n c_k c_{n-k}$ and use Rules 1 and 3; $\frac{C(x)-1}{x} = C(x)^2$

(2) RHS as a convolution of c_n with c_{n-1} , i.e. $C(x) \cdot xC(x)$

(3) RHS as a convolution of c_n with c_n shifted by Rule 1, i.e. $x \cdot C(x)^2$

(4) rewriting through sums and changing the order of summation:

$$\sum_{n \geq 1} x^n \sum_{k=0}^{n-1} c_k c_{n-1-k} = \sum_{k=0}^{\infty} x^k c_k \sum_{n \geq k+1} c_{n-1-k} x^{n-k} = \sum_{k=0}^{\infty} x^k c_k x C(x) = x C(x) \cdot C(x)$$

- consequently, $C(x) - 1 = x C(x)^2$ and thus $C(x) = \frac{1 \pm \sqrt{1-4x}}{2x} = \frac{1}{2x} (1 - \sqrt{1-4x})$
- we want C continuous and $C(0) = 1$, so we choose the minus sign (note that the resulting function below is analytical since $\binom{2n}{n}/(n+1) < 2^{2n}$; it would be analytical also if we chose the plus sign)
- binomial theorem yields

$$\begin{aligned} \sqrt{1-4x} &= (1-4x)^{1/2} = \sum_{k \geq 0} \binom{1/2}{k} (-4x)^k = 1 + \sum_{k \geq 1} \frac{1}{2k \cdot (-4)^{k-1}} \binom{2k-2}{k-1} (-4)^k x^k \\ &= 1 - \sum_{k \geq 1} \frac{2}{k} \binom{2k-2}{k-1} x^k \end{aligned}$$

- we used $\binom{1/2}{k} = \frac{1/2}{k} \binom{-1/2}{k-1} = \frac{1}{2k(-4)^{k-1}} \binom{2k-2}{k-1}$ because $\binom{-1/2}{m} = \frac{1}{(-4)^m} \binom{2m}{m}$

- therefore,

$$C(x) = \frac{1}{2x} \sum_{k \geq 1} \frac{2}{k} \binom{2k-2}{k-1} x^k = \sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} x^n$$

Exercises

1. Assume that $A(x) \overset{\text{ogf}}{\longleftrightarrow} (a_n)$. Express the generating function for $\sum_{n \geq 0} a_{3n} x^n$ in terms of $A(x)$.
[$\frac{1}{3}(A(x^{1/3}) + A(\omega x^{1/3})) + A(\omega^2 x^{1/3})$, where $\omega = e^{2\pi i/3}$]
2. Compute $S_n = \sum_{n \geq 0} F_{3n} \cdot 10^{-n}$ (by plugging a suitable value into the generating function for F_{3n}).
[The gf is $\frac{2x}{1-4x-x^2}$ and $S_n = 20/59$.]
3. Compute $\sum_k \binom{n}{4k}$. [$2^{\frac{n}{2}-2} (2^{\frac{n}{2}} + \cos(\frac{1}{4}n\pi) + (-1)^n \cos(\frac{3}{4}n\pi))$]
4. Compute $\sum_k \binom{6m}{3k+1}$. [Compute it for general n and then plug in $n = 6m$; $(2^{6m} - 1)/3$]
5. Evaluate $S_n = \sum_{k=0}^n (-1)^k k^2$. [$f(x) = \frac{-x}{(1+x)^3}$; $S_n = \frac{1}{2}(-1)^n n(n+1)$]
6. Find ogf for $H_n = 1 + 1/2 + 1/3 + \dots$. [$-\ln(1-x)/(1-x)$]
7. Find the number of ways of cutting a convex n -gon with labelled vertices into triangles.
[C_{n-2} (shifted Catalan numbers)]

4 Snake Oil

The Snake Oil method [Wilf 118, chapter 4.3] – external method vs. internal manipulations within a sum.

1. identify the free variable and give the name to the sum, e.g. $f(n)$
2. let $F(x) = \sum f(n)x^n$
3. interchange the order of summation; solve the inner sum in closed form
4. find coefficients of $F(x)$

- Example 0

– let's evaluate $f(n) = \sum_k \binom{n}{k}$; after Step 2, $F(x) = \sum_{n \geq 0} x^n \sum_k \binom{n}{k}$

$$- F(x) = \sum_k \sum_n \binom{n}{k} x^n = \sum_{k \geq 0} \frac{x^k}{(1-x)^{k+1}} = \frac{1}{1-x} \cdot \frac{1}{1-\frac{x}{1-x}} = \frac{1}{1-2x}$$

- Example 1 [Wilf 121]

– let's evaluate $f(n) = \sum_{k \geq 0} \binom{k}{n-k}$; after Step 2, $F(x) = \sum_n x^n \sum_{k \geq 0} \binom{k}{n-k}$

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

– so $f(n) = F_{n+1}$

- Example 2 [Wilf 122]

– let's evaluate $f(n) = \sum_k \binom{n+k}{m+2k} \binom{2k}{k} \frac{(-1)^k}{k+1}$, where m, n are nonnegative integers

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

– so $f(n) = \binom{n-1}{m-1}$

- Example 6 [Wilf 127]

– prove that $\sum_k \binom{m}{k} \binom{n+k}{m} = \sum_k \binom{m}{k} \binom{n}{k} 2^k$, where m, n are nonnegative integers

– the ogf of the left-hand side is

$$L(x) = \sum_k \binom{m}{k} x^{-k} \sum_{n \geq 0} \binom{n+k}{m} x^{n+k} = \frac{(1+x)^m}{(1-x)^{m+1}}$$

– we get the same for the right-hand side

1. Prove that $\sum_k k \binom{n}{k} = n2^{n-1}$ via the snake oil method. [$L(x) = P(x) = \frac{x}{(1-2x)^2}$]

2. Evaluate $f(n) = \sum_k k^2 \binom{n}{k} 3^k$.

$$\left[F(x) = \frac{3x(1+2x)}{(1-4x)^3} = \frac{3/8}{1-4x} - \frac{3/2}{(1-4x)^2} + \frac{9/8}{(1-4x)^3}; f(n) = 3 \cdot 4^{n-2} \cdot n(1+3n) \right]$$

3. Find a closed form for $\sum_{k \geq 0} \binom{k}{n-k} t^k$. [W4.11(a)]

$$\left[F(x) = 1/(1-tx-tx^2) \right]$$

4. Evaluate $f(n) = \sum_k \binom{n+k}{2k} 2^{n-k}$, $n \geq 0$. [Wilf 125, Example 4]

$$\left[F(x) = \frac{1-2x}{(1-x)(1-4x)} = \frac{2}{3(1-4x)} + \frac{1}{3(1-x)}; f(n) = (2^{2n+1} + 1)/3 \right]$$

5. Evaluate $f(n) = \sum_{k \leq n/2} (-1)^k \binom{n-k}{k} y^{n-2k}$. [Wilf 122, Example 3]

$$\left[F(x) = 1/(1-xy+x^2) \right]$$

6. Evaluate $f(n) = \sum_k \binom{2n+1}{2p+2k+1} \binom{p+k}{k}$. [W4.11(c)]

$$\left[\text{replace } 2n+1 \text{ by } m \text{ and solve for } f(m) = \binom{m-p-1}{p} 2^{m-2p-1}; f(2n+1) = \binom{2n-p}{p} 4^{n-p}; \right.$$

$$F(x) = \frac{x}{(1-x)^2} \sum_{k \geq 0} \binom{p+k}{p} \left(\frac{x}{1-x} \right)^{2(p+k)} = \frac{x^{p+1}}{2^p} \cdot \frac{(2x)^p}{(1-2x)^{p+1}} \left. \right]$$

7. Try to prove that $\sum_k \binom{n}{k} \binom{2n}{n+k} = \binom{3n}{n}$ via the snake oil method in three different ways: consider the sum

$$\sum_k \binom{n}{k} \binom{m}{r-k}$$

and the free variable being one of n , m , r .

5 Asymptotic estimates

- Purpose of asymptotics [Knuth 439]
 - sometimes we do not have a closed form or it is hard to compare it to other quantities
 - $S_n = \sum_{k=0}^n \binom{3n}{k} \sim 2 \binom{3n}{n}$; $S_n = \binom{3n}{n} \left(2 - \frac{4}{n} + O\left(\frac{1}{n^2}\right) \right)$
 - how to compare it with F_{4n} ? we need to approximate the binomial coefficient
 - purpose is to find *accurate* and *concise* estimates:
 H_n is $\sum_{k=1}^n 1/k$ vs. $O(\log n)$ vs. $\ln n + \gamma + O(n^{-1})$
- Hierarchy of log-exp functions [Hardy, see Knuth 442]
 - the class \mathcal{L} of logarithmico-exponential functions: the smallest class that contains constants, identity function $f(n) = n$, difference of any two functions from \mathcal{L} , e^f for every $f \in \mathcal{L}$, $\ln f$ for every $f \in \mathcal{L}$ that is “eventually positive”
 - every such function is identically zero, eventually positive or eventually negative
 - functions in \mathcal{L} form a hierarchy (every two of them are comparable by \prec or \asymp)
- Notations
 - $f(n) = O(g(n))$ iff $\exists c : |f(n)| \leq c|g(n)|$ (alternatively, for $n \geq n_0$ for some n_0)
 - $f(n) = o(g(n))$ iff $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$
 - $f(n) = \Omega(g(n))$ iff $\exists c : |f(n)| \geq c|g(n)|$ (alternatively, ...)
 - $f(n) = \Theta(g(n))$ iff $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$
 - basic manipulation: $O(f) + O(g) = O(|f| + |g|)$, $O(f)O(g) = O(fg) = fO(g)$ etc.
 - meaning of O in sums
 - *relative* vs. *absolute* error
- Warm-ups
 1. Prove or disprove: $O(f + g) = f + O(g)$ if f and g are positive. [K9.5] [false]
 2. Multiply $\ln n + \gamma + O(1/n)$ by $n + O(\sqrt{n})$. [K9.6] [$n \ln n + \gamma n + O(\sqrt{n} \ln n)$]
 3. Compare $n^{\ln n}$ with $(\ln n)^n$. [\prec]
 4. Compare $n^{\ln \ln \ln n}$ with $(\ln n)!$. [\prec]
 5. Prove or disprove: $O(x + y)^2 = O(x^2) + O(y^2)$. [K9.11] [true]
- Common tricks
 - cut off series expansion (works for convergent series, Knuth 451)
 - substitution, e.g. $\ln(1 + 2/n^2)$ with precision of $O(n^{-5})$ [$\frac{2}{n^2} - \frac{4}{n^4} + O(n^{-6})$]
 - factoring (pulling the large part out), e.g. $\frac{1}{n^2+n} = \frac{1}{n^2} \frac{1}{1+\frac{1}{n}} = \frac{1}{n^2} - \frac{1}{n^3} + O(n^{-4})$
 - division, e.g. $\frac{H_n}{\ln(n+1)} = \frac{\ln n + \gamma + O(n^{-1})}{(\ln n)(1 + O(n^{-1}))} = 1 + \frac{\gamma}{\ln n} + O(n^{-1})$
 - exp-log, i.e. $f(x) = e^{\ln f(x)}$
- Typical situations for approximation
 - Stirling formula: $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + O(n^{-3})\right)$
 - harmonic numbers: $H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + O(n^{-4})$
 - rational functions, e.g. $\frac{n+1}{n+2} = \frac{1+\frac{1}{n}}{1+\frac{2}{n}} = \left(1 + \frac{1}{n}\right) \left(1 - \frac{2}{n} + \frac{4}{n^2} + O(n^{-3})\right) = 1 - \frac{1}{n} + \frac{2}{n^2} + O(n^{-3})$

– exponentials: $e^{H_n} = ne^\gamma e^{O(1/n)} = ne^\gamma(1 + O(1/n)) = ne^\gamma + O(1)$

– rational function powered to n , e.g.

$$\left(\frac{n-3}{n}\right)^n = e^{n \ln(1-\frac{3}{n})} = \exp\left(n\left(\frac{-3}{n} + O(n^{-2})\right)\right) = e^{-3+O(n^{-1})} = \frac{1}{e^3} + O(n^{-1})$$

– binomial coefficient, e.g. $\binom{2n}{n}$: factorials and Stirling formula

$$\binom{2n}{n} = \frac{\sqrt{4\pi n} \left(\frac{2n}{e}\right)^{2n} (1 + O(n^{-1}))}{2\pi n \left(\frac{n}{e}\right)^{2n} (1 + O(n^{-1}))^2} = \frac{2^{2n}}{\sqrt{\pi n}} (1 + O(n^{-1}))$$

• Exercises

1. Estimate $\ln(1 + 1/n) + \ln(1 - 1/n)$ with abs. error $O(n^{-3})$ [$-1/n^2 + O(n^{-4})$]
2. Estimate $\ln(2 + 1/n) - \ln(3 - 1/n)$ with abs. error $O(n^{-2})$ [$\ln \frac{2}{3} + \frac{5}{6n} + O(n^{-2})$]
3. Estimate $\lg(n - 2)$, abs. error $O(n^{-2})$ [$\frac{\ln n}{\ln 2} - \frac{2}{n \ln 2} + O(n^{-2})$]
4. Evaluate H_n^2 with abs. error $O(n^{-1})$. [$(\ln n)^2 + 2\gamma \ln n + \gamma^2 + (\ln n)/n + O(1/n)$]
5. Estimate $n^3/(2 + n + n^2)$ with abs. error $O(n^{-3})$ [$n - 1 - \frac{1}{n} + \frac{3}{n^2} + O(n^{-3})$]
6. Prove or disprove: [K9.20] (b) $e^{(1+O(1/n))^2} = e + O(1/n)$ (c) $n! = O(((1 - 1/n)^n)^n)$ [yes, no]
7. Evaluate $(n + 2 + O(n^{-1}))^n$ with rel. error $O(n^{-1})$. [K9.13] [$n^n \cdot e^2(1 + O(n^{-1}))$]
8. Compare H_{F_n} with $F_{[H_n]}^2$ [K9.2] [$H_{F_n} \sim n \ln \varphi$, $F_{[H_n]}^2 = O(n^{\ln \varphi^2}) = o(n)$]
9. Estimate $\sum_{k \geq 0} e^{-k/n}$ with abs. error $O(n^{-1})$. [K9.7] [$n + 1/2 + O(n^{-1})$]
10. Estimate $H_n^5 / \ln(n + 5)$ with abs. error $O(n^{-2})$. [$2 + \frac{\gamma}{\ln n} - \frac{6}{n \ln n} - \frac{3\gamma}{n \ln^2 n} + O(n^{-2})$]
11. Estimate $\binom{2n}{n}$ with relative error $O(n^{-2})$. [A1] [$\frac{2^{2n}}{\sqrt{\pi n}} (1 - \frac{1}{8n} + O(n^{-2}))$]
12. Estimate $\binom{2n+1}{n}$ with relative error $O(n^{-2})$. [A2] [$\frac{2^{2n+1}}{\sqrt{\pi n}} (1 - \frac{1}{5n} + O(n^{-2}))$]
13. Compare $(n!)!$ with $((n-1)!)! \cdot (n-1)!^{n!}$. [K9.2c] (Homework if not enough time is left.)

6 Estimates of sums and products

• Warm-ups

1. Let $f(n) = \sum_{k=1}^n \sqrt{k}$. Show that $f(n) = \Theta(n^{3/2})$. Find $g(n)$ such that $f(n) = g(n) + O(\sqrt{n})$. [$\int_0^n \sqrt{x} dx \leq S_n \leq \int_1^{n+1} \sqrt{x} dx$; $g(n) = \frac{2}{3}n\sqrt{n}$]
2. Estimate $(n-2)!/(n-1)!$ with abs. error $O(n^{-2})$. [TODO consider $\frac{n!}{n(n-1)^2}$]
3. For a constant integer k , estimate n^k/n^k with abs. error $O(n^{-3})$. [A5]

$$\left[1 - \binom{k}{2} \frac{1}{n} + \frac{3k^4 - 10k^3 + 9k^2 - 2k}{24} \frac{1}{n^2} + O\left(\frac{1}{n^3}\right) \right]$$

- Find a good estimate of $P_n = \frac{(2n-1)!!}{n!}$.

– obviously $1.5^{n-1} \leq \frac{1}{1} \cdot \frac{3}{2} \cdot \frac{5}{3} \cdot \dots \cdot \frac{(2n-1)}{n} \leq 2^{n-1}$

– we split the product into a “small” part (first k terms, each at least $3/2$ except the first one) and a “large” part (remaining $n - k$ terms); then

$$P_n \geq \left(\frac{2k+1}{k+1}\right)^{n-k} \cdot 1.5^{k-1} = Q_n \cdot 1.5^{k-1}; \text{ we estimate } Q_n$$

– if we try $k = \alpha n$, then

$$Q_n = 2^{n-\alpha n} \exp\left((n-\alpha n) \ln\left(1 - \frac{1}{2(\alpha n + 1)}\right)\right) = 2^{n(1-\alpha)} e^{\frac{\alpha-1}{2\alpha}} (1 + O(n^{-1})),$$

$$\text{so } P_n \geq (2^{1-\alpha} \cdot 1.5^\alpha)^n \Theta(1)$$

– if we try $k = \ln n$, then

$$Q_n = \exp \left((n - \ln n) \left[\ln 2 + \ln \left(1 - \frac{1}{2(1 + \ln n)} \right) \right] \right);$$

if we expand \ln into Taylor series, the error will be $1/\ln^k n = \omega(n^{-1})$, so we can get relative error $O(1)$ at best;

anyway, if we carry it through, we get $P_n = \Omega(2^n n^{-c} e^{-0.5n/\ln n})$

– if we try $k = \sqrt{n}$, then

$$\begin{aligned} Q_n &= \exp \left((n - \sqrt{n}) \left[\ln 2 + \ln \left(1 - \frac{1}{2(1 + \sqrt{n})} \right) \right] \right) \\ &= 2^{n-\sqrt{n}} \exp \left((n - \sqrt{n}) \left[\frac{-1}{2\sqrt{n}} + \frac{3}{8n} - \frac{7}{24n^{3/2}} + O(n^{-2}) \right] \right) \\ &= 2^{n-\sqrt{n}} \exp \left(-\frac{\sqrt{n}}{2} + \frac{7}{8} - \frac{2}{3\sqrt{n}} + O(n^{-1}) \right), \end{aligned}$$

thus $P_n \geq 2^n \cdot 0.75^{\sqrt{n}} \cdot e^{-\frac{\sqrt{n}}{2} + \frac{7}{8} - \frac{2}{3\sqrt{n}}} (1 + O(n^{-1})) = \Omega \left(2^n c^{\sqrt{n}} \right)$ for $c \in (0, 1)$.

– TODO compare with previous estimate from $k = \ln n$; which is better?

– another approach: $P_n = \frac{(2n)!}{n!2^n n!} = \binom{2n}{n}/2^n = \frac{2^n}{\sqrt{\pi n}} (1 + O(n^{-1}))$

- Estimate $S_n = \sum_{k=1}^n \frac{1}{n^2+k}$ with absolute error (a) $O(n^{-3})$, (b) $O(n^{-7})$. [Knuth 458/Problem 4]
First approach: $\frac{1}{n^2+k} = \frac{1}{n^2(1+k/n^2)}$ etc.; second approach: $S_n = H_{n^2+n} - H_n$. (DU)

- Sums — gross bound on the tail: $S_n = \sum_{0 \leq k \leq n} k! = n! \left(1 + \frac{1}{n} + \frac{1}{n(n-1)} + \dots \right)$, all the terms except the first two are at most $1/n(n+1)$, so $S_n = n! \left(1 + \frac{1}{n} + n \frac{1}{n(n-1)} \right) = n! (1 + O(n^{-1}))$

- Sums — make the tail infinite:

$$\begin{aligned} n! \sum_{k=0}^n \frac{(-1)^k}{k!} &= n! \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} - \sum_{k \geq n+1} \frac{(-1)^k}{k!} \right) \\ &= n! \left(e^{-1} - O \left(\frac{1}{(n+1)!} \right) \right) = \frac{n!}{e} + O(n^{-1}) \end{aligned}$$

- Estimate $S_n = \sum_{k=0}^n \binom{3n}{k}$ with relative error $O(n^{-2})$. We split the sum into a “small” and a “large” part at b (which is yet to be determined).

$$\begin{aligned} \sum_{k=0}^n \binom{3n}{k} &= \sum_{k=0}^n \binom{3n}{n-k} = \sum_{0 \leq k < b} \binom{3n}{n-k} + \sum_{b \leq k \leq n} \binom{3n}{n-k}. \\ \binom{3n}{n-k} &= \binom{3n}{n} \frac{n(n-1) \cdots 1}{(2n+1)(2n+2) \cdots (2n+k)} = \\ &= \binom{3n}{n} \cdot \frac{n^k}{(2n)^k} \frac{\prod_{j=0}^{k-1} (1 - \frac{j}{n})}{\prod_{j=1}^k (1 + \frac{j}{2n})} = \binom{3n}{n} \cdot \frac{1}{2^k} \cdot \left[1 - \frac{3k^2 - k}{4n} + O \left(\frac{k^4}{n^2} \right) \right]. \\ \sum_{b \leq k \leq n} \binom{3n}{n-k} &\leq n \cdot \binom{3n}{n-b} = \binom{3n}{n} \cdot \frac{1}{2^b} O(n) = \binom{3n}{n} \cdot O(n^{-2}) \text{ if } \sqrt{n} \succ b \geq 3 \lg n. \\ \sum_{0 \leq k < 3 \lg n} \frac{1}{2^k} &= 2 - \frac{1}{2^{3 \lg n}} = 2 + O(n^{-3}). \\ -\frac{3}{4n} \sum_{0 \leq k < 3 \lg n} \frac{k^2}{2^k} &= \frac{-9}{2n} + O(n^{-3}). \\ +\frac{1}{4n} \sum_{0 \leq k < 3 \lg n} \frac{k}{2^k} &= \frac{1}{2n} + O(n^{-3}). \\ O(n^{-2}) \cdot \sum_{0 \leq k < 3 \lg n} \frac{k^4}{2^k} &= O(n^{-2}) \end{aligned}$$

$$\sum_{k=0}^n \binom{3n}{k} = \binom{3n}{n} \cdot \left[2 - \frac{4}{n} + O(n^{-2}) \right]$$

- Estimate $S_n = \sum_{k=0}^n \binom{4n+1}{k+1}$ with relative error $O(n^{-2})$.

$$\binom{4n+1}{k+1} = \binom{4n}{k+1} + \binom{4n}{k};$$

$$S_n = \sum_{k=0}^n \binom{4n+1}{k+1} = \sum_{k=0}^n \binom{4n}{k+1} + \sum_{k=0}^n \binom{4n}{k} = \sum_{k=0}^n \binom{4n}{k} + \sum_{k=1}^{n+1} \binom{4n}{k};$$

$$S_n = 2 \sum_{k=0}^n \binom{4n}{k} + \binom{4n}{n+1} - \binom{4n}{0}.$$

$$Q_n = \sum_{k=0}^n \binom{4n}{k} = \sum_{k=0}^n \binom{4n}{n-k};$$

$$\binom{4n}{n-k} = \binom{4n}{n} \cdot \frac{\prod_{j=0}^{k-1} (n-j)}{\prod_{j=1}^k (3n+j)} = \binom{4n}{n} \cdot \left(\frac{1}{3}\right)^k \cdot \frac{\prod_{j=0}^{k-1} (1-j/n)}{\prod_{j=1}^k (1+j/3n)}$$

$$Q_n = \sum_{0 \leq k \leq 2 \log_3 n} \binom{4n}{n-k} + \sum_{2 \log_3 n \leq k < n} \binom{4n}{n-k}$$

$$\sum_{2 \log_3 n \leq k < n} \binom{4n}{n-k} = O\left(n \cdot \binom{4n}{n - \lceil 2 \log_3 n \rceil}\right) = O\left(\binom{4n}{n} \cdot \frac{1}{n}\right).$$

$$\frac{\prod_{j=0}^{k-1} (1-j/n)}{\prod_{j=1}^k (1+j/3n)} = \frac{1 - \frac{1}{n} \cdot \sum_{0 \leq j < k} j + O\left(\frac{k^4}{n^2}\right)}{1 + \frac{1}{3n} \cdot \sum_{1 \leq j \leq k} j + O\left(\frac{k^4}{n^2}\right)} = 1 + \frac{2k^2 + k}{3n} + O\left(\frac{\log^n}{n^2}\right),$$

$$\begin{aligned} \sum_{0 \leq k \leq 2 \log_3 n} \binom{4n}{n-k} &= \binom{4n}{n} \cdot \sum_{0 \leq k \leq 2 \log_3 n} \left(\frac{1}{3}\right)^k \cdot \left[1 + \frac{2k^2 + k}{3n} + O\left(\frac{\log^n}{n^2}\right)\right] = \\ &= \frac{3}{2} \cdot \binom{4n}{n} (1 + O(n^{-1})). \end{aligned}$$

$$\binom{4n}{n+1} = \binom{4n}{n} \cdot \frac{3n}{n+1} = 3 \cdot \binom{4n}{n} (1 + O(n^{-1}));$$

$$S_n = 6 \cdot \binom{4n}{n} (1 + O(n^{-1})).$$

- How many bits are needed to represent a binary tree with n internal nodes?

- we need just the internal vertices to capture the structure; what is the relation between the number of internal vertices and total number of vertices?
- imagine labeling the vertices by $1, 2, \dots, n$ in such a way that we get a binary search tree (descendants in the left subtree are smaller, in the right subtree are larger); by summing over possible roots of the tree we get $t_n = \sum_{i=1}^n t_{i-1} t_{n-i}$; $t_0 = 1$
- this is the same as for Catalan numbers, so $t_n = \binom{2n}{n} \frac{1}{n+1}$
- and so we need $\log_2 t_n \sim 2n - 1.5 \lg n - 0.5 \lg \pi + O(n^{-1})$ bits