GRAZ, July 2006

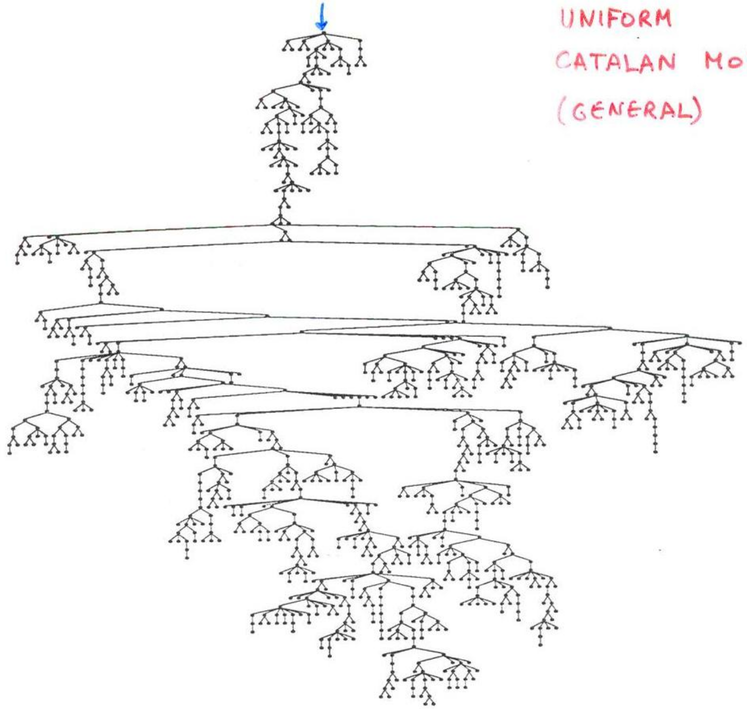
Singularities and Random Combinatorial Structures

Philippe FLAJOLET
INRIA, France.

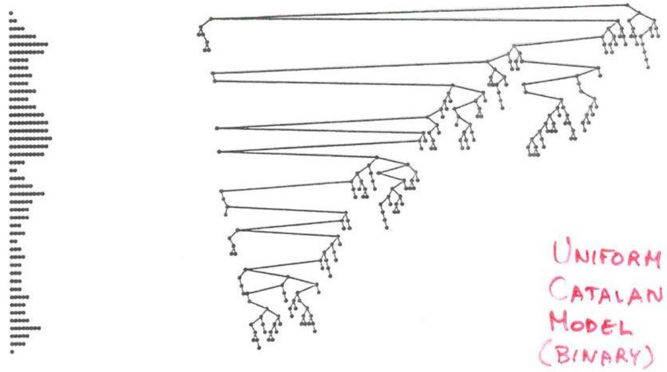
ANALYTIC COMBINATORICS

- Symbolic methods B1
- Complex asymptotic methods B2-B3
- Random structures B4.

Book by P.F + R. Sedgewick (2007)



§5.4 TREES 231



UNIFORM CATALAN MODEL (BINARY)

Figure 5.5 A random binary tree with 256 internal nodes

BINARY SEARCH TREE

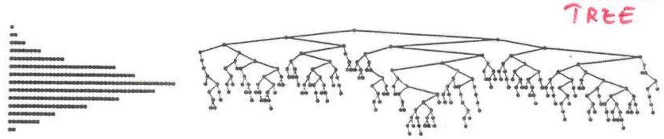


Figure 5.11 A binary search tree built from 256 randomly ordered keys

Sedgewick & Flajolet: An Introduction to the Analysis of Algorithms

1, 1, 2, 5, 14, ...

?

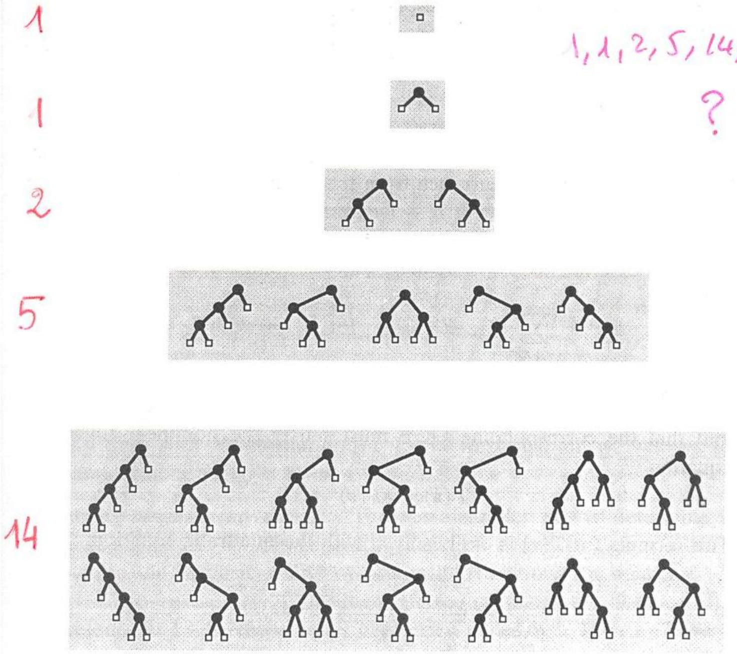
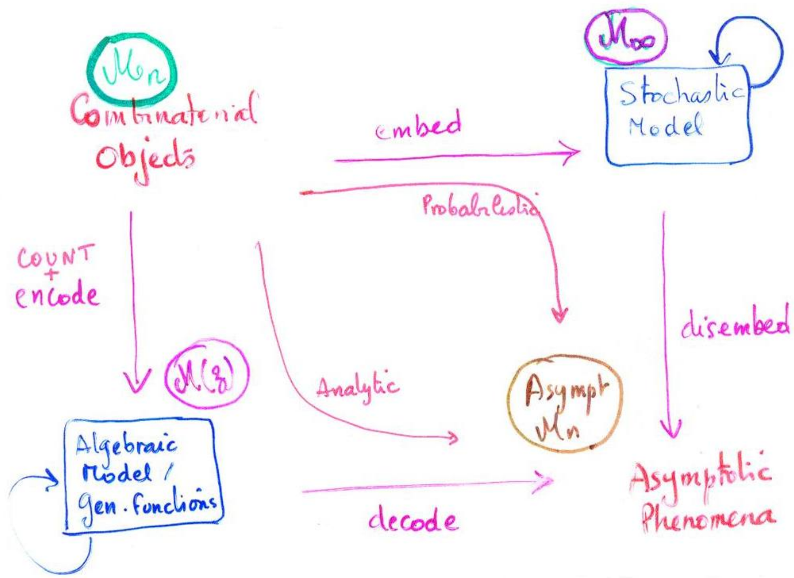
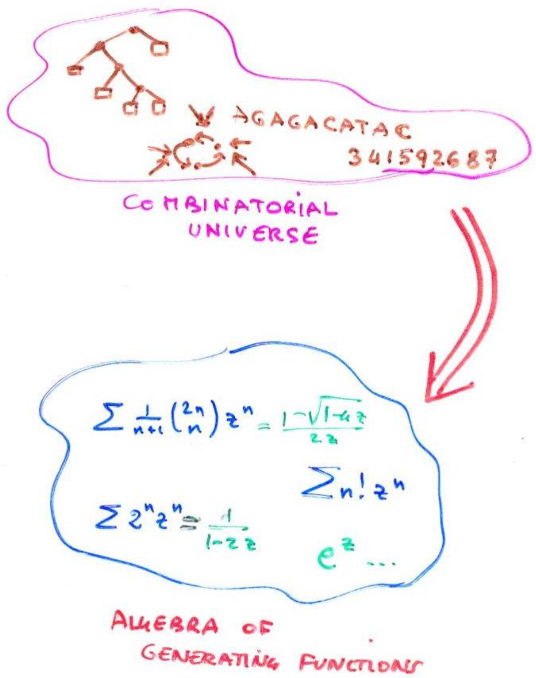


Figure 3.1 All binary trees with 1, 2, 3, 4, and 5 external nodes



PART I. SYMBOLIC METHODS



CHAPTER 1

Unlabelled Structures and Ordinary generating functions

Ordinary Generating Function OGF

$$(f_n) \longrightarrow f(z) = \sum_{n=0}^{\infty} f_n z^n$$

number sequence
e.g., counting sequence

(Later: Exponential generating function (EGF)

$$(f_n) \longrightarrow \sum f_n \frac{z^n}{n!}$$)

$\mathcal{C}$ = a combinatorial class (at most denumerable set w/ SIZE function)

$\mathcal{C}_n$ = subclass of objects of size n

C_n = # objects of size n = card$(\mathcal{C}_n)$

$$C(z) = \sum_{n=0}^{\infty} C_n z^n = \sum_{\gamma \in \mathcal{B}} z^{|\gamma|}$$

Naturally: Count up to combinatorial isomorphism

($\mathcal{C} \equiv \mathcal{D}$ if $\exists$ size-preserving bijections).

How many binary trees B_N with N external nodes?

$$\mathcal{B} = \square +$$ 

```
type BinTree = Union(
                 ExternalNode,
                 Product(BinTree, BinTree));
```

Classical (naïve?) approach

$$B_N = \sum_{k=1}^{N-1} B_k B_{N-k} \qquad (N>1)$$

$$B(z) := \sum_N B_N z^N$$
$$= \sum_{t\in\mathcal{B}} z^{|t|}$$

$$\sum_{t\in\mathcal{B}} z^{|t|} = \sum_{t=\square} z^{|t|} + \sum_{t_1,t_2} z^{|t_1|+|t_2|}$$
$$= z^{|\square|} + \underbrace{\sum_{t_1\in\mathcal{B}} z^{|t_1|}} \times \underbrace{\sum_{t_2\in\mathcal{B}} z^{|t_2|}}$$

$$B(z) = z + (B(z)\times B(z))$$

$$B = z + B^2$$

Solve the quadratic equation

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

The number of binary trees with $\underline{N}$ external nodes is

$$B_N = \frac{1}{N}\binom{2N-2}{N-1} \qquad \text{(Catalan numbers)}$$

OUTLINE

Define a collection of set-theoretic combinatorial CONSTRUCTIONS

union, product, sequence, set, cycle (...)

allowing RECURSIVE DEFINITIONS

Then:

meta-THM1: Generating functions are automatically computable (equations!)

meta-THM2: Counting sequences are also automatically computable $O(N^2)$ $\hookrightarrow O(N^{1+\epsilon})$

meta-THM3: Random generation is fast $O(N \log N)$ arithmetic complexity

Theorem 1.1 There exists a dictionary

{Combinatorial Constructions} $\longleftrightarrow$ {OGF operations}

UNLABELLED

Construction	Translation (OGF)
$\mathcal{F} = \mathcal{G} \uplus \mathcal{H}$	$F(z) = G(z) + H(z)$
$\mathcal{F} = \mathcal{G} \times \mathcal{H}$	$F(z) = G(z) \cdot H(z)$
$\mathcal{F} = \texttt{sequence}(\mathcal{G}) = \mathcal{G}^*$	$F(z) = \frac{1}{1-G(z)}$
$\mathcal{F} = \texttt{set}(\mathcal{G})$	$F(z) = \exp(G(z) - \frac{1}{2}G(z^2) + \frac{1}{3}G(z^3) - \cdots)$
$\mathcal{F} = \texttt{multiset}(\mathcal{G})$	$F(z) = \exp(G(z) + \frac{1}{2}G(z^2) + \frac{1}{3}G(z^3) + \cdots)$
$\mathcal{F} = \texttt{cycle}(\mathcal{G})$	$F(z) = \log(1 - G(z))^{-1} + \cdots$
$\cancel{\mathcal{F} = \mathcal{G}[\mathcal{H}]}$	$\cancel{F(z) = G(H(z))}$

$\mathcal{E}$ or 1 : neutral class formed with one element of size 0 [cf empty word] $\rightarrow$ $E(z) \equiv 1$.

$\mathcal{Z}$: atomic class formed with one element of size 1, an ATOM. $\rightarrow$ $Z(z) \equiv z$

Proof (s)

$$\mathcal{A} \longmapsto A(z) = \sum A_n z^n$$

$$A(z) = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|}$$

UNION

$$\mathcal{C} = \mathcal{A} \cup \mathcal{B} \qquad \boxed{C(z) = A(z) + B(z)}$$

$$\sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_{\alpha \in \mathcal{A}} z^{|\alpha|} + \sum_{\beta \in \mathcal{B}} z^{|\beta|}$$

CARTESIAN PRODUCT

$$\boxed{C(z) = A(z) \times B(z)}$$

$$\mathcal{C} = \mathcal{A} \times \mathcal{B}$$

$$\sum_{\gamma \in \mathcal{C}} z^{|\gamma|} = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} z^{|\alpha|+|\beta|}$$

SEQUENCE

$$\mathcal{C} = \mathrm{Seq}(\mathcal{A}) \qquad \boxed{C(z) = \frac{1}{1-A(z)}}$$

(i) $\mathcal{C} = \mathbf{1} + (\mathcal{A} \times \mathcal{C})$ (ii) $\frac{1}{1-f} = 1 + f + f^2 + \cdots$

SET

$$\mathcal{C} = \mathrm{Set}(\mathcal{A}) \qquad \boxed{C(z) = \exp\left(A(z) - \frac{A(z^2)}{2} + \cdots\right)}$$

$$\cong \prod_{\alpha \in \mathcal{A}} (1 + \{\alpha\})$$

PÓLYA OPERATOR

$$C(z) = \prod_{n=1}^{\infty} (1 + z^n)^{A_n} = \exp\left(\sum_n A_n \log(1 + z^n)\right) = \cdots$$

MULTISET

$$\mathcal{C} = \mathrm{MSet}(\mathcal{A}) \cong \prod_{\alpha \in \mathcal{A}} \mathrm{Seq}(\{\alpha\}).$$

$$C(z) = \prod_{n=1}^{\infty} (1 - z^n)^{-A_n} = \exp\left(A(z) + \frac{1}{2} A(z^2) + \cdots\right).$$

End of Proof ■

This theorem permits us to write AUTOMATICALLY for binary trees

$$\mathcal{B} = \blacksquare + (\mathcal{B} \bullet \mathcal{B})$$

$$\Downarrow$$

$$B(z) = z + B(z)^2$$

EXAMPLE 1 WORDS

$$\boxed{\mathcal{W} = \mathrm{Seq}(\underline{a} + \underline{b})} \qquad W(z) = \frac{1}{1-2z}$$

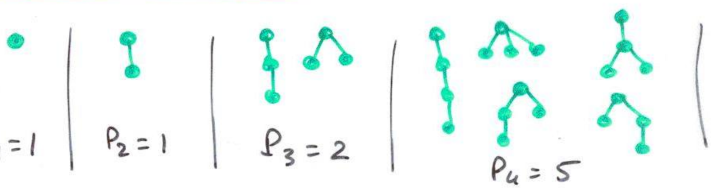
starting with $\underline{b}$, never more than 4 consec. "$\underline{a}$"

$$\mathcal{W}^{\bullet} = \mathrm{Seq}\left(\underline{b} \times (1 + \underline{a} + \underline{aa} + \underline{aaa} + \underline{aaaa})\right)$$

$$W^{\bullet}(z) = \frac{1}{1-(z+z^2+z^3+z^4+z^5)}$$

EXAMPLE 2 PLANE TREES ("general")

$$\boxed{\mathcal{P} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{P})}$$



$P_1 = 1$ | $P_2 = 1$ | $P_3 = 2$ | $P_4 = 5$

$$P(z) = \frac{z}{1-P(z)} \Rightarrow P(z) = \frac{1-\sqrt{1-4z}}{2} \qquad P_n = \frac{1}{n}\binom{2n-2}{n-1}$$

EXAMPLE 3 NONPLANE TREES

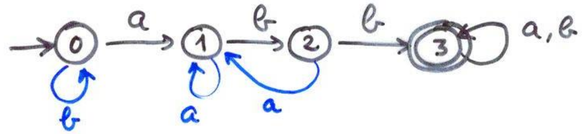
$$\boxed{\mathcal{N} = \mathcal{Z} \times \mathrm{Set}(\mathcal{N})}$$

$U_1 = 1$, $U_2 = 1$, $U_3 = 2$, $U_4 = 4$ ...

$$U(z) = z.\exp\left(\frac{U(z)}{1} + \frac{U(z^2)}{2} + \frac{U(z^3)}{3} + \dots\right)$$

CAYLEY → Recurrences; Pólya → Asymptotics

■ - Words containing a pattern

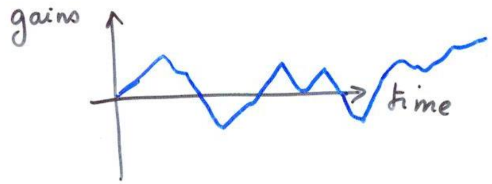


$\mathcal{L}_j$: language accepted from state j

$$\begin{cases} \dots \\ \mathcal{L}_1 = a \cdot \mathcal{L}_1 + b\,\mathcal{L}_2 \\ \dots \end{cases}$$

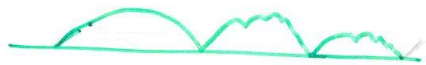
$$\boxed{\mathrm{Reg} \mapsto \mathbb{Q}(z)}$$
rational

■ - Walks and Excursions



$$\boxed{\text{Walks} \mapsto \mathbb{Q}(z, \sqrt{1-4z^2})}$$

Excursion =



$= \mathrm{Seq}(\nearrow \text{Excursion} \searrow)$

Positive path = Excursion × Seq($\nearrow$ Excursion)

Draw game = Seq($\searrow \overline{\text{Excursion}} \nearrow + \nearrow \text{Excursion} \searrow$) etc

Exercise: Integer compositions

- a sequence $(x_1, \ldots, x_k)$ with $x_j \in \mathbb{Z}_{\geq 1}$ is a composition of n if $x_1 + \cdots + x_k = n$. (Order of summands counts)

Q: How many compositions of integer n, for $n = 1, 2, 3, 4$
– for general n?

Hint: Argue that $\mathcal{C} = \mathrm{Seq}(\mathcal{N})$; $\mathcal{N} = \mathcal{Z} \times \mathrm{Seq}(\mathcal{Z})$

Exercise: Denumerants. In how many ways can you arrange to give a change for n Euro-cents knowing that you have an arbitrary supply of coins of 1¢, 2¢, 5¢, 10¢?

[order of coins does NOT matter!]

Q1: Prove that OGF is $\dfrac{1}{(1-z)(1-z^2)(1-z^5)(1-z^{10})}$?

Q2: generalize to Ω. **Q3** How to obtain coefficients? (Complexity?)

Exercise: Unary-binary trees are (plane) trees where nodes have degree either 0, 1, or 2.

Q1: How many U-B trees of size 1, 2, 3, 4?

Using
$$\begin{cases} \mathcal{U} = \bullet \times (1 + \mathcal{U} + \mathcal{U} \times \mathcal{U}) \\ \mathcal{U} = \bullet + \underset{\mathcal{U}}{\bullet} + \underset{\mathcal{U}\ \mathcal{U}}{\bullet} \end{cases}$$

find equation for OGF and solve it..

Q2: Generalization: Let $\Omega \ni 0$ and let $\mathcal{Y}$ be the class of trees with nodes having (out)degrees constrained by Ω. Establish for the OGF the equation

$$(Y \equiv Y(z)) \qquad Y = z\,\phi(Y).$$

Lagrange Inversion Theorem:

$$[z^n]\,Y(z) = \frac{1}{n}\,[w^{n-1}]\,\phi(w)^n.$$

Q3: Ω finite $\Rightarrow Y_n$ is a sum of products of ~~binomial~~ multinomial coefficients.

Exercise: • Binary trees with n internal nodes $\hat{\mathcal{B}}_n = \frac{1}{n+1}\binom{2n}{n}$

• general (plane) trees with $n+1$ internal nodes $\mathcal{G}_{n+1} = \frac{1}{n+1}\binom{2n}{n}$

• simple excursions ↑↓ of length $2n$ $\mathcal{E}_{2n} = \frac{1}{n+1}\binom{2n}{n}$

are equinumerous.

Question why?

[Find a complete set of bijections]