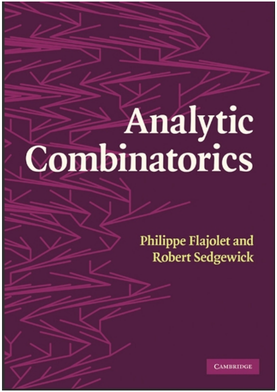
Una introducción a la combinatoria analítica

Eda Cesaratto

Universidad Nacional de Gral. Sarmiento y Conicet

SODyG,
Diciembre de 2018

Analytic Combinatorics,
Philippe Flajolet, INRIA y
Robert Sedgewick, Princeton University
Premio Leroy P. Steele 2019
otorgado por
The American Mathematical Society.



¿De qué se trata esta charla?

Un problema de conteo

Contar la cantidad de árboles binarios que tienen una cantidad prefijada de nodos internos.

Solución de Euler 1751 y Catalan 1838:

Se define una serie asociada al problema de conteo.

¿De qué se trata esta charla?

Un problema de conteo

Contar la cantidad de árboles binarios que tienen una cantidad prefijada de nodos internos.

Solución de Euler 1751 y Catalan 1838:

Se define una serie asociada al problema de conteo.

Se reescribe la serie usando propiedades combinatorias.

¿De qué se trata esta charla?

Un problema de conteo

Contar la cantidad de árboles binarios que tienen una cantidad prefijada de nodos internos.

Solución de Euler 1751 y Catalan 1838:

Se define una serie asociada al problema de conteo.

Se reescribe la serie usando propiedades combinatorias.

Se piensa a la serie como una función de variable compleja y se determinan sus singularidades.

¿De qué se trata esta charla?

Un problema de conteo

Contar la cantidad de árboles binarios que tienen una cantidad prefijada de nodos internos.

Solución de Euler 1751 y Catalan 1838:

Se define una serie asociada al problema de conteo.

Se reescribe la serie usando propiedades combinatorias.

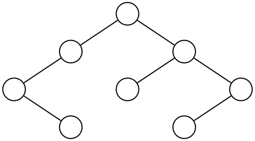
Se piensa a la serie como una función de variable compleja y se determinan sus singularidades.

La ubicación y la naturaleza de las singularidades dan respuesta al problema de conteo inicial.

Conteo de árboles binarios

Definiciones

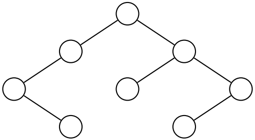
Un árbol libre es un grafo no dirigido acíclico conexo.



Conteo de árboles binarios

Definiciones

Un árbol libre es un grafo no dirigido acíclico conexo.

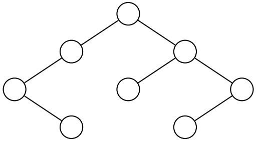


Todo par de vértices está conectado por exactamente un camino.

Conteo de árboles binarios

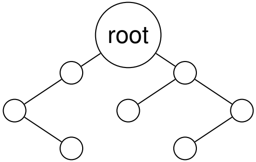
Definiciones

Un árbol libre es un grafo no dirigido acíclico conexo.



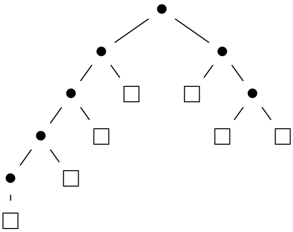
Todo par de vértices está conectado por exactamente un camino.

Un árbol libre con raíz es un árbol libre en el cual un vértice se distingue del resto. Este vértice es la raíz.
Un árbol libre con raíz planar es un árbol libre con raíz en el cual se distingue la derecha de la izquierda.



Elementos de un árbol

- Nodo: es el término usado para referirse a un vértice de un árbol con raíz.
- Si (x, y) es el último eje o arco en el camino desde la raíz hacia y, entonces x es el padre de y e y es el hijo de x.
- Un nodo sin hijos es un nodo externo u hoja
- Los nodos no hojas son nodos internos.



Árbol binario

Un árbol binario es un árbol en el cual cada nodo interno tiene a lo sumo dos hijos. Los nodos internos son aquellos que tienen dos hijos y los externos son los que no tienen ninguno.

Conteo de árboles binarios

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

Conteo de árboles binarios

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

$n = 0$

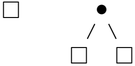
$\square$

Conteo de árboles binarios

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

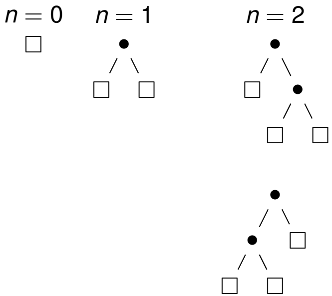
$n = 0 \quad n = 1$



Conteo de árboles binarios

$n =$ cantidad de nodos internos

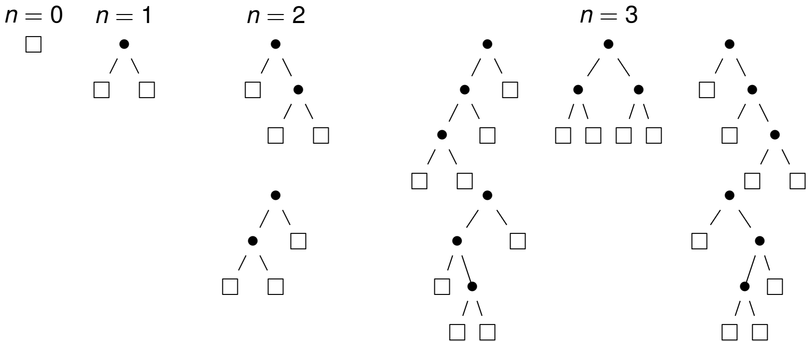
$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.



Conteo de árboles binarios

$n =$ cantidad de nodos internos

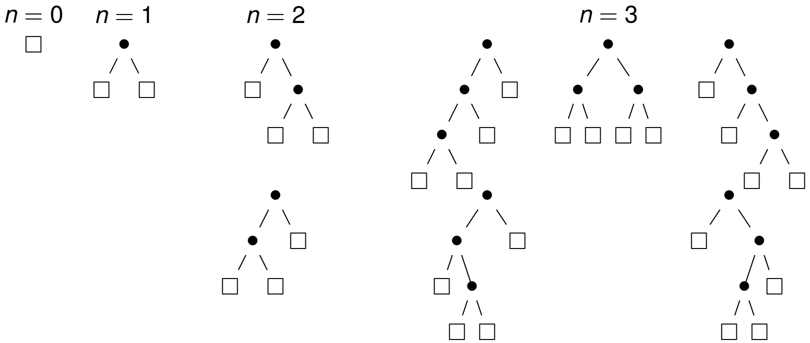
$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.



Conteo de árboles binarios

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.



$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5,$

PROBLEMA: Estimar C_n cuando n tiende a infinito.

Conteo de árboles binarios con n nodos internos

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

Conteo de árboles binarios con n nodos internos

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

Conteo de árboles binarios con n nodos internos

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5,$

Conteo de árboles binarios con n nodos internos

$n =$ cantidad de nodos internos

$C_n =$ cantidad de áboles binarios planares con raíz de n nodos internos.

$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5,$

Idea: Usar series [Euler, alrededor de 1750]

Serie generatriz o función generatriz asociada a la sucesión de conteo C_n

$$C(z) = \sum_{n=0}^{\infty} C_n z^n \,.$$

Sabemos que $C(z) = 1 + z + 2z^2 + 5z^3 + \ldots$

Combinatoria analítica

Definición

Una clase combinatoria, o simplemente una clase, es un conjunto finito o numerable en el cual está definida una función talla que satisface las siguientes condiciones:

1. la talla de un elemento es un número no negativo;
2. el número de elementos de una talla dada es finito.

Definición

Una clase combinatoria, o simplemente una clase, es un conjunto finito o numerable en el cual está definida una función talla que satisface las siguientes condiciones:

1. la talla de un elemento es un número no negativo;
2. el número de elementos de una talla dada es finito.

Si $\mathcal{A}$ es una clase, la talla de un elemento $\alpha \in \mathcal{A}$ se designa por $|\alpha|$ o $|\alpha|_{\mathcal{A}}$

Definición

Una clase combinatoria, o simplemente una clase, es un conjunto finito o numerable en el cual está definida una función talla que satisface las siguientes condiciones:

1. la talla de un elemento es un número no negativo;
2. el número de elementos de una talla dada es finito.

Si $\mathcal{A}$ es una clase, la talla de un elemento $\alpha \in \mathcal{A}$ se designa por $|\alpha|$ o $|\alpha|_{\mathcal{A}}$

Se designa por A_n a la sucesión de conteo que es el cardinal del conjunto de elementos de talla n, para cada n natural.

Combinatoria analítica

Definición

Una clase combinatoria, o simplemente una clase, es un conjunto finito o numerable en el cual está definida una función talla que satisface las siguientes condiciones:

1. la talla de un elemento es un número no negativo;
2. el número de elementos de una talla dada es finito.

Si $\mathcal{A}$ es una clase, la talla de un elemento $\alpha \in \mathcal{A}$ se designa por $|\alpha|$ o $|\alpha|_{\mathcal{A}}$

Se designa por A_n a la sucesión de conteo que es el cardinal del conjunto de elementos de talla n, para cada n natural.

Ejemplo

Tomamos como talla de un árbol binario el número de nodos.

Definición

Una clase combinatoria, o simplemente una clase, es un conjunto finito o numerable en el cual está definida una función talla que satisface las siguientes condiciones:

1. la talla de un elemento es un número no negativo;
2. el número de elementos de una talla dada es finito.

Si $\mathcal{A}$ es una clase, la talla de un elemento $\alpha \in \mathcal{A}$ se designa por $|\alpha|$ o $|\alpha|_{\mathcal{A}}$

Se designa por A_n a la sucesión de conteo que es el cardinal del conjunto de elementos de talla n, para cada n natural.

Ejemplo

Tomamos como talla de un árbol binario el número de nodos.

Definición

La serie generatriz o función generatriz asociada a la sucesión de conteo se define por

$$A(z) = \sum_{n \in \mathbb{N}} A_n z^n \; .$$

Árboles binarios

Clase combinatoria

$\mathcal{C}$ = conjunto de árboles binarios con un nodo externo "suelto".

Talla de un árbol $t \in \mathcal{C}$ es $|t|$ = la cantidad de nodos internos.

Queremos calcular o estimar la suceción de conteo C_n,

$$C_n = \#\{t \in \mathcal{C} : \; |t| = n\}$$

Árboles binarios

Clase combinatoria

$\mathcal{C}$ = conjunto de árboles binarios con un nodo externo “suelto”.

Talla de un árbol $t \in \mathcal{C}$ es $|t|$ = la cantidad de nodos internos.

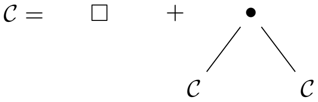
Queremos calcular o estimar la suceción de conteo C_n,

$$C_n = \#\{t \in \mathcal{C} : \; |t| = n\}$$

Serie generatriz asociada a la clase combinatoria

$$C(z) = \sum_{n \geq 0} C_n z^n$$

Propiedad: Descomposición de la clase combinatoria



$$\mathcal{C} = \{\square\} \cup \{\bullet\} \times \mathcal{C} \times \mathcal{C} \,. \tag{1}$$

Diccionario

Sean $(\mathcal{A}, |\ |_{\mathcal{A}})$ y $(\mathcal{B}, |\ |_{\mathcal{B}})$ clases combinatorias.

Sucesiones de conteo:
$A_n = \#\{x \in \mathcal{A} : \ |x|_{\mathcal{A}} = n\}, \quad B_n = \#\{y \in \mathcal{B} : \ |y|_{\mathcal{B}} = n\}$

Series generatrices

$$A(z) = \sum_{n \geq 0} A_n z^n, \quad B(z) = \sum_{n \geq 0} B_n z^n$$

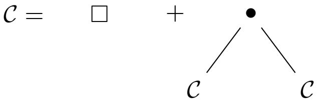
Teorema

Sean $\mathcal{A}$ y $\mathcal{B}$ clases combinatorias con tallas $|\ |_{\mathcal{A}}$ y $|\ |_{\mathcal{B}}$ respectivamente.

Operación	*Talla*	*Serie generatriz*
$\mathcal{C} = \mathcal{A} + \mathcal{B}$	$|x|_{\mathcal{C}} = \begin{cases} |x|_{\mathcal{A}} & si\ x \in \mathcal{A} \\ |x|_{\mathcal{B}} & si\ x \in \mathcal{B} \end{cases}$	$A(z) + B(z)$
$\mathcal{C} = \mathcal{A} \times \mathcal{B}$	$|(x, y)|_{\mathcal{C}} = |x|_{\mathcal{A}} + |y|_{\mathcal{B}}$	$A(z)B(z)$

Cálculo de la serie generatriz para árboles binarios

Descomposición



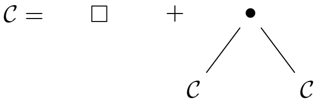
$$\mathcal{C} = \{\Box\} \cup \{\bullet\} \times \mathcal{C} \times \mathcal{C} \ .$$

La serie generatriz $C(z) = \sum_{n \geq 0} C_n z^n$ satisface la ecuación

$$C(z) = 1 + zC(z)^2.$$

Cálculo de la serie generatriz para árboles binarios

Descomposición



$$\mathcal{C} = \{\square\} \cup \{\bullet\} \times \mathcal{C} \times \mathcal{C} \; .$$

La serie generatriz $C(z) = \sum_{n \geq 0} C_n z^n$ satisface la ecuación

$$C(z) = 1 + zC(z)^2.$$

Entonces, $C(z)$ es igual a una de las dos funciones que se dan a continuación

$$\frac{1+\sqrt{1-4z}}{2z} \quad \text{o} \quad \frac{1-\sqrt{1-4z}}{2z}.$$

De la ecuación funcional resulta $B(0) = 1$ que es igual a $\lim_{z \to 0} \frac{1-\sqrt{1-4z}}{2z}$. Esta igualdad nos dice que

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

Extracción de coeficientes

Fórmula de Newton generalizada:

$$(1+z)^x = \sum_{k\geq 0} \binom{x}{k} z^k,$$

donde $\binom{x}{k}$ se entiende formalmente como una generalización del binomio de Newton para naturales y se define de la siguiente manera:

$$\binom{x}{k} = \frac{x(x-1)(x-2)\ldots(x-k+1)}{k!} \quad \text{y} \quad \binom{x}{0} = 1.$$

Aplicando esta fórmula se tiene que

$$(1-4z)^{1/2} = \sum_{n\geq 0} \binom{1/2}{n} (-4z)^n,$$

donde

$$\binom{1/2}{n} = \frac{1}{n!}(1/2)(1/2-1)\ldots(1/2-n+1) = \frac{2(-1)^{n-1}}{n4^n}\binom{2(n-1)}{n-1}$$

Finalmente, la serie generatriz $C(z)$ satisface que

$$C(z) = \frac{1-\sqrt{1-4z}}{2z} = \sum_{n\geq 1} \frac{1}{n+1}\binom{2n}{n} z^n .$$

Equivalencia asintótica

la serie generatriz $C(z)$ satisface que

$$C(z) = \frac{1-\sqrt{1-4z}}{2z} = \sum_{n\geq 1} \frac{1}{n+1}\binom{2n}{n} z^n .$$

Esto implica que

$$[z^n]C(z) = \frac{1}{n+1}\binom{2n}{n} z^n$$

La equivalencia asintótica se obtiene de la fórmula de Stirling para el factorial:

$$n! \sim (n^n/e^n)\sqrt{2\pi n},$$

lo que completa la prueba de la siguiente estimación

$$C_n \sim \frac{4^n}{n\sqrt{\pi n}}.$$

Conteo de árboles binarios

Teorema

El número de árboles binarios completos con n nodos internos está dado por los números de Catalan[1] *:*

$$C_n = [z^n]\frac{1}{2z}(1 - \sqrt{1-4z}) = \frac{1}{n+1}\binom{2n}{n}$$

Asintóticamente, para $n \to \infty$, $C_n \sim \frac{4^n}{n\sqrt{\pi n}}$.

[1] En honor al matemático belga Eugéne Catalan.

Conteo de árboles binarios

Teorema

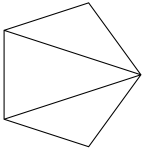
El número de árboles binarios completos con n nodos internos está dado por los números de Catalan[1] :

$$C_n = [z^n]\frac{1}{2z}(1 - \sqrt{1-4z}) = \frac{1}{n+1}\binom{2n}{n}$$

Asintóticamente, para $n \to \infty$, $C_n \sim \frac{4^n}{n\sqrt{\pi n}}$.

Una pregunta de Euler carta a Goldbach, 1751

“Recientemente encontré una pregunta que me parece destacada. Concierne al número de maneras en las que un polígono [convexo] puede ser dividido en triángulos por líneas diagonales.”

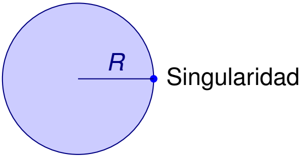


[1] En honor al matemático belga Eugéne Catalan.

Las series generatrices como objetos analíticos

Radio de convergencia R de una serie $F(z) = \sum_{n\geq 0} a_n z^n$

R es el supremo de los r tales que si $|z| < r$ entonces $F(z)$ converge absolutamente, i.e. $\sum_{n\geq 0} |a_n||z|^n < \infty$.



Criterio de Cauchy para convergencia de series:

Dada la serie $F(z) = \sum_{n\geq 0} a_n z^n$ donde a_n es una sucesión de números (complejos o reales) consideremos el límite

$$L = \limsup \sqrt[n]{|a_n|}.$$

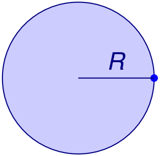
Vale que:

$$R = \frac{1}{L} \quad \text{donde} \quad L = \limsup \sqrt[n]{|a_n|}.$$

Las series generatrices como objetos analíticos

Radio de convergencia R de una serie $F(z) = \sum_{n\geq 0} a_n z^n$

R es el supremo de los r tales que si $|z| < r$ entonces $F(z)$ converge absolutamente, i.e. $\sum_{n\geq 0} |a_n||z|^n < \infty$.



Criterio de Cauchy para convergencia de series:

Dada la serie $F(z) = \sum_{n\geq 0} a_n z^n$ donde a_n es una sucesión de números (complejos o reales) consideremos el límite

$$L = \limsup \sqrt[n]{|a_n|}.$$

Vale que:

$$R = \frac{1}{L} \quad \text{donde} \quad L = \limsup \sqrt[n]{|a_n|}.$$

Heurística: $|a_n| \cong L^n$

Las series generatrices como objetos analíticos

La función $C(z) = \frac{1-\sqrt{1-4z}}{2z}$ tiene una singularidad en $z = 1/4$ y

$$C_n \simeq 4^n$$

Consecuencia de la definición de radio de convergencia de una serie y del criterio de Cauchy.

Teorema de transferencia [Flajolet y Odlyzco, 1982 hasta 1990]

Teorema

Si, en un entorno de $z = 1/\rho$,

$$f(z) \sim \frac{1}{(1-(\rho z))^{\alpha}} \left(\frac{1}{\rho z} \ln^{\beta} \left(\frac{1}{(1-\rho z)} \right) \right)$$

con $n \notin \{-1, -2, \dots\}$, *entonces,*

$$[z^n] f(z) \sim \rho^n \frac{n^{\alpha - 1}}{\Gamma(\alpha)} \ln^{\beta} n \, .$$

Γ *es la función gama que generaliza el factorial.*

Teorema de transferencia [Flajolet y Odlyzco, 1982 hasta 1990]

Teorema

Si, en un entorno de $z = 1/\rho$*,*

$$f(z) \sim \frac{1}{(1-(\rho z))^{\alpha}} \left(\frac{1}{\rho z} \ln^{\beta}\left(\frac{1}{(1-\rho z)}\right)\right)$$

con $n \notin \{-1, -2, \dots\}$*, entonces,*

$$[z^n]f(z) \sim \rho^n \frac{n^{\alpha-1}}{\Gamma(\alpha)} \ln^{\beta} n \,.$$

Γ *es la función gama que generaliza el factorial.*

Aplicación al conteo de árboles binarios

Serie generatriz

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

Tenemos $\rho = 4, \quad \alpha = -1/2, \quad \beta = 0$

Resulta

$$C_n = [z^n]\frac{1}{2z}(1-\sqrt{1-4z}) \sim 4^n \frac{n^{-1/2-1}}{\Gamma(-1/2)} = \frac{4^n}{n\sqrt{\pi n}}$$

pues $\Gamma(-1/2) = -2\sqrt{\pi}$.

Aplicaciones a grafos

Conteo

¿Cuántos grafos hay con n vértices que son

- Acíclicos y conexos?
- Acíclicos?
- Unicíclicos?
- Conexos y con un exceso k?

Aplicaciones a grafos

Conteo

¿Cuántos grafos hay con n vértices que son

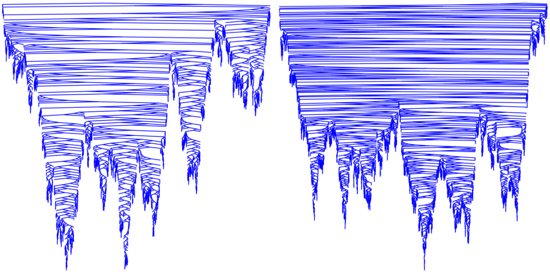
- Acíclicos y conexos?
- Acíclicos?
- Unicíclicos?
- Conexos y con un exceso k?

Respuestas: página 134 del libro de Flajolet y Sedgewick.

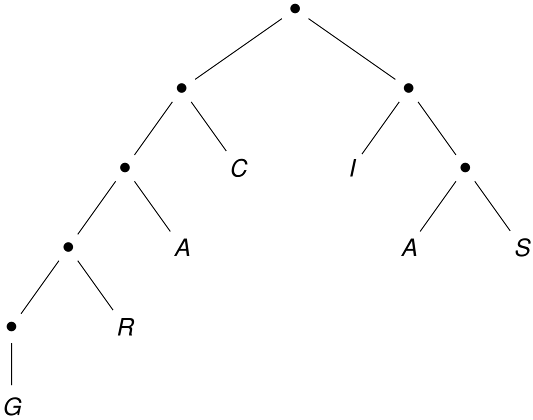
Generación aleatoria de estructuras: método de Boltzman

Boltzmann Samplers for the Random Generation of Combinatorial Structures, Duchon, Flajolet, Louchard, Schaeffer, 2004

Árboles binarios de talla 4469 y 8255



Gentileza de Carine Pivoteau.


C
I
A
A
S
R
G