CUBO, Vol. 3. Octubre 1987
Págs. 37-55

APLICACIONES DE FUNCIONES GENERATRICES Y LENGUAJES FORMALES A PROBLEMAS DE COMBINATORIA Y ESTRUCTURA DE DATOS *

Luis F. Hevia R.**

RESUMEN:

Se presentan propiedades de las funciones generatrices ordinarias y de los lenguajes formales, interrelacionándolos de forma que, dado un problema combinatorial, (los cuales son frecuentes en estructura de datos), si se logra representar sus objetos por un lenguaje y si este lenguaje puede ser generado por una gramática, entonces es más fácil obtener la función generatriz para la familia de objetos que se estudian.

** Departamento de Matemática, Universidad Técnica Federico Santa María, Valparaíso.

* Proyecto 861203, financiado por la Dirección General de Investigación y Postgrado de la Univ. Técnica Federico Santa María.

1.- FUNCIONES GENERATRICES ORDINARIAS.

Se usará el término configuración para denotar cualquier objeto matemático, por ejemplo, una permutación, un árbol, una función, etc. Si S es un conjunto de configuraciones, entonces la función

$$\omega : S \longrightarrow \{0,1,2,\dots\} = \mathbb{N}$$
$$\sigma \longmapsto \omega(\sigma)$$

es llamada función peso en S y para cada $\sigma \in S$, $\omega(\sigma)$ se llama el peso de σ.

1.1.- DEFINICION:

La función generatriz de un conjunto S de configuraciones distintas, ordinaria con respecto a la función peso ω en S, es

$$\Phi_S^{\omega}(x) := \sum_{\sigma \in S} x^{\omega(\sigma)}$$

x está indeterminada y se dice que marca a ω, y que $\Phi_S^{\omega}(x)$ enumera a S con respecto a ω.

1.2.- DEFINICION:

Sean S un conjunto de configuraciones, E una relación de equivalencia en S, y σ_1, σ_2 elementos de S, entonces se dice que σ_1 es indistinguible de σ_2 ssi $\sigma_1 \mathrel{E} \sigma_2$ y que σ_1 es distinguible de σ_2 ssi $[\sigma_1] \neq [\sigma_2]$.

Sea S un conjunto de distintas configuraciones y una función peso en S. Entonces el problema enumerativo general que nos interesa es encontrar el número C_n definido por:

$C_n = \# \{\sigma \in S : \quad (\;) = n \}$.

En otras palabras, encontrar el número de distintas configuraciones de peso n. Esta función peso, puede ser el largo de una secuencia, el número de árboles binarios, el número de secuencias que empiezan con el 1, etc.

1.3.- DEFINICION:

Sea S, un conjunto de configuraciones y ω una función peso en S, se llamará s-objetos a los elementos de S que pertenezcan a alguna configuración de interés.

En la solución de cualquier problema, la primera tarea es identificar el conjunto S, la función peso ω , y la relación de equivalencia E asociada con la indistinguibilidad. Normalmente se tratarán situaciones donde E es la identidad.

1.4.- EJEMPLO:

Encontrar el número de secuencias de largo 2 que estén formadas por los símbolos 0,1 y 2.

Solución:

Aquí, S será el conjunto de todas las secuencias de cualquier longitud, formadas con los símbolos 0,1 y 2.

La subconfiguración de interés será la que corresponde a la secuencias de largo 2. Así $S = \{ 0,1,2 \}^*$, con

$$\{0,1,2\}^* = \phi \,\dot{\cup}\, \{0,1,2\} \,\dot{\cup}\, \{0,1,2\}^2 \,\dot{\cup}\, \{0,1,2\}^3 \,\dot{\cup}\, \ldots$$
$$= \{0,1,2\} \,\dot{\cup}\, \{0,1,2\} \times \{0,1,2\} \,\dot{\cup}\, \{0,1,2\} \times \{0,1,2\} \times$$
$$\times \{0,1,2\} \,\dot{\cup}\, \ldots$$

En este caso los s-objetos son: 00, 01, 02, 10, 11, 12, 20, 21, 22,

E es la relación identidad sobre S, y la función peso es $\omega(\sigma) = (n_1 n_2 \ldots n_k) = K$. Por lo tanto, la respuesta es $C_2 = 9$.

La función generatriz permitirá contar cuantos elementos existen sin contarlos uno por uno.

1.5.- LEMAS DE CONTEO ELEMENTALES.

Si F y G son conjuntos de configuraciones distintas con f y g funciones generatrices ordinarias respectivas, entonces

1) La unión disjunta $F \,\dot{\cup}\, G$

2) El producto cartesiano $F \times G$

3) La composición $F \circ G$

tendrán como funciones generatrices a $f + g$, $f \cdot g$, $f \circ g$ respectivamente.

1.6.- LEMA DE LA SUMA.

Sea ω la función de peso en el conjunto S de configuraciones distintas. Sean A, B subconjuntos disjuntos de S. Entonces

$$\Phi^{\omega}_{A\dot{\cup}B}(x) = \Phi^{\omega}_{A}(x) + \Phi^{\omega}_{B}(x) .$$

Demostración:

$$\Phi^{\omega}_{A\dot{\cup}B}(x) = \sum_{\sigma\in A\dot{\cup}B} x^{\omega(\sigma)}$$

$$= \sum_{\sigma\in A} x^{\omega(\sigma)} + \sum_{\sigma\in B} x^{\omega(\sigma)}$$

$$= \Phi^{\omega}_{A}(x) + \Phi^{\omega}_{B}(x)$$, ya que la unión es disjunta. ∎

1.7.- LEMA DEL PRODUCTO.

Sean α, β y ω las funciones pesos en los conjuntos A, B, A x B de configuraciones distintas. Si $\omega((a,b)) = \alpha(a) + \beta(b)$, para todo $(a,b) \in A \times B$, entonces

$$\Phi^{\omega}_{A\times B}(x) = \Phi^{\alpha}_{A}(x) \quad \Phi^{\beta}_{B}(x)$$

Demostración:

De la definición 1 se tiene que:

$$\Phi^{\omega}_{A\times B}(x) = \sum_{\sigma\in A\times B} x^{\omega(\sigma)} = \sum_{(a,b)\in A\times B} x^{\omega((a,b))}$$

$$= \sum_{(a,b)\in A\times B} x^{\alpha(a)+\beta(b)}$$

Dejando a fijo, recorriendo todos los elementos de b de B y repitiendo el proceso para todos los elementos a de A se tiene que:

$$\Phi^{\omega}_{A\times B}(x) = \sum_{a\in A} x^{\alpha(a)} \cdot \sum_{b\in B} x^{\beta(b)}$$

$$= \Phi^{\alpha}_{A}(x) \cdot \Phi^{\beta}_{B}(x)$$

1.8.- COROLARIO:

Sea $S = A^{*} = A^{0} \dot{\cup} A^{1} \dot{\cup} A^{2} \dot{\cup} \ldots$, y supóngase que ω cumple con la propiedad aditiva de 1.7, entonces

$$\Phi^{\omega}_{S}(x) = [1 - \Phi^{\omega}_{A}(x)]^{-1}$$

Demostración:

$$\Phi^{\omega}_{s}(x) = \Phi^{\omega}_{A^{0}\dot{\cup}A^{1}\dot{\cup}A^{2}\dot{\cup}\ldots}(x)$$

$$= \sum_{k\geq 0} \Phi^{\omega}_{A^{k}}(x) \quad \text{(lema de la suma)}$$

$$= \sum_{k\geq 0} (\Phi^{\omega}_{A}(x))^{k} \quad \text{(lema del producto)}$$

$$= 1 + \Phi^{\omega}_{A}(x) + (\Phi^{\omega}_{A}(x))^{2} + \ldots$$

$$= [1 - \Phi^{\omega}_{A}(x)]^{-1} \quad ■$$

1.9.- DEFINICION. (Descomposición del Conjunto).

Para encontrar la función generatriz que describe un problema se pueden utilizar básicamente 3 métodos:

Sea $A \tilde{\rightarrow} f(B_0, B_1, \ldots, B_2)$

i) La descomposición es directa para A si $B_i \neq A$ $\forall i \in [0,\ldots,n] = \{0,1,2,\ldots,n\}$.

ii) La descomposición es indirecta para todo B_i con $i \in [0,\ldots,n]$ si $B_i \neq A$.

iii) La descomposición es recursiva para A si existe $B_i = A$, tal que, $i \in [0,\ldots,n]$.

1.10.- EJEMPLO:

Encontrar el número de secuencias en el alfabeto $\{0,1,2\}$ que comienzan con 2.

a) Sea A el conjunto que describe el evento

$A \tilde{\rightarrow} \{2\} \times \{0,1,2\}^*$

Claramente la descomposición es directa para A

$$\Phi_A(x) = \Phi_{\{2\}}(x)\ \Phi_{\{0,1,2\}^*}(x)$$

donde $\omega(n_1,\ldots,n_k) = K$

Sea $S = A$

$$\Phi_S(x) = x^{\omega(2)}(1 + x^{\omega(0)} + x^{\omega(1)} + x^{\omega(2)} + \dots)$$

$$= x(1 + 3x + 9x^2 + 27x^3 + \dots\dots\dots)$$

$$= x(1 - 3x)^{-1}$$

$$= \sum_{n \geq 1} 3^{n-1} x^n$$

b) Sea $A = \{0,1,2\}^*$

$$B_1 = \{1\} \times \{0,1,2\}^*$$

$$B_2 = \{2\} \times \{0,1,2\}^*$$

$$B_3 = \{3\} \times \{0,1,2\}^*$$

$$B_0 = \{\ \}$$

$$A \xrightarrow{\sim} B_0 \dot{\cup} B_1 \dot{\cup} B_2 \dot{\cup} B_3$$

La descomposición es indirecta para B_0, B_1, B_2, B_3

$$\Phi_A(x) = 1 + \Phi_{B_1}(x) + \Phi_{B_2}(x) + \Phi_{B_3}(x)$$

Por la simetría del problema,

$$\Phi_{B_1}(x) = \Phi_{B_2}(x) = \Phi_{B_3}(x) = \Phi_S(x)$$

Por lo tanto

$$\Phi_{\{0,1,2\}^*}(x) = 1 + 3\,\Phi_S(x)$$

$$(1-3x)^{-1} = 1 + 3\,\Phi_S(x)$$

$$\Phi_S(x) = \frac{1}{3}\,[\,(1-3x)^{-1} - 1\,]$$

$$= x(1-3x)^{-1}$$

c) Sea $A = \{0,1,2\}^*$

$B_0 = \{\ \}$

$B_1 = \{0\}$

$B_2 = \{1\}$

$B_3 = \{2\}$

$B_4 = \{0,1,2,\}^*$

$$A \overset{\sim}{\to} B_0 \,\dot{\cup}\, B_1 \times B_4 \,\dot{\cup}\, B_2 \times B_4 \,\dot{\cup}\, B_3 \times B_4$$

La descomposición es recursiva para A ($A = B_4$)

Como $\Phi_A(x) = \Phi_{B_4}(x) = (1-3x)^{-1}$

$\Phi_{B_1}(x) = \Phi_{B_2}(x) = \Phi_{B_3}(x)$

Sea $S = B_1 \times B_4$

$$(1-3x)^{-1} = 1 + 3\,\Phi_S(x)$$

$$\Phi_S(x) = x(1-3x)^{-1}$$

1.11.- DEFINICION:

Sean S y T conjuntos de configuraciones combinatoriales.

a) Si existe una biyección $\Omega : S \to T$, entonces de dice que Ω es una descomposición de S y se anota

$$S \underset{\omega}{\tilde{\to}} T.$$

b) Si existe una función peso ω' en T tal que $\omega(\sigma) = \omega'(\Omega(\sigma))$ para todo $\sigma \in S$, entonces se dice que Ω es una descomposición de S que preserva el peso ω y se anota $S \tilde{\to} T$.

1.12.- PROPOSICION:

Si $S \underset{\omega}{\tilde{\to}} T$, donde S y T son conjuntos de configuraciones combinatoriales, y ω es una función peso para S, entonces existe una función peso ω' en T tal que

$$\Phi_S^{\omega}(x) = \Phi_T^{\omega'}(x).$$

Demostración:

Resulta inmediato de las definiciones 1.1 y 1.12.

1.13.- EJEMPLO:

¿Cuántos subconjuntos de k elementos tiene un conjunto de n elementos?.

Sea $A = \{ a_1, a_2, \ldots, a_n \}$

$S = P(A)$

$C_{n,k} = \# \{ \sigma \in S \ / \ \omega(\sigma) = k \}$

$\omega(\sigma)$ es el número de elementos de G

Sea $S \overset{\sim}{\to} \{0,1\}^n = T$

$\sigma \to (\alpha_1, \alpha_2, \ldots, \alpha_n)$

$$\text{donde } \alpha_j = \begin{cases} 1 & \text{si } a_j \in \sigma \\ 0 & \text{si } a_j \notin \sigma \end{cases}$$

Veamos un caso:

Si $A = \{ a_1, a_2, a_3 \}$

$S = \{\Phi, \{a_1\}, \{a_2\}, \{a_3\}, \{a_1, a_2\}, \ldots, \{a_1, a_2, a_3\}\}$

$\{0,1\}^3 = \{000, 100, 010, 001, 110, \ldots, 111\}$

Luego $\Phi_S^{\omega}(x) = \Phi^{\omega'}_{\{0,1\}^n}(x)$

donde $\omega'(\alpha_1, \ldots, \alpha_n) = \sum_{i=1}^{n} \alpha_i$

$\Phi_S^{\omega}(x) = [\Phi^{\omega'}_{\{0,1\}}(x)]^n$ por lema del producto

$= [x^{\omega'(0)} + \omega^{\omega'(1)}]^n = (1+x)^n$

$= \sum_{k=0}^{n} \binom{n}{k} x^k$

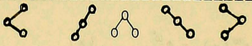
Así el número de subconjuntos de k elementos de un conjunto de n elementos es:

$$[x^k]\ \Phi^{\omega}_{S}(x) = \binom{n}{k}$$

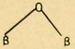
1.4.- EJEMPLO:

Determinar el número de árboles binarios con n nodos internos.

Veamos una tabla con los primeros árboles binarios

Si n	Nº árboles	árboles
0	1	□ (nodo externo)
1	1	0
2	2	
3	5	

Sea β el conjunto de los árboles binarios ω (β) cuenta el nº de nodos internos de β

$\beta \overset{\sim}{\to} \{\square\} \ \dot{\cup}\ \{$ 

$\}$ descomposición recursiva

(forma pictórica de la relación)

$$\beta \overset{\sim}{\to} \{\square\} \ \dot{\cup}\ \{\,0\,\} \times (\beta \times \beta)$$

Así $\Phi^{\omega}_{\beta}(x) = \Phi^{\omega}_{\{\square\}}(x) + \Phi^{\omega}_{\{0\}}(x)\,[\Phi^{\omega}_{\beta}(x)]^{2}$

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

Como $\lim\limits_{x \to 0} \Phi_{\beta}(x) = 1$ seleccionamos el signo-, y desarrollando en serie la expresión resultante es:

$$\Phi_{\beta}(x) = \sum_{n \geq 0} \frac{1}{n+1} \binom{2n}{n} x^{n}$$

De donde $C_n = \frac{1}{n+1}\binom{2n}{n}$, el conocido número de Catalán.

Observación:

En virtud de la proposición 1.12 se puede construir una biyección entre el problema de los árboles binarios y otros problemas que tienen como respuesta el número de Catalán.

2.- GRAMATICAS.-

2.1- DEFINICIONES BASICAS.

Sea Σ un conjunto de símbolos, a los cuales llamaremos terminales.

Un lenguaje será un conjunto finito o infinito de palabras sobre Σ al cual llamaremos lenguaje formal (en contraposición con los lenguajes naturales de comunicación humana).

Una gramática es la tetrada (N, Σ, P, S) donde

N es un conjunto de símbolos no terminales

Σ es un conjunto de símbolos terminales

S es el símbolo de partida

P es un conjunto de producciones de la forma

$$\alpha \to \beta \quad \alpha \in V^* \; N \; V^*$$

$$\beta \in V^* \text{ con } V = \Sigma \cup N$$

La gramática es de libre contexto (GLC) si $\forall P, \alpha \to \beta$ donde $\alpha \in N$ $\beta \in V^+$ ($V^+ = V^* - \{\varepsilon\}$)

ε palabra vacía.

La gramática genera un lenguaje formado por las derivaciones (Transitivas) del símbolo inicial, según las producciones, que se definen como:

$$L(G) = \{s \in \Sigma^* \;/\; S \overset{*}{\Rightarrow} s\}$$

Aplicando las producciones al símbolo inicial, se obtiene un árbol de derivación, y si para una palabra existe más de un árbol de derivación diremos que la gramática de libre contexto es ambiqua.

2.2.- EJEMPLO:

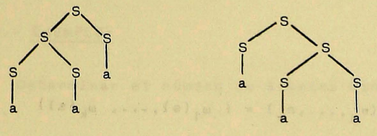
Sea $N = \{S\}$

$\Sigma = \{a\}$, $V = \{S, a\}$

$P = \{S \to SS, S \to a\}$

$= \{S \to SS \mid a\}$

(N, Σ, P, S) es una GLC ambigua, ya que veamos dos ărboles de derivación diferente para a^3.



2.3.- EL ENFOQUE DE LOS LENGUAJES FORMALES.

La conexión entre lenguajes formales y problemas de enumeración surgió por primera vez en los trabajos de Schutzenberger y Chomsky [1]. Ellos encontraron que existían series de poten - cias útiles para la clasificación de Lenguajes Formales. Gross [2] encontró propiedades de dichas series para interpretar los coeficientes como todas aquellas palabras que tienen una misma composición de terminales. En consecuencia, sí se representa los objetos mediante un lenguaje y si este lenguaje puede ser generado por una gramática, entonces es más fácil obtener la función generatriz para la familia de objetos.

Flajolet [3] muestra que es posible estudiar problemas com putacionales evitando recurrencias y sumas a través de derivaciones generatrices combinatoriales y exponenciales.

Interesará utilizar los lenguajes formales para el análi - sis de algoritmos, tal que, a partir de una descripción de la gramática en términos de la forma normal de Chomsky, ellos se implementan para contar y enumerar objetos combinatoriales tí - picos en estructura de datos.

2.4.- GRAMATICAS Y FUNCIONES GENERATRICES.

Sea $G = (N, \Sigma, P, S, \omega)$ una G.L.C. con una función tamaño ω, donde

$$\omega : L(G) \to N^k$$

$$S \to \omega(s) = (n_1, \ldots, n_k) = (\omega_1(s), \ldots, \omega_k(s))$$

ω es una función peso o de tamaño que permite particionar $L = L(G)$

$$L = \bigcup_{(n_1, \ldots, n_k) \in \mathbb{N}^k} L_{n_1, \ldots, n_k}$$

donde $L_{n_1, \ldots, n_k} = \{s \in L / \omega(s) = (n_1, \ldots, n_k)\}$

Siendo las clases disjuntas escribimos

$$L = \sum_{n_1, \ldots, n_k} L_{n_1, \ldots, n_k}$$

Nos interesa determinar $\ell_{n_1, \ldots, n_k} = \# L_{n_1, \ldots, n_k}$

A la gramática G definida en este punto, se le asociará un sistema de ecuaciones de series de potencias, a partir de las producciones, reemplazando

$\to$ por $=$

ε por 1

no terminales A por A(x)

los terminales no se modifican

no terminales A por A(x)

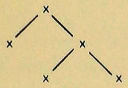
los terminales no se modifican

2.5.- EJEMPLO:

Determinar el número de árboles binarios con n nodos internos.

Los objetos serán codificados mediante secuencias en el alfabeto {x,b,s}* según un recorrido en preorden

Así se tiene: (b: baja, s: sube)



xbxsbxbxsbxss

La gramática que describe el evento es

$B \rightarrow xbBsbBs \;/\; \varepsilon$

cuya ecuación es: $B(x, \varepsilon, b, s) = xb^2s^2B^2 + 1$

(no nos interesa el número de bajadas y subidas

$b = s = 1$)

Luego $B(x) = xB^2 + 1$, que como veíamos en el ejemplo 1.14, conduce al número de Catalán.

2.6.- EJEMPLO:

Determinar el número de k elementos de un conjunto de n elementos.

Por ejemplo 1.13, sabíamos que $S = P(A)$

$\Sigma = \{0,1\}$

$N = \{S\}$

$S \rightarrow 0S|1S|\varepsilon$

Contamos con x el número de elementos del conjunto A, con y el número de elementos del subconjunto de A

$S(x,y) = xS + xyS + 1$

Desarrollando

$$S(x,y) = (1-(x+xy))^{-1}$$

$$= \sum_{n\geq 0} x^n (1+y)^n$$

$$= \sum_{n\geq 0} x^n \left(\sum_{k=0}^{n} \binom{n}{k} y^k \right)$$

$$= \sum_{n,k} \binom{n}{k} y^k x^n$$

3.- BIBLIOGRAFIA

[1] N. CHOMSKY : MP Schutzenberger. "The Algebraic Theory of Contact Free Languaje" Studies in Logic and Foundations of Ma - thematics 118-159. North-Holland, 1963.

[2] M. GROSS : "Aplications Geométriques des Languages Formels.. International Computation Cen- tre Bulletin 5, 141-167, 1966.

[3] P. FLAJOTEL, J. FRANCON, L. VUILLEMIN. : "Sequence of Ope- rations Analysis for Dynamic Data Struc - tures". Journal of Algorithms, 1 : 111 - 141, 1980.