

TAMAÑO MEDIO DE ARBOLES SIMPLIFICADOS

M.I. FERNANDEZ CAMACHO

Universidad Complutense de Madrid

Estudiamos la media y varianza del tamaño de los árboles simplificados obtenidos a partir de árboles p -arios y generales. Se demuestra que ambos parámetros son lineales en el tamaño del input. Pero lo que creemos más interesante del artículo es la metodología seguida, que pensamos podría aplicarse al análisis de gran parte de algoritmos de simplificación sobre árboles. Tal técnica involucra un estudio de ciertas series de potencias formales, y la aplicación del teorema de Darboux-Polya para aproximaciones asintóticas, así como otros resultados del análisis complejo.

Average size of simplified trees

Keywords: Complexity Algorithms, Generating Functions, Complex Analysis.

1. INTRODUCCION

En los últimos años, la constatación de que la complejidad en el peor de los casos no refleja adecuadamente las observaciones experimentales, ha provocado un gran auge en el estudio de la complejidad en media de algoritmos. Con frecuencia, tal análisis conduce a un estudio combinatorio de ciertos parámetros relacionados con el algoritmo, y a la utilización de métodos clásicos del análisis real, lo que queda ampliamente ilustrado en [4] Kn 68.69.73. También resulta especialmente útil la aplicación del análisis complejo para la obtención de estimadores asintóticos de parámetros. En [2] F1 84 y [7] St 84 puede encontrarse un estudio sistemático de tales herramientas.

—M.I. Fernández Camacho - Universidad Complutense de Madrid - Facultad de C. Matemáticas - Dpto. Informática y Automática - 28040 Madrid.

—Article rebut el gener de 1987.

En el presente artículo estudiamos un sistema de reescritura cuyo output devuelve una única copia de un subárbol cuando todos los subárboles son idénticos. El input son árboles p -arios (apartado 2) o árboles generales (apartado 3) etiquetados en las hojas. Los parámetros a estudiar son pues definidos de manera recursiva según una filosofía bottom-up (de las hojas a la raíz), que va modificando los subárboles del árbol original según el resultado del test de igualdad. Dichos parámetros (media y varianza del tamaño del output) resultan ser lineales en el tamaño del input.

El problema es una generalización del estudiado en [1] CS 86 para el caso binario y la metodología seguida es la propuesta en dicho trabajo: traducir a un conjunto infinito de ecuaciones sobre series de potencias los esquemas de definición recursivos, y a partir de las mismas obtener una ecuación que nos permita resolver el problema, ayudándonos del teorema de Darboux.

En un determinado punto del proceso se introducen una sucesión de funciones generatrices que se definen de forma recursiva mediante un sistema de ecuaciones funcionales. El estudio de la analiticidad de tales funciones a partir de dicho sistema se complica bastante con respecto al caso binario, dada la mayor complejidad del sistema de ecuaciones. El método utilizado en el presente artículo se basa en el estudio de determinadas funciones sobre el eje real sugeridas por la estructura del sistema.

Pero el resultado que nos parece más interesante de este artículo es el recogido en la proposición 2.3.: Dadas una sucesión $\{f_i(z)\}$ de series de potencias con coeficientes positivos e igual radio de convergencia; una sucesión $\{\alpha_i\}$ de números reales positivos y dado un número natural $p > 1$, la serie $\sum \alpha_i f_i^p(z)$ tiene un radio de convergencia estrictamente superior al de $\sum \alpha_i f_i(z)$, siempre y cuando se verifiquen determinadas condiciones técnicas que pueden expresarse informalmente de la siguiente forma: i) los α_i crecen lentamente; ii) el radio de convergencia $\sum \alpha_i f_i(z)$ es estrictamente inferior al de las $f_i(z)$ y iii) las $f_i(z)$ están acotadas superiormente, en norma, de manera exponencial decreciente en i , con exponentes que asintóticamente crecen según una determinada tasa.

Este problema de convergencia surge en muchas situaciones a la hora de encontrar una expresión, para una función generatriz, a la cual podamos aplicar el método de Darboux-Polya. Además las condiciones técnicas anteriores se dan en todos los problemas de simplificación estudiados hasta el momento. Todo ello nos hace pensar que puede ser una herramienta bastante general para el estudio de algoritmos que operan recursivamente sobre árboles modificando su estructura según una filosofía bottom-up.

Por todo esto el presente trabajo nos parece un nuevo paso hacia un método general para el estudio de una amplia clase de algoritmos de manipulación de árboles que aparecen en inteligencia artificial, cálculo formal, etc...

Más en concreto, el criterio de simplificación que hemos seguido se basa en la idempotencia de los operadores (ocultos) sobre los nodos internos de los árboles estudiados. Así podríamos concretar nuestros resultados para la simplificación de términos del cálculo proposicional o equivalentemente del álgebra conjuntista (construidos con unión e intersección).

Otros criterios de simplificación, probablemente más complejos, nos permitirán estudiar, siguiendo presumiblemente la misma técnica, algoritmos más potentes de simplificación; en particular nos parece interesante el estudio de la complejidad media del algoritmo de Knuth–Bendix, partiendo de conjuntos de axiomas completos (véase [5] KB 70), para diversas teorías concretas sencillas (teorías de grupos, etc...)

En el último apartado del presente artículo se ofrecen tablas con aproximaciones de los coeficientes teóricos de proporcionalidad, respecto al tamaño del input, de los parámetros estudiados, para diversas familias de árboles; así como de las constantes de proporcionalidad obtenidas a partir de simulaciones, pudiéndose observar la semejanza de ambas tablas.

2. ARBOLES P-ARIOS

2.1. EL ALGORITMO DE SIMPLIFICACIÓN

Consideremos la familia \mathcal{Q} de los árboles p-arios con q posibles etiquetas para sus hojas ($p, q \geq 2$). El algoritmo de simplificación con el que trabajaremos transforma dichos árboles reemplazando cada subárbol con p subárboles radicales idénticos por una única copia de uno de ellos, procediendo según una filosofía de tipo bottom-up (de las hojas a la raíz). A continuación escribimos este procedimiento en pseudo-PASCAL, señalando previamente que en general, para cualquier familia de árboles, dado un árbol T de la familia,

$T.degree$ denota la aridad de la raíz de T (número de sus hijos).

$T.son(i)$ para $1 \leq i \leq T.degree$, es el i -ésimo hijo de la raíz de T (i -ésimo subárbol radical de T).

$|T|$ denota el tamaño de T , que según la familia de árboles que se maneje se definirá como el número de nodos internos, o como el número total de nodos (internos o no) del árbol.

Function SIMP ($T : \mathcal{Q}$) : \mathcal{Q} ;

if $T.degree = 0$ then SIMP := T

else for $i := 1$ to p do $t(i) := \text{SIMP}(T. \text{son}(i))$ od

$i := 1$; CON := true;

While($i \leq p - 1$) and CON do

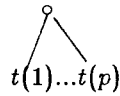
if EQUAL ($t(i), t(i + 1)$)

then $i := i + 1$

else CON := false od

if CON then SIMP := $t(1)$

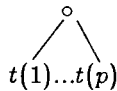
else SIMP :=



(0)

end SIMP.

La notación



representa un árbol formado por un nodo raíz que

tiene por hijos el conjunto de subárboles $\{t(1), \dots, t(p)\}$.

En particular todo árbol obtenido a partir de un árbol p -ario por sustitución de todas sus hojas por el mismo árbol t de \mathcal{Q} se simplifica al simplificado de t .

Nuestro objetivo aquí es el estudio del tamaño medio de un árbol simplificado con respecto a su tamaño inicial, entendiendo como tamaño de un árbol el número de sus nodos internos.

Antes de pasar a estudiar la función generatriz del tamaño de los árboles simplificados consideremos la ecuación que define la familia \mathcal{Q} :

$$\mathcal{Q} = \sum_{i=1}^q e_i + \circ(\underbrace{\mathcal{Q}, \dots, \mathcal{Q}}_p) \text{ donde } \{e_1, \dots, e_q\} \text{ es el conjunto de posibles eti-}$$

quetas para las hojas de los árboles de la familia, y la notación $\circ(\underbrace{\mathcal{Q}, \dots, \mathcal{Q}}_p)$

representa el conjunto de los árboles formados por un nodo raíz y p subárboles radicales pertenecientes a \mathcal{Q} , es decir el conjunto de los árboles de \mathcal{Q} que no son hojas. Entonces la función generatriz $Q(z) = \sum q_n z^n$ asociada a \mathcal{Q} , donde q_n es el número de árboles de tamaño n en \mathcal{Q} , verifica la ecuación:

$$(1) \quad Q(z) = q + zQ^p(z)$$

A partir del teorema de la función implícita y del teorema de inversión de Lagrange se obtiene que el radio de convergencia de $Q(z)$ es $\rho = \frac{(p-1)^{p-1}}{p^p q^{p-1}}$, que $Q(z)$ sólo admite una singularidad, en $z = \rho$, sobre su círculo de convergencia y que en un entorno de $z = \rho$, $Q(z)$ admite el siguiente desarrollo:

$$(2) \quad Q(z) = (1 - z/\rho)^{1/2} g(z) + h(z)$$

siendo $g(z)$ y $h(z)$ analíticas en un entorno de $z = \rho$; con

$$g(\rho) = \frac{-q}{p-1} \sqrt{\frac{2p}{p-1}} \quad \text{y} \quad h(\rho) = \frac{pq}{p-1}$$

Con lo cual podemos concluir que

$$(3) \quad q_n = \frac{g(\rho)}{-2\sqrt{\pi}} \rho^{-n} n^{-3/2} (1 + o(1/n))$$

puesto que estamos en condiciones de aplicar el teorema de Darboux–Polya, que enunciaremos tras indicar que dada la serie $f(z) = \sum_{n \geq 0} f_n z^n$, su n -ésimo coeficiente será denotado por $[z^n]f(z) = f_n$.

Teorema (Darboux–Polya) [3] He 74:

Sea $f(z)$ una función analítica en un entorno del origen de radio ρ , con una única singularidad sobre su círculo de convergencia en $z = \rho$.

Supongamos además que $f(z)$ admite en un entorno de $z = \rho$ el desarrollo siguiente:

$$f(z) = g(z)(1 - z/\rho)^{-s} + h(z)$$

donde

- 1) $s \in \mathbf{R} - \{0, -1, -2, \dots\}$
- 2) g y h son analíticas en un entorno de $z = \rho$
- 3) $g(\rho) \neq 0$

Sea $g(z) = \sum_{i=0}^{\infty} g_i (z - \rho)^i$ el desarrollo de Taylor de $g(z)$ en $z = \rho$.

Entonces el n -ésimo coeficiente de Taylor de $f(z)$ en el origen satisface:

$$[z^n]f(z) \simeq \rho^{-n} \sum_{i=0}^{\infty} (-1)^i \rho^i g_i (-1)^n \binom{-s+i}{n}$$

cuando $n \rightarrow \infty$

Normalmente es suficiente la aproximación por el primer término de la serie, lo que permite reescribir la conclusión del teorema de la siguiente forma:

$$[z^n]f(z) = \rho^{-n} g(\rho) n^{s-1} \frac{1}{\Gamma(s)} (1 + O(1/n))$$

donde $\Gamma(x)$ es la función gamma de Euler.

No obstante en el cálculo de las varianzas para las variables estudiadas en este artículo, esta última aproximación no es suficiente y es necesario considerar los dos primeros términos de la serie.

2.2. FUNCIONES GENERATRICES PARA EL TAMAÑO DE LOS ÁRBOLES SIMPLIFICADOS

Sea \mathcal{S} la clase de todos los árboles de \mathcal{Q} irreducibles, es decir que no pueden ser simplificados por el algoritmo anterior. Denotamos por $S(z)$ la función generatriz del tamaño de los árboles obtenidos a partir del citado algoritmo.

$$S(z) = \sum_{t \in \mathcal{Q}} |\text{SIMP}(t)| z^{|t|} = \sum s_n z^n$$

No resulta fácil encontrar una ecuación funcional cómoda para $S(z)$ que permita razonar como en el caso de $Q(z)$. Por ello utilizaremos aquí la misma técnica desarrollada en [1] CS por R. Casas y J. M. Steyaert para la clase de los árboles binarios simplificados con dos posibles etiquetas para sus hojas.

Para cada $t \in \mathcal{S}$, definimos $\mathcal{Q}_t = \{T \in \mathcal{Q} / \text{SIMP}(T) = t\}$ y es claro que \mathcal{Q} es la unión disjunta de las clases \mathcal{Q}_t sobre todo $t \in \mathcal{S}$.

Dado $t \in \mathcal{S}$, $Q_t(z) = \sum_{T \in \mathcal{Q}/t = \text{SIMP}(T)} z^{|T|}$ es la función generatriz de todos los árboles T de \mathcal{Q} cuyo simplificado es t . Entonces $Q(z) = \sum_{t \in \mathcal{S}} Q_t(z)$.

Por otra parte es fácil comprobar que

$$(4) \quad \forall e \in \{e_1, \dots, e_q\} \quad Q_e(z) = 1 + zQ_e^p(z) = P(z)$$

donde $P(z)$ es la función generatriz asociada a la familia \mathcal{P} de los árboles p -arios.

Además, para toda p -upla (t_1, t_2, \dots, t_p) de árboles simplificados en la que al menos dos de ellos son distintos se tiene que:

$$(5) \quad Q_{\begin{matrix} \circ \\ / \quad \backslash \\ t_1 \quad t_2 \dots t_p \end{matrix}}(z) = z \prod_{i=1}^p Q_{t_i}(z) + z Q^p_{\begin{matrix} \circ \\ / \quad \backslash \\ t_1 \quad t_2 \dots t_p \end{matrix}}(z)$$

ya que si un árbol tiene como simplificado $\begin{matrix} \circ \\ / \quad \backslash \\ t_1 \dots t_p \end{matrix}$ es o bien porque su subárbol radical i -ésimo tiene como simplificado $t_i, i \in \{1, \dots, p\}$, o bien porque sus p subárboles radicales tienen cada uno como simplificado



Con vistas a la evaluación de $S(z)$, introduciremos la siguiente función generatriz de dos variables:

$$X(y, z) = \sum_{t \in S} y^{|t|} Q_t(z)$$

Obsérvese que el valor en $y = 1$, de la derivada de X con respecto a y , es justamente $S(z)$.

$$S(z) = X'_y(1, z)$$

A partir de (5) y para todo árbol simplificado $t = \begin{matrix} \circ \\ / \quad \backslash \\ t_1 \dots t_p \end{matrix}$ de tamaño

superior a cero se tiene que:

$$y^{|t|} Q_t(z) = yz \prod_{i=1}^p y^{|t_i|} Q_{t_i}(z) + zy^{|t|} Q_t^p(z)$$

Sumando todas estas ecuaciones obtenemos que:

$$(6) \quad X(y, z) = q + yz X^p(y, z) - z \sum_{t \in S} (y^{p|t|+1} - y^{|t|}) Q_t^p(z)$$

sin más que notar que:

$$\sum_{t_1 \dots t_p \in \mathcal{S}} \prod_{i=1}^p y^{|t_i|} Q_{t_i}(z) = X^p(y, z) - \sum_{t \in \mathcal{S}} y^{|t|} Q_t^p(z)$$

tales que $\exists i, j \in \{1, \dots, p\}$
con $t_i \neq t_j$

Derivando en (6) con respecto a y , tomando $y = 1$ y resolviendo la ecuación lineal resultante, llegamos a que

$$(7) \quad S(z) = \frac{zQ^p(z) - z \sum_{t \in \mathcal{S}} ((p-1)|t| + 1)Q_t^p(z)}{1 - pzQ^{p-1}(z)}$$

sin más que observar que $X(1, z) = Q(z)$.

Por otra parte, derivando en (1) se deduce que

$$(8) \quad Q'(z) = \frac{Q^p(z)}{1 - pzQ^{p-1}(z)}$$

que aplicado en (7) nos conduce a lo siguiente:

Proposición 2.1.:

La función generatriz $S(z)$ del tamaño de los árboles simplificados viene dada por:

$$(9) \quad S(z) = zQ'(z) - z \sum_{t \in \mathcal{S}} ((p-1)|t| + 1)Q_t^p(z) \cdot \frac{Q'(z)}{Q^p(z)}$$

Dado que $\forall z \in \mathbf{R}^+$, $S(z) = \sum_{T \in \mathcal{Q}} |\text{SIMP}(T)|z^{|T|} \leq \sum_{T \in \mathcal{Q}} |T|z^{|T|}$, y como se trata de series de coeficientes positivos, sabemos que el radio de convergencia de $S(z)$ es mayor o igual que el radio de convergencia $\rho = \frac{(p-1)^{p-1}}{p^p q^{p-1}}$ de $Q(z)$. Pero por otra parte, $S(z)$ tiene coeficientes superiores a los de $\sum_{T \in \mathcal{Q} / |\text{SIMP}(T)| > 0} z^{|T|}$ cuyo radio de convergencia es ρ , lo que nos permite concluir que el radio de convergencia de $S(z)$ es ρ . Además $z = \rho$ es la única singularidad de $S(z)$ sobre su círculo de convergencia ya que la aplicación del teorema de la función implícita a $Q(z)$ nos asegura que ρ es el único cero de $1 - pzQ^{p-1}(z)$ (denominador de (7)); y el segundo término del numerador tiene un radio de convergencia superior a ρ como veremos a continuación, lo que nos permitirá además abordar la determinación del comportamiento local de $S(z)$ en un entorno de $z = \rho$.

Proposición 2.2.:

Existe $\varepsilon > 0$ tal que $\sum_{t \in S} ((p-1)|t|+1)Q_t^p(z)$ es analítica en $\{z/|z| < \rho + \varepsilon\}$.

Para probar esta proposición demosntremos previamente la siguiente:

Proposición 2.3.

Sea $\{f_i(z)\}$ una sucesión de series de potencias con coeficientes positivos y radio de convergencia ω .

Supongamos que existe una sucesión $\{\alpha_i\}$ de números reales positivos tal que:

i) $\lim_{i \rightarrow \infty} \sup \frac{\alpha_{i+1}}{\alpha_i} \leq 1$.

ii) $\sum \alpha_i f_i(z)$ tiene radio de convergencia β , con $\beta < \omega$.

Si además se verifica:

iii) Dado $p > 1, p \in \mathbf{N}$, existe $\varepsilon > 0$ y existen $c, k, \epsilon \in \mathbf{R}^+$, con $\gamma^p < k$ de manera que $\forall i$ y $\forall z$ con $|z| < \beta + \varepsilon$, se tiene que $|f_i(z)| < c\gamma^{\alpha_i}$ donde $\alpha_i \in \mathbf{N}$ y $\text{card}(\{n/a_n = i\}) \leq M.k^{-i}$ para alguna constante $M \in \mathbf{R}^+$.

Entonces:

a) Dado $p > 1, p \in \mathbf{N}$, $\sum \alpha_i f_i^p(z)$ tiene radio de convergencia superior a β .

b) Si además, las constantes ε, c, γ y k de la hipótesis iii) son independientes de p , y se cumple que $\gamma < 1$ y $c\gamma < 1$, entonces $\sum_{p \geq 2} \sum_{i \geq 0} \alpha_i f_i^p(z)$ tiene radio de convergencia superior a β

Demostración:

a) Sea $p > 1, p \in \mathbf{N}$, y sea $f(z) = \sum \alpha_i f_i^p(z)$.

A partir de iii) tenemos que $|\alpha_i f_i^p(z)| < c^p \alpha_i \gamma^{p\alpha_i}$ para todo z con $|z| < \beta + \varepsilon$.

Ahora:

$\sum_{i \in \mathbf{N}} \alpha_i \gamma^{p\alpha_i} \leq M \sum_{j \in \mathbf{N}} \alpha_j k^{-j} \gamma^{pj}$ que es convergente por cumplirse i) y ser $\gamma^p < k$.

Luego $f(z) = \sum \alpha_i f_i^p(z)$ converge absoluta y uniformemente sobre $\{z/|z| < \beta + \varepsilon\}$. (Criterio de Weierstrass para la convergencia uniforme).

Además $f(z)$ es analítica en $\{z/|z| < \beta + \varepsilon\}$ por la completitud del espacio de las funciones holomorfas, es decir por ser analítica cada $\alpha_i f_i^p(z)$ y converger $f(z)$ uniformemente sobre todo disco cerrado incluido en $\{z/|z| < \beta + \varepsilon\}$. (Teorema de Weierstrass)

b) Sea $g(z) = \sum_{p \geq 2} \sum_{i \in \mathbf{N}} \alpha_i f_i^p(z)$.

Dado $p > 1, p \in \mathbf{N}$, sabemos que $|\sum_{i=0}^{\infty} \alpha_i f_i^p(z)| < \sum \alpha_i c^p \gamma^{p a_i}$ para todo z con $|z| < \beta + \varepsilon$. También sabemos que $\sum_{i=0}^{\infty} \alpha_i c^p \gamma^{p a_i}$ es convergente. Por lo tanto, en concreto para $p = 2$ tenemos que para todo $\varepsilon' > 0$ existe un j tal que

$$(10) \quad \sum_{i=j}^{\infty} \alpha_i (c\gamma^{a_i})^2 = A < \varepsilon' < 1$$

Supongamos, sin pérdida de generalidad, que los α_i son mayores o iguales a 1. Entonces para todo $p \geq 3$ tenemos que $\alpha_i (c\gamma^{a_i})^p \leq \alpha_i^{p/2} (c\gamma^{a_i})^p$. Con lo cual, a partir de (10), podemos afirmar que existe un j tal que $\forall p \geq 3$

$$\sum_{i=j}^{\infty} \alpha_i (c\gamma^{a_i})^p \leq \sum_{i=j}^{\infty} \alpha_i^{p/2} (c\gamma^{a_i})^p < A^{p/2} < (\varepsilon')^{p/2} < 1.$$

Ahora:

$$\begin{aligned} \sum_{p \geq 2} \sum_{i \in \mathbf{N}} \alpha_i (c\gamma^{a_i})^p &= \sum_{p \geq 2} \sum_{i=0}^{j-1} \alpha_i (c\gamma^{a_i})^p + \sum_{p \geq 2} \sum_{i=j}^{\infty} \alpha_i (c\gamma^{a_i})^p < \\ &< \sum_{i=0}^{j-1} \alpha_i \sum_{p \geq 2} (c\gamma^{a_i})^p + \sum_{p \geq 2} A^{p/2} \end{aligned}$$

que es convergente por ser $c\gamma < 1, \gamma < 1$ y $A < 1$.

Entonces, la parte b) de la proposición queda demostrada sin más que aplicar de nuevo el criterio de Weierstrass para la convergencia uniforme, y la completitud del espacio de las funciones holomorfas.(1)

Ahora para demostrar la proposición 2.2. basta comprobar que se verifican las hipótesis de la proposición 2.3.

En primer lugar $\sum_{t \in S} ((p-1)|t|+1)Q_t(z)$ tiene radio de convergencia ρ , puesto que $\sum_{t \in S} ((p-1)|t|+1)Q_t(z) = (p-1) \sum_{t \in S} |t|Q_t(z) + \sum_{t \in S} Q_t(z) = (p-1)S(z) + Q(z)$.

Además $\sum_{t \in S} ((p-1)|t|+1)Q_t(z) = \sum_i \sum_{t \in S/|t|=i} ((p-1)i+1)Q_t(z)$ con lo que la hipótesis i) se verifica trivialmente.

Queda pues por probar la hipótesis iii).

Para ello comenzamos por demostrar lo siguiente:

Lema 2.1.:

Para todo $p \geq 3$, para $z_0 = \frac{(p-1)^{p-1}}{p^p q}$ y $\gamma = \frac{p-1}{pq} \cdot \frac{1}{\sqrt[p]{2}}$ se verifica que para todo $t \in \mathcal{S}$,

$$Q_t(z_0) < \frac{p}{p-1} \gamma^{|t|}$$

Demostración:

Nótese que $\rho = \frac{(p-1)^{p-1}}{p^p q^{p-1}} < z_0 < \frac{(p-1)^{p-1}}{p^p}$, y que $\gamma^p < \rho$.

Razonemos por inducción sobre el tamaño de t .

Por (4) sabemos que para todo $e \in \{e_1, \dots, e_q\}$ $Q_e(z) = P(z)$.

Con lo que, $Q_e(z_0) < \frac{p}{p-1}$, pues $P(z)$ es una serie de potencias con coeficientes positivos, radio de convergencia $\frac{(p-1)^{p-1}}{p^p}$ (teorema de la función implícita) y $P\left(\frac{(p-1)^{p-1}}{p^p}\right) = \frac{p}{p-1}$.

Sea $t \in \mathcal{S}$ con $|t| \geq 1$, y supongamos que para todo $t^* \in \mathcal{S}$ con $|t^*| < |t|$ se verifica que $Q_{t^*}(z_0) < \frac{p}{p-1} \gamma^{|t^*|}$.

A partir de (5) y de la hipótesis de inducción tenemos que:

$$Q_t(z_0) - z_0 Q_t^p(z_0) = z_0 \prod_{i=1}^p Q_{t_i}(z_0) < z_0 \left(\frac{p}{p-1}\right)^p \gamma^{|t|-1}$$

Lo que nos permite concluir que $Q_t(z_0) < \frac{p}{p-1} \gamma^{|t|}$ sin más que observar que:

1º) La función $y(x) = x - z_0 x^p$, definida sobre \mathbf{R}^+ es creciente en $[0, 1/\sqrt[p-1]{pz_0}]$, con $y(1/\sqrt[p-1]{pz_0}) > 1$.

2º) $Q_t(z_0) < 1/\sqrt[p-1]{pz_0}$

3º) $\frac{p}{p-1} \gamma^{|t|} - z_0 \left(\frac{p}{p-1}\right)^p \gamma^{p|t|} \geq z_0 \left(\frac{p}{p-1}\right)^p \gamma^{|t|-1}$

Para probar 2º) nótese que $z_0 \left(\frac{p}{p-1}\right)^p \gamma^{|t|-1} \leq \frac{1}{p-1}$. Por lo tanto $Q_t(z_0) - z_0 Q_t^p(z_0) = \alpha_0 < \frac{1}{p-1}$.

Pero $\frac{2}{p-1}$ es la mejor raíz positiva de

$$x - z_0 x^p = \frac{2}{p-1} - z_0 \frac{2^p}{(p-1)^p} < y (1/\sqrt[p-1]{pz_0})$$

$$y \frac{1}{p-1} < \frac{2}{p-1} - z_0 \left(\frac{2}{p-1}\right)^p$$

Luego por 1º), llegamos a que $Q_t(z_0) < \frac{2}{p-1} < 1/\sqrt[p-1]{pz_0}$.

Para comprobar 3º) hemos de verificar que

$$\frac{p}{p-1} - z_0 \left(\frac{p}{p-1}\right)^p \gamma^{(p-1)|t|} \geq z_0 \left(\frac{p}{p-1}\right)^p \frac{1}{\gamma}, \text{ para lo cual}$$

es suficiente ver que $\frac{p}{p-1} - z_0 \left(\frac{p}{p-1}\right)^p \gamma^{p-1} \geq z_0 \left(\frac{p}{p-1}\right)^p \frac{1}{\gamma}$
 puesto que $|t| \geq 1$. y $\gamma < 1$.

Ahora:

$$\frac{p}{p-1} - z_0 \left(\frac{p}{p-1}\right)^p \gamma^{p-1} \geq z_0 \left(\frac{p}{p-1}\right)^p \frac{1}{\gamma} \iff 1 \geq \frac{\sqrt[p]{2}}{p-1} (\gamma^p + 1)$$

lo cual es cierto por ser $p \geq 3$ y $\gamma^p + 1 < 1 + \frac{1}{8\sqrt[p]{8}}$

Lema 2.2.:

Sea $p = 2$. Para todo $z_0 \in \mathbf{R}^+$ tal que $\frac{1}{4q} < z_0 < \frac{\sqrt{q}}{1+4q}$ existe $\gamma \in \mathbf{R}^+$ con $\gamma < \frac{1}{2\sqrt{q}}$ tal que para todo $t \in \mathcal{S}$ se verifica que $Q_t(z_0) < 2\gamma^{|t|}$.

Demostración:

Análoga a la del lema 2.1., por lo que nos limitaremos a señalar la relación entre γ y z_0 .

$$\gamma = \frac{1}{2\sqrt{q}} \varepsilon \text{ con } 2\sqrt{q}(q - \sqrt{q^2 - 1}) < \varepsilon < 1 \text{ cumpliéndose la ecuación}$$

$$z_0 = \frac{\sqrt{q} \cdot \varepsilon}{\varepsilon^2 + 4q}$$

Luego a partir de los lemas 2.1. y 2.2. sabemos que dado $p > 1, p \in \mathbf{N}$, existe $\varepsilon > 0$ y existen $c, \gamma \in \mathbf{R}^+$, con $\gamma^p < \rho$ tales que para todo $t \in \mathcal{S}$ y para todo z con $|z| < \rho + \varepsilon$ se tiene que $|Q_t(z)| < c\gamma^{|t|}$.

Además para todo $\varepsilon > 0$ $\text{card} \{t \in S / |t| = i\} \leq \rho^{-i}(1 + \varepsilon)^i$ en casi todo punto.

Con lo que podemos concluir la parte a) de la proposición 2.3., sin más que notar que, por continuidad, $\gamma^p < \rho$ implica que existe $\varepsilon' > 0$ tal que $\gamma^p(1 + \varepsilon') < \rho$.

Luego queda demostrada la proposición 2.2. (2) y (3)

Nos encontramos pues en condiciones de demostrar el

Teorema 2.1.1.:

Los coeficientes de Taylor de $S(z)$ pueden aproximarse asintóticamente, cuando n tiende a infinito por:

$$s_n = \delta \rho^{-n} n^{-1/2} (1 + o(1/n))$$

siendo

$$\rho = \frac{(p-1)^{p-1}}{p^p q^{p-1}} \quad y \quad \delta = \frac{1}{2\sqrt{\pi}} \sqrt{\frac{2p}{p-1}} \left[\frac{q}{p-1} - \alpha(\rho) \right] \quad \text{donde}$$

$$\alpha(\rho) = \rho \sum_{t \in S} ((p-1)|t| + 1) Q_t^p(\rho).$$

Demostración:

A partir de (2) obtenemos

$$(11) \quad Q'(z) = g_1(z)(1 - z/\rho)^{-1/2} + g_2(z)(1 - z/\rho)^{1/2} + h_1(z)$$

donde $g_1(z)$, $g_2(z)$ y $h_1(z)$ son analíticas en un entorno de $z = \rho$, con $g_2(\rho) = -\frac{11p+2}{24p} \frac{p+1}{p-1} \frac{g(\rho)}{\rho}$, $h_1(\rho) = -\frac{2}{3} \frac{p+1}{p-1} \frac{h(\rho)-q}{\rho}$ y $g_1(\rho) = -\frac{1}{2\rho} g(\rho)$.

Entonces, a partir de las proposiciones anteriores y sustituyendo en (9) $Q(z)$ y $Q'(z)$ por sus desarrollos locales (2) y (11), y z y $\alpha(z)$ por sus desarrollos de Taylor en $z = \rho$, llegamos a que $S(z)$ admite, en un entorno de $z = \rho$, el desarrollo:

$$(12) \quad S(z) = A(z)(1 - z/\rho)^{-1/2} + B(z)(1 - z/\rho)^{1/2} + C(z)$$

con $A(z)$, $B(z)$ y $C(z)$ analíticas en un entorno de $z = \rho$ y

$$\begin{aligned}
A(\rho) &= \frac{1}{2} \sqrt{\frac{2p}{p-1}} \cdot \left[\frac{q}{p-1} - \alpha(\rho) \right] \\
B(\rho) &= \frac{-g(\rho)}{p-1} + \frac{\rho\alpha'(\rho)}{2} \sqrt{\frac{2p}{p-1}} + \frac{2-p}{12p} \frac{p+1}{p-1} A(\rho) \\
C(\rho) &= \frac{p-2}{3(p-1)} \left[\frac{q}{p-1} - \alpha(\rho) \right] - \frac{1}{p-1} h(\rho)
\end{aligned}$$

satisfaciéndose pues las hipótesis del teorema de Darboux.

Teorema 2.1.2.:

El tamaño medio de un árbol simplificado obtenido a partir de un árbol aleatorio de tamaño n de \mathcal{Q} satisface

$$\bar{m}_n = \frac{s_n}{q_n} = \lambda n(1 + O(1/n))$$

con $\lambda = 1 - \frac{p-1}{q} \alpha(\rho)$

Demostración:

Inmediata a partir de (3) y del teorema 2.1.1.

Para obtener la varianza V_n del tamaño de los árboles simplificados a partir de árboles de \mathcal{Q} de tamaño n , observemos que:

$$\begin{aligned}
(13) \quad V_n &= \sum_{m=0}^n m^2 \frac{i_{m,n}}{q_n} - \left(\sum_{m=0}^n m \frac{i_{m,n}}{q_n} \right)^2 = \\
&= \frac{[z^n](X_y''(1, z) + X_y'(1, z))}{q_n} - \left(\frac{[z^n](X_y'(1, z))}{q_n} \right)^2
\end{aligned}$$

donde $i_{m,n}$ es el número de árboles de \mathcal{Q} de tamaño n cuyo simplificado es de tamaño m .

Ahora, derivando en (6) con respecto a y dos veces, haciendo $y = 1$, y resolviendo la ecuación lineal a la que se llega obtenemos:

$$(14) \quad X_y''(1, z) = \frac{2pzQ^{p-1}(z)S(z) + p(p-1)zQ^{p-2}(z)(S(z))^2 - \beta(z)}{1 - pzQ^{p-1}(z)}$$

donde

$$\beta(z) = (p+1)z \sum_{t \in S} |t|((p-1)|t|+1)Q_t^p(z)$$

Por un razonamiento análogo al efectuado para $S(z)$, y recordando que $X_y''(1, z) = \sum_{T \in \mathcal{Q}} |\text{SIMP}(T)|(|\text{SIMP}(T)|-1)z^{|\text{T}|}$, concluimos que ρ es el radio de convergencia de $X_y''(1, z)$, así como su única singularidad sobre su círculo de convergencia, para lo cual ha de notarse, además, que $\beta(z)$ tiene radio de convergencia superior a ρ , por verificarse las hipótesis de la proposición 2.3.

Ahora, aplicando (1) y (8) en (14) llegamos a

$$X_y''(1, z) = \frac{zQ'(z)}{Q(z)-q} \left[p(Q(z)-q) \frac{S(z)}{Q(z)} \left(2 + (p-1) \frac{S(z)}{Q(z)} \right) - \beta(z) \right]$$

y sustituyendo aquí, $Q(z)$, $Q'(z)$ y $S(z)$ por sus respectivos desarrollos locales (2), (11) y (12) así como los desarrollos de Taylor de z y $\beta(z)$ en $z = \rho$, obtenemos:

$$X_y''(1, z) = D(z)\left(1 - \frac{z}{\rho}\right)^{-\frac{3}{2}} + E(z)\left(1 - \frac{z}{\rho}\right)^{-\frac{1}{2}} + \\ F(z)\left(1 - \frac{z}{\rho}\right)^{\frac{1}{2}} + J(z)\left(1 - \frac{z}{\rho}\right)^{-1} + H(z)$$

donde $D(z)$, $E(z)$, $F(z)$, $J(z)$ y $H(z)$ son analíticas en un entorno de $z = \rho$, lo que nos permite concluir:

Teorema 2.2:

La varianza del tamaño de los árboles simplificados obtenidos a partir de los árboles de \mathcal{Q} de tamaño n tiene el equivalente asintótico :

$$V_n \simeq \nu \cdot n \quad n \rightarrow \infty$$

donde

$$\nu = -\frac{p-1}{q^2} \left(\frac{q}{p-1} - \alpha(\rho) \right)^2 + \left(2 \left(\frac{p-1}{q} \right)^2 \rho \alpha'(\rho) + \right. \\ \left. + \frac{p+1}{q} \right) \left(\frac{q}{p-1} - \alpha(\rho) \right) - \frac{p-1}{q} \beta(\rho) - \frac{p}{p-1}$$

Demostración:

A partir de (13) y de la aplicación del teorema de Darboux a los desarrollos locales de $X_y''(1, z) + X_y'(1, z)$, $Q(z)$ y $X_y'(1, z)$, tras un gran número de cálculos tediosos.

3. ARBOLES GENERALES

3.1. ALGORITMO DE SIMPLIFICACIÓN

Consideremos ahora la familia $\mathcal{G}Q$ de los árboles generales con q posibles etiquetas para sus hojas ($q \geq 2$); para la que vamos a considerar el mismo tipo de algoritmo de simplificación que en el apartado anterior. Para $\mathcal{G}Q$, tal algoritmo puede escribirse en pseudo-Pascal como sigue:

Function SIMP ($T : \mathcal{G}Q$) : $\mathcal{G}Q$;

if T .degree = 0 then SIMP := T

else for $i := 1$ to T .degree do $t(i) := \text{SIMP}(T. \text{son}(i))$ od

$i := 1$; CON := true;

While ($i \leq T$.degree) and CON do

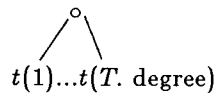
if EQUAL ($t(i), t(i + 1)$)

then $i := i + 1$

else CON := false od

if CON then SIMP := $t(1)$

else SIMP :=



end SIMP.

Nuestro objetivo es, de nuevo, el estudio del tamaño medio de un árbol simplificado con respecto a su tamaño inicial, entendiendo ahora como tamaño de un árbol el número total de sus nodos.

El desarrollo a seguir es absolutamente análogo al del apartado anterior por lo que aquí nos limitaremos a esbozarlo enunciando a la vez los resultados obtenidos.

3.2. FUNCIONES GENERATRICES

La familia $\mathcal{G}Q$ viene definida por:

$$\mathcal{G}Q = \sum_{i=1}^q e_i + \sum_{p=1}^{\infty} \circ(\underbrace{\mathcal{G}Q, \dots, \mathcal{G}Q}_p) \text{ donde } \{e_1, \dots, e_q\} \text{ es el conjunto de posi-}$$

bles etiquetas para las hojas de los árboles de la familia. A partir de lo cual puede deducirse que el radio de convergencia de su función generatriz asociada, $GQ(z)$, es $\rho = \frac{1}{(1+\sqrt{q})^2}$; y que

$$(15) \quad GQ(z) = (1 - z/p)^{1/2} g(z) + h(z)$$

siendo $h(z)$ entera y $g(z)$ analítica en un entorno del origen de radio

$$\frac{1}{(1-\sqrt{q})^2}, \text{ con } g(\rho) = \frac{-q^{1/4}}{1+\sqrt{q}} \text{ y } h(\rho) = \frac{\sqrt{q}}{1+\sqrt{q}}$$

Entonces por el teorema de Darboux-Polya:

$$(16) \quad g_{q,n} = \rho^{-n} n^{-3/2} \frac{q^{1/4}}{2\sqrt{\pi}(1+\sqrt{q})} (1 + o(1/n))$$

La función generatriz asociada a la familia \mathcal{G} de los árboles generales es $G(z) = \frac{1-\sqrt{1-4z}}{2}$ cuyo radio de convergencia es $1/4$.

Llamamos \mathcal{S} a la clase de todos los árboles simplificados obtenidos según el algoritmo anterior, y sea

$$\mathcal{S}(z) = \sum_{t \in \mathcal{G}Q} |\text{SIMP}(t)| z^{|t|} = \sum s_n z^n$$

la función generatriz del tamaño de los árboles de \mathcal{S} .

Para todo $t \in \mathcal{S}$ definimos

$$\mathcal{G}Q_t = \{T \in \mathcal{G}Q / \text{SIMP}(T) = t\} \quad \mathcal{G}Q = \bigcup_{t \in \mathcal{S}} \mathcal{G}Q_t \text{ (unión disjunta).}$$

Luego

$$GQ(z) = \sum_{t \in \mathcal{S}} GQ_t(z) \text{ donde } GQ_t(z) = \sum_{T \in \mathcal{G}Q / \text{SIMP}(T)=t} z^{|T|}$$

Para todo $t \in \mathcal{S}$, $GQ_t(z)$ queda definida recursivamente por el sistema de ecuaciones:

$$(17) \quad \left\{ \begin{array}{l} \forall e \in \{e_1, \dots, e_q\} \quad GQ_e(z) = G(z) \\ \text{Dado } p \geq 2, \forall (t_1, \dots, t_p) \text{ tal que } t_i \in \mathcal{S} \text{ y existen } i, j \in \{1, \dots, p\} \\ \text{tales que } t_i \neq t_j \end{array} \right. \quad GQ_{\begin{array}{c} \circ \\ / \quad \backslash \\ t_1 \dots t_p \end{array}}(z) = z \prod_{i=1}^p GQ_{t_i}(z) + z \sum_{j \geq 1} GQ^j_{\begin{array}{c} \circ \\ / \quad \backslash \\ t_1 \dots t_p \end{array}}(z)$$

(Obsérvese que no existen árboles simplificados cuya raíz tenga un único hijo (caso $p = 1$)).

$$\text{Sea } X(y, z) = \sum_{t \in \mathcal{S}} y^{|t|} GQ_t(z).$$

$$S(z) = X'_y(1, z) \text{ y } X(1, z) = GQ(z).$$

A partir de esto, de (17) y de que

$$(18) \quad GQ'(z) = \frac{GQ(z)}{z(1 - z \sum_{p \geq 1} p GQ^{p-1}(z))}$$

concluimos que:

Proposición 3.1.:

La función generatriz $S(z)$ del tamaño de los árboles simplificados viene dada por:

$$(19) \quad S(z) = (1 - z)zGQ'(z) - \alpha(z)z \frac{GQ'(z)}{GQ(z)}$$

$$\text{con } \alpha(z) = z \sum_{p \geq 2} \sum_{t \in \mathcal{S}} ((p-1)|t| + 1) GQ_t^p(z)$$

Ahora bien, el radio de convergencia de $S(z)$ es $\rho = \frac{1}{(1+\sqrt{q})^2}$, que es además su única singularidad sobre su círculo de convergencia, lo que se comprueba a partir del teorema de la función implícita y de que, como veremos a continuación, $\alpha(z)$ es analítica en un entorno del origen de radio superior a ρ . Y para determinar su comportamiento local en un entorno de $z = \rho$, se comienza por probar:

Proposición 3.2.:

Existe $\varepsilon > 0$ tal que $\sum_{p \geq 2} \sum_{t \in \mathcal{S}} ((p-1)|t|+1)GQ_t^p(z)$ es analítica en $\{z/|z| < \rho + \varepsilon\}$.

Para demostrar esto basta ver que se cumplen las hipótesis de la proposición 2.3.

La comprobación de la hipótesis iii) se apoya en el siguiente resultado:

Lema 3.1.:

$$\text{Sea } z_0 = \frac{(1 + \sqrt{2})^2}{5} \frac{1}{(1 + \sqrt{q})^2}, \text{ existe } \gamma < \frac{1}{1 + \sqrt{q}}$$

tal que para todo $t \in \mathcal{S}$ se cumple que $GQ_t(z_0) < \gamma^{|t|}$.

Demostración:

El lema se cumple para todo $\gamma = \frac{\varepsilon}{1 + \sqrt{q}}$

$$\text{con } \frac{1 + \sqrt{q} - \sqrt{(1 + \sqrt{q})^2 - (1 + \sqrt{2})^2 \cdot 4/5}}{2} < \varepsilon < 1$$

La demostración es análoga a la del lema 2.1., sin más que observar que:

1º) La función $y(x) = x - z_0 \frac{x}{1-x}$, definida sobre $[0, 1]$, es creciente en $[0, 1 - \sqrt{z_0}]$, con $y(1 - \sqrt{z_0}) = (1 - \sqrt{z_0})^2$.

2º) $GQ_t(z_0) < 1 - \sqrt{z_0}$

3º) $\gamma^{|t|} - z_0 \frac{\gamma^{|t|}}{1 - \gamma^{|t|}} > z_0 \gamma^{|t|-1}$ para todo $t \in \mathcal{S}$ con $|t| \geq 2$

La comprobación de las restantes hipótesis de la proposición 2.3. es análoga a la realizada en la sección anterior y no ofrece especial dificultad.

Finalmente estamos en condiciones de enunciar el siguiente teorema:

Teorema 3.1.1.

Los coeficientes de Taylor de $S(z)$ pueden aproximarse asintóticamente, cuando n tiende a infinito, por :

$$s_n = \delta \rho^{-n} n^{-1/2} (1 + o(1/n))$$

$$\text{donde } \rho = \frac{1}{(1 + \sqrt{q})^2}, \quad \delta = \frac{1}{2\sqrt{\pi}} \frac{q^{1/4}}{(1 + \sqrt{q})} \left(1 - \rho - \frac{1 + \sqrt{q}}{\sqrt{q}} \alpha(\rho)\right)$$

$$\text{y } \alpha(\rho) = \rho \sum_{p \geq 2} \sum_{t \in S} ((p-1)|t| + 1) G Q_t^p(\rho)$$

Teorema 3.1.2.:

El tamaño medio de un árbol simplificado obtenido a partir de un árbol aleatorio de tamaño n de \mathcal{GQ} satisface:

$$\bar{m}_n = \frac{s_n}{g_{q,n}} = \lambda n (1 + o(1/n))$$

con

$$\lambda = \left(1 - \rho - \frac{1 + \sqrt{q}}{\sqrt{q}} \alpha(\rho)\right)$$

Demostración:

El teorema 3.1.1. se deduce a partir de las proposiciones 3.1. y 3.2. que nos conducen a un desarrollo local de $S(z)$ en un entorno de $z = \rho$ de la forma:

$$(20) \quad S(z) = A(z)(1 - z/\rho)^{-1/2} + B(z)(1 - z/\rho)^{1/2} + C(z)$$

donde $A(z)$, $B(z)$ y $C(z)$ son analíticas en un entorno de $z = \rho$, siendo

$$A(\rho) = -\frac{g(\rho)}{2} \left(1 - \rho - \frac{\alpha(\rho)}{h(\rho)}\right)$$

$$B(\rho) = \frac{1}{2} \left((3\rho g'(\rho) + g(\rho)) \left(1 - \rho - \frac{\alpha(\rho)}{h(\rho)}\right) - \rho g(\rho) \left(1 + \frac{\alpha'(\rho)}{h(\rho)}\right) + \frac{\alpha(\rho)}{h(\rho)} \left(\frac{3g(\rho)\rho h'(\rho)}{h(\rho)} + \frac{g^3(\rho)}{h^2(\rho)} \right) \right)$$

$$C(\rho) = \rho h'(\rho) \left(1 - \rho - \frac{\alpha(\rho)}{h(\rho)}\right) - \frac{\alpha(\rho)g^2(\rho)}{2h^2(\rho)}$$

con

$$h'(\rho) = \frac{q-1}{2} \quad \text{y} \quad g'(\rho) = \frac{(1 - \sqrt{q})(1 - q)}{8q^{1/4}}$$

Con lo que estamos en situación de aplicar el teorema de Darboux.

El teorema 3.1.2. es inmediato a partir de (16) y del teorema 3.1.1.

Para la obtención de la varianza del tamaño de los árboles simplificados a partir de árboles de \mathcal{GQ} de tamaño n se sigue el mismo procedimiento que en el apartado 2 por lo que nos limitaremos a enunciar los resultados.

Proposición 3.3.:

$$X_y''(1, z) = D(z)(1 - z/\rho)^{-3/2} + E(z)(1 - z/\rho)^{-1/2} + F(z)(1 - z/\rho)^{-1} + H(z)(1 - z/\rho)^{1/2} + J(z)$$

satisfaciéndose las hipótesis del teorema de Darboux, y siendo:

$$D(\rho) = -\frac{A^2(\rho)}{g(\rho)}$$

$$E(\rho) = -\frac{2A^2(\rho)}{g^2(\rho)} \left(g(\rho) - \frac{3g^3(\rho)}{2h^2(\rho)} - \frac{3g(\rho)\rho h'(\rho)}{h(\rho)} + \frac{\rho g'(\rho)}{2} \right) - A(\rho) \left((1 - \rho) \left(1 - \frac{2g^2(\rho)}{h^2(\rho)} - \frac{3\rho h'(\rho)}{h(\rho)} \right) + 1 - \frac{\rho \alpha'(\rho)}{h(\rho)} \right) + \frac{(1 - \rho)^2 g^3(\rho)}{4h^2(\rho)} + \frac{g(\rho)}{2h(\rho)} \beta(\rho)$$

$$\beta(z) = z \sum_{p \geq 2} \sum_{t \in S} (p + 1) |t| ((p - 1) |t| + 1) GQ_i^p(z)$$

Ahora, tras la aplicación del teorema de Darboux al desarrollo anterior, así como a (20) y (15), concluimos que la varianza es lineal en el tamaño del input:

Teorema 3.2.:

La varianza del tamaño de los árboles simplificados obtenidos a partir de árboles de \mathcal{GQ} de tamaño n cumple:

$$V_n = \nu n(1 + O(1/n)) \quad \text{cuando } n \rightarrow \infty$$

con

$$\nu = \frac{6\sqrt{q} + q - 3}{8\sqrt{q}} \lambda^2 + \left(3\rho - \frac{\sqrt{q}}{1 + \sqrt{q}} + \frac{2}{\sqrt{q}(1 + \sqrt{q})} \alpha'(\rho) \right) \lambda - \frac{\sqrt{q}(2 + \sqrt{q})^2}{2} \rho^2 - \frac{1 + \sqrt{q}}{\sqrt{q}} \beta(\rho)$$

4. APROXIMACIONES NUMERICAS PARA DIVERSAS FAMILIAS DE ARBOLES. SIMULACIONES.

A continuación se da una tabla con los valores aproximados de las constantes de proporcionalidad teóricas obtenidas, λ y ν , para diversas familias de árboles (Tabla 1.a.), así como de las constantes resultado de la simulación del proceso de simplificación sobre las mismas familias (Tabla 1.b.).

Como puede observarse hay una gran semejanza entre tales cantidades.

Para las simulaciones se ha utilizado el método de generación de árboles aleatorios propuesto por [6] Re 80. En cuanto a las aproximaciones, se ha procedido al cálculo cuidadoso de los términos de $\alpha(\rho)$ y $\beta(\rho)$, correspondientes a los simplificados de tamaño menor o igual a seis, guardando rastro de la estructura de los simplificados en combinación con el sistema recursivo que caracteriza la función generatriz asociada al conjunto de árboles que se simplifican a uno dado.

TABLA 1

Valores de las constantes para árboles p -arios con q posibles etiquetas para sus hojas.

Tabla 1.a. Resultados por
aproximación

Tabla 1.b. Resultados por
simulación

p	q	λ	ν	λ	ν
2	2	0.81959764	0.216602258	0.81836666	0.217610777
2	3	0.89523471	0.118241731	0.89423333	0.127301444
2	4	0.92614127	0.080893486	0.92617777	0.079439407
3	2	0.91617279	0.087425557	0.91598888	0.097828852
3	3	0.96533246	0.035276246	0.96558888	0.035307074
3	4	0.98094646	0.019236459	0.98156666	0.019252555
4	2	0.95819829	0.042384292	0.95765555	0.044295518
4	3	0.98809411	0.011953141	0.98745555	0.012179963
4	4	0.99502326	0.004984991	0.99534444	0.004799774

5. REFERENCIAS BIBLIOGRAFICAS

- [1] CASAS R. STEYAERT J.M. "Bottom-up Recursion in Trees". CAAP-86 LNCS 214, 1986, pp 172-182.
- [2] FLAJOLET P. "Mathematical Methods in the Analysis of Algorithms and Data Structures". Lecture Notes for A Graduate Course on Computation Theory. Udine Sept-Oct 1984.
- [3] HENRICI P. "Applied and Computational Complex Analysis." Vol 1 y 2. J. Wiley, New York 1974.
- [4] KNUTH D.E. "The Art of Computer Programming". Vol. 1, 2 y 3. Addison Wesley, Reading 1968, 1969,1973.
- [5] KNUTH D.E., BENDIX P. "Simple Word Problems in Universal Algebra". Computational Problems in Abstract Algebra (ed. J. Leech). Pergamon Press 1970. pp 263-297.
- [6] REMY J.L. "Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire". Rapport de Recherche n° 80- P-53, CRIN, Nancy, 1974.
- [7] STEYAERT J.M. "Structure et Complexité des Algorithmes". Thèse Paris 7. 1984.

6. NOTAS

- (0) EQUAL testea la igualdad de dos árboles, en orden prefijo.
- (1) -La proposición 2.3. sigue siendo cierta aunque no se conozca exactamente ω siempre y cuando se verifique iii).
-i) puede cambiarse por i') $\limsup_{i \rightarrow \infty} \sqrt[i]{\alpha_i} \leq 1$ y los resultados siguen siendo válidos
- (2) Dada una serie de potencias $f(z)$ con coeficientes positivos, si existe $z_0 \in \mathbf{R}^+$ tal que $f(z_0)$ converge, entonces para todo $z \in \mathbf{C}$ tal que $|z| < z_0$ se cumple que $|f(z)| < f(z_0)$.
- (3) Siguiendo las mismas técnicas desarrolladas en los lemas 2.1. y 2.2. puede demostrarse que para todo $t \in \mathcal{S}$, el radio de convergencia de $Q_t(z)$ es $\frac{(p-1)^{p-1}}{p^p}$. Para ello se demuestra que $\forall z/|z| < \frac{(p-1)^{p-1}}{p^p}$ y para todo $t \in \mathcal{S}$ $Q_t(z) < \frac{p}{p-1} \cdot \left(\frac{3}{4}\right)^{|t|}$
- (4) También puede afirmarse que para todo $t \in \mathcal{S}$, $GQ_t(z)$ tiene radio de convergencia $1/4$, a partir del hecho de que para todo $z \in \mathbf{R}^+$ con $\rho < z < 1/4$, $GQ_t(z) < \frac{1}{2} \left(\frac{1}{8}\right)^{|t|}$