Über Abzählprobleme von unabhängigen Teilmengen in Bäumen

Michael Drmota (Wien)

Es bezeichne $\mathcal{F}$ eine Familie von Wurzelbäumen und y_n die Anzahl der Bäume dieser Familie mit n Knoten. $\mathcal{F}$ heißt einfach erzeugt, falls die erzeugende Funktion $Y(x) = \sum_n y_n x^n$ eine Funktionalgleichung der Gestalt $Y = x\varphi(Y)$ erfüllt. Die Koeffizienten der Potenzreihenentwicklung von $\varphi(t) = 1 + \varphi_1 t + \varphi_2 t^2 + \ldots$ beschreiben dabei den Aufbau der einfach erzeugten Familie. ($\varphi(t) = 1 + t^2$ für binäre Bäume, $\varphi(t) = 1/(1-t)$ für ebene Bäume, $\varphi(t) = e^t$ für markierte Bäume.) Geht man umgekehrt von einem vorgegebenem $\varphi(t)$ aus und bezeichne $\omega(T_n) = \prod_i \varphi_i^{D_i(T_n)}$ das Gewicht eines Wurzelbaumes mit n Knoten, wobei $D_i(T_n)$ die Anzahl der Knoten in T_n ist, von denen i Unterbäume abzweigen, so erfüllt die erzeugende Funktion Y von $y_n = \sum \omega(T_n)$ die Funktionalgleichung $Y = X\varphi(Y)$.

Eine Teilmenge I eines Baumes (oder allgemeiner eines Graphen) heißt unabhängig, falls keine zwei Knoten aus I durch eine Kante verbunden sind. Eine Teilmenge I eines Wurzelbaumes soll i-unabhängig heißen ($i \geq 1$), falls sie unabhängig ist und jeder Knoten, der nicht in I liegt, durch einen Weg der Länge $\leq i$ verbunden werden kann, wobei die Orientierung des Weges auf jeder Teilkante entweder nur zur Wurzel hin- oder nur wegführen darf. 1-unabhängige Teilmengen sind gerade die maximal unabh. Teilmengen. Es soll nun die durchschnittliche Anzahl $e^{(i)}(n)$ und die durchschnittliche Größe $\mu^{(i)}(n)$ von i-unabh. Teilmengen in Wurzelbäumen mit n Knoten einer einfach erzeugten Familie bestimmt werden. (Für 1-unabh. Teilmengen haben dies bereits Meir und Moon [3] durchgeführt.)

Bezeichne $a_k^{(i)}(T_n)$ die Anzahl der i-unabh. Teilmengen der Größe k eines Wurzelbaumes T_n mit n Knoten, die die Wurzel enthalten, $b_k^{(i)}(T_n)$ die entsprechende Anzahl, die die Wurzel nicht enthalten, und $g_k^{(i)}(T_n) = a_k^{(i)}(T_n) + b_k^{(i)}(T_n)$.

Sei weiters $a_{nk}^{(i)} = \sum \omega(T_n) a_k^{(i)}(T_n)$, $b_{nk} = \sum \omega(t_n) b_k^{(i)}(T_n)$,

$g_{nk}^{(i)} = a_{nk}^{(i)} + b_{nk}^{(i)}$, $a_n^{(i)} = \sum_k a_{nk}^{(i)}$, $b_n^{(i)} = \sum_k b_{nk}^{(i)}$,

$g_n^{(i)} = a_n^{(i)} + b_n^{(i)}$. Betrachtet man überdies die erzeugenden Funktionen

$$A^{(i)}(x,z) = \sum_{n,k} a_{nk}^{(i)} z^k x^n , \quad B^{(i)}(x,z) = \sum_{n,k} b_{nk}^{(i)} z^k x^n ,$$

$$G^{(i)} = A^{(i)} + B^{(i)} ,$$

so ist $G^{(i)}(x,1) = \sum_n g_n^{(i)} x^n$, woraus sich $e^{(i)}(n)$ durch $e^{(i)}(n) = g^{(i)}(n)/y_n$ berechnen läßt. $\mu^{(i)}(n)$ kann man aus $G_z^{(i)}(x,1) = \sum_n \mu^{(i)}(n) g_n^{(i)} x^n$ ermitteln.

Für diese erzeugenden Funktionen lassen sich Funktionalgleichungen angeben.

Satz 1.

(i) $A^{(1)} = xz\varphi(x\ (G^{1)}))$

$B^{(1)} = x\varphi(g^{(1)} - x\varphi(B^{(1)}))$ (Meir, Moon [3])

(ii) $A^{(2)} = xz\varphi(x\varphi(A^{(2)} + x\varphi(G^{(2)})))$

$B^{(2)} = x\varphi(G^{(2)}) - x\varphi(B^{(2)} - x\varphi(G^{(2)}) + x\varphi(B^{(2)}))$

(iii) $A^{(3)} = xz\varphi(x\varphi(A^{(3)} + x\varphi(A^{(3)} + x\varphi(G^{(3)}))))$

$B^{(3)} = x\varphi(G^{(3)}) - x\varphi(B^{(3)} - x\varphi(G^{(3)}) + x\varphi(B^{(3)} - x\varphi(G^{(3)}) +$

$+ x\varphi(B^{(3)})))$.

Unter geeigneten Analytizitätsbedingungen an $\varphi(t)$ ergibt eine Singularitätenanalyse, daß $A^{(i)}(x,1)$, $B^{(i)}(x,1)$, $G^{(i)}(x,1)$ einen gemeinsamen Konvergenzradius r_i besitzen, um den eine Entwicklung in Potenzen von $\sqrt{r_i - x}$ besteht. Mit Hilfe des Satzes von Darboux lassen sich daraus asymptotische Formeln für $e^{(i)}(n)$ und $\mu^{(i)}(n)$ ableiten.

Satz 2. $\log e^{(i)}(n) \sim E^{(i)}n$, $\mu^{(i)}(n) \sim S^{(i)}n$.

Für $i = 1$ ist dieses Ergebnis in [3] erzielt worden. Zum Unterschied zum Fall $i = 1$ kann man für $i > 1$ die beiden Funktionalgleichungen für $A^{(i)}$ und $B^{(i)}$ nicht auf eine reduzieren. Die Berechnung der Konstanten $E^{(i)}$ und $S^{(i)}$ soll daher am Beispiel $i = 2$ demonstriert werden. Es bezeichne

$$F_1(x,z,a,b) = xz\varphi(x\varphi(a + x\varphi(a + b))) - a$$

$$F_2(x,z,a,b) = x\varphi(a + b) - x\varphi(b - x\varphi(a + b) + x\varphi(b)) - b$$

und $r > 0$, α, β die Lösung des Gleichungssystems

$$F_1(x,1,a,b) = 0 \quad , \quad F_2(x,1,a,b) = 0$$

$$(F_{1_a}F_{2_b} - F_{1_b}F_{2_a})(x,1,a,b) = 0 \ .$$

Daraus berechnet man $E^{(2)} = \log r/\rho$, wobei mit $\tau\varphi'(\tau) = \varphi(\tau)$ $\rho = \tau/\varphi(\tau)$ ist, und

$$S^{(2)} = \frac{1}{r}\,\frac{(F1_z(F2_b - F2_a) - F2_z(F1_b - F1_a))(r,1,\alpha,\beta)}{(F1_x(F2_b - F2_a) - F2_x(F1_b - F1_a))(r,1,\alpha,\beta)}\,.$$

$E^{(i)}$ und $S^{(i)}$ erhält man analog, wenn man F_1 und F_2 den Funktionalgleichungen aus Satz 1 entsprechend anpaßt.

Nach Kirschenhofer et al. [1] besitzen die durchschnittliche Anzahl und die durchschnittliche Größe aller u.a. Teilmengen eine analoge Asymptotik, ebenso die Größe der größten 1-unabh. Teilmenge [2]. Die entsprechenden Konstanten sind daher in die folgende Tabelle aufgenommen worden. (Für den binären Baum muß die Asymptotik als $\log e^{(i)}(2n+1) \sim E^{(i)}(2n+1)$ und $\mu^{(i)}(2n+1) \sim S^{(i)}(2n+1)$ verstanden werden, da nur Bäume mit ungerader Anzahl von Knoten auftreten können.)

$E^{(i)}/S^{(i)}$	alle unabh.	3-unabh.	2-unabh.	1-unabh.	größte 1-unabh.
bin. B.	0.505/0.309	0.426/0.401	0.360/0.438	0.223/0.482	-/0.585
eben. B.	0.523/0.333	0.447/0.410	0.374/0.416	0.215/0.5	-/0.618
mark. B.	0.504/0.307	0.446/0.380	0.388/0.393	0.242/0.463	-/0.567

(Die numerischen Werte sind auf drei Stellen gerundet).

[1] P. Kirschenhofer, H. Prodinger and R.F. Tichy, Fibonacci Numbers of Graphs: II, Fibonacci Q. 21 (1983), 219-229.

[2] A. Meir and J.W. Moon, Packing and Covering Constants for Certain Families of Trees, I, J. Graph Th. 1 (1977), 157-174.

[3] A. Meir and J.W. Moon, On Maximal Independent Sets of Nodes in Trees, J. Graph Theory, submitted.