Computing 28, 363—366 (1982)

Computing

Short Communications / Kurze Mitteilungen

A Note on a Result of R. Kemp on R-Tuply Rooted Planted Plane Trees

H. Prodinger, Wien

Received July 23, 1981

Abstract — Zusammenfassung

A Note on a Result of R. Kemp on R-Tuply Rooted Planted Plane Trees. R. Kemp has shown that the average height of r-tuply rooted planted plane trees is

$$\sqrt{\pi n}-\frac{1}{2}(r-2)+O\left(\log(n)/n^{1/2-\varepsilon}\right),\ \varepsilon>0,\ n\to\infty,$$

assuming that all such trees with n nodes are equally likely. We give a quite short proof of this result (with an error term of $O(1)$).

Eine Bemerkung zu einem Resultat von R. Kemp über r-fach gewurzelte Bäume. R. Kemp hat gezeigt, daß die mittlere Höhe von r-fach gewurzelten Bäumen

$$\sqrt{\pi n}-\frac{1}{2}(r-2)+O\left(\log(n)/n^{1/2-\varepsilon}\right),\ \varepsilon>0,\ n\to\infty$$

ist, falls man annimmt, daß alle solchen Bäume mit n Knoten gleich wahrscheinlich sind. Wir geben für dieses Resultat (mit einem Fehler von $O(1)$) einen ziemlich kurzen Beweis.

1. Introduction

A *planted plane tree* is a rooted tree which has been embedded in the plane so that the relative order of subtrees at each branch is part of its structure. Kemp [2] has defined an *r-tuply rooted planted plane tree* to be a planted plane tree, such that the root has degree r (r a fixed parameter $\in \mathbb{N}$).

The *height* of a planted plane tree is defined to be the maximal number of nodes on a path from the root to a leave.

Kemp [2] has shown in a rather lengthy paper that the average height of an r-tuply rooted planted tree with n nodes is given by

$$\sqrt{\pi n}-\frac{1}{2}(r-2)+O\left(\log(n)/n^{1/2-\varepsilon}\right),\ \varepsilon>0,\ n\to\infty \tag{1}$$

0010-485X/82/0028/0363/$01.00

provided that all such trees with n nodes are assumed to be equally likely. In this note we show that the weaker result

$$\sqrt{\pi n}+O(1) \tag{2}$$

can be obtained quite quickly.

First note that the existence of the root is rather superficial and uncomfortable for the computations. Dropping this extra node, Kemp's result reads:

The average height of an r-tuple of planted plane trees with together n nodes is

$$\sqrt{\pi n}-\frac{r}{2}+O(\log(n)/n^{1/2-\varepsilon}),\ \varepsilon>0,\ n\to\infty. \tag{3}$$

If the height of a planted plane tree t is denoted by $h(t)$, the height of an r-tuple $t=(t_1,\ldots,t_r)$ of trees is defined by

$$h_r(t):=\max\{h(t_1),\ldots,h(t_r)\}. \tag{4}$$

We define a further notion $h_{r,1}$ of "height" as follows (this is inspired by [3]):

$$h_{r,1}(t):=\max\{h(t_1),h(t_2)+1,\ldots,h(t_r)+r-1\}. \tag{5}$$

We observe that the rather obvious estimate

$$h_{r,1}(t)-(r-1)\le h_r(t)\le h_{r,1}(t) \tag{6}$$

holds for all *r-tuples t*. Thus we have similar inequalities for the averages $\overline{h_r}(n)$, $\overline{h_{r,1}}(n)$ of r-tuples of n nodes:

$$\overline{h_{r,1}}(n)-(r-1)\le\overline{h_r}(n)\le\overline{h_{r,1}}(n) \tag{7}$$

and

$$h_r(n)=h_{r,1}(n)+O(1). \tag{8}$$

So if we show

Theorem 1: *The average "$h_{r,1}$-height" of an r-tuple with $n+r$ nodes is*

$$\overline{h_{r,1}}(n+r)=\sqrt{\pi n}+O(1),\ n\to\infty,$$

we are done.

2. The Average $h_{r,1}$-Height of R-Tuples of Trees

Note that the number $T_{n,r}$ of r-tuples of trees with together n nodes is given by

$$T_{n,r}=\frac{r}{n}\binom{2n-r-1}{n-1}; \tag{9}$$

the generating function of $\{T_{n,r}\}_{n\ge1}$ is just

$$z^r C^r(z)=\sum_{n\ge0} T_{n,r}z^n \tag{10}$$

with $C(z)=(1-u)/2z$, where $u=(1-4z)^{1/2}$ (see [2]).

It is known [1] that the generating function $A_h(z)$ of trees with height $\leq h$ is

$$A_h(z) = 2z \frac{(1+u)^h - (1-u)^h}{(1+u)^{h+1} - (1-u)^{h+1}}. \tag{11}$$

Thus the generating function $D_h(z)$ of r-tuples with $h_{r,1}$-height $\leq h$ is $(h \geq r)$

$$\begin{aligned} D_h(z) &= A_h(z) \cdot A_{h-1}(z) \ldots A_{h-(r-1)}(z) \\ &= 2^r z^r \frac{(1+u)^{h-(r-1)} - (1-u)^{h-(r-1)}}{(1+u)^{h+1} - (1-u)^{h+1}}. \end{aligned} \tag{12}$$

Observing $(1+u)^{-1} = \frac{1}{2} C(z)$ and $(1-u)/(1+u) = z\, C^2(z)$, we find

$$D_h = z^r C^r \frac{1 - z^{h-r+1} C^{2(h-r+1)}}{1 - z^{h+1} C^{2(h+1)}} \tag{13}$$

and therefore $E_h(z) = \sum_{h \geq 1} E_{n,h} z^n$, the generating function of r-tuples with $h_{r,1}$-height $> h$ fulfills by (10) and (13)

$$\begin{aligned} E_h &= z^r C^r - D_h = z^r C^r \frac{z^{h+1-r} C^{2(h+1-r)} - z^{h+1} C^{2(h+1)}}{1 - z^{h+1} C^{2(h+1)}}; \ (h \geq r-1) \\ E_{h-1} &= \sum_{\lambda \geq 1} z^{h\lambda} C^{2h\lambda - r} - \sum_{\lambda \geq 1} z^{h\lambda + r} C^{2h\lambda + r}. \end{aligned} \tag{14}$$

Since the coefficients of the powers of $C(z)$ are well known (compare (9) and (10)) we have

$$E_{n+r,h-1} = \sum_{\lambda \geq 1} \left[\binom{2n+r-1}{n-h\lambda} - \binom{2n+r-1}{n-h\lambda-1} - \binom{2n+r-1}{n+r-h\lambda} + \binom{2n+r-1}{n+r-h\lambda-1} \right]. \tag{15}$$

Now the average $h_{r,1}$-height is

$$\overline{h_{r,1}}(n+r) = T_{n+r,r}^{-1} \cdot \left\{ \sum_{h \geq r-1} E_{n+r,h} + (r-1)\, T_{n+r,r} \right\}; \tag{16}$$

if we replace $E_{n+r,h}$ for $h = 0, \ldots, r-2$ by the righthand side of (15), and start the sum in (16) by $h = 0$, we make again an error of $O(1)$, yielding

$$\overline{h_{r,1}}(n+r) = T_{n+r,r}^{-1} \cdot \xi + O(1) \tag{17}$$

with

$$\xi = \sum_{k \geq 1} d(k) \left[\binom{2n+r-1}{n-k} - \binom{2n+r-1}{n-k-1} - \binom{2n+r-1}{n-k+r} + \binom{2n+r-1}{n-k+r-1} \right] \tag{18}$$

where $d(k)$ denotes the number of divisors of k.

The following approximation of the binomial coefficients is well known [1], [2]: $(\varepsilon > 0)$

$$\binom{2n}{n+a-k}=\binom{2n}{n}\cdot\begin{cases}\exp(-k^2/n)\,[f_a(n,k)+O(n^{-2+\varepsilon})] & |k-a|<n^{1/2+\varepsilon}\\ O(\exp(-n^{2\varepsilon})) & \text{otherwise}\end{cases} \tag{19}$$

with

$$f_a(n,k)=1-\frac{a^2}{n}+\left[\frac{2a}{n}-\frac{2a^3+a}{n^2}\right]k+\frac{4a^2+1}{2n^2}k^2+\frac{4a^3+5a}{3n^3}k^3-\frac{1}{6n^3}k^4-\frac{a}{3n^4}k^5. \tag{20}$$

Now we use (19) where n is replaced by $N:=n+\frac{r-1}{2}$ and $\binom{2n}{n}$ by $\pi^{-1/2}\,2^{2n}\,n^{-1/2}$ $(1+O(n^{-1}))$:

We do this for the 4 summands in (18) with

$$a=\frac{r+1}{2};\ \frac{r-1}{2};\ -\frac{r-1}{2};\ -\frac{r+1}{2}$$

yielding

$$\xi=\frac{2^{2n+r-1}}{\sqrt{\pi n}}\left(1+O\left(\frac{1}{n}\right)\right)\sum_{k\geq 1}d(k)\,e^{-k^2/N}\left[\frac{-2r}{N}+\frac{4rk^2}{N^2}+O\left(\log(N)/N^{-3/2+\varepsilon}\right)\right]. \tag{21}$$

Now

$$T_{n+r,r}=\frac{r}{n}\cdot\frac{1}{\sqrt{\pi n}}\cdot 2^{2n+r-1}\left(1+O\left(\frac{1}{n}\right)\right) \tag{22}$$

and

$$\sum_{k\geq 1}d(k)\,e^{-k^2/N}\left[-\frac{2}{N}+\frac{4k^2}{N^2}+O\left(\log(N)/N^{-3/2+\varepsilon}\right)\right]=\left(\frac{\pi}{N}\right)^{1/2}+O\left(\frac{1}{N}\right) \tag{23}$$

(compare [1]), so that finally

$$\overline{h_{r,1}}\,(n+r)=\sqrt{\pi n}+O(1), \tag{24}$$

as desired.

References

[1] De Bruijn, N. G., Knuth, D. E., Rice, S. O.: The average height of planted plane trees, in: Graph theory and computing (Read, R. C., ed.), pp. 15—22. New York-London: Academic Press 1972.

[2] Kemp, R.: The average height of r-tuply rooted planted plane trees. Computing *25*, 209—232 (1980).

[3] Kirschenhofer, P., Prodinger, H.: On the average height of monotonically labelled binary trees, presented at: 6th Hungarian Colloquium on Combinatorics, Eger, 1981.

H. Prodinger
Institut für Algebra und Diskrete Mathematik
Technische Universität Wien
Gusshausstrasse 27-29
A-1040 Wien
Austria