Comment on “A note on the distribution of the three types of nodes in uniform binary trees” by Helmut Prodinger

C. Krattenthaler

Institut für Mathematik der Universität Wien,
Strudlhofgasse 4, A-1090 Wien, Austria.
e-mail: KRATT@Pap.Univie.Ac.At
WWW: `http://radon.mat.univie.ac.at/People/kratt`

Abstract. It is shown that the summations in Prodinger's paper “A note on the distribution of the three types of nodes in uniform binary trees” follow from standard hypergeometric summation formulas. Second, a bijection is given which directly proves the enumeration results of Prodinger's paper.

In [3], Prodinger showed that Mahmoud's [2] expressions for the number of uniform binary trees with a fixed number of internal nodes having ℓ internal nodes as successors, $\ell \in \{0,1,2\}$ fixed, can in fact be simplified to closed form expressions. Let N_ℓ denote this number, $\ell \in \{0,1,2\}$. Mahmoud's first two expressions are

$$N_0 = \sum_{i=0}^{n-j} (-1)^{n-j-i} \frac{1}{(i+1)} \binom{2i}{i} \binom{i+1}{n-i} \binom{n-i}{j}, \tag{1}$$

$$N_1 = \sum_{i=0}^{n-j} (-1)^{n-j-i} 2^{n-i} \frac{1}{(i+1)} \binom{2i}{i} \binom{n-1}{i-1} \binom{n-i}{j}. \tag{2}$$

Prodinger uses Zeilberger's algorithm to evaluate these two binomial sums. The purpose of this comment is, first, to point out that standard hypergeometric summations explain these two simplifications, and, second, to show how the simplified formulas can be obtained directly, using a simple bijection.

In terms of the usual hypergeometric notation

$$ {}_rF_s\left[\begin{matrix} a_1,\dots,a_r \\ b_1,\dots,b_s \end{matrix}; z\right] = \sum_{k=0}^{\infty} \frac{(a_1)_k \cdots (a_r)_k}{k!\,(b_1)_k \cdots (b_s)_k} z^k \ ,$$

the shifted factorial $(a)_k$ being defined by $(a)_k := a(a+1)\cdots(a+k-1)$, $k \ge 1$, $(a)_0 := 1$, the sum in (1), with the order of summation reversed (i.e., i is replaced by $n-j-i$) reads

$$\frac{(2n-2j)!}{j!\,(n-2j+1)!\,(n-j)!}\, {}_2F_1\left[\begin{matrix} -\frac{1}{2}+j-\frac{n}{2},\, j-\frac{n}{2} \\ \frac{1}{2}+j-n \end{matrix}; 1\right].$$

The ${}_2F_1$ series can be summed by means of the Chu–Vandermonde summation (cf. [4, (1.7.7); Appendix (III.4)]),

$$ {}_2F_1\left[\begin{matrix} a, -m \\ c \end{matrix}; 1\right] = \frac{(c-a)_m}{(c)_m}, \tag{3} $$

where m is a nonnegative integer. Little simplification then leads to Prodinger's result (see [3, Theorem 2])

$$ N_0 = 2^{n+1-2j} \frac{(n-1)!}{j!\,(j-1)!\,(n+1-2j)!}. \tag{4} $$

The sum in (2), again with order of summation reversed, in hypergeometric terms can be written as

$$ \frac{2^j\,(n-1)!\;(2n-2j)!}{j!\;(n-j-1)!\;(n-j)!\;(n-j+1)!}\,{}_2F_1\left[\begin{matrix} -1+j-n, 1+j-n \\ \frac{1}{2}+j-n \end{matrix}; \frac{1}{2}\right]. $$

This ${}_2F_1$ series can be summed by means of Gauß second ${}_2F_1$ summation (cf. [4, (1.7.1.9); Appendix (III.6)])

$$ {}_2F_1\left[\begin{matrix} 2a, 2b \\ \frac{1}{2}+a+b \end{matrix}; \frac{1}{2}\right] = \frac{\Gamma\left(\frac{1}{2}\right)\,\Gamma\left(\frac{1}{2}+a+b\right)}{\Gamma\left(\frac{1}{2}+a\right)\Gamma\left(\frac{1}{2}+b\right)}. \tag{5} $$

Some simplification, using well-known properties of the gamma function, finally gives

$$ N_1 = \chi(n+j \equiv 1 \mod 2)\, 2^j \frac{(n-1)!}{j!\,\left(\frac{n}{2}-\frac{j}{2}-\frac{1}{2}\right)!\,\left(\frac{n}{2}-\frac{j}{2}+\frac{1}{2}\right)!}. \tag{6} $$

Here, $\chi(\mathcal{A})$=1 if $\mathcal{A}$ is true and $\chi(\mathcal{A})$=0 otherwise. Formula (6) is easily seen to be a uniform way of writing Prodinger's respective formulas (see [3, Theorem 2]).

In passing, we note that the summations that lead to Prodinger's expressions [3, Theorem 3] for the factorial moments related to N_0, N_1, N_2 are also special cases of Chu–Vandermonde summation (3).

Finally, we show how to bijectively prove (4) and (6). In view of Prodinger's [3] observation that, in fact, the three enumeration problems of finding N_0, N_1, N_2, respectively, are basically the same, it suffices to give a bijective proof of (4).

The number N_0 that we want to find is actually the number of binary trees of size n with j leaves. We are going to construct a bijection between the set of all these trees and Motzkin paths from $(0,0)$ to $(n-1,0)$ (i.e., lattice paths with up-steps $(1,1)$, down-steps $(1,-1)$, and level-steps $(1,0)$, which start in $(0,0)$, end in $(n-1,0)$, and never go below the x-axis) with exactly $j-1$ up-steps, exactly $j-1$ down-steps, and with *two different types* of level-steps (the total number of level steps being necessarily $n+1-2j$).

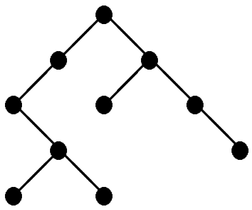

Figure 1

Consider such a binary tree, see Figure 1. Starting from the root and heading left, move along the "outside" of the tree, until returning to the root from the far-right side of the tree. Sequentially, form a Motzkin path as follows. Whenever an edge is met that was not yet considered before, translate it into a step of the Motzkin path according to the following rules:

(1) If this edge is a left edge, and there is a right edge incident to the upper node of the left edge, then translate it into an up-step.
(2) If this edge is a right edge, and there is a left edge incident to the upper node of the right edge, then translate it into a down-step.
(3) If this edge is a left edge, and there is no right edge incident to the upper node of the left edge, then translate it into a level-step labelled l.
(4) If this edge is a right edge, and there is no left edge incident to the upper node of the right edge, then translate it into a level-step labelled r.

Figure 2 shows the image of the binary tree in Figure 1 under this mapping.

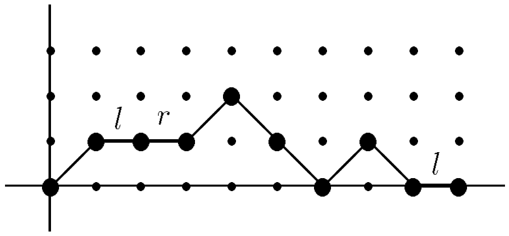

Figure 2

It is straight-forward to check that this mapping is a bijection between the above described sets. Using the reflection principle (see e.g. [1, p. 22]), the number of these Motzkin paths is easily obtained as

$$2^{n+1-2j}\frac{(n-1)!}{(j-1)!\,(j-1)!\,(n+1-2j)!}-2^{n+1-2j}\frac{(n-1)!}{j!\,(j-2)!\,(n+1-2j)!},$$

which easily simplifies to the expression in (4).

Of course, one suspects that the results (4), (6), and the simple bijection that explains them, should already be known for long. However, I was not able to find them in earlier literature.

References

1. L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, Holland, 1974.
2. H. M. Mahmoud, *The joint distribution of the three types of nodes in uniform binary trees*, Algorithmica **13** (1995), 313–323.
3. H. Prodinger, *A note on the distribution of the three types of nodes in uniform binary trees*, Séminaire Lotharingien Combin. **38** (1996), paper B38b.
4. L. J. Slater, *Generalized hypergeometric functions*, Cambridge University Press, Cambridge, 1966.

Institut für Mathematik der Universität Wien, Strudlhofgasse 4, A-1090 Wien, Austria.