CAYLEY'S FORMULA: A PAGE FROM THE BOOK

ARNON AVRON AND NACHUM DERSHOWITZ

Abstract. A simple proof of Cayley's formula is given.

We give a short elementary proof of Cayley's famous formula for the enumeration T_n of free, unrooted trees with $n \geq 1$ labeled nodes. We first count $F_{n,k}$, the number of n-node forests composed of k rooted, directed trees, $1 \leq k \leq n$. For the history of the formula, including Jim Pitman's use of directed forests, see [1, pp. 201–206].

The crux of the proof is a simple double counting. There are two equivalent ways of counting the number of k-tree forests with one designated internal (non-root) node, showing that, for all $k = 1, \ldots, n-1$,

$$(n-k)F_{n,k} \quad = \quad knF_{n,k+1} \; . \qquad (*)$$

- For the left side of (*): Consider one of the $F_{n,k}$ forests with k trees. Designate any one of its $n-k$ internal nodes.
- For the right side: Consider one of the $F_{n,k+1}$ forests with $k+1$ trees. Choose any one of the n nodes, and hang from it any one of the k trees not containing that node. The root of that grafted subtree is the designated internal node.

Iterating (*) $n-1$ times gives:

$$F_{n,1} = \tfrac{1}{n-1}\, nF_{n,2} = \tfrac{1}{n-1}\tfrac{2}{n-2}\, n^2 F_{n,3} = \cdots = \tfrac{1}{n-1}\tfrac{2}{n-2}\cdots\tfrac{n-1}{1}\, n^{n-1}F_{n,n} \; .$$

The k and $n-k$ factors all cancel each other out. Because there is precisely one way of turning n nodes into n distinct trees (each root being a whole tree), we have $F_{n,n} = 1$. Thus, the number $F_{n,1}$ of n-node rooted trees is n^{n-1}. Since any of the n nodes in a tree can be the root, $F_{n,1} = nT_n$, and Cayley's formula, $T_n = n^{n-2}$, follows.

Applying (*) only $n-k$ times, yields $F_{n,k} = \binom{n}{k} kn^{n-k-1}$, for $k = 1, \ldots, n$.

Alternatively, the relation $(k+1)R_{n,k} = knR_{n,k+1}$ for the number $R_{n,k}$ of n-node forests with k designated roots leads to $R_{n,k} = kn^{n-k-1}$ and to $T_n = R_{n,1} = n^{n-2}$.

As a final remark, there are $(n+1)^{n-1}$ rooted trees with $n+1$ nodes that all share the same root. Each corresponds to a rooted forest with n nodes—just chop off the root node. Therefore, the limit of the ratio of rooted labeled forests to rooted labeled trees, as their size grows, is $\lim_{n\to\infty}(n+1)^{n-1}/n^{n-1} = e$.

Acknowledgment. We thank everyone who read a draft and especially Ed Reingold for his suggestions.

References

[1] Martin Aigner and Günter M. Ziegler, *Proofs from THE BOOK*, 4th ed., Springer-Verlag, Berlin, 2010.