A formula for some integer sequences that can be described by generating trees

Chams Lahlou

January 27, 2019

Irwin [1] proposed to compute sequence A002449 by formula

$$\texttt{A002449}(n+2) = \sum_{i_1=1}^{2} \sum_{i_2=1}^{2i_1} \cdots \sum_{i_{n-1}=1}^{2i_{n-2}} \sum_{i_n=1}^{2i_{n-1}} 2i_n, \text{ for } n \geq 1. \tag{1}$$

We show that (1) can be generalised to some sequences that can be described by an enumerative technique known as "generating tree". We give a proof for the general formula, which establishes the correctness of Irwin's conjecture.

1 A generating tree for A002449

West showed in [2] that generating trees are useful for counting because when a sequence is represented by a generating tree the number of nodes on level n of the tree is the n^{th} element of the sequence.

Definition 1 (West [2]) *A generating tree is a rooted, labelled tree having the property that the labels of the set of children of each node x can be determined from the label of x itself.*

In order to describe a sequence by a generating tree, it is then necessarily to understand how the objects counted by the sequence can be enumerated recursively. By definition, A002449(n) is the "number of different types of binary trees of height n".

First, observe that trees of height $n+1$ can be generated from the ones of height n by adding children to the leaves of the last level, i.e. to the *terminal leaves*. For example, trees of height 0 to 3 are obtained as follows (see Figure 1): tree (b) is obtained by adding two nodes to tree (a) ; trees (c) and (d) are obtained by adding respectively 2 and 4 children to tree (b); similarly, trees (e) and (f) are deduced from tree (c), and trees (g)-(j) from tree (d).

Second, it is important to notice that *two trees have the same type if they have the same number of nodes at each level.* Otherwise, we would have more than A002449(3) trees of height 3. This fact is illustrated in Figure 2: tree (h') could have been derived from tree (d), but it has the same type as tree (h). Thus, without loss of generality, children can be added from left to right.

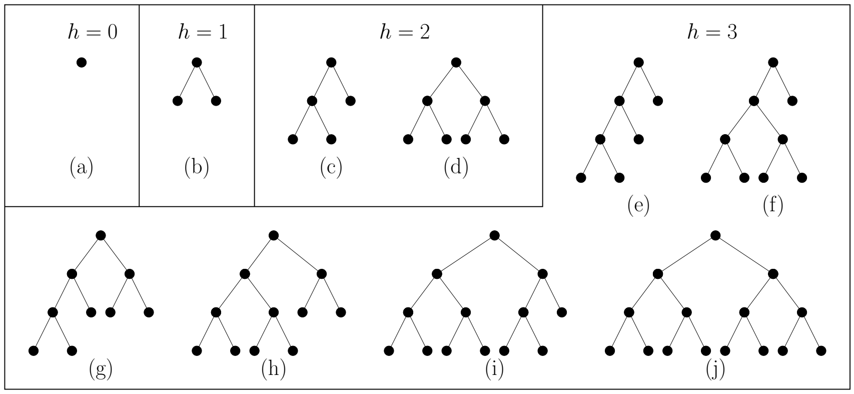

Figure 1: A002449: the binary trees of height 0, 1, 2 and 3.

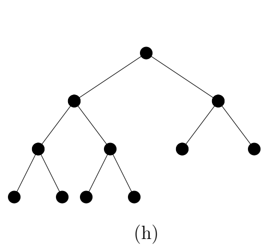

Figure 2: A002449: two binary trees of height 3 which have the same type.

As a consequence, we can enumerate all trees of A002449 by repeating the following rule: duplicate each tree with k terminal leaves k times, and add $2, 4, \ldots, 2k$ children to the terminal leaves of the first, second, ..., k^{th} copy, respectively. This enumeration can be described by a tree itself, where each node represents a tree of A002449, as in Figure 3. This tree can be turned into a generating tree by labelling each node by the number of its children. Using notation of West [2] it is defined as follows:

Root: (1)
Rule: $(k) \rightarrow (2)(4)\cdots(2k)$

The generating tree is illustrated in Figure 4 for heights 0 to 4: the root representing tree (a) has label (1) since it has one child; its child, which represents tree (b), has label (2) since it has two children (c) and (d); and so on.

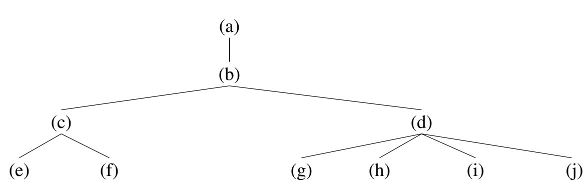

Figure 3: A002449 represented as a tree.

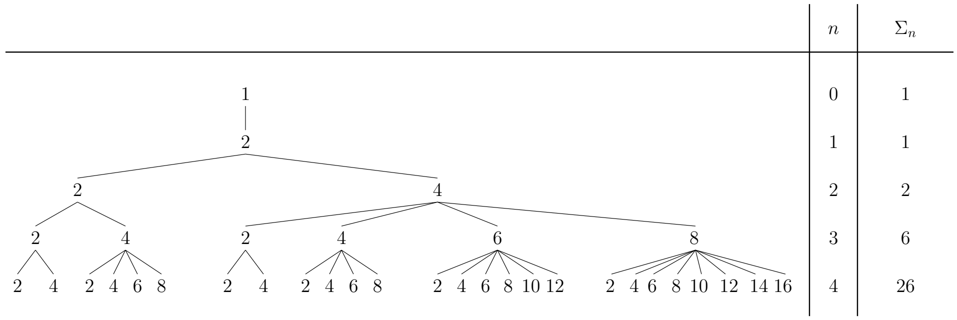

Figure 4: The generating tree of sequence A002449, for $0 \leq n \leq 4$.

2 A general formula

It is possible to generalise Formula (1) for some sequences:

Theorem 1 *If A is an integer sequence which can be described by the generating tree*

Root: (l_1)
Rule: $(k) \rightarrow (f(1))(f(2)) \cdots (f(k))$,

where f is a function from $\mathbb{N}$ to $\mathbb{N}$, then the n^{th} element of A is obtained by

$$A(n) = \begin{cases} 1 & \text{if } n = 0, \\ l_1 & \text{if } n = 1, \\ \sum_{i_2=1}^{l_1} \sum_{i_3=1}^{f(i_2)} \cdots \sum_{i_n=1}^{f(i_{n-1})} f(i_n) & \text{otherwise.} \end{cases} \tag{2}$$

Proof. Let T denote the generating tree associated with A, and let Σ_n be the number of nodes on level n of T. Because $A(n) = \Sigma_n$, as explained above, we shall determine the number of nodes on each level of T.

Since T has one root, with l_1 children, we have $\Sigma_0 = 1$ and $\Sigma_1 = l_1$, and thus (2) is true for $0 \leq n \leq 1$. Now, observe that, by definition of T, the value of Σ_n is the sum of the labels on level $n-1$. Thus, it suffices to prove the proposition *"Terms in the sum of* (2) *are the labels on level* $n-1$*"*, which is easily done by induction. The succession rule of T implies that the nodes on level 1 are labelled $(f(1))$, $(f(2))$, $\cdots, (f(l_1))$. Since $\sum_{i_2=1}^{l_1} f(i_2) = f(1) + f(2) + \cdots + f(l_1)$, the proposition is true for $n = 2$.

Assuming that it is also true for $2 \leq k \leq n$, by hypothesis of induction each term $f(i_n)$ of the sum

$$\sum_{i_2=1}^{l_1} \sum_{i_3=1}^{f(i_2)} \cdots \sum_{i_n=1}^{f(i_{n-1})} f(i_n)$$

is the label of a node on level $n-1$. According to the succession rule, each node generates $f(i_n)$ nodes on level n that have labels $(f(1)), (f(2)), \ldots, (f(f(i_n)))$ respectively. Hence, the sum of labels on level n is

$$\sum_{i_2=1}^{l_1} \sum_{i_3=1}^{f(i_2)} \cdots \sum_{i_n=1}^{f(i_{n-1})} \sum_{i_{n+1}=1}^{f(i_n)} f(i_{n+1}),$$

where each term is the label of a node on level n. Hence the proposition is true for $n \geq 2$, which completes the proof. $\square$

The proof of Formula (1) is then straightforward: renumbering the indices gives $\texttt{A002449}(n) = \sum_{i_3=1}^{2} \sum_{i_4=1}^{2i_3} \cdots \sum_{i_{n-1}=1}^{2i_{n-2}} \sum_{i_n=1}^{2i_{n-1}} 2i_n$, for $n \geq 3$, which is Formula (2) when we set $l_1 = 1$, $f(k) = 2k$ and $n \geq 3$.

For two well-known sequences the parameters of (2) are:

OEIS sequence	Generating tree	l_1	$f(k)$
A000045 (Fibonacci numbers)	Root: (1) Rules: $(1) \to (2)$; $(2) \to (1)(2)$	1	$3-k$
A000108 (Catalan numbers)	Root: (1) Rule: $(k) \to (2)(3)\cdots(k+1)$	1	$k+1$

Conversely, setting parameters of (2) define sequences. In some cases, they seem to be known, as for example:

l_1	$f(k)$	Generating tree	OEIS sequence verified for the given table
1	$4-k$	Root: (1) Rules: $(1) \to (3)$; $(2) \to (2)(3)$; $(3) \to (1)(2)(3)$	A077998
2	$4-k$	Root: (2) Rules: $(1) \to (3)$; $(2) \to (2)(3)$; $(3) \to (1)(2)(3)$	A106805
3	$5-k$	Root: (3) Rules: $(1) \to (4)$; $(2) \to (3)(4)$; $(3) \to (2)(3)(4)$ $(4) \to (1)(2)(3)(4)$	A076264
4	$5-k$	Root: (4) Rules: $(1) \to (4)$; $(2) \to (3)(4)$; $(3) \to (2)(3)(4)$ $(4) \to (1)(2)(3)(4)$	A006357
1	$k+2$	Root: (1) Rule: $(k) \to (3)(4)\cdots(k+2)$	A001764
1	$\lfloor \frac{k}{2} \rfloor + 2$	Root: (1) Rule: $(k) \to (2)(3)(3)\cdots(i)(i)\cdots(\lfloor \frac{k}{2} \rfloor + 2)$	A001519
2	$2\lfloor \frac{k}{2} \rfloor + 2$	Root: (2) Rule: $(k) \to (2)(4)(4)\cdots(i)(i)\cdots(2\lfloor \frac{k}{2} \rfloor + 2)$	A006318
2	$\lfloor \frac{k}{3} \rfloor + 3$	Root: (2) Rule: $(k) \to (3)(3)(4)(4)(4)\cdots(i)(i)(i)\cdots(\lfloor \frac{k}{3} \rfloor + 3)$	A006012
3	$3\lfloor \frac{k}{3} \rfloor + 3$	Root: (3) Rule: $(k) \to (3)(3)(6)(6)(6)\cdots(i)(i)(i)\cdots(3\lfloor \frac{k}{3} \rfloor + 3)$	A047891
3	$\lfloor \sqrt{k} \rfloor \lfloor \sqrt{k} \rfloor + 1$	Root: (3) Rule: $(k) \to (2)\cdots(\lfloor \sqrt{k} \rfloor \lfloor \sqrt{k} \rfloor + 1)$	A003945

References

[1] Benedict W. J. Irwin, The On-Line Encyclopedia of Integer Sequences. Published electronically at http://oeis.org/A002449, Nov 16 2016.

[2] Julian West, *Generating trees and forbidden subsequences.* Discrete Mathematics, 157, p. 363–374, 1996.