ON THE SUM OF THE SIZES OF BINARY SUBTREES OF A
PERFECT BINARY TREE

CYRIL BANDERIER, PERSONAL NOTE, 8, JUNE 2000
ALGORITHMS PROJECT INRIA
HTTP://ALGO.INRIA.FR/BANDERIER/

This note answers to Florent Madeleine’s question (private communication) “Can the sum of the sizes of binary subtrees of a perfect binary tree of size \( n \) be polynomially bounded?”

I will show that the answer is no.

A binary tree is a tree whose each internal node have

- either two child which are internal nodes,
- or two child which are leaves.

Let \( b_k \) be the number of binary trees with \( k \) internal nodes. The binary trees satisfies the following recursive definition

\[
B = L \text{ on } B \times N \times B
\]

(a binary tree is either a leaf \( L \), or a node \( N \) with two binary subtrees \( B \)) hence, giving weight 0 to leaves and 1 to internal nodes, one gets a functional equation for the generating function \( B(z) \):

\[
B(z) = 1 + zB(z)^2
\]

and solving it gives the closed form formula

\[
B(z) = \frac{1 - \sqrt{1 - 4z}}{2z}
\]

The smallest binary tree, which consist of a “root which is also a leaf”, has height 0. Let \( B_h(x) = \sum_{k \geq 0} b_{k,h} x^k \) be the generating function (in fact that’s a polynomial!) for binary trees of height \( \leq h \) (and where \( n \) codes the size of the tree, i.e. its number of internal nodes).

Then, one has the following recurrence:

\[
B_h(x) = 1 + B_{h-1}(x)^2 \quad B_0(x) = 1.
\]

A perfect binary tree (some authors say also “complete” binary tree) is a binary tree for which all leaves are at the same height, say \( h \). Consequently, a perfect binary tree has \( 2^h - 1 \) internal nodes.

The number of binary subtrees of this tree is the number of binary trees of height \( \leq h \), that is \( B_h(1) \). The sum of the size of the binary subtrees \( \mathcal{A} \) is given by

\[
\sum_{\mathcal{A} \subseteq \mathcal{B}(h)} |\mathcal{A}| = \sum_{k \geq 0} kb_{k,h} = B_h(1).
\]

The asymptotics of the \( b_{k,h} \)'s is a non trivial problem, nevertheless it has been solved by Flajolet & Odlyzko in 1984. It is related to the behavior of the Mandelbrot
fractal $z_{j+1} = z_j^2 + 1$ at $z_0 = 1$. The authors give
\[
\max_{k \geq 0} b_{h,k} \sim 2^{-h/2} \exp(2^h 0.407354 \ldots) 0.685517 \ldots
\]
Taking $n := 2^h - 1$ (input size, i.e., size of the perfect binary tree) leads to the upper bound
\[
B'_h(1) \leq n \max_{k \geq 0} b_{h,k} = \frac{C \exp(0.407354 \ldots n)}{\sqrt{n}}
\]
In conclusion, $B'_h(1) \approx B_h(1) \approx C_1^{2^n}$ where $C_1 = 1.50283680104975649975293642373 \ldots$
(in fact one has $B'_h(1) \sim C_2 2^h B_h(1)$ where $C_2 = 1.58900495515926 \ldots$)
Finally, one gets the following bounds for the number of subtrees of complete binary tree of size $n$ when $n > 2$
\[
1.5^n < B'_h(1) < 1.51^n.
\]
This excludes any polynomial bound.

The following Maple lines easily compute the first few terms
\[
B := 1; L := \emptyset; L2 := \emptyset;
\]
for $n$ do $B := 1 + x \cdot B^2$; $L := \text{op}(L), \text{subs}(=1,B)$; $L2 := \text{op}(L2), \text{subs}(=1, \text{diff}(B, n))$
do:

Thus, the value for $h \geq 0$ of $B_h(1)$ are: 1, 2, 5, 26, 677, 458330, 210066388901, 44127887745906175987802, 3947270476915529649559703445408348930452791205, 3791862310265920682868235028027893277370233152247388584761734150717768254410341175325352026,

and for $B'_h(1)$: 1, 1, 8, 61, 136, 8766473, 8245941529080, 350851820795115793746961, 3115942957467884941700662098030628486462075222648, 121730849123090662509298880681439040673894356171860018013054502913055217344025410753271773705.

Reference:
http://algo.inria.fr/flajolet/Publications/FlOd84.ps.gz