Height-Ratio-Balanced Trees

Gaston H. Gonnet
Data Structuring Group, Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
Henk J. Olivié
Unit for Mathematics and Computer Science, IHAM, Paardenmarkt 94, B–2000, Antwerp, Belgium
Derick Wood
Data Structuring Group, Department of Computer Science, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

We introduce a new class of binary search trees, the height-ratio-balanced binary search trees, as the height based analogy of weight (-ratio) balanced binary search trees. They form a proper subclass of the class of binary search trees, but not a logarithmic one, indeed an n node height-ratio balanced tree of order α, $0 \leq \alpha \leq \frac{1}{3}$, has a worst case height of $\mu e^{\mu + O(1)}$, where $\mu = \{-2\ln(\alpha/(1-a))\ln(n)\}^{1/2}$. This result indicates that these naturally defined trees should not be used to implement the DICTIONARY operations, in practical situations.

INTRODUCTION

Since the AVL or height-balanced binary search trees were introduced by Adelson-Velskii and Landis[1] in 1962, there have been surprisingly few new classes of 'logarithmically-balanced' search trees introduced. The only ones known to the authors are the weight-balanced trees,[2] k-height-balanced trees,[3] one-sided height-balanced trees,[4] half-balanced trees,[5] and α-balanced trees.[6] All these classes allow updating to be carried out in $O(\log n)$ time, when the starting tree has n nodes and the resulting tree is in the same class. Furthermore searching a tree of n nodes in any of these classes is also an $O(\log n)$ time operation. Typically whenever these so called DICTIONARY operations[7] need to be implemented with $O(\log n)$ time complexity, one of these classes of trees is chosen (typically the AVL-trees).

In each of these classes of trees mentioned above,[1–6] the notion of a balanced node is defined which depends on either the height or the weight of the node's subtrees (additionally the trees in Refs 5 and 6 require the shortest path to a leaf from the node). Hence a natural question arises, namely, when can the roles of height and weight be interchanged leaving a logarithmically-balanced class of trees. This paper considers the weight-balanced trees of Nievergelt and Reingold[2] as such a candidate.

We prove that these height-ratio-balanced trees give a non-logarithmic class of trees, but of more interest is the worst case height of a height-ratio-balanced tree of n nodes: $h = \mu e^{\mu + O(1)}$, where $\mu = \{-2\ln(\alpha/(1-\alpha))\ln(n)\}^{1/2}$.

HEIGHT-RATIO-BALANCED TREES

Before introducing our central notion we require some preliminary definitions.

A *binary tree of n nodes*, T_n is the empty tree T_0 if $n = 0$ and otherwise is a triple (T_l, u, T_r) where $l + r + 1 = n$, T_l and T_r are binary trees, u is the *root* of T_n, T_l is the left subtree of u and T_r is the right subtree of u. For the purposes of this paper we define the *height* of a tree T_n, denoted by $ht(T_n)$, as follows:

$$ht(T_n) = 1 \text{ if } n = 0 \text{ and } 1 + \max(ht(T_l), ht(T_r)) \text{ otherwise.}$$

The height is defined as one larger than usual to simplify the balancing formula.

The particular balancing measure we study is captured in the following definition.

Definition

Let $n \geq 1$ and $T_n = (T_l, u, T_r)$. Then the *balance of u*, denoted by $\beta(u)$, is defined by

$$\beta(u) = \frac{ht(T_l)}{ht(T_l) + ht(T_r)}$$

This in turn leads to our central notion:

Definition

Let α be a number, $0 \leq \alpha \leq \frac{1}{2}$. A tree T_n is said to be *height-ratio-balanced of order* α, α-*hrb*, if either $n = 0$ or $n \geq 1$, $T_n = (T_l, u, T_r)$, $\alpha \leq \beta(u) \leq 1 - \alpha$ and both T_l and T_r are α-*hrb*.

With any notion of balance it must be demonstrated that there is a tree of every size satisfying the balancing criterion. In the present case we do this in two stages, we first show that not all values of α in $[0, \frac{1}{2}]$ are viable and second we show that for viable α there exist trees of every size. Observe that by definition, the class of 0-*hrb*-trees equals the class of binary trees, and that not all α are viable, that is similar to the case of weight-balanced trees[2] there is a 'gap' lemma.

Lemma 1

For all α, $\frac{1}{3} < \alpha \leq \frac{1}{2}$, the class of α-*hrb* trees does not contain any trees with an even number of nodes.

Proof. Let T_n be α-*hrb*, for some α, $\frac{1}{3} < \alpha \leq \frac{1}{2}$. This implies that $\alpha \leq \beta(u) \leq 1 - \alpha$, for all nodes u in T_n. That is, letting x be the height of u's left subtree and y the height of u's right subtree, $\alpha \leq x/(x+y) \leq 1 - \alpha$. Since $\alpha > \frac{1}{3}$, this

CCC–0010–4620/83/0026–0106 $01.50

implies $x < 2y < 4x$, must have integral solutions for y for all integral values of $x \geq 1$. In particular $1 < 2y < 4$ implies $y = 1$, that is $\beta(u) = \frac{1}{2}$. But if n is even there must be at least one node with both an empty subtree and a non-empty one, that is with balance at most $\frac{1}{3}$. This proves the result. ∎

Note that this gap result is not as strong as the one of Ref. 2, since in their case, there are only completely balanced trees in the gap. Our result says that there are no trees in the gap with n even. Because of Lemma 1 we will only treat viable α in the remainder of the paper, that is $0 \leq \alpha \leq \frac{1}{3}$.

Lemma 2

For all α, $0 \leq \alpha \leq \frac{1}{3}$ and for all $n \geq 0$, there exists a T_n which is α-*hrb*.

Proof. Let T_n be a minimal height tree with n nodes, then for every node u in T_n, the difference between the height of u's subtrees is at most 1. Letting h_l denote the height of the left subtree of u, then $\beta(u) =$ either $\frac{1}{2}$ or $h_l/(2h_l + 1)$, without any loss of generality. In the latter case $\beta(u) \geq \frac{1}{3}$ implies $h_l \geq 1$, which is trivially true. Hence in both cases $\frac{1}{3} \leq \beta(u) \leq \frac{1}{2}$, as desired. ∎

To demonstrate that the class of α-*hrb* trees is, indeed, balanced, we need to prove that insertions and deletions can be performed in $O(ht(T))$ time, for all T in the class, yielding, perhaps by way of some restructuring, a tree T' in the same class. However, because of the worst case analysis of the height, which we now present, this is left to the interested reader.

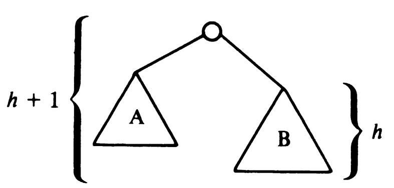
Theorem 3

Let α be viable and T_n be an α-*hrb*-tree, then

$$ht(T_n) \leq \mu\, e^{\mu + O(1)}$$

where $a = \alpha/(1 - \alpha)$ and $\mu = \{-2\ln(a)\ln(n)\}^{1/2}$.

Proof. To prove this theorem we will find the smallest tree (least number of nodes) of a given height. The tree may be represented as



Let $ht(B) \geq ht(A)$. If this tree has the least number of nodes, then B also has the least number of nodes, that is it is in the same class. From the balancing condition we conclude that

$$\frac{ht(B)}{ht(B) + ht(A)} \leq 1 - \alpha$$

or

$$\frac{\alpha}{1 - \alpha} ht(B) \leq ht(A).$$

Letting $a = \alpha/(1 - \alpha)$ and noticing that the height is always an integer

$$ht(A) \geq \lfloor a \cdot ht(B) \rfloor$$

Since the number of nodes for this class is clearly monotone in the height, we will select A to be the smallest possible tree with the least number of nodes, and also in the same class.

Consequently we have a recurrence relation in the minimal number of nodes $N(h)$ of a tree with height h:

$$N(h + 1) = N(h) + N(\lceil a \cdot h \rceil) + 1$$

Let $h(n)$ be the smallest h such that $N(h + 1) > n$. Then it is easy to see that the height of any tree with n nodes is bounded from above by $h(n)$. If $N^{-1}(n)$ denotes the inverse function of $N(h)$ then it is easy to see that $h(n) = \lfloor N^{-1}(n) \rfloor$.

For example with $\alpha = \frac{1}{3}$ and $a = \frac{1}{2}$ we obtain

h	10	20	30	40	50	60	70
$N(h)$	29	194	729	2061	4913	10 398	20 133

h	80	90	100	150	200
$N(h)$	36 450	62 573	102 928	782 153	3 694 785

Then we can define

$$N^*(h + 1) = N^*(h) + N^*(ah) + 1$$

a functional equation defined for real h. Using standard techniques we can show that $\ln(N^*(h))$ has a proper asymptotic expansion in terms of $\omega(h)$, the first few terms being:

$$\ln(N^*(h)) = \frac{-1}{\ln(a)}\{\omega(h)^2 + c \cdot \omega(h) + \ln(a) \cdot \ln(\omega(h)) + O(1)\} \quad (1)$$

where

$$c = -\ln(a) + 2\ln(\ln(a)) + 2$$

and $\omega(h)$ is the transcendental function defined by $\omega(h)\, e^{\omega(h)} = h$.

We can also invert the asymptotic series to obtain h in terms of N (the inverse of the function $N^*(h)$):

$$h^*(N) = e^{\mu - c/2}\left(\mu - \frac{\ln(a)\ln(\mu)}{2} + O(1)\right) \quad (2)$$

where

$$\mu = \{-2\ln(a)\ln(n)\}^{1/2}$$

Intuitively, $N(h)$ should be close to $N^*(h)$, the only difference being the ceiling function in one of the arguments.

To prove that the relation $N(h)/N^*(h)$ is bounded we will first introduce the function $N^+(h)$,

$$N^+(h + 1) = N^+(h) + N^+(\lfloor ah \rfloor) + 1$$

with the same initial conditions as $N(h)$. Then it is not difficult to show that

$$N(h) \geq N^*(h) \geq N^+(h)$$

A careful study of the difference $N(h) - N^+(h)$ shows that

$$\lim_{h\to\infty} \frac{N(h)}{N^+(h)} \leq \text{constant}$$

The relation between $N(h)$ and $N^*(h)$ follows immediately. ▮

The final step is to relate $h^*(n)$ to $h(n)$ (the inverses of $N^*(h)$ and $N(h)$). The previous theorem says that

$$h(N) = h^*(KN)$$

in some bounded constant, K. Since

$$\mu(KN) = \mu(N)\left(1 + O\left(\frac{1}{\ln n}\right)\right)$$

we finally conclude that the height of an n node tree is greater than or equal to

$$h(N) = e^{\mu - c/2}\left(\mu - \frac{\ln(a)\ln\mu}{2} + O(1)\right). \quad ▮$$

There is an interesting relation between $N(h)$ and $P(h)$, a partition number. $P(h)$ of index r is the number of different solutions (number of different ordered sets of values $h_0, h_1, h_2, \ldots$) of

$$h_0 + h_1 r + h_2 r^2 + \cdots \leq h$$

This latter problem was solved by Mahler[8] and de Bruijn[9] in great detail as was kindly pointed out to us by A. Odlyzko (private communication).

$P(h)$ satisfies the functional equation

$$P(h+1) = P(h) + P\left(\left\lfloor \frac{h+1}{r} \right\rfloor\right).$$

It is easy to verify that the binary partition problem (Mahler's partition problem for $r = 2$) satisfies exactly the same functional equation as the *hrb*-tree for $\alpha = \frac{1}{3}$. Owing to different initial conditions,

$$N(h) = P(h)/2 - 1.$$

In any case $P(h)$ always satisfies the same asymptotic expression (1) with $a = 1/r$.

It is interesting to note that $N(h)$ has a much simpler solution in terms of $\omega(h)$ than in terms of $\ln(h)$ and $\ln(\ln(h))$, cf. Mahler[8] and de Bruijn.[9]

Acknowledgement

The work of the first and third authors was carried out under Natural Sciences and Engineering Research Council of Canada Grants Nos A-3353 and A-7700.

REFERENCES

1. G. M. Adelson-Velskii and E. M. Landis, An information organization algorithm. *Dokl. Akad. Nauk SSSR* **146**, 263 (1962).
2. J. Nievergelt and E. M. Reingold, Binary search trees of bounded balance. *SIAM Journal of Computing* **2**, 33 (1973).
3. C. C. Foster, A generalization of AVL trees. *Communications of the ACM* **16**, 513 (1973).
4. D. E. Knuth, *The Art of Computer Programming, Vol. 3: Sorting and Searching,* Addision-Wesley, Reading, Massachusetts (1973).
5. H. J. Olivié, Half-balanced binary search trees, *RAIRO Informatique Théoretique*, to appear (1981).
6. H. J. Olivié, On α-balanced binary trees. Theoretical Computer Science Conference. *Lecture Notes in Computer Science* **104**, pp. 98–108. Springer-Verlag, Heidelberg (1981).
7. A. V. Aho, J. E. Hopcroft and J. D. Ullman, *The Design and Analysis of Computer Algorithms,* Addison-Wesley, Reading, Massachusetts (1974).
8. K. Mahler, On a special Functional equation. *J. London Math. Soc.* **15**, 115 (1940).
9. N. G. de Bruijn, On Mahler's partition problem. *Neth. Aka. Wetenschef. Proc.* **51**, 659 (1948).
10. F. Luccio and L. Pagli, Power Trees. *Communications of the ACM* **21**, 941 (1978).

Received December 1981