



planetmath.org

Math for the people, by the people.

weight-balanced binary trees are ultrametric

Canonical name	WeightbalancedBinaryTreesAreUltrametric
Date of creation	2013-03-22 13:28:31
Last modified on	2013-03-22 13:28:31
Owner	mathcam (2727)
Last modified by	mathcam (2727)
Numerical id	9
Author	mathcam (2727)
Entry type	Result
Classification	msc 05C05

Let X be the set of leaf nodes in a weight-balanced binary tree. Let the distance between leaf nodes be identified with the weighted path length between them. We will show that this distance metric on X is ultrametric.

Before we begin, let the join of any two nodes x, y , denoted $x \vee y$, be defined as the node z which is the most immediate common ancestor of x and y (that is, the common ancestor which is farthest from the root). Also, we are using weight-balanced in the sense that

- the weighted path length from the root to each leaf node is equal, and
- each subtree is weight-balanced, too.

Lemma: two properties of weight-balanced trees

Because the tree is weight-balanced, the distances between any node and each of the leaf node descendants of that node are equal. So, for any leaf nodes x, y ,

$$d(x, x \vee y) = d(y, x \vee y) \quad (1)$$

Hence,

$$d(x, y) = d(x, x \vee y) + d(y, x \vee y) = 2 * d(x, x \vee y) \quad (2)$$

Back to the main proof

We will now show that the ultrametric three point condition holds for any three leaf nodes in a weight-balanced binary tree.

Consider any three points a, b, c in a weight-balanced binary tree. If there exists a renaming of a, b, c such that $d(a, b) \leq d(b, c) = d(a, c)$, then the three point condition holds. Now assume this is not the case. Without loss of generality, assume that $d(a, b) < d(a, c)$.

Applying Eqn. ??,

$$\begin{aligned} 2 * d(a, a \vee b) &< 2 * d(a, a \vee c) \\ d(a, a \vee b) &< d(a, a \vee c) \end{aligned}$$

Note that both $a \vee b$ and $a \vee c$ are ancestors of a . Hence, $a \vee c$ is a more distant ancestor of a and so $a \vee c$ must be an ancestor of $a \vee b$.

Now, consider the path between b and c . to get from b to c is to go from b up to $a \vee b$, then up to $a \vee c$, and then down to c . Since this is a tree, this is the only path. The highest node in this path (the ancestor of both b and c) was $a \vee c$, so the distance $d(b, c) = 2 * d(b, a \vee c)$.

But by Eqn. ?? and Eqn. ?? (noting that b is a descendent of $a \vee c$), we have

$$d(b, c) = 2 * d(b, a \vee c) = 2 * d(a, a \vee c) = d(a, c)$$

To summarize, we have $d(a, b) \leq d(b, c) = d(a, c)$, which is the desired ultrametric three point condition. So we are done.

Note that this means that, if a, b are leaf nodes, and you are at a node outside the subtree under $a \vee b$, then $d(\text{you}, a) = d(\text{you}, b)$. In other words, (from the point of view of distance between you and them,) the of any subtree that is not your own doesn't matter to you. This is expressed in the three point condition as "if two points are closer to each other than they are to you, then their distance to you is equal".

(above, we have only proved this if you are at a leaf node, but it works for any node which is outside the subtree under $a \vee b$, because the paths to a and b must both pass through $a \vee b$).