L9: Minimal AVL Trees

Definitions. Please refer to A7 for the full definition of AVL trees. We will use $\bullet$ as a shorthand for an empty AVL tree, and $N(d, t_1, t_2)$ as a shorthand for a non-empty AVL tree. We define the size of an AVL tree as the number of internal nodes:

$$\begin{array}{rcl} S(\bullet) & = & 0 \\ S(N(d, t_1, t_2)) & = & S(t_1) + S(t_2) + 1 \end{array}$$

Also recall the definition of Fibonacci numbers:

$$\begin{array}{rcl} F_0 & = & 0 \\ F_1 & = & 1 \\ F_n & = & F_{n-1} + F_{n-2} \end{array}$$

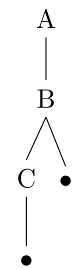
For example:

$$\begin{array}{rcl} F_0 & = & 0 \\ F_1 & = & 1 \\ F_2 & = & 1 \\ F_3 & = & 2 \\ F_4 & = & 3 \\ F_5 & = & 5 \\ F_6 & = & 8 \end{array}$$

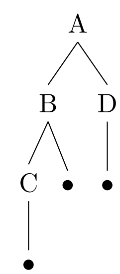
Examples. It turns out that there is an interesting connection between the minimal size of an AVL tree and the Fibonacci numbers. As an example, let's consider an AVL tree of height 3 and reason about its size. Since the height is 3, the tree must have a path like this:

$$\begin{array}{c} \text{A} \\ | \\ \text{B} \\ | \\ \text{C} \\ | \\ \bullet \end{array}$$

This is not a valid AVL tree however. What is the minimum number of additional nodes we need to add to make a valid AVL tree? We need, at every level that the difference of heigths be no more than 1. For the node C this is satisfied. For the node B, we need another subtree of height 1.



The same reasoning applies at node A: we need another subtree of height 2.



In other words, the minimal size of an AVL tree of height 3 is 4. As the pattern above suggests, the minimal size of an AVL tree of height h is the sum of the minimal sizes of trees of height $h-1$ and $h-2$. This is of course the definition of Fibonacci numbers and suggests the following property.

Theorem. Let g_h be the minimal size of an AVL tree of height h. Prove $g_h = F_{h+2} - 1$.
Prove the theorem by induction.