UNIVERSIDADE FEDERAL DO CEARÁ
Campus de Quixadá
Prof. Arthur Araruna
QXD0115- Estrutura de Dados Avançada

Árvores AVL são balanceadas

1 Conceitos

Vamos relembrar o nossos conceitos que serão necessários para o resultado a ser demonstrado:

Definição 1.1. Uma árvore binária T tal que $h(T) = O(\lg n)$ é dita *balanceada*.

Definição 1.2. O *balanço* de um nó binário n é definido pela função $bal(n) := h(n.dir) - h(n.esq)$.

Definição 1.3. Uma árvore binária T é *AVL* quando, para todo nó $n \in T$, $|bal(n)| \leq 1$.

2 Demonstração

Note que devemos mostrar que $h \leq c \cdot \lg n$ para algum $c > 0$, com h sendo a altura de uma árvore AVL qualquer e n sendo seu tamanho. Entretanto, nossa estratégia será um pouco mais restrita, mas permitirá concluirmos a mesma coisa.

Em vez de analisarmos uma árvore qualquer, vamos observar, para um dado h arbitrário, uma árvore AVL com altura h que contenha o mínimo possível de nós. Se essa árvore respeitar a desigualdade anterior, poderemos concluir que *todas* as AVL com essa mesma altura também a respeitarão (já que para essa tal árvore o lado direito da desigualdade é o menor possível, todas as outras apresentarão valores ainda maiores, o que mantém a desigualdade).

2.1 Estrutura da árvore

Para entendermos como construir uma árvore como desejamos, veja que sempre que, em uma árvore AVL T com altura h, encontrarmos um nó n com $bal(n) = 0$, poderíamos substituir uma de suas sub-árvores por outra AVL com uma altura 1 unidade menor que a que foi retirada sem invalidar a propriedade AVL de T. Isso constrói uma nova árvore T' com a mesma altura que a anterior T e certamente tamanho menor. Dessa forma, podemos concluir que a árvore desejada não pode ter nenhum nó com balanço nulo.

2.2 Contagem de nós da árvore

Vamos definir a função $N(h)$ de forma a representar o menor número de nós possível para uma árvore AVL de altura h, que chamaremos de T.

Como vimos anteriormente, as sub-árvores de T devem possuir alturas $h-1$ e $h-2$, não necessariamente nesta ordem. Além disso, podemos afirmar que ambas se tratam de árvores AVL com o mínimo de nós para a respectiva altura, caso contrário não faria sentido T possuir tamanho mínimo (já que elas poderiam ser trocadas por outras árvores de mesma altura e menos nós).

Dessa forma, podemos afirmar que o valor de $N(h)$ pode ser calculado como

$$N(h) = \begin{cases} N(h-1) + N(h-2) + 1, & \text{se } h \geq 2; \\ 1, & \text{se } h = 1; \\ 0, & \text{se } h = 0. \end{cases}$$

Veja que existe uma enorme semelhança entre a função N e a função F a seguir, que define a *Sequência de Fibonacci*:

$$F(h) = \begin{cases} F(h-1) + F(h-2), & \text{se } h \geq 2; \\ 1, & \text{se } h = 1; \\ 0, & \text{se } h = 0. \end{cases}$$

De fato, para qualquer valor de h, é o caso que $N(h) \geq F(h)$. Vamos usar o valor f_h do h-ésimo termo dessa sequência para obtermos uma relação entre h e o número de nós $N(h)$ da árvore, que denotaremos simplesmente por n. Podemos calcular esse termo de forma não-recursiva como[1]

$$f_h = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^h - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^h.$$

Como $h > 0$, podemos concluir[2] que

$$\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^h \leq 1 \implies f_h \geq \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^h - 1.$$

3 Conclusão

Daí, considerando $a = \frac{1+\sqrt{5}}{2}$ e lembrando que $n = N(h) \geq F(h) = f_h$, podemos, ao reescrever a desigualdade anterior, concluir que

$$\begin{aligned} n \geq f_h \geq & \quad \frac{1}{\sqrt{5}}a^h - 1 & \implies \\ n + 1 \geq & \quad \frac{1}{\sqrt{5}}a^h & \implies \\ \log_a(n+1) \geq & \quad h - \log_a \sqrt{5} & \implies \\ \log_a(n+1) + \log_a \sqrt{5} \geq & \quad h. & \end{aligned}$$

Isso significa que $h \leq \log_a(n+1) + \log_a \sqrt{5} = O(\lg n)$, como desejávamos.

Apêndice

[1] Veja `https://pt.wikipedia.org/wiki/Sequência_de_Fibonacci#Fórmula_explícita`

[2] Ver apêndice para a argumentação mais detalhada.