Alberi binari di ricerca e AVL

Visita in-ordine di BST

Nel seguito dimostriamo la proposizione seguente:

Una visita in ordine di un albero binario di ricerca visita i nodi in ordine crescente rispetto alle chiavi.

Prova. Ricordiamo innanzi tutto la visita in ampiezza di un BST a partire da un certo nodo v:

```
inorder(T, v) { // T è l'albero, v è un nodo di T
    if(v == null)
        return;
    inorder(T, v.left);
    visita(v);
    inorder(T, v.right)
}
```

La prova procede per induzione rispetto all'altezza di T. Se T è vuoto oppure ha altezza 1 (T consiste di un solo nodo) l'affermazione è chiaramente vera. Supponiamo che l'affermazione sia vera per tutte le altezze $\leq h$ e facciamo vedere che vale anche per BST di altezza $\leq h+1$. Un BST di altezza al più $h+1$ è costituito da una radice e da due BST di altezze $\leq h$. Inoltre, per la proprietà di BST, la chiave associata alla radice è maggiore o uguale della chiave associata a qualunque nodo del sottoalbero sinistro e minore o uguale della chiave associata a qualsiasi nodo del sottoalbero destro. Siano T_1 e T_2 i sottoalberi sinistro e destro di T. T_1 e T_2 hanno altezza al più h, in quanto T ha altezza al più $h+1$ Allora:

i) tutte le chiavi di T_1 sono visitate in ordine crescente per ipotesi induttiva; ii) dopo aver visitato T_1 si visita v per la definizione di visita in ordine e la chiave associata a v è maggiore o uguale di tutte le chiavi associate ai nodi di T_1 in quanto T è un BST; iii) infine si visita T_2 in ordine crescente rispetto alle chiavi, di per ipotesi induttiva. Per concludere la prova ricordiamo di nuovo che la chiave associata a v è minore o uguale della chiave associata a quasiasi nodo di T_2 poiché T è un BST.

Altezza di un AVL

Vogliamo dimostrare che, per un albero AVL contenente n nodi, l'altezza è $O(\log n)$. E' più semplice procedere dimostrando che il numero *minimo* di nodi contenuto in un AVL di altezza h è *almeno* ca^h per costanti positive a e c. Ciò implica che se abbiamo un albero AVL di n nodi, la sua altezza h deve soddisfare le seguenti relazioni:

$$n = (numero\ di\ nodi) \geq (\ numero\ minimo\ di\ nodi\ per\ altezza\ h) \geq ca^h$$

Ciò implica $h \leq \log_a \frac{n}{c}$.

Costruzione bottom-up di un AVL con il numero minimo di nodi.

Nel seguito di questa prova, chiamiamo altezza il numero di *archi* del percorso più lungo dalla radice a una delle foglie. Ciò ha il solo scopo di semplificare leggermente la notazione. Per costruire un AVL di altezza h con il minimo numero di nodi usiamo due AVL di altezza $h-1$ e $h-2$, ciascuno con il numero minimo di nodi per la rispettiva altezza. L'altezza del nuovo AVL è h. Il numero di nodi è:

$$n(h) = 1 + n(h-1) + n(h-2) > 2n(h-2).$$

Dall'uguaglianza precedente possiamo ottenere la seguente catenza di disuguaglianze:

$$n(h) > 2n(h-2) > 2 \cdot 2n(h-2-2) > 2 \cdot 2 \cdot 2 \cdot n(h-2-2-2) \ldots = 2^i n(h-2i)$$

Ci dobbiamo fermare al minimo valore di i tale che $h - 2i \geq 1$, cioè quando $i = \left\lceil \frac{h}{2} \right\rceil - 1$. Poiché abbiamo $n\left(h - 2\left(\left\lceil \frac{h}{2} \right\rceil - 1\right)\right) \geq n(0) = 1$, otteniamo $n(h) > 2^{\left\lceil \frac{h}{2} \right\rceil - 1} \geq 2^{\frac{h}{2} - 1}$, che a sua volta implica $\frac{h}{2} \leq \left\lceil \frac{h}{2} \right\rceil < \log_2 n(h) + 1$ e quindi $h \leq 2 \log_2 n(h) + 2$. In sostanza, abbiamo mostrato che $n(h) \geq ca^h$, con $a = \sqrt{2}$ e $c = 1/2$.

Perché alberi di Fibonacci.

Riprendiamo l'equazione di ricorrenza iniziale:

$n(h) = 1 + n(h-1) + n(h-2)$. Dimostriamo che, per ogni k, $n(k) = F_{k+3} - 1$, dove F_i è l'i-esimo numero di Fibonacci.

Prova. Si procede per induzione. Ricordiamo una possibile definizione dei numeri di Fibonacci: $F_0 = 0, F_1 = 1, F_2 = F_1 + F_0 = 1, F_3 = F_2 + F_1 = 2, \ldots, F_k = F_{k-1} + F_{k-2}$, per $k \geq 2$. D'altra parte abbiamo: $n(0) = 1, n(1) = 2, n(2) = 4, n(3) = 7$. Vediamo quindi che l'affermazione è vera per $k = 0, 1, 2$ e ciò costituisce la base dell'induzione. Supponiamo ora l'affermazione vera per $1, 2, \ldots, k-1$ e mostriamo che l'affermazione è vera per k. Abbiamo:

$$\begin{aligned} n(k) = 1 + n(k-1) + n(k-2) = 1 + (F_{k-1+3} - 1) + (F_{k-2+3} - 1) &= 1 + (F_{k+2} - 1) + (F_{k+1} - 1) \\ &= F_{k+2} + F_{k+1} - 1 = F_{k+3} - 1, \end{aligned}$$

dove la prima uguaglianza segue dall'ipotesi induttiva, mentre l'ultima segue ricordando che $F_{k+3} = F_{k+2} + F_{k+1}$ per definizione.

Questo è il motivo per il quale tali alberi sono noti come alberi di Fibonacci. Sappiamo che, per ogni k, $F_k = \left[\frac{\phi^k}{\sqrt{5}}\right]$, con ϕ la sezione aurea e dove $[\cdot]$ denota l'intero più vicino. Ciò implica peraltro che

$$n(h) > \frac{\phi^{h+3}}{\sqrt{5}} - 1,$$

dove ϕ è la *sezione aurea*. Ciò consente di ricavare un limite superiore all'altezza di un AVL con n nodi che è più accurato di quello che abbiamo calcolato prima.