Baumstrukturen[1]

1 Definitionen

- Eine *Baumstruktur* vom Grundtyp T ist entweder

  1. die leere Struktur oder
  2. ein *Knoten* (node) vom Typ T mit einer endlichen Zahl verknüpfter, voneinander verschiedener Baumstrukturen vom Grundtyp T, sogenannter *Teilbäume* (subtrees).

- Ein *geordneter Baum* ist ein Baum, dessen Verzweigungen in jedem Knoten geordnet sind.

- Jedem Knoten eines Baumes kann eine *Stufe* zugeordnet werden. Die *Wurzel* des Baumes, d.h. der Knoten ohne Vorgänger, besitzt die Stufe 1; alle Knoten besitzen eine um 1 größere Stufe als ihre Vorgänger oder *Vater*. Die größte Stufe eines Baumes heißt seine *Höhe*.

- Hat ein Knoten keine Nachfolger oder *Söhne*, so heißt er *Endknoten* oder *Blatt* (leaf); ein Element, das nicht Endknoten ist, wird *innerer Knoten* genannt.

- Die Zahl der Söhne eines inneren Knotens nennt man den *Grad* des Knotens. Der höchste Grad unter allen Knoten ist der *Grad des Baumes*. Bäume vom Grad 2 heißen *Binärbäume*, Bäume vom Grad größer als 2 heißen *Vielweg-Bäume*.

- Die Zahl der Kanten von der Wurzel bis zum Knoten x heißt *Weglänge* von x. Die Wurzel hat Weglänge 0. Die *Weglänge eines Baumes* oder *innere Weglänge* ist definiert als die Summe der Weglängen aller seiner Knoten.

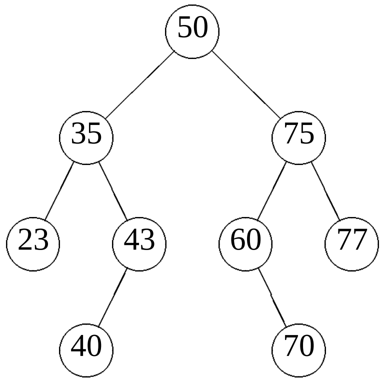
- Ein Baum ist *vollständig ausgeglichen*, wenn sich für jeden Knoten die Zahlen der Knoten in seinen Teilbäumen um höchstens 1 unterscheiden.

- Ein Baum ist genau dann *ausgeglichen* oder *AVL–ausgeglichen*, wenn sich für jeden Knoten die Höhen der zugehörigen Teilbäume um höchstens 1 unterscheiden. Bäume mit dieser Eigenschaft heißen *AVL–Bäume* nach ihren Schöpfern Adelson-Velskii und Landis.

[1] aus Niklaus Wirth, Algorithmen und Datenstrukturen

2 Übungen

1. Der nebenstehend dargestellte Baum ist nach allen Merkmalen zu untersuchen. Wie ist am sinnvollsten die *mittlere Weglänge* eines Baumes zu definieren? Bestimme die mittlere Weglänge des dargestellten Baumes.



2. Zeichne vollständig ausgeglichene Binärbäume mit $n = 1, 2, 3, 4, 5, 6, 7, 8, 9$ Knoten.

3. Zeichne einen AVL–ausgeglichenen Baum, der nicht vollständig ausgeglichen ist.

4. Welche Beziehung besteht bei vollständig ausgeglichenen Bäumen zwischen der Anzahl der Knoten und der Höhe des Baumes?

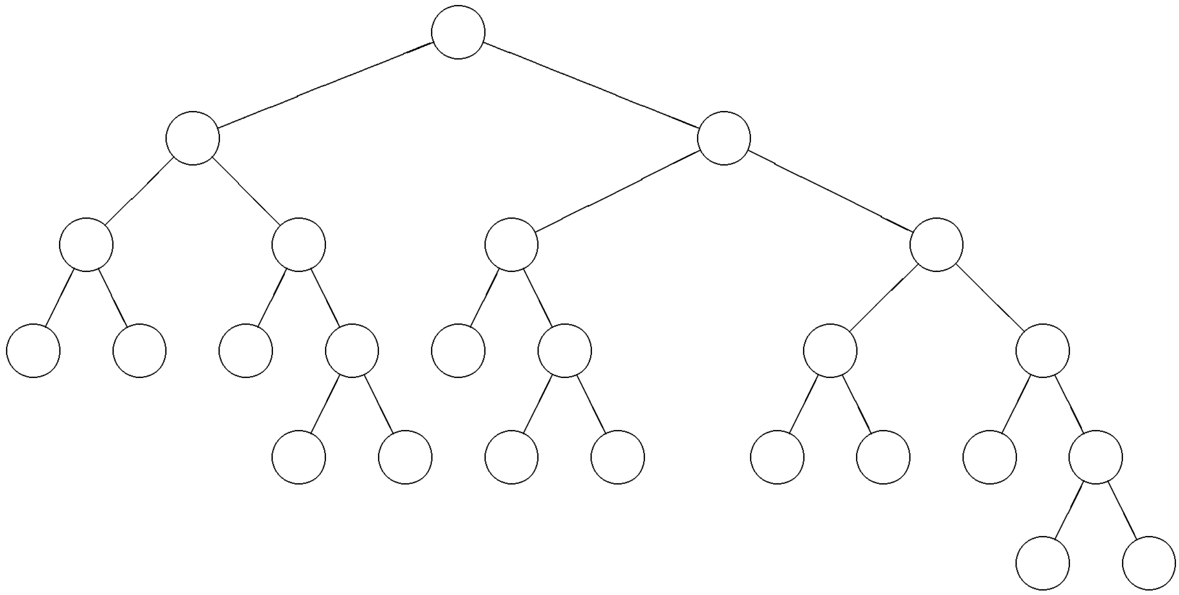
5. Der am schlechtesten AVL–ausgeglichene Baum mit der Höhe h sei mit T_h bezeichnet. Stelle die Bäume T_1 bis T_5 graphisch dar.

6. Welche der Bäume T_1 bis T_5 sind noch vollständig ausgeglichen?

7. Welche Struktur ist zwischen T_h, T_{h-1} und T_{h-2} erkennbar? Gib eine Definition für die betrachteten Bäume an.

8. Mit N_h sei die Anzahl der Knoten von T_h bezeichnet. Stelle eine Rekursionsvorschrift für die Zahlen N_h auf. Lasse Dir die auch *Leonardo–Zahlen* genannten Zahlen durch ein kleines Programm bis $h = 20$ berechnen.

9. Welche Änderungen müssen an untenstehendem Baum vorgenommen werden, um ihn zu einem T_h–Baum zu machen? Welche mittlere Weglänge ergibt sich dann bei dem veränderten Baum (auch *Fibonacci-Baum* genannt)?



10. Stelle die Struktur der rekursiven Aufrufe für die bekannte Funktion FIBO zur Berechnung der Fibonacci–Zahlen als Baum dar. Als Beispiel sei FIBO (5) genommen.

11. Mit W_h sei die Weglänge von T_h bezeichnet. Gib eine Rekursionsvorschrift für die Zahlen W_h an. Lasse Dir auch hier die Zahlen durch ein kleines Programm darstellen.

12. Entwickle einen geschlossenen Ausdruck für die Weglänge $WVA(n)$ in einem vollständig ausgeglichenen Baum mit n Knoten.

13. Bestimme mit Hilfe eines Programmes das Verhältnis

$$\frac{W_{AVL}(n)}{W_{VA}(n)} = \frac{\text{mittl. Weglänge im AVL-Baum mit n Knoten}}{\text{mittl. Weglänge im vollst. ausgegl. Baum mit n Knoten}}$$

zur Abschätzung der in einem AVL-ausgeglichenen Baum im Vergleich zum entsprechenden vollständig ausgeglichenen Baum vergrößerten mittleren Weglänge.

14. **Theorem von Adelson-Velskii und Landis:** Die Höhe eines AVL–ausgeglichenen Baumes ist um höchstens $44,04\%$ größer als diejenige seines vollständig ausgeglichenen Gegenstückes.
Anleitung :

    (a) Die von BINET 1843 gefundene explizite Darstellung der Fibonacci–Zahlen lautet

$$\begin{aligned} Fibo(n) &= \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right] \\ &= \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \underbrace{\left[1-(-1)^n\left(\frac{\sqrt{5}-1}{\sqrt{5}+1}\right)^n\right]}_{\alpha}. \end{aligned}$$

    α kann einfach nach unten und oben abgeschätzt werden.

(b) Für die Anzahl der Knoten im am schlechtesten AVL–ausgeglichenen Baum T_h der Höhe h gilt

$$N(h) = \sum_{i=1}^{h} Fibo(i) = Fibo(h+2) - 1.$$

(c) Für die Höhe h eines vollständig ausgeglichenen Baumes mit N Knoten gilt

$$h = [\log_2 N] + 1$$

bzw.

$$\log_2 N < h \leq \log_2 N + 1.$$

(d) Aus der Mittelstufe bekannt ist

$$\log_2 x = \frac{\ln x}{\ln 2}, \quad \log_a(x \cdot y) = \log_a x + \log_a y \quad \text{und} \quad e^{\ln x} = x.$$