Saturierte Binärbäume

Informell sind saturierte Binärbäume Bäume, bei denen jeder Knoten entweder zwei Nachfolgeknoten hat (*innere Knoten*) oder gar keinen Nachfolgenknoten (*Blatt*). Formaler:

1) Ein einzelner Knoten ist *saturierter Binärbaum der Höhe 0*. Er hat keine inneren Knoten, aber ein Blatt (zugleich die Wurzel!).
2) Sind T_1 und T_2 saturierte Binärbäume mit Knotenmengen V_1 bzw. V_2, wobei $V_1 \cap V_2 = \emptyset$ und $v \notin V_1 \cup V_2$, dann bildet das Tripel (v, T_1, T_2) einen neuen saturierten Binärbaum mit der Wurzel v, linkem Teilbaum T_1 und rechtem Teilbaum T_2. Innere Knoten des neuen Baums sind v und die inneren Knoten von T_1 und T_2, die Höhe ist 1 plus das Maximum der Höhen beider Teilbäume. Blätter sind diejenigen beider Teilbäume.

Achtung: Die Reihenfolge der Teilbäume ist hier von Bedeutung.

Allgemeine Binärbäume

Ein *Binärbaum* ist ein Baum, bei dem jeder Knoten maximal zwei Nachfolgeknoten hat. Formal kann man das so definieren:

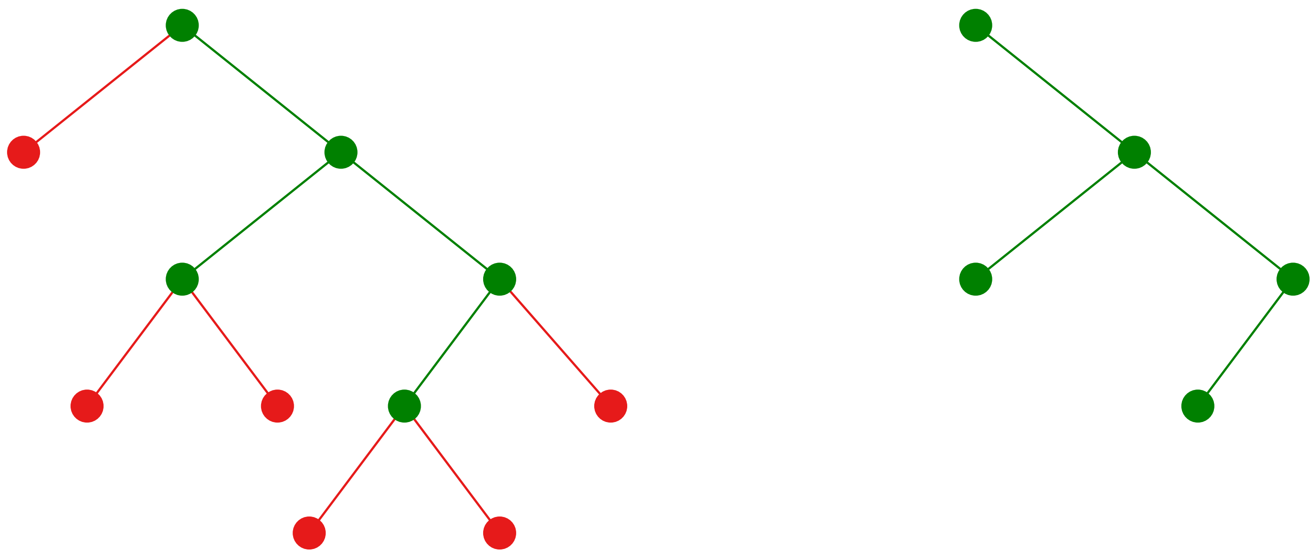
Ein Binärbaum entsteht aus einem saturierten Binärbaum durch Weglassen aller Blätter. Insbesondere ist auch der leere Graph ein Binärbaum.

Geht man von den saturierten Binärbäumen auf diese Weise zu allgemeinen Binärbäumen über und fügt dann wieder an jeden „freien Platz“ ein Blatt an, so erhält man den ursprünglichen Binärbaum. Umgekehrt geht das ebenso. Daraus erhalten wir eine 1:1 Korrespondenz zwischen Binärbäumen mit n Knoten und saturierten Binärbäumen mit n inneren Knoten.

Beachte zusätzlich: Ein saturierter Binärbaum mit n inneren Knoten hat $n+1$ Blätter, insgesamt also $2n+1$ Knoten.

Beispiele

Wir zeigen einen saturierten Binärbaum mit zugehörigem allgemeinem Binärbaum:



Bitte beachten: Die folgenden Bäume zählen wir als 4 verschiedene Binärbäume.

Anzahl verschiedener Binärbäume

Wir wollen eine Formel, die besagt, wieviele verschiedene Binärbäume mit n Knoten es gibt. Wegen der 1:1 Korrespondenz ist es hinreichend zu untersuchen, wieviele saturierte Binärbäume mit n inneren Knoten (und $n+1$ Blättern) es gibt.

Satz: Die Anzahl saturierter Binärbäume mit n inneren Knoten ist C_n.

Hierbei ist C_n die n-te Catalan-Zahl, d.h. $C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{2n+1}\binom{2n+1}{n}$.

Für den Beweis benutzen wir die Tatsache, dass $C_n = |E_n|$ gilt, wobei E_n die Menge der Wörter der Länge $2n+1$ ist, die aus Dyck-Wörtern durch Anhängen eines b entstehen.

Beweis des Satzes

Wir definieren induktiv eine Abbildung von der Menge E_n in die Menge der saturierten Binärbäume mit n inneren Knoten. Auf Grund der Bijektivität dieser Abbildung ist damit der Beweis komplett.

$n = 0$: Das einzige Wort in E_0 ist das Wort b, dem wir den einzigen saturierten Binärbaum ohne innere Knoten, bestehend nur aus der Wurzel, zuordnen.

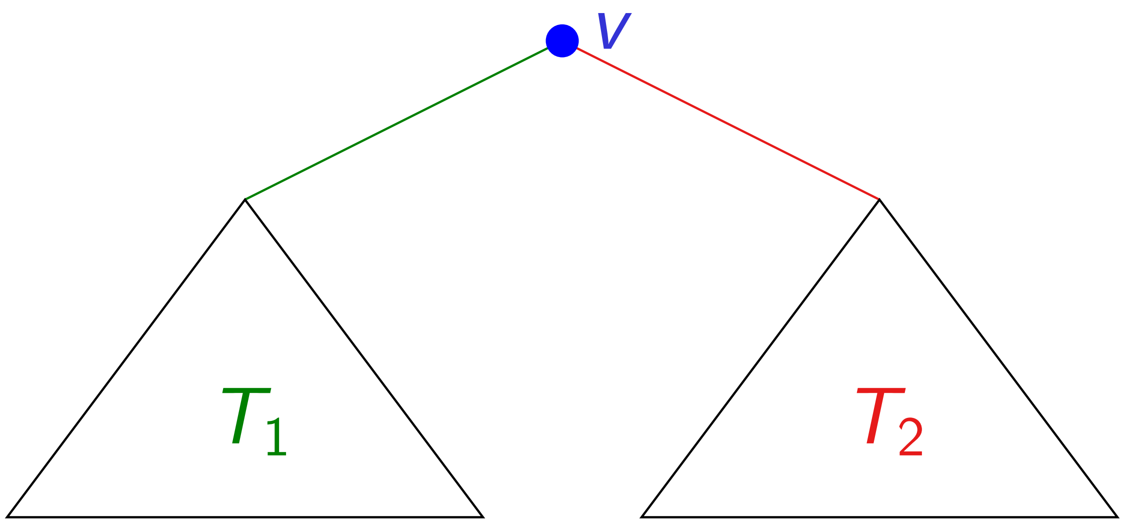
Nun sei also $n > 0$:

Jedes Wort in E_n ist von der Form $aubvb$ für Dyck-Wörter u und v; und diese Zerlegung ist eindeutig. (Bitte überprüfen Sie die Richtigkeit dieser Behauptung!)

Jetzt liegen die Teilwörter ub und vb auch jeweils in einem E_m mit $m < n$. Sei also $ub \in E_{m_1}$ und $vb \in E_{m_2}$. Dann gehört zu ub induktiv bereits ein saturierter Binärbaum T_1 mit m_1 inneren Knoten, zu vb ein T_2 mit m_2 inneren Knoten.
Wir beachten noch, dass ub die Länge $2m_1 + 1$ und vb die Länge $2m_2 + 1$ hat.
Das gesamte Wort $aubvb$ hat damit Länge $2(m_1 + m_2) + 3 = 2(m_1 + m_2 + 1) + 1$.
Daher gilt $n = m_1 + m_2 + 1$.

Abschluss des Beweises

Dem Wort *aubvb* ordnen wir jetzt den saturierten Binärbaum (v, T_1, T_2) zu (mit einem neuen Wurzelknoten v und disjunkten Teilbäumen T_1 und T_2):



Die Wohldefiniertheit der konstruierten Abbildung folgt aus der Eindeutigkeit der Zerlegung des Ausgangswortes in *aubvb* mit Dyck-Wörtern u und v. Injektivität und Surjektivität sind leicht mit Induktion einzusehen.