

Kapitel 5

Graphen und Bäume

Graphen sind ein grundlegendes Konzept, das in der Programmierung an vielen Stellen auftaucht. Da Graphen anschaulich grafisch dargestellt werden können, ist das Konzept des Graphen leicht zu verstehen. Andererseits können Graphen einfach durch mathematische Objekte dargestellt werden. Graphen bieten uns daher die willkommene Gelegenheit, das Wechselspiel zwischen intuitivem Konzept und formaler Definition zu studieren. Dieses Wechselspiel interessiert uns, da beim Programmieren routinemäßig intuitive Konzepte in formale Definitionen umgesetzt werden müssen.

Bäume sind spezielle Graphen. Für Bäume werden wir drei verschiedene Formalisierungen angeben. Jede dieser Formalisierungen eröffnet eine neue Sichtweise auf das Konzept des Baumes.

Wir werden verschiedene Eigenschaften von Bäumen präzise formulieren und beweisen. Dafür benötigen wir induktive Beweistechniken, mit deren Einführung wir dieses Kapitel beginnen.

5.1 Induktionsbeweise

Induktion ist eine Beweistechnik, mit der man allquantifizierte Aussagen der Form

$$\forall x \in X: A(x)$$

beweisen kann. Dabei soll $A(x)$ eine Aussage für $x \in X$ sein. Hier ist ein Beispiel für eine Aussage, die man mit Induktion beweisen kann:

$$\forall n \in \mathbb{N}: 2^n \geq n^2 - 1$$

Induktionsregel
für $\mathbb{N}$

Es gibt verschiedene Induktionsregeln. Wir beginnen mit einer einfachen Regel für Aussagen $A(n)$ über $n \in \mathbb{N}$:

$$\frac{A(0) \quad \forall n \in \mathbb{N}, n > 0: A(n-1) \Rightarrow A(n)}{\forall n \in \mathbb{N}: A(n)} \quad (5.1)$$

Mit dieser Regel können Aussagen der Form

$$\forall n \in \mathbb{N}: A(n)$$

dadurch bewiesen werden, dass man die folgenden zwei Aussagen beweist:

$$\begin{array}{ll} A(0) & \text{Induktionsanfang} \\ \forall n \in \mathbb{N}, n > 0: A(n-1) \Rightarrow A(n) & \text{Induktionsschritt} \end{array}$$

Ein Beweis für den Induktionsschritt kann die Aussage

$$n \in \mathbb{N} \wedge n > 0 \wedge A(n-1) \quad \text{Induktionsannahme}$$

als gültig annehmen und darauf aufbauend die Aussage

$$A(n) \quad \text{Induktionsbehauptung}$$

zeigen (für das n aus der Induktionsannahme).

Wir erproben die Induktionsregel an der folgenden Proposition.

Proposition 5.1.1 Sei $f \in \mathbb{N} \rightarrow \mathbb{N}$ rekursiv definiert durch

$$f(n) = \text{if } n = 0 \text{ then } 0 \text{ else } n + f(n-1)$$

Dann gilt für alle $n \in \mathbb{N}$:

$$f(n) = \frac{n}{2}(n+1)$$

Beweis Durch Induktion über n .

Sei $n = 0$. Dann

$$\begin{aligned} f(n) &= f(0) \\ &= 0 && \text{(nach Definition von } f) \\ &= \frac{n}{2}(n+1) \end{aligned}$$

Sei $n > 0$. Dann

$$\begin{aligned}
 f(n) &= n + f(n-1) && \text{(nach Definition von } f) \\
 &= n + \frac{n-1}{2}(n-1+1) && \text{(nach Induktionsannahme)} \\
 &= \frac{n^2 + n}{2} \\
 &= \frac{n}{2}(n+1)
 \end{aligned}$$

□

Die Anwendung der Induktionsregel im Beweis folgt der Kürze und Übersichtlichkeit halber einem Schema. Die Einleitung

*Schematisches
Aufschreiben von
Induktionsbeweisen*

Durch Induktion über n

sagt, dass die die Behauptung $\forall n \in \mathbb{N}: A(n)$ der Proposition mit Hilfe einer Induktionsregel bewiesen wird. Die Aussage $A(n)$ ist $f(n) = \frac{n}{2}(n+1)$. Die Phrase

Sei $n = 0$

leitet den Beweis des Induktionsanfangs $A(0)$ ein. Die Phrase

Sei $n > 0$

leitet den Beweis der Induktionsbehauptung $A(n)$ ein und sagt implizit, dass dafür die Induktionsannahme

$$n \in \mathbb{N} \wedge n > 0 \wedge A(n-1)$$

als gültig vorausgesetzt wird.

Die gerade bewiesene Proposition ist für die Programmierung interessant. Sie sagt nämlich, dass die beiden Prozeduren

*Relevanz für
Programmierung*

```
fun g n = if n=0 then 0 else g(n-1) + n
```

```
fun h n = (n*(n+1)) div 2
```

für Argumente $n \in \mathbb{N}$ die gleichen Ergebnisse liefern.¹

Warum ist die Induktionsregel 5.1 korrekt? Korrektheit bedeutet, dass aus der Gültigkeit der Prämissen die Gültigkeit der Konklusion der Regel folgt.

*Korrektheit der
Induktionsregel*

¹ Wir idealisieren hier etwas. Wegen der Größenbeschränkungen für ganze Zahlen wird die Prozedur `g` für einige Argumente noch Ergebnisse liefern, für die die Prozedur `h` bereits die Ausnahme `Overflow` wirft.

Proposition 5.1.2 Sei $A(n)$ eine Aussage für $n \in \mathbb{N}$ und seien die beiden Aussagen

$$A(0) \\ \forall n \in \mathbb{N}, n > 0: A(n-1) \Rightarrow A(n)$$

gültig. Dann gilt $\forall n \in \mathbb{N}: A(n)$.

Beweis Durch Widerspruch. Annahme: Es gibt mindestens ein $n \in \mathbb{N}$, sodass $A(n)$ nicht gilt. Dann gibt es ein kleinstes $n_0 \in \mathbb{N}$, sodass $A(n_0)$ nicht gilt. Da $A(0)$ gilt, folgt $n_0 \neq 0$. Da n_0 kleinstmöglich gewählt ist, wissen wir, dass $A(n_0 - 1)$ gilt. Mit der Annahme

$$\forall n \in \mathbb{N}, n > 0: A(n-1) \Rightarrow A(n)$$

folgt, dass $A(n_0)$ gilt. Das ist ein Widerspruch. $\square$

Hier ist ein zweites Beispiel für einen Induktionsbeweis.

Proposition 5.1.3 Sei X eine endliche Menge mit n Elementen. Dann ist $\mathcal{P}(X)$ endlich und hat 2^n Elemente.

Beweis Durch Induktion über n .

Sei $n = 0$. Dann ist $X = \emptyset$. Also ist $\mathcal{P}(X) = \{\emptyset\}$. Folglich ist $\mathcal{P}(X)$ endlich und hat $2^0 = 1$ Elemente.

Sei $n > 0$. Wir wählen ein $x \in X$. Die Menge $X - \{x\}$ hat $n - 1$ Elemente. Also ist $\mathcal{P}(X - \{x\})$ nach Induktionsannahme endlich und hat 2^{n-1} Elemente. Mit der disjunkten Zerlegung

$$\mathcal{P}(X) = \mathcal{P}(X - \{x\}) \cup \{Y \cup \{x\} \mid Y \in \mathcal{P}(X - \{x\})\}$$

folgt, dass $\mathcal{P}(X)$ endlich ist und $2^n = 2^{n-1} + 2^{n-1}$ Elemente hat. $\square$

Wir haben bereits darauf hingewiesen, dass die Induktionsregel 5.1 nur eine von vielen ist. Hier sind zwei allgemeinere Induktionsregeln, die wir später benötigen werden.

Proposition 5.1.4 Sei $M \subseteq \mathbb{Z}$ eine Menge von ganzen Zahlen mit dem kleinsten Element $m_0 \in M$. Sei $A(n)$ eine Aussage für $n \in M$. Dann gilt die Regel:

$$\frac{A(m_0) \quad \forall n \in M, n > m_0: (\forall k \in M, k < n: A(k)) \Rightarrow A(n)}{\forall n \in M: A(n)} \quad (5.2)$$

Induktionsregel
für $\mathbb{Z}$

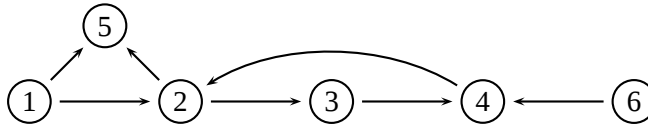


Abbildung 5.1: Ein gerichteter Graph

Für diese Regel vereinbaren wir die folgenden Sprechweisen:

$A(m_0)$	Induktionsanfang
$\forall n \in M, n > m_0: (\forall k \in M, k < n: A(k)) \Rightarrow A(n)$	Induktionsschritt
$n \in M \wedge n > m_0 \wedge \forall k \in M, k < n: A(k)$	Induktionsannahme
$A(n)$	Induktionsbehauptung

Proposition 5.1.5 Sei $f \in X \rightarrow M$ und $M \subseteq \mathbb{Z}$. Sei $m_0 \in M$ das kleinste Element von M und sei $A(x)$ eine Aussage für $x \in X$. Dann gilt die Regel:

Induktionsregel
für X

$$\frac{\begin{array}{l} \forall x \in X, f(x) = m_0: A(x) \\ \forall x \in X, f(x) > m_0: (\forall y \in X, f(y) < f(x): A(y)) \Rightarrow A(x) \end{array}}{\forall x \in X: A(x)} \quad (5.3)$$

Für diese Regel vereinbaren wir die folgenden Sprechweisen:

$\forall x \in X, f(x) = m_0: A(x)$	Induktionsanfang
$\forall x \in X, f(x) > m_0: (\forall y \in X, f(y) < f(x): A(y)) \Rightarrow A(x)$	Induktionsschritt
$x \in X \wedge f(x) > m_0 \wedge \forall y \in X, f(y) < f(x): A(y)$	Induktionsannahme
$A(x)$	Induktionsbehauptung

5.2 Graphen

Abbildung 5.1 zeigt die grafische Darstellung eines gerichteten Graphen. Dieser Graph besteht aus 6 Knoten und 7 gerichteten Kanten. Die Kanten sind als Pfeile eingezeichnet, die jeweils zwei Knoten verbinden.

Die Formalisierung von gerichteten Graphen ist einfach.²

Definition 5.2.1 Ein gerichteter Graph ist ein Paar (V, E) , sodass V und E endliche Mengen mit $E \subseteq V \times V$ sind. Die Elemente von V heißen Knoten (englisch „vertices“) und die Elemente von E heißen Kanten (englisch „edges“). Wir sagen, dass eine Kante $(v_1, v_2) \in E$ vom Knoten v_1 zum Knoten v_2 geht.

² Unter der Formalisierung eines intuitiven Konzepts versteht man seine Darstellung mithilfe von mathematischen Objekten.

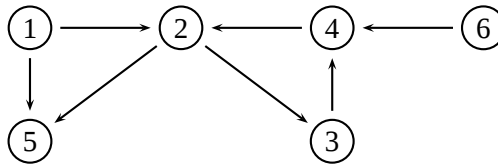


Abbildung 5.2: Eine andere Darstellung des Graphen aus Abbildung 5.1

Abbildung 5.1 stellt also einen gerichteten Graphen $G = (V, E)$ wie folgt dar:

$$V = \{1, 2, 3, 4, 5, 6\}$$

$$E = \{(1, 2), (1, 5), (2, 3), (2, 5), (3, 4), (4, 2), (6, 4)\}$$

Dieser Graph lässt sich auch wie in Abbildung 5.2 darstellen. Obwohl sich die beiden grafischen Darstellungen unterscheiden, stellen sie doch denselben Graphen dar.

Graphen als Beispiel für Formalisierung

Das Konzept des Graphen hat sich aus den anschaulichen grafischen Darstellungen entwickelt. Unsere Definition formalisiert Graphen als mathematische Objekte. Die formale Definition von Graphen regelt vollständig, was unter einem Graphen zu verstehen ist und bildet eine tragfähige Grundlage für die Formulierung und den Beweis von Grapheigenschaften.

Informatiker stehen oft vor der Aufgabe, intuitive Konzepte präzise formulieren zu müssen. Im Idealfall gelingt es, eine einfache formale Definition wie bei der Formalisierung von gerichteten Graphen zu finden.

Neben den gerade eingeführten Graphen mit gerichteten Kanten gibt es auch Graphen mit ungerichteten Kanten. Da wir aber nur gerichtete Graphen betrachten, sagen wir im Folgenden einfach Graph und meinen stets gerichteter Graph.

Sei $G = (V, E)$ ein Graph. Wir vereinbaren die folgenden Sprechweisen:

- Ein Knoten v' heißt *Nachfolger* eines Knotens v genau dann, wenn $(v, v') \in E$.
- Ein Knoten v heißt *Vorgänger* eines Knotens v' genau dann, wenn $(v, v') \in E$.
- Ein Knoten heißt *Quelle* genau dann, wenn er keinen Vorgänger hat.
- Ein Knoten heißt *Senke* genau dann, wenn er keinen Nachfolger hat.
- Ein Knoten heißt *innerer Knoten* genau dann, wenn er weder eine Senke noch eine Quelle ist.
- Der *Eingrad* eines Knotens ist die Anzahl seiner Vorgänger.

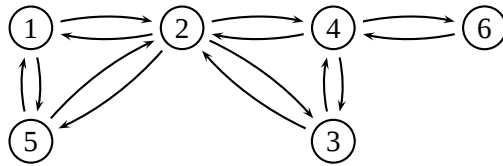


Abbildung 5.3: Der symmetrische Abschluss des Graphen aus Abbildung 5.2

- Der *Ausgrad* eines Knotens ist die Anzahl seiner Nachfolger.
- Ein *Pfad* ist ein nichtleeres Tupel $\langle v_1, \dots, v_n \rangle$, dessen Komponenten durch Kanten verbundene Knoten sind. Genauer:
 1. Für jeden Knoten v ist das Tupel $\langle v \rangle$ ein Pfad.
 2. Für $n \geq 2$ Knoten $v_1, \dots, v_n$ mit $(v_i, v_{i+1}) \in E$ für alle $i \in \{1, \dots, n-1\}$ ist das Tupel $\langle v_1, \dots, v_n \rangle$ ein Pfad.
 Dabei heißt $n-1$ die *Länge*, v_1 der *Ausgangspunkt* und v_n der *Endpunkt* des Pfades. Wir sagen auch, dass der Pfad von v_1 nach v_n geht.
- Ein Pfad heißt *einfach* genau dann, wenn er keinen Knoten mehrfach enthält.
- Ein Pfad $\langle v_1, \dots, v_n \rangle$ heißt *Zyklus* genau dann, wenn $n \geq 2$ und $v_1 = v_n$.
- Ein Knoten v' ist von einem Knoten v aus *erreichbar* genau dann, wenn ein Pfad von v nach v' existiert.
- Wir schreiben $v \rightarrow_G v'$ genau dann, wenn $(v, v') \in E$.
- Wir schreiben $v \rightarrow_G^* v'$ genau dann, wenn v' von v aus erreichbar ist.

Für den Graphen in Abbildung 5.1 gilt:

Ein Beispiel

- Der Graph hat die Quellen 1 und 6, die Senke 5 und die inneren Knoten 2, 3 und 4.
- Der Knoten 2 hat den Eingrad 2 und den Ausgrad 2.
- Der Knoten 6 hat den Eingrad 0 und den Ausgrad 1.
- Vom Knoten 2 aus sind die Knoten 2, 5, 3 und 4 erreichbar.
- Vom Knoten 1 aus sind alle Knoten bis auf 6 erreichbar.
- Der Pfad $\langle 2, 3, 4, 2 \rangle$ ist ein Zyklus.

Wir brauchen noch mehr Begriffe für Graphen:

- Die *Größe* eines Graphen ist die Anzahl seiner Knoten.

- Ein Graph heißt *nichtleer* genau dann, wenn er mindestens einen Knoten hat. Der Graph $(\emptyset, \emptyset)$ heißt der *leere Graph*.
- Die *Tiefe* eines nichtleeren Graphen ist die maximale Länge seiner einfachen Pfade.
- Ein Graph heißt *zyklisch* genau dann, wenn er einen Zyklus enthält.
- Ein Graph heißt *azyklisch* genau dann, wenn er keinen Zyklus enthält.
- Der *symmetrische Abschluss* eines Graphen $G = (V, E)$ ist der Graph $\overleftrightarrow{G} = (V, E \cup E^{-1})$ mit $E^{-1} = \{(v, v') \mid (v', v) \in E\}$
- Ein Graph heißt *stark zusammenhängend* genau dann, wenn jeder seiner Knoten von jedem seiner Knoten aus erreichbar ist.
- Ein Graph (V, E) heißt *zusammenhängend* genau dann, wenn $\overleftrightarrow{G}$ stark zusammenhängend ist.

Abbildung 5.3 zeigt den symmetrischen Abschluss des Graphen in Abbildung 5.2. Der Graph in Abbildung 5.2 ist nichtleer, zyklisch und zusammenhängend, aber nicht stark zusammenhängend. Er hat die Größe 6 und die Tiefe 3. Der Graph in Abbildung 5.3 ist nichtleer, zyklisch, zusammenhängend und stark zusammenhängend. Er hat die Größe 6 und die Tiefe 5.

Proposition 5.2.2 (Azyklische Graphen) Für jeden Graph G sind die folgenden Aussagen äquivalent:

1. G ist azyklisch.
2. G hat nur einfache Pfade.
3. G hat nur endlich viele Pfade.

Weiter gilt für jeden azyklischen Graphen G : Zu jedem Knoten v existiert eine Quelle, von der aus v erreichbar ist.

Beweis Die Äquivalenzen sind offensichtlich. Sei v ein Knoten in einem azyklischen Graphen G . Wir wählen einen einfachen Pfad maximaler Länge, dessen Endpunkt v ist. Der Ausgangspunkt dieses Pfades muss eine Quelle sein. $\square$

Teilgraphen

Ein Graph $G = (V, E)$ heißt *Teilgraph* eines Graphen $G' = (V', E')$ genau dann, wenn $V \subseteq V'$ und $E \subseteq E'$.

Sei $G = (V, E)$ ein Graph und $v \in V$. Der von v aus *erreichbare Teilgraph* von G ist wie folgt definiert:

$$G^v = (V', E \cap V' \times V') \quad \text{mit} \quad V' = \{v' \in V \mid v \rightarrow_G^* v'\}$$

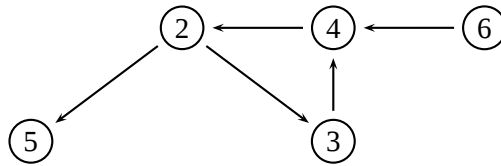


Abbildung 5.4: Der von 6 aus erreichbare Teilgraph des Graphen in Abb. 5.2

Sei G der Graph in Abbildung 5.3. Dann gilt $G^v = G$ für jeden Knoten v von G . Abbildung 5.4 zeigt den von 6 aus erreichbaren Teilgraphen des Graphen in Abbildung 5.2.

Proposition 5.2.3 *Sei G ein Graph mit einem Knoten v . Dann ist der von v aus erreichbare Teilgraph von G zusammenhängend.*

5.3 Bäume

Ein *Wald* ist ein azyklischer Graph, in dem jeder Knoten höchstens einen Vorgänger hat. Die Quellen eines Waldes heißen auch *Wurzeln* und die Senken heißen auch *Blätter*.

Ein *Baum* ist ein Wald, der höchstens eine Wurzel hat.

Abbildung 5.5 zeigt einen Baum in der bereits bekannten Graphdarstellung und zusätzlich in einer abkürzenden Darstellung speziell für Bäume. Wenn man die Grafiken jeweils um 180 Grad dreht, erinnern sie tatsächlich an Bäume. Die Wurzel 1 erscheint jeweils als der oberste Knoten. Die Blätter sind die Knoten 3, 4, 6, 9, 10 und 11. Die inneren Knoten sind 2, 5, 7 und 8. Die Größe des Baumes ist 11 und die Tiefe ist 3. Formal handelt es sich um einen Baum wie folgt:

$$B = (V, E)$$

$$V = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$$

$$E = \{(1, 2), (1, 5), (1, 7), (2, 3), (2, 4), (5, 6), (7, 8), (7, 11), (8, 9), (8, 10)\}$$

Wenn man einen Wald grafisch darstellt, bekommt man tatsächlich eine Ansammlung von Bäumen.

Proposition 5.3.1 (Wälder) *Sei G ein azyklischer Graph. Dann sind die folgenden Aussagen äquivalent:*

1. G ist ein Wald.

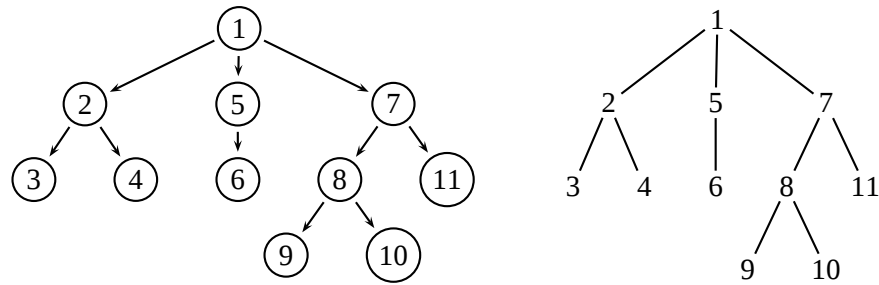


Abbildung 5.5: Ein Baum in Graph- und Baumdarstellung

2. Jeder Knoten hat höchstens einen Vorgänger.
3. Zwischen zwei Knoten gibt es höchstens einen Pfad.

Weiter gilt für jeden Wald W : Zu jedem Knoten v existiert genau eine Wurzel, von der aus v erreichbar ist.

Beweis Folgt mit Proposition 5.2.2. □

Proposition 5.3.2 (Bäume) Sei B ein nichtleerer Baum. Dann gilt:

1. B hat genau eine Wurzel.
2. Zu jedem Knoten v gibt es genau einen Pfad von der Wurzel nach v .
3. B ist zusammenhängend.

Beweis Folgt mit Proposition 5.3.1. □

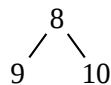
Teilbäume

Ein *Teilbaum* eines Baums B ist ein Teilgraph von B , der ein Baum ist.

Proposition 5.3.3 Sei v ein Knoten eines Baums B . Dann ist der von v aus erreichbare Teilgraph von B ein Teilbaum von B mit der Wurzel v .

Folglich nennt man den von v aus erreichbaren Teilgraphen B^v eines Baums B auch den von v aus erreichbaren *Teilbaum* von B .

Der vom Knoten 8 aus erreichbare Teilbaum des Baums in Abbildung 5.5 ist



Unterbäume

Die *Unterbäume* eines Baums B sind die von den Nachfolgern der Wurzel von B aus erreichbaren Teilbäume. Der Baum in Abbildung 5.5 hat drei Unterbäume:

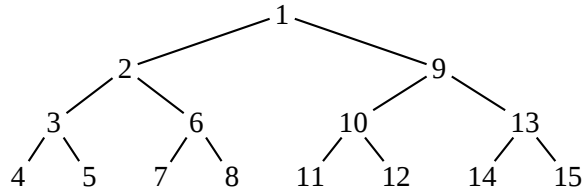
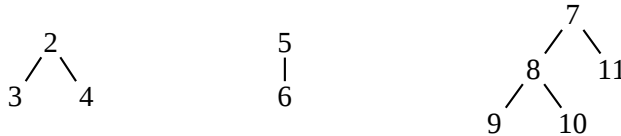


Abbildung 5.6: Ein balancierter Binärbaum



Gegeben ein Baum B und ein Knoten v von B , werden die Unterbäume des von v aus erreichbaren Teilbaums als *Unterbäume von v* bezeichnet.

Ein Baum heißt *linear* genau dann, wenn wenn jeder seiner Knoten höchstens einen Nachfolger hat. Offensichtlich ist ein Baum genau dann linear, wenn er höchstens ein Blatt hat. *Linear*

Ein Baum heißt *binär* genau dann, wenn jeder seiner Knoten, der kein Blatt ist, genau zwei Nachfolger hat. *Binär*

Ein Baum heißt *balanciert* genau dann, wenn für jeden Knoten v gilt, dass alle Unterbäume des von v aus erreichbaren Teilbaums die gleiche Tiefe haben. Ein nichtleerer Baum ist also genau dann balanciert, wenn alle Pfade von der Wurzel zu einem Blatt die gleiche Länge haben. *Balanciert*

Abbildung 5.6 zeigt ein Beispiel für einen balancierten Binärbaum.

Balanciert und binär

5.4 Größenverhältnisse in Bäumen

Proposition 5.4.1 Sei $B = (V, E)$ ein nichtleerer Baum. Dann gilt $|V| = |E| + 1$.

Beweis Durch Induktion über $|V|$. (Wir benutzen die Induktionsregel 5.3.)

Sei $|V| = 1$. Dann hat B keine Kante, da Bäume azyklisch sind. Also gilt $|V| = |E| + 1$.

Sei $|V| > 1$. Seien $B_1 = (V_1, E_1), \dots, B_n = (V_n, E_n)$ die $n \geq 1$ Unterbäume

von B . Dann gilt:

$$\begin{aligned}
 |E| &= n + |E_1| + \cdots + |E_n| && (E_1, \dots, E_n \text{ paarweise disjunkt}) \\
 &= n + |V_1| - 1 + \cdots + |V_n| - 1 && (\text{nach Induktionsannahme}) \\
 &= |V_1| + \cdots + |V_n| \\
 &= |V| - 1 && (V_1, \dots, V_n \text{ paarweise disjunkt}) \quad \square
 \end{aligned}$$

$g(B)$, $t(B)$ und $b(B)$ Sei B ein Baum. Dann bezeichnen wir mit $g(B)$ die Größe, mit $t(B)$ die Tiefe und mit $b(B)$ die Anzahl der Blätter von B .

Proposition 5.4.2 Sei B ein nichtleerer Baum, bei dem jeder Knoten, der kein Blatt ist, mindestens zwei Nachfolger hat. Dann $2b(B) > g(B)$.

Beweis Durch Induktion über $g(B)$.

Sei $g(B) = 1$. Dann ist der einzige Knoten von B ein Blatt. Also $2b(B) > g(B)$.

Sei $g(B) > 1$. Seien $B_1, \dots, B_n$ die $n \geq 2$ Unterbäume von B . Da $B_1, \dots, B_n$ keine gemeinsamen Knoten haben, gilt:

$$\begin{aligned}
 2b(B) &= 2(b(B_1) + \cdots + b(B_n)) \\
 &\geq (g(B_1) + 1) + \cdots + (g(B_n) + 1) && (\text{mit Induktionsannahme}) \\
 &= g(B_1) + \cdots + g(B_n) + n \\
 &> g(B) && (\text{da } n \geq 2) \quad \square
 \end{aligned}$$

Proposition 5.4.3 Für jeden nichtleeren, balancierten Binärbaum B gilt: $b(B) = 2^{t(B)}$.

Beweis Durch Induktion über $t(B)$.

Sei $t(B) = 0$. Dann hat B genau einen Knoten, der ein Blatt ist. Also gilt $b(B) = 2^{t(B)}$.

Sei $t(B) > 0$. Seien B_1 und B_2 die beiden Unterbäume von B . Da B balanciert ist, gilt $t(B_1) = t(B_2) = t(B) - 1$. Also

$$\begin{aligned}
 b(B) &= b(B_1) + b(B_2) \\
 &= 2^{t(B_1)} + 2^{t(B_2)} && (\text{nach Induktionsannahme}) \\
 &= 2^{t(B)-1} + 2^{t(B)-1} \\
 &= 2^{t(B)} \quad \square
 \end{aligned}$$

Proposition 5.4.4 Für jeden nichtleeren, balancierten Binärbaum B gilt: $g(B) = 2^{t(B)+1} - 1$.

Beweis Durch Induktion über $g(B)$.

Sei $g(B) = 1$. Dann gilt $t(B) = 0$ und folglich $g(B) = 2^{t(B)+1} - 1$.

Sei $g(B) > 1$. Seien B_1 und B_2 die beiden Unterbäume von B . Da B balanciert ist, gilt $t(B_1) = t(B_2) = t(B) - 1$. Also

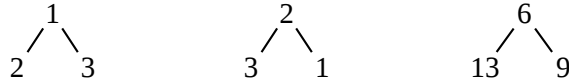
$$\begin{aligned}
 g(B) &= g(B_1) + g(B_2) + 1 \\
 &= 2^{t(B_1)+1} - 1 + 2^{t(B_2)+1} - 1 + 1 \quad (\text{nach Induktionsannahme}) \\
 &= 2^{t(B)} + 2^{t(B)} - 1 \\
 &= 2^{t(B)+1} - 1
 \end{aligned}$$

□

5.5 Geordnete Bäume

Wir haben Bäume etwas indirekt als spezielle Graphen formalisiert. Wir geben jetzt eine direkte Formalisierung für Bäume mit einer zusätzlichen Ordnungseigenschaft an.

Wenn wir die grafische Darstellung eines Baumes betrachten, können wir zwischen *Form* und *Markierung* unterscheiden. Beispielsweise haben die Bäume



alle die gleiche Form und unterscheiden sich nur in der Markierung. Die neue Definition von Bäumen wird im Gegensatz zur alten eine explizite Unterscheidung zwischen Form und Markierung machen.

Wir beginnen mit der Definition von Baumbereichen. Baumbereiche stellen nur die Form von Bäumen dar und verzichten völlig auf die Markierung.

Definition 5.5.1 Ein Baumbereich ist eine Menge D wie folgt:

1. D ist eine endliche Teilmenge von $\mathcal{L}(\mathbb{N}^+)$.
2. $\forall u \in \mathcal{L}(\mathbb{N}^+) \forall i \in \mathbb{N}: u @ [i] \in D \Rightarrow u \in D$.
3. $\forall u \in \mathcal{L}(\mathbb{N}^+) \forall i \in \mathbb{N}: u @ [i] \in D \wedge i \geq 2 \Rightarrow u @ [i - 1] \in D$.

Die Elemente von D werden als die Knoten oder Pfade des Baumbereichs bezeichnet.

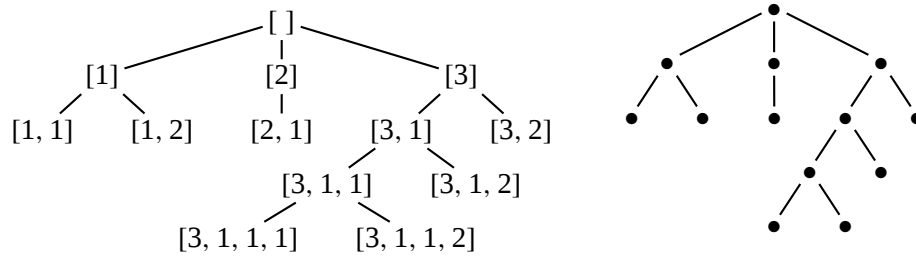


Abbildung 5.7: Zwei grafische Darstellung eines Baumbereichs

Hier ist ein Beispiel für einen Baumbereich:

$$\{ [],$$

$$[1], [1, 1], [1, 2],$$

$$[2], [2, 1],$$

$$[3], [3, 1], [3, 1, 1], [3, 1, 1, 1], [3, 1, 1, 2], [3, 1, 2], [3, 2] \}$$

Dieser Baumbereich ist in Abbildung 5.7 zweifach dargestellt. Aus beiden Darstellungen kann man den Baumbereich eindeutig rekonstruieren. Die beiden Darstellungen machen den Zusammenhang zwischen Baumbereichen und der Form von Bäumen klar.

Es ist nicht schwer, die Knoten eines Baumbereichs mit Markierungen zu versehen.

Definition 5.5.2 Sei X eine Menge. Ein geordneter Baum über X ist eine endliche Funktion $b \in \mathcal{L}(\mathbb{N}^+) \xrightarrow{\text{fin}} X$, sodass $\text{Dom } b$ ein Baumbereich ist.

Sei b eine geordneter Baum über X . Dann bezeichnen wir die Elemente von $\text{Dom } b$ als die Knoten von b und $b(v)$ als die Marke des Knotens $v \in \text{Dom } b$.

Die Menge aller geordneten Bäume über X bezeichnen wir mit $\mathcal{B}(X)$.

Abbildung 5.8 zeigt ein Beispiel für einen geordneten Baum über $\mathbb{N}$. Die Tabelle links beschreibt eine endliche Funktion und stellt damit den geordneten Baum formal dar. Das Bild rechts stellt den geordneten Baum grafisch dar. Beachten Sie, dass mehrere Knoten die gleiche Markierung tragen.

Eine geordneter Baum b legt im Gegensatz zu einem Baum $B = (V, E)$ (als Graph) für jeden seiner Knoten v eine Ordnung für die Nachfolger von v fest. Dies geschieht mithilfe des Baumbereichs $\text{Dom } b$, der die Nachfolger von v beginnend mit 1 durchnummeriert (siehe Abbildung 5.7).

[]	1
[1]	2
[1,1]	3
[1,2]	1
[2]	2
[2,1]	3
[3]	1
[3,1]	2
[3,1,1]	3
[3,1,1,1]	1
[3,1,1,2]	2
[3,1,2]	3
[3,2]	1

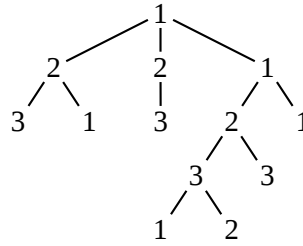


Abbildung 5.8: Formale und grafische Darstellung eines geordneten Baums

Definition 5.5.3 Sei b ein geordneter Baum über X . Dann ist der b zugeordnete Graph $G(b) = (V, E)$ wie folgt definiert:

$$V = b$$

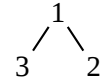
$$E = \{(v, v') \in b \times b \mid \exists u, i, x, x': v = (u, x) \wedge v' = (u @ [i], x')\}$$

Proposition 5.5.4 Seien b und b' geordnete Bäume über einer Menge X . Dann gilt:

1. Der b zugeordnete Graph $G(b)$ ist ein Baum.
2. Die Wurzel von $G(b)$ ist $([], b([]))$.
3. Der geordnete Baum b hat genau so viele Knoten wie der Graph $G(b)$.
4. Die Tiefe von $G(b)$ ist $\max\{|u| \mid u \in \text{Dom } b\}$.
5. Wenn $G(b) = G(b')$, dann $b = b'$.

Da wir jedem geordneten Baum eindeutig einen Graphen zuordnen, der ein Baum ist, lassen sich die Konzepte Wurzel, Blatt, innerer Knoten, Vorgänger und Nachfolger, Größe und Tiefe sowie Teilbaum und Unterbaum von Bäumen (als Graphen) auf geordnete Bäume übertragen. Selbstverständlich kann man diese Konzepte auch direkt auf geordneten Bäumen definieren.

Betrachten Sie die grafischen Baumdarstellungen



Beide Darstellungen beschreiben denselben Graphen

$$V = \{1, 2, 3\}$$

$$E = \{(1, 2), (1, 3)\}$$

Graphen legen also keine Ordnung für die Nachfolger eines Knoten fest.

Wenn wir die obigen Baumdarstellungen als grafische Darstellungen von geordneten Bäumen auffassen, bekommen wir dagegen zwei unterschiedliche geordnete Bäume:

[]	1
[1]	2
[2]	3

[]	1
[1]	3
[2]	2

Jedem der zwei geordneten Bäume ist per Definition ein Graph zugeordnet. Diese Graphen sind unterschiedlich. Hier ist der Graph für den linken geordneten Baum:

$$V = \{([], 1), ([1], 2), ([2], 3)\}$$

$$E = \{([], 1), ([1], 2), ([1], 1), ([2], 3)\}$$

Daraus lernen wir, dass grafische Darstellungen ganz unterschiedlich interpretiert werden können. Diese Interpretationsfreiheit ist für formale Darstellungen ausgeschlossen.

5.6 Terme

Wir geben jetzt eine zweite Formalisierung für nichtleere, geordnete Bäume an.

Definition 5.6.1 Sei X eine Menge. Die Menge $\mathcal{T}(X)$ der Terme über X ist wie folgt rekursiv definiert:

$$\frac{x \in X \quad n \geq 0 \quad t_1 \in \mathcal{T}(X) \quad \dots \quad t_n \in \mathcal{T}(X)}{\langle x, t_1, \dots, t_n \rangle \in \mathcal{T}(X)}$$

Gegeben ein Term $t = \langle x, t_1, \dots, t_n \rangle \in \mathcal{T}(X)$, heißt t_i der i -te Unterterm von t ($i \in \{1, \dots, n\}$). Weiter heißt x die Wurzelmarke von t .

Den geordneten Baum in Abbildung 5.8 können wir durch den folgenden Term über $\mathbb{N}$ darstellen:

$$\begin{aligned} &\langle 1, \langle 2, \langle 3 \rangle, \langle 1 \rangle \rangle, \\ &\quad \langle 2, \langle 3 \rangle \rangle, \\ &\quad \langle 1, \langle 2, \langle 3, \langle 1 \rangle, \langle 2 \rangle \rangle, \\ &\quad \quad \langle 3 \rangle \rangle, \\ &\quad \langle 1 \rangle \rangle \end{aligned}$$

Sei X eine Menge. Den Zusammenhang zwischen Termen und geordneten Bäumen über X formalisieren wir durch eine Funktion

$$B \in \mathcal{T}(X) \rightarrow \mathcal{B}(X) - \{\emptyset\}$$

die wie folgt rekursiv definiert ist:

$$B\langle x, t_1, \dots, t_n \rangle = \{([], x)\} \cup \{(i :: u, y) \mid (u, y) \in B(t_i), 1 \leq i \leq n\}$$

Man überzeugt sich leicht davon, dass B eine Bijektion von $\mathcal{T}(X)$ nach $\mathcal{B}(X) - \{\emptyset\}$ ist. Das bedeutet, dass sich Terme und nichtleere, geordnete Bäume über X eineindeutig entsprechen.

Sei $t \in \mathcal{T}(X)$. Die *Knoten* oder *Pfade* von t sind die Listen in $\text{Dom } B(t)$.

Jedem Term $t \in \mathcal{T}(X)$ ist der Baum $G(B(t))$ zugeordnet. Damit lassen sich die Konzepte Größe, Tiefe, Teilbaum und Unterbaum von Bäumen (als Graphen) auf Terme übertragen. Selbstverständlich kann man diese Konzepte auch direkt auf Termen definieren.

Definition 5.6.2 Sei X eine Menge. Dann sind die Funktionen γ und τ wie folgt rekursiv definiert:

$$\gamma \in \mathcal{T}(X) \rightarrow \mathbb{N}$$

$$\gamma(\langle x \rangle) = 1$$

$$\gamma(\langle x, t_1, \dots, t_n \rangle) = 1 + \gamma(t_1) + \dots + \gamma(t_n)$$

$$\tau \in \mathcal{T}(X) \rightarrow \mathbb{N}$$

$$\tau(\langle x \rangle) = 0$$

$$\tau(\langle x, t_1, \dots, t_n \rangle) = 1 + \max\{\tau(t_1), \dots, \tau(t_n)\}$$

Proposition 5.6.3 Sei $t \in \mathcal{T}(X)$. Dann gilt:

1. Die Größe von $G(B(t))$ ist $\gamma(t)$.
2. Die Tiefe von $G(B(t))$ ist $\tau(t)$.
3. Wenn $t = \langle x, t_1, \dots, t_n \rangle$, dann sind

$$G(B(t_1)), \dots, G(B(t_n))$$

die Unterbäume von $G(B(t))$.

Übungen

Aufgabe 5.1 (Induktion) Gegeben sei die rekursiv definierte Funktion

$$f \in \mathbb{N} \rightarrow \mathbb{N}$$

$$f(n) = \text{if } n = 0 \text{ then } 1 \text{ else } f(n-1) + 2n + 1$$

Beweisen Sie durch Induktion über n :

$$\forall n \in \mathbb{N}: f(n) = (n+1)^2$$

Aufgabe 5.2 (Induktion) Gegeben sei die rekursiv definierte Funktion

$$f \in \mathbb{N} \rightarrow \mathbb{N}$$

$$f(n) = \text{if } n = 0 \text{ then } 0 \text{ else } f(n-1) + n^2$$

Beweisen Sie durch Induktion über n :

$$\forall n \in \mathbb{N}: f(n) = \frac{n}{6}(2n^2 + 3n + 1)$$

Aufgabe 5.3 (Induktion) Gegeben sei die rekursiv definierte Funktion

$$f \in \mathbb{R} \times \mathbb{N} \rightarrow \mathbb{R}$$

$$f(q, n) = \text{if } n = 0 \text{ then } 1 \text{ else } f(q, n-1) + q^n$$

Beweisen Sie durch Induktion über n :

$$\forall n \in \mathbb{N} \forall q \in \mathbb{R}, q \neq 1: f(q, n) = \frac{1 - q^{n+1}}{1 - q}$$

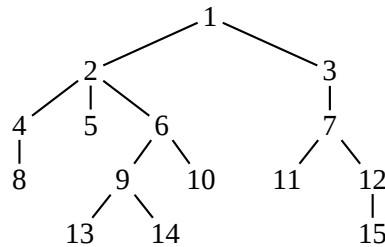
Aufgabe 5.4 (Graphen) Seien die folgenden Graphen gegeben:

1. $V = \{1, 2, 3, 4, 5, 6\}$
 $E = \{(1, 5), (2, 1), (2, 3), (2, 4), (3, 5), (6, 2), (6, 3)\}$
2. $V = \{1, 2, 3, 4, 5, 6, 7\}$
 $E = \{(2, 7), (3, 1), (3, 6), (4, 2), (4, 3), (7, 5)\}$
3. $V = \{1, 2, 3, 4, 5, 6, 7, 8\}$
 $E = \{(2, 1), (2, 6), (3, 8), (4, 4), (5, 2), (5, 7), (6, 1), (7, 6)\}$
4. $V = \{1, 2, 3, 4, 5\}$
 $E = \{(1, 4), (2, 4), (3, 2), (3, 5), (4, 3), (5, 1)\}$

Zeichnen Sie diese Graphen und beantworten Sie für jeden Graphen die folgenden Fragen:

- Welche Größe und welche Tiefe hat der Graph?
- Welche Quellen und welche Senken hat der Graph?
- Ist der Graph zyklisch? Wenn ja, geben Sie einen Zyklus an.
- Ist der Graph zusammenhängend? Stark zusammenhängend?
- Ist der Graph ein Wald? Ein Baum?

Aufgabe 5.5 (Bäume) Gegeben sei die grafische Darstellung eines Baums:



1. Geben Sie den Baum als Graph (V, E) an.
2. Geben Sie den Baum als geordneten Baum in Tabellenform an.
3. Geben Sie den Baum als Term an.
4. Geben Sie die Größe und Tiefe des Baums an.
5. Geben Sie die Wurzel, die Blätter, und die inneren Knoten des Baums an.
6. Geben Sie für jedes Blatt v einen Pfad an, der von der Wurzel zu v führt.
7. Zeichnen Sie alle Teilbäume des Baums, die den Knoten 7 als Wurzel haben.

8. Zeichnen Sie den vom Knoten 2 aus erreichbaren Teilbaum des Baums.
9. Zeichnen Sie alle Unterbäume des vom Knoten 2 aus erreichbaren Teilbaums des Baums.

Aufgabe 5.6 (Geordnete Bäume) Gegeben sei der geordnete Baum

[]	4
[1]	3
[2]	0
[2,1]	4
[2,2]	2
[2,2,1]	1
[2,2,2]	6
[2,3]	7
[3]	2
[3,1]	4
[3,1,1]	0
[3,2]	5

1. Zeichnen Sie den Baum.
2. Geben Sie die Größe und Tiefe des Baums an.
3. Wieviele Unterbäume hat der Baum?
4. Geben Sie den von [3] aus erreichbaren Teilbaum in Tabellenform an.
5. den von [2,1] aus erreichbaren Teilbaum in Tabellenform an.

Aufgabe 5.7 (Größenverhältnis in Binärbäumen) Beweisen Sie durch Induktion über $t(B)$, dass für jeden nichtleeren Binärbaum B gilt: $b(B) \leq 2^{t(B)}$.

Aufgabe 5.8 (Größenverhältnisse in ternären Bäumen) Ein *ternärer* Baum ist ein Baum, bei dem jeder Knoten, der kein Blatt ist, genau 3 Unterbäume hat. Beweisen Sie durch Induktion über $t(B)$, dass für jeden nichtleeren, balancierten ternären Baum B gilt:

1. $b(B) = 3^{t(B)}$
2. $g(B) = \frac{1}{2}(3^{t(B)+1} - 1)$

Aufgabe 5.9 (Balancierte Binärbäume (schwer)) Beweisen Sie durch Induktion über $t(B)$, dass für jeden nichtleeren Binärbaum B gilt:

$$b(B) = 2^{t(B)} \Rightarrow B \text{ balanciert}$$

Verwenden Sie dabei die Ungleichung aus Aufgabe 5.7.