7

Graphen, Bäume, Terme

Graphen stellen spezielle Relationalstrukturen dar, wie sie etwa bei der abstrakten Repräsentation der Struktur von Programmen, Arbeitsabläufen oder Wegnetzen in natürlicher Weise auftreten. Da der Begriff sehr allgemein ist, gibt es eine Vielzahl spezieller Klassen von Graphen. In Abschnitt 7.1 stellen wir in Form eines kurzen Überblicks die wichtigsten Arten vor, ohne allerdings hier tiefer auf mathematische Zusammenhänge einzugehen. In Abschnitt 7.2 kommen wir dann ausführlicher auf die wichtigste Klasse von Graphen, die sogenannten Bäume zu sprechen. Bäume begegnen uns in vielen alltäglichen Zusammenhängen als Gliederungsschema. Sie werden darüberhinaus in der Informatik und Logik insbesondere zur Repräsentation der Struktur von Termen und Formeln, in der Linguistik zur Repräsentation der Struktur natürlich-sprachlicher Sätze verwendet. Für Bäume einer speziellen Form ist eine einfache algebraische Beschreibung möglich, auf die wir in Abschnitt 7.3 eingehen. In Kapitel 7.4 führen wir Terme ein und geben wechselseitige Übersetzungen zwischen Bäumen und Termen an. Die Darstellung einiger wichtiger algebraischer Eigenschaften von Baum- und Termalgebra beendet den Kern des Kapitels.

Die Ergänzungen am Kapitelende beginnen mit einer kurzen Darstellung der sogenannten ,,Unifikation" von Termen. Diese Operation stellt den zentralen Rechenschritt der Programmiersprache Prolog dar, sie wird uns auch später im Bereich der Logik beim resolutionsbasierten Theorembeweisen begegnen. Schließlich kehren wir in Abschnitt 7.5.2 nochmals zu den Graphen zurück und betrachten das bekannte Problem, in einem endlichen Graphen einen sogenannten Eulerschen Kreis zu finden. Die Hauptgegenstände dieses

Kapitels, Graphen und Bäume, sind aufgrund ihrer großen Bedeutung in der Informatik in vielen Büchern detaillierter besprochen, auch mit Hinblick auf algorithmische Probleme. Enstprechende Literaturangaben finden sich am Kapitelende.

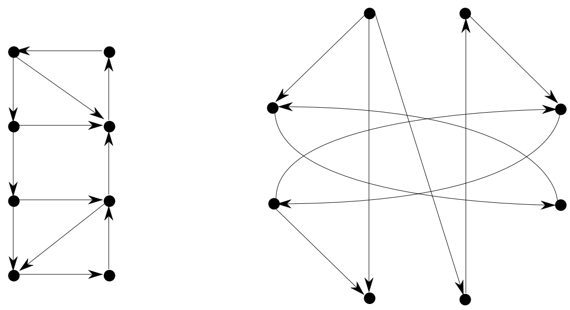
7.1 Beispiele und Typen von Graphen

Bei den Graphen unterscheidet man die Teilklassen der gerichteten und der ungerichteten Graphen. Wir stellen zunächst das Grundkonzept eines gerichteten Graphen vor.

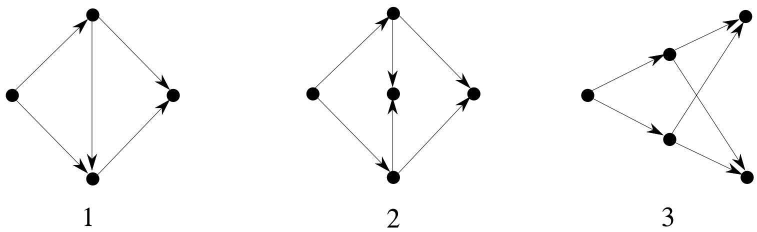
Gerichtete Graphen

Definition 7.1 Ein *gerichteter Graph* (engl. „directed graph" oder „digraph") ist ein Paar $\mathcal{G} = \langle N, E\rangle$ wo N eine nichtleere Menge ist und $E \subseteq N \times N$ eine binäre Relation auf N. Die Elemente von N werden *Knoten* (engl. „nodes" oder „vertices") genannt, die Elemente von E heißen *Kanten* (engl. „edges"). Ist $e = \langle n, m\rangle \in E$ eine Kante, so bezeichnet $start(e)$ (resp. $ende(e)$) den Knoten n (resp. m).

Gerichtete Graphen sind also einfache Relationalstrukturen im Sinn von Definition 6.17. Der Begriff ist sehr allgemein, da im Prinzip jede binäre Relation auf einer nichtleeren Menge einen Graphen definiert. Wenn man von gerichteten Graphen redet, hat man allerdings fast immer eine Form der bildlichen Repräsentation im Hinterkopf, bei der die Knoten durch Punkte repräsentiert sind und die Kanten als Pfeile zwischen diesen Punkten. Die folgende Abbildung repräsentiert zwei gerichtete Graphen auf diese Weise.



Bemerkung 7.2 Auf einige Unterschiede zwischen den Graphen selbst und ihrer bildlichen Repräsentation sei hingewiesen. In den Bildern werden die Punkte meist nicht mit Elementbezeichnungen versehen. Die tatsächliche Knotenmenge bleibt damit im Dunkeln, bei den vorausgegangenen Abbildungen kommt jeweils jede Grundmenge mit 8 Elementen in Frage. Damit legt diese Art der bildlichen Repräsentation den Graphen nur bis auf Isomorphie fest und abstrahiert von konkreten mengentheoretischen Darstellung. Da man isomorphe Strukturen ohnehin im Grunde gerne identifizieren möchte, ist dies eher ein Vorteil. In anderer Weise sind Abbildungen aber spezifischer als Graphen. Sie positionieren die Knoten in willkürlicher Weise im zweidimensionalen Raum. Dies führt teilweise zu eher unerwünschten Phänomenen. So kommen uns in der nachfolgenden Abbildung die Graphen 1 und 2 ähnlicher vor als 2 und 3, obwohl 2 und 3 isomorphe Graphen repräsentieren, wie wir in Aufgabe 7.1 sehen.



Es ist nicht schwer, Bilder von Graphen anzugeben, wo man ohne maschinelle Hilfe praktisch nicht mehr entscheiden kann, ob die zugehörigen Graphen isomorph sind. Sind die Graphen genügend groß, kann die Frage unter Umständen selbst mit Computern nicht in praktisch akzeptabler Zeit beantwortet werden.

Beispiel 7.3 Jedes Hasse-Diagramm (vgl. Beispiel 4.28 Nr. 3) repräsentiert einen endlichen gerichteten Graphen, wenn wir uns in Erinnerung rufen, daß die Kanten stets von unten nach oben gerichtet sind.

Beispiel 7.4 Bei geeigneter Abstraktion kann man auch viele Arten von Diagrammen als gerichtete Graphen ansehen. Wenn wir den textlichen Inhalt von Flussdiagrammen ignorieren, kommen wir zu Figuren wie der nachfolgenden, die Graphen darstellen.

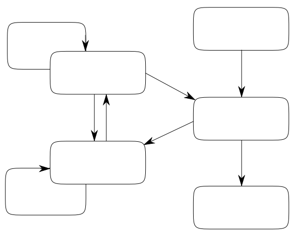


Diagramme ihrerseits werden häufig verwendet, um die Struktur von Programmen zu beschreiben. Graphen können damit die Struktur von Programmen auf einer abstrakten Ebene widerspiegeln.

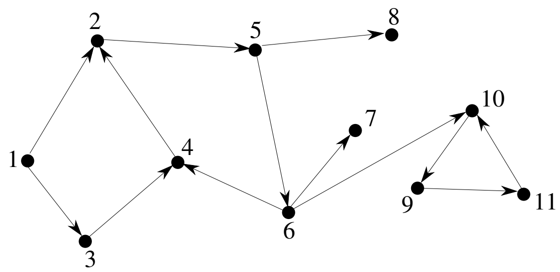
Definition 7.5 Sei $\mathcal{G} = \langle N, E\rangle$ ein gerichteter Graph. Der Knoten m heißt *direkter Nachfolger* des Knotens n genau dann, wenn $E(n,m)$ gilt. Umgekehrt heißt in dieser Situation n ein *direkter Vorgänger* von m. Eine Folge von Kanten $e_1, \ldots, e_k$ $(k \geq 1)$ heißt *Pfad* der Länge k genau dann, wenn $ende(e_i) = start(e_{i+1})$ für $1 \leq i < k$ gilt. Ist $m := start(e_1)$ und $n := ende(e_k)$, so heißt $e_1, \ldots, e_k$ ein *Pfad von n nach m*. Ein Pfad $e_1, \ldots, e_k$ heißt *geschlossen* genau dann, wenn $ende(e_k) = start(e_1)$ gilt. Ein Pfad der Länge $k \geq 1$ heißt *einfach* genau dann, wenn er k verschiedene Kanten enthält. Der Knoten m ist ein *Nachfolger* des Knotens n genau dann, wenn es einen Pfad $e_1, \ldots, e_k$ $(k \geq 1)$ gibt mit $start(e_1) = n$ und $ende(e_k) = m$. In dieser Situation heißt n ein *Vorgänger* von m.

Man beachte, daß ein Knoten durchaus ein Nachfolger oder Vorgänger von sich selbst sein kann. Die Kantenrelation E ist identisch zur direkten Nachfolger-Beziehung. Die Nachfolger-Relation ist gerade die transitive Hülle E^+ der Kantenrelation im Sinn von Definition 4.38. Im speziellen Fall, wo E^+ eine partielle Ordnung „$\leq$" ist, stimmen der obige (direkte)

Nachfolger- und Vorgängerbegriff mit den entsprechenden Begriffen aus Definition 4.27 überein. Der Pfadbegriff ist verwandt zum Begriff der R-Kette (vgl. Definition 3.28). Während R-Ketten jedoch Folgen von Elementen sind, stellen Pfade Folgen geordneter Paare dar. Es gilt jedoch offenkundig die folgende Reformulierung von Lemma 4.44.

Lemma 7.6 *Es sei $\mathcal{G} = \langle N, E\rangle$ ein gerichteter Graph. Dann gilt $E^+ = \{\langle n, m\rangle \in N \times N \mid \exists$ Pfad der Länge $n \geq 1$ von n nach $m\}$.*

Beispiel 7.7 Um die Begriffe zu illustrieren, labeln wir in der folgenden Darstellung eines Graphen $\mathcal{G}$ die Knoten mit den Elementnamen. Die direkten Nachfolger des Knotens 6 sind $4, 7, 10$. Die Vorgänger von 4 sind $1, 3, 6, 5, 2$ sowie 4 selbst, die Vorgänger von 8 sind $5, 2, 1, 4, 3, 6$. Es ist $\langle 1, 2\rangle, \langle 2, 5\rangle, \langle 5, 6\rangle, \langle 6, 10\rangle, \langle 10, 9\rangle, \langle 9, 11\rangle$ ein einfacher Pfad der Länge 6 von $\mathcal{G}$.



Der Pfad $\langle 4, 2\rangle, \langle 2, 5\rangle, \langle 5, 6\rangle, \langle 6, 4\rangle$ ist einfach und geschlossen. Die Pfade $\langle 2, 5\rangle, \langle 5, 6\rangle, \langle 6, 4\rangle, \langle 4, 2\rangle, \langle 2, 5\rangle$ und $\langle 9, 11\rangle, \langle 11, 10\rangle, \langle 10, 9\rangle, \langle 9, 11\rangle, \langle 11, 10\rangle$ sind *nicht* einfach.

Lemma 7.8 *Der Knoten n des gerichteten Graphen $\mathcal{G}$ sei ein Nachfolger (resp. Vorgänger) des Knotens m. Dann ist jeder Nachfolger (resp. Vorgänger) von n auch ein Nachfolger (resp. Vorgänger) von m.*

Beweis. Übung (vgl. Aufgabe 7.3). ∎

Gerichtete azyklische Graphen

Im allgemeinen ist auch bei endlichen Graphen die Kantenrelation E nicht wohlfundiert im Sinn von Bemerkung 4.50. Sobald nämlich ein geschlossener Pfad existiert, gibt es eine unendlich absteigende E-Kette. Wie das in 4.50 dargestellte Prinzip der wohlfundierten Induktion verdeutlicht, sind fundierte Relationen beweistechnisch — aber auch algorithmisch — leichter zu handhaben. Es liegt daher nahe, die Teilklasse derjenigen Graphen zu betrachten, die keine geschlossenen Pfade enthalten. Dies führt auf folgende Definition.

Definition 7.9 Ein gerichteter Graph $\mathcal{G}$ heißt *azyklisch* genau dann, wenn $\mathcal{G}$ keinen geschlossenen Pfad enthält. Gerichtete azyklische Graphen werden *DAGs* (für „Directed Acyclic Graph") genannt.

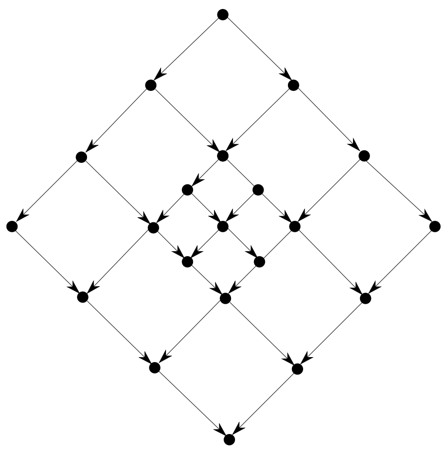
Lemma 7.10 *Ein gerichteter Graph $\mathcal{G}$ ist ein DAG genau dann, wenn für jeden Pfad $e_1, \dots, e_k$ von $\mathcal{G}$ die Knoten $start(e_1)$, $start(e_2)$, …, $start(e_k)$, $ende(e_k)$ stets paarweise verschieden sind.*

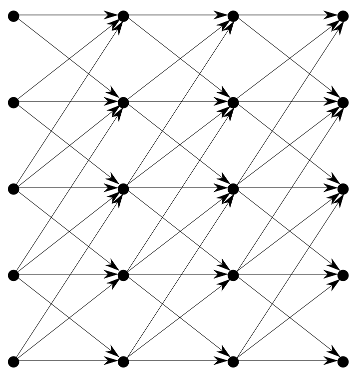
Beweis. Wären zwei Knoten der Folge $start(e_1)$, $start(e_2)$, …, $start(e_k)$, $ende(e_k)$ identisch, so hätte $\mathcal{G}$ offenkundig einen geschlossenen Pfad und wäre kein DAG. Sind für jeden Pfad $e_1, \dots, e_k$ von $\mathcal{G}$ die Knoten $start(e_1), \dots, start(e_k), ende(e_k)$ stets paarweise verschieden, so kann $\mathcal{G}$ keinen geschlossenen Pfad haben und ist ein DAG. ■

Aus Lemma 7.10 ergibt sich sofort

Korollar 7.11 *Ist $\mathcal{G}$ ein DAG mit n Knoten, so beträgt die Länge eines beliebigen Pfads in $\mathcal{G}$ höchstens n. Es gibt nur endliche viele verschiedene Pfade.*

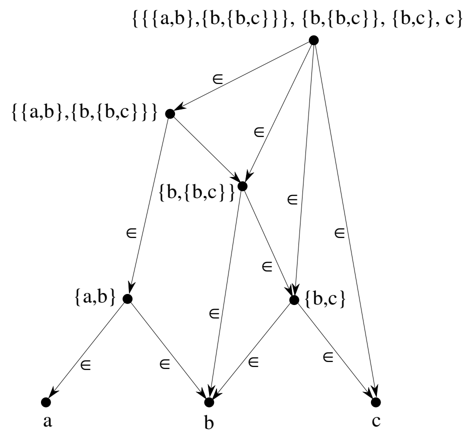
Beispiele 7.12 Die nachfolgende Abbildung zeigt zwei DAGs.





Jedes Hasse-Diagramm (vgl. Beispiele 4.28 Nr. 3) repräsentiert einen azyklischen gerichteten Graphen.

Beispiel 7.13 Wenn wir vom im Abschnitt 2.7.3 geschilderten Fundiertheitsprinzip für Mengen ausgehen, so können wir jede Menge M in kanonischer Weise als einen DAG darstellen, wo die Kantenrelation gerade der Elementbeziehung entspricht. Die Knoten selbst repräsentieren M, die Elemente von M, die Elemente von Elementen etc. Dabei wird jede auftretende Menge nur einmal repräsentiert. Betrachten wir als ein Beispiel die Menge $\{\{\{a,b\},\{b,\{b,c\}\}\},\{b,\{b,c\}\},\{b,c\},c\}$. Der zugehörige DAG hat folgende Form:



Lemma 7.14 *Ein gerichteter Graph $\mathcal{G} = \langle N, E\rangle$ ist ein DAG genau dann, wenn die Nachfolgerrelation E^+ auf $\mathcal{G}$ irreflexiv (damit eine strikte partielle Ordnung) ist.*

Beweis. Sei $\mathcal{G} = \langle N, E\rangle$ ein DAG, also azyklisch. Angenommen es gibt einen Knoten $n \in N$ mit $R(n, n)$, wo $R = E^+$ die Nachfolgerrelation ist. Dann gibt es gemäß Lemma 7.6 einen Pfad $e_1, \ldots, e_k$ der Länge $k \geq 1$ in $\mathcal{G}$ von n nach n. Da dieser Pfad geschlossen ist, erhalten wir einen Widerspruch. Folglich ist R irreflexiv. Da $R = E^+$ stets transitiv ist, ist dann E^+ eine strikte partielle Ordnung. Die Umkehrung erfolgt analog. ■

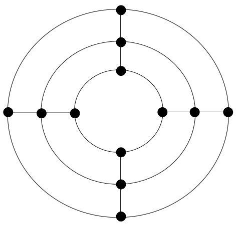
Lemma 7.15 *Es sei $\mathcal{G} = \langle N, E\rangle$ ein DAG mit endlicher Grundmenge N. Dann besitzt $\mathcal{G}$ zumindest ein Element ohne direkten Vorgänger (direkten Nachfolger).*

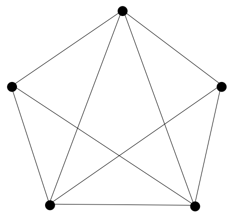
Beweis. Nach Lemma 7.14 ist E^+ eine strikte partielle Ordnung. Insbesondere folgt aus $E(m, n)$ stets $m \neq n$ für alle $m, n \in N$. Es ist $E^* = E^+ \cup Id_N$ eine partielle Ordnung auf der endlichen Menge N. Mit Lemma 4.32 folgt, daß es ein bezüglich E^* minimales (maximales) Element m gibt. Hätte m einen direkten Vorgänger (Nachfolger) n, so gilt $E(n, m)$ (resp. $E(m, n)$). Da andererseits $m \neq n$ folgt, widerspräche dies der Minimalität (Maximalität) von m bezüglich E^*. ■

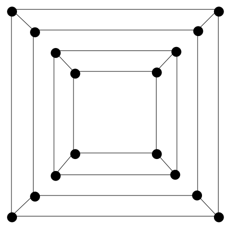
Symmetrische Graphen

Definition 7.16 Ein *symmetrischer Graph* ist ein gerichteter Graph $\mathcal{G} = \langle N, E\rangle$ wo die Kantenrelation E symmetrisch ist.

Symmetrische Graphen werden manchmal auch *ungerichtete Graphen* genannt. Mit $\langle m, n\rangle$ liegt auch stets $\langle n, m\rangle$ in der Kantenmenge. Ein Paar $\{\langle m, n\rangle, \langle n, m\rangle\}$ kann alternativ als Menge $\{m, n\}$ repräsentiert werden und wird oft als *ungerichtete Kante* bezeichnet. Eine ungerichtete Kante von n nach n hat damit die Form $\{n\}$. Bei der bildlichen Darstellung symmetrischer Graphen läßt man bei den Kanten die Richtungsangaben weg und liest alle Verbindungen als bidirektionale Kanten. Die nachfolgende Abbildung gibt einige Beispiele.



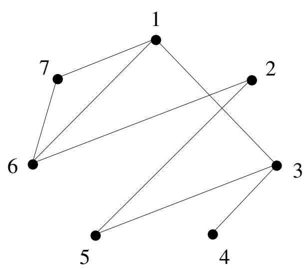




Symmetrische Graphen können zur formalen Analyse von Planungsproblemen und Arbeitsabfolgen verwendet werden. Typischerweise treten hierbei keine Kanten auf, die einen Knoten mit sich selbst verbinden.

Beispiel 7.17 In einer Firma sollen sieben Aufträge abgearbeitet werden. Jeder Auftrag dauert einen halben Tag, durch paralleles Abarbeiten soll die Lieferzeit möglichst kurz gehalten werden. Die Maschinen der Firma können gleichzeitig immer nur für einen Auftrag eingesetzt werden. Einige der vorhandenen Maschinen werden sowohl für Auftrag 1 als auch für Auftrag 3 benötigt. Ähnliches gilt für die Auftragspaare $\{1,6\}$, $\{1,7\}$, $\{2,5\}$, $\{2,6\}$, $\{3,4\}$, $\{3,5\}$ und $\{6,7\}$. Weitere Bedingungen sollen nicht existieren. Wieviele Halbtage sind mindestens notwendig, um die Aufträge abzuarbeiten?

Wir repräsentieren die Aufträge als Knoten eines symmetrischen Graphen. Zwei Knoten i und j werden mit einer Kante verbunden genau dann, wenn es eine Maschine gibt, die für Auftrag i und Auftrag j gebraucht wird.



Wir färben nun die Knoten derart, daß kein Knotenpaar derselben Farbe durch eine Kante verbunden ist. Dabei verwenden wir eine minimale Anzahl von Farben. Eine Lösung, die mit drei Farben auskommt, beruht auf der Zerlegung in die Teilmengen $\{1,2\}$, $\{3,6\}$ und $\{4,5,7\}$. Es ist nicht möglich, mit nur zwei Farben auszukommen. Wenn wir an einem Halbtag die Aufträge

1 und 2 parallel bearbeiten, an einem anderen die Aufträge 3 und 6, und am dritten Halbtag die Aufträge $4, 5$ und 7, so wird keine Maschine zeitgleich für zwei Jobs benötigt.

Definition 7.18 Eine *Knotenfärbung* vom Grad $n \in \mathbb{N}$ eines symmetrischen Graphen $\mathcal{G} = \langle V, E\rangle$ ist eine Abbildung f von V in eine Menge C mit n Elementen, so daß für jede ungerichtete Kante $\{m, n\}$ stets $f(m) \neq f(m)$ gilt.

Es ist ein bekanntes algorithmisches Problem, motiviert durch Anwendungen wie die oben geschilderte, zu einem gegebenen symmetrischen Graphen (wo E irreflexiv ist) eine Knotenfärbung minimalen Grades zu berechnen.

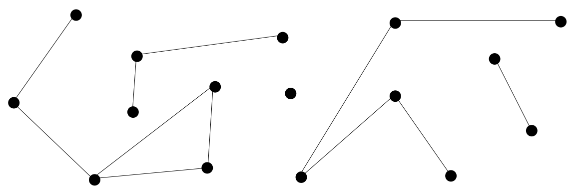
Definition 7.19 Ein symmetrischer Graph $\mathcal{G} = \langle N, E\rangle$ heißt *zusammenhängend* (engl. „connected") genau dann, wenn es zu je zwei verschiedenen Knoten $m, n \in N$ stets einen Pfad von m nach n gibt.

Man beachte, daß ein zusammenhängender Graph gemäß Definition 7.19 stets symmetrisch ist.

Lemma 7.20 *Sei $\mathcal{G} = \langle N, E\rangle$ ein symmetrischer Graph. Dann ist die reflexiv-transitive Hülle E^* von E eine Äquivalenzrelation auf der Knotenmenge N.*

Beweis. Wie Aufgabe 4.31 zeigt, ist die reflexiv-transitive Hülle E^* der symmetrischen Relation E selbst wieder symmetrisch. Natürlich ist E^* sowohl reflexiv als auch transitiv. Damit ist E^* eine Äquivalenzrelation. ■

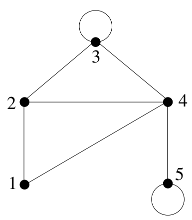
Als Übungsaufgabe (vgl. Aufgabe 7.7). werden wir zeigen, daß ein symmetrischer Graph $\mathcal{G}$ zusammenhängend ist genau dann, wenn die Quotientenmenge N/E^* genau ein Element hat. Die Äquivalenzklassen von E^* werden auch die *Zusammenhangskomponenten* des Graphen $\mathcal{G}$ genannt. Die Zusammenhangskomponenten sind genau die maximalen Teilmengen, wo je zwei Knoten stets durch einen Pfad verbunden sind. Der nachfolgend dargestellte Graph besitzt fünf Zusammenhangskomponenten.



Definition 7.21 Sei $\mathcal{G} = \langle N, E\rangle$ ein symmetrischer Graph. Der *Grad* $w(n)$ des Knotens $n \in N$ ist

1. die Zahl der ungerichteten Kanten der Form $\{n, m\}$ mit $n \neq m$, plus
2. 2 (resp. 0) falls es eine (keine) ungerichtete Kante der Form $\{n\}$ gibt.

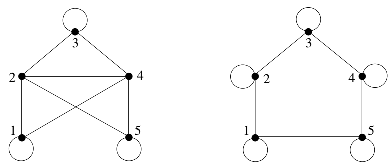
Bei der üblichen bildlichen Darstellung eines symmetrischen Graphen entspricht dem Grad eines Knotens damit die Zahl der Linien, die von dem Knoten ausgehen. Im nachfolgenden Beispiel haben die Knoten 2 und 5 den Grad 3, die Knoten 3 und 4 haben den Grad 4, der Knoten 1 hat den Grad 2.



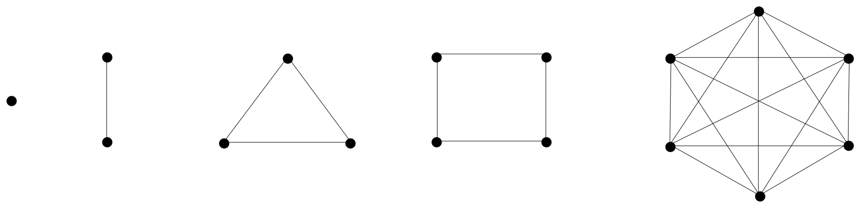
Da jede ungerichtete Kante eines symmetrischen Graphen entweder bei einem Knoten zwei Beiträge zum Grad oder bei zwei Knoten genau einen Beitrag zum Grad liefert, gilt das folgende Lemma.

Lemma 7.22 *In einem endlichen symmetrischen Graphen beträgt die Summe der Grade aller Knoten gerade* $2e$*, wo* e *die Zahl der ungerichteten Kanten angibt.*

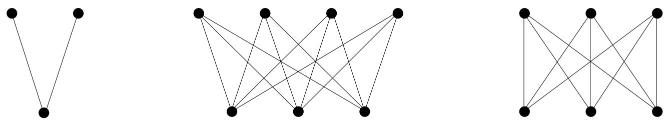
Beispiel 7.23 Ein symmetrischer Graph $\mathcal{G}$ heißt *regulär* genau dann, wenn jeder Knoten von $\mathcal{G}$ denselben Grad hat. Die nachfolgende Abbildung stellt zwei reguläre Graphen dar, wo jeder Knoten den Grad 4 besitzt.



Hat der reguläre Graph $\mathcal{G}$ genau k Knoten, hat jeder Knoten den Grad $k-1$, und sind je zwei verschiedene Knoten stets durch eine Kante verbunden, so spricht man von einem *vollständigen* Graphen. Für jede natürliche Zahl $k \geq 1$ gibt es bis auf Isomorphie genau einen vollständigen Graphen. Für $k = 1, \ldots 4, 6$ sind diese in der nachfolgenden Abbildung wiedergegeben.



Beispiel 7.24 Ein symmetrischer Graph $\mathcal{G} = \langle N, E\rangle$ heißt *Bipartitionsgraph* (engl. „bipartite graph“) genau dann, wenn es eine Partition der Knotenmenge N der Form $\{N_1, N_2\}$ gibt, so daß für jede gerichtete Kante $\langle n, m\rangle \in E$ stets entweder $n \in N_1$ und $m \in N_2$ oder aber $n \in N_2$ und $m \in N_1$ gilt. Dies ist äquivalent dazu, daß es eine Knotenfärbung vom Grad 2 für $\mathcal{G}$ gibt. Wenn darüberhinaus für alle $n \in N_1$ und $m \in N_2$ stets $\langle n, m\rangle \in E$ gilt, so heißt $\mathcal{G}$ ein *vollständiger Bipartitionsgraph*. Die nachfolgende Abbildung zeigt drei vollständige Bipartitionsgraphen.



Gelabelte Graphen

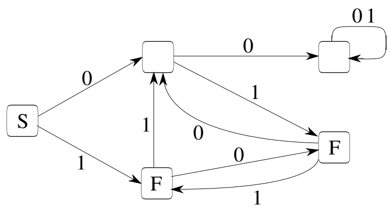
Am Ende dieses Teilkapitels wollen wir noch kurz auf einige Erweiterungen des Grundkonzepts eines Graphen eingehen. Häufig werden etwa *gelabel-*

te Graphen (machmal auch „etikettierte Graphen“ genannt) behandelt. Es können hierbei entweder

- nur die Knoten,
- nur die Kanten, oder
- sowohl Knoten als auch Kanten

gelabelt sein, die Etikettierung kann jeweils total oder nur partiell sein. Sind die Kanten gelabelt, so kann es mehrere Kanten von einem Knoten n zu einem Knoten m geben, man spricht gegebenenfalls von einem *Multigraphen*. Sowohl gerichtete als auch symmetrische Graphen können gelabelt sein. Nachfolgend zwei Beispiele gerichteter gelabelter Graphen.

Beispiel 7.25 Die nachfolgende Abbildung repräsentiert einen sogenannten „endlichen Automaten“ in Form eines gelabelten gerichteten Graphen.



Endliche Automaten werden verwendet, um die Wörter einer bestimmten Sprache L über dem Kantenalphabet, das im Beispiel die Form $\{0, 1\}$ hat, zu erkennen. Genau einer der Knoten ist als Startzustand ausgezeichnet und trägt das Label S. Ein Teil der Knoten ist als Menge der Finalzustände ausgezeichnet, Finalzustände haben Label F. Wir können die Kantenlabels als „Tickets“ interpretieren, die zum Passieren der Kante benötigt werden. Ein „Eingabewort“ w über dem Kantenalphabet kann nun als „Ticketfolge“ interpretiert werden. Das Wort w gehört zu der durch den Automaten charakterisierten Sprache L genau dann, wenn es einen Pfad $e_1, \ldots, e_n$ durch den Graphen gibt, der beim Startzustand beginnt und bei einem Finalzustand endet, so daß w genau die zum Passieren von $e_1, \ldots, e_n$ benötigte Ticketfolge ist. Beim obigen Beispielautomaten gehören etwa die Wörter 1, 111, 10, 101 zur charakterisierten Sprache L, alle mit dem Präfix 00 beginnenden

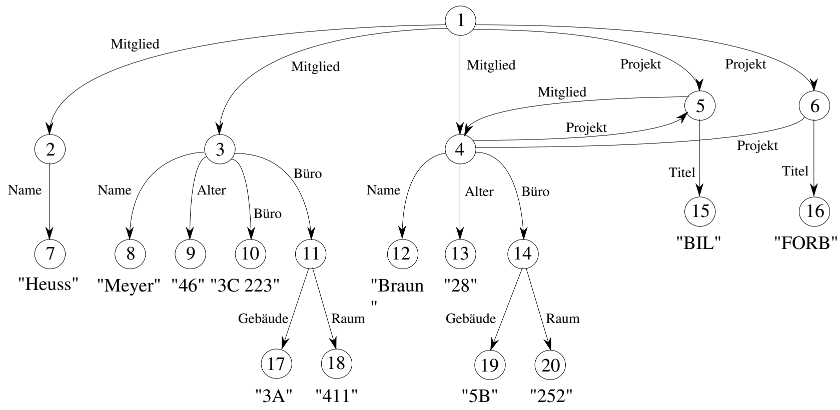


Abbildung 7.1: Datenbank als gelabelter Graph.

Wörter hingegen nicht. Zum algebraischen Hintergrund endlicher Automaten vergleiche man die Diskussion in Beispiel 6.53. Weiter Literaturangaben zu Automaten befinden sich am Kapitelende.

Beispiel 7.26 Gelabelte Graphen werden auch verwendet, um sogenannte semi-strukturierte Daten zu modellieren. Figur 7.1 gibt ein vereinfachtes Beispiel, wie ein Ausschnitt einer semi-strukturierten Datenbank aussehen kann.

7.2 Bäume als Relationalstrukturen

Die in diesem Abschnitt zu besprechenden Bäume stellen eine Teilklasse der DAGs dar, die aufgrund ihrer besonderen Eigenschaften eine zentrale Rolle als Datenstruktur in der Informatik und in der Linguistik spielen. Die unten angefügten Beispiele decken nur einen Teil all derjenigen Bereiche und Situationen ab, wo Bäume in natürlicher Weise auftreten. Zu unterscheiden sind die Teilklassen der ungeordneten und die der geordneten Bäume. Bei der letzteren Klasse gibt es zwischen den direkten Nachfolgern eines Knotens eine feste links-rechts Abfolge. Wir beginnen zunächst mit einer Darstellung der ungeordneten Bäume, wo eine solche Ordnung der direkten Nachfolger fehlt.[1]

Ungeordnete Bäume

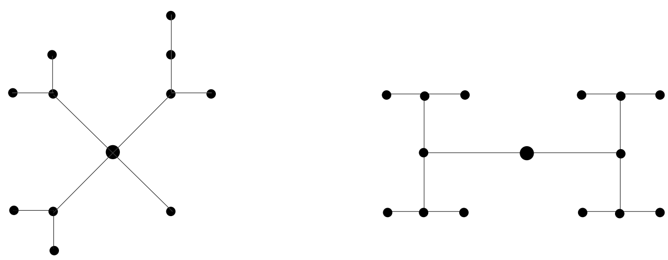
Ungeordnete Bäume mit Wurzeln kann man sowohl als symmetrische Graphen als auch als gerichtete Graphen formalisieren. Wir stellen zunächst die erste Alternative dar.

Definition 7.27 Ein *ungeordneter Baum* ist ein Tupel der Form $\mathcal{T} = \langle N, E, n_0 \rangle$, wo gilt

1. $\langle N, E \rangle$ ist ein zusammenhängender (damit symmetrischer) Graph. Zu je zwei verschiedenen Knoten m und n gibt es genau einen einfachen Pfad von m nach n,

2. $n_0 \in N$ ist ein ausgezeichnetes Element, das die *Wurzel* des Baums $\mathcal{T}$ genannt wird.

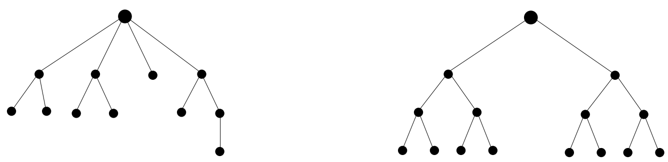
Die nachfolgende Abbildung zeigt zwei ungeordnete Bäume. Die Wurzel ist jeweils hervorgehoben.



Der Übergang zur Formalisierung als *gerichteter* Graph ergibt sich wie folgt. In einem ungeordneten Baum können wir jedem Knoten eine eindeutig bestimmte *Entfernung* zur Wurzel zuordnen. Für die Wurzel selbst ist die Entfernung 0, für jeden anderen Knoten n ist die Entfernung gerade die

[1] Die nachfolgend dargestellten Bäume werden manchmal auch als „verwurzelte Bäume“ (englisch „rooted tree“) bezeichnet. Wir gehen nicht auf Bäume ohne Wurzeln ein.

Länge des eindeutig bestimmten einfachen Pfads von der Wurzel zu n. Damit können wir jeder Kante eine *Vorzugsrichtung* geben, vom Knoten geringerer Entfernung zur Wurzel hin zum Knoten größerer Entfernung. Bei der Darstellung endlicher ungeordneter Bäume können wir die Wurzel als oberstes Element zeichnen und alle Kanten gemäß ihrer Vorzugsrichtung von oben nach unten einzeichnen. Auf diese Weise erhalten wir für die oben dargestellten Bäume nun Darstellungen der folgenden Form.



Wir werden im nachfolgenden abweichend von Definition 7.27 ungeordnete Bäume stets als *gerichtete* Graphen betrachten, wobei wir annehmen, daß jede Kante stets in der Vorzugsrichtung weg von der Wurzel ausgerichtet ist (in Zeichnungen von oben nach unten). Insbesondere nehmen wir bei Begriffen wie „Vorgänger“, „Nachfolger“ und „Pfad“ immer nur auf die Kanten in Vorzugsrichtung Bezug. Das nachfolgende Lemma charakterisiert ungeordnete Bäume als gerichtete Graphen und könnte als eine alternative Definition verwendet werden.

Lemma 7.28 *Es sei $\mathcal{T}$ ein ungeordneter Baum. Dann gilt:*

1. *Die Wurzel besitzt keinen direkten Vorgänger. Jeder von der Wurzel verschiedene Knoten von $\mathcal{T}$ besitzt genau einen direkten Vorgänger,*
2. *$\mathcal{T}$ ist ein DAG.*

Beweis. 1. Da direkte Vorgänger nach Definition der Vorzugsrichtung stets näher an der Wurzel sind, kann die Wurzel keinen direkten Vorgänger haben. Es sei n ein von der Wurzel verschiedener Knoten. Wir zeigen zunächst die *Existenz* eines direkten Vorgängers. Nach Bedingung 1 aus Definition 7.27 gibt es einen einfachen Pfad mit Kanten in Vorzugsrichtung von der Wurzel bis zu n. Ist e die letzte Kante, so gilt $ende(e) = n$, damit ist $start(e)$ ein direkter Vorgänger von n. Um die *Eindeutigkeit* zu verifizieren, nehmen wir an, daß n zwei verschiedene direkte Vorgänger n_1 und

n_2 hat. Der Knoten n_i kann entweder die Wurzel sein (dies kann nur auf einen Knoten zutreffen), oder es gibt wiederum einen einfachen Pfad (mit Kanten in Vorzugsrichtung) von der Wurzel bis zu n_i $(i = 1, 2)$. In beiden Fällen repräsentiert die Verlängerung mit der Kante $\langle n_i, n\rangle$ einen einfachen Pfad von der Wurzel zu n. Damit gibt es offensichtlich zwei verschiedene einfache Pfade von der Wurzel zu n. Dies widerspricht Bedingung 1 von Definition 7.27.

2. Es ist zu zeigen, daß $\mathcal{T}$ keinen geschlossenen Pfad enthält. Wäre $e_1, \ldots, e_k$ ein geschlossener Pfad von $\mathcal{T}$, so gäbe es (einen und damit) unendlich viele Pfade von der Wurzel von $\mathcal{T}$ zu $start(e_1)$, da mit jedem solchem Pfad die Verlängerung mit $e_1, \ldots, e_k$ ein Pfad derselben Art wäre. Da somit $\mathcal{T}$ keinen geschlossenen Pfad enthält, ist $\mathcal{T}$ ein DAG. ■

Lemma 7.29 *Es sei $\mathcal{T}$ ein ungeordneter Baum. Ist $e_1, \ldots, e_k$ ein Pfad von $\mathcal{T}$, so sind die Knoten $start(e_1)$, $start(e_2)$, …, $start(e_k)$, $ende(e_k)$ paarweise verschieden.*

Beweis. Dies folgt aus Teil 2 des vorangegangenen Lemmas mit Lemma 7.10. ■

Definition 7.30 Es sei n ein Knoten des ungeordneten Baums $\mathcal{T}$. Ein direkter Nachfolger von n wird auch *Kind* (oder *Sohn, Tochter*) von n genannt. Der *Verzweigungsgrad* von n ist die Zahl der verschiedenen Kinderknoten von n. Ist n von der Wurzel verschieden, so heißt der nach Lemma 7.28 eindeutig bestimmte direkte Vorgänger von n auch der *Elternknoten* (oder *Vaterknoten, Mutterknoten*) von n. Ein Knoten, der keinen Kinderknoten besitzt, wird *Blatt* von $\mathcal{T}$ genannt. Jeder Pfad von der Wurzel bis zu einem Blatt wird ein *Ast* des Baums genannt.

Beispiel 7.31 Die Menge aller Dateien und Ordner auf einem Computer ist üblicherweise in der Form eines ungeordneten Baums organisiert. Die in einem Ordner enthaltenen Dateien und Teilordner stellen die Kinder des Ordners dar. Dateien und leere Ordner repräsentieren die Blätter dieses Baums.

Definition 7.32 Es seien $\mathcal{T} = \langle N, E, n_0\rangle$ und $\mathcal{T}_1 = \langle N_1, E_1, n_1\rangle$ ungeordnete Bäume. $\mathcal{T}_1$ heißt *Teilbaum* von $\mathcal{T}$ genau dann, wenn gilt:

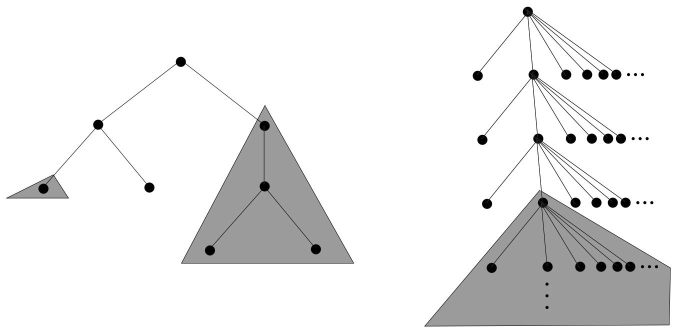


Abbildung 7.2: Drei Teilbäume.

1. $n_1 \in N$,
2. $N_1 = \{n_1\} \cup \{n \in N \mid n$ ist ein Nachfolger von n_1 in $\mathcal{T}\}$,
3. $E_1 = \{e \in E \mid start(e) \in N_1 \wedge ende(e) \in N_1\}$.

In dieser Situation heißt $\mathcal{T}_1$ der *durch* n_1 *bestimmte Teilbaum* von $\mathcal{T}$. Ist n_1 ein Kind der Wurzel n_0, so heißt $\mathcal{T}_1$ ein *unmittelbarer Teilbaum* von $\mathcal{T}$.

Insbesondere ist stets $\mathcal{T}$ selbst ein Teilbaum von $\mathcal{T}$, der durch die Wurzel bestimmt ist. Abbildung 7.2 stellt zwei Bäume dar, in denen Teilbäume unterlegt sind. Wie der rechte Baum illustriert, kann ein Knoten durchaus unendlichen Verzweigungsgrad haben, Bäume können auch unendliche Pfade haben. Dies führt auf die folgenden Unterscheidungen.

Definition 7.33 Ein Baum $\mathcal{T}$ heißt *endlich* genau dann, wenn $\mathcal{T}$ nur endlich viele Knoten besitzt, andernfalls heißt $\mathcal{T}$ *unendlich*. Der Baum $\mathcal{T}$ heißt *endlich verzweigend* genau dann, wenn jeder Knoten von $\mathcal{T}$ endlichen Verzweigungsgrad besitzt. Der Baum $\mathcal{T}$ hat *uniform endlichen Verzweigungsgrad* genau dann, wenn es eine Zahl $k \in \mathbb{N}$ gibt, so daß jeder Knoten von $\mathcal{T}$ höchstens Verzweigungsgrad k besitzt.

$\mathcal{T}$ ist ein *rationaler* Baum genau dann, wenn $\mathcal{T}$ endlich verzweigend ist und bis auf Isomorphie[2] nur endlich viele Teilbäume besitzt.

[2]Beim Isomorphiebegriff von Teilbäumen beziehen wir uns auf den Isomorphiebegriff für Strukturen, vgl. Definition 6.40.

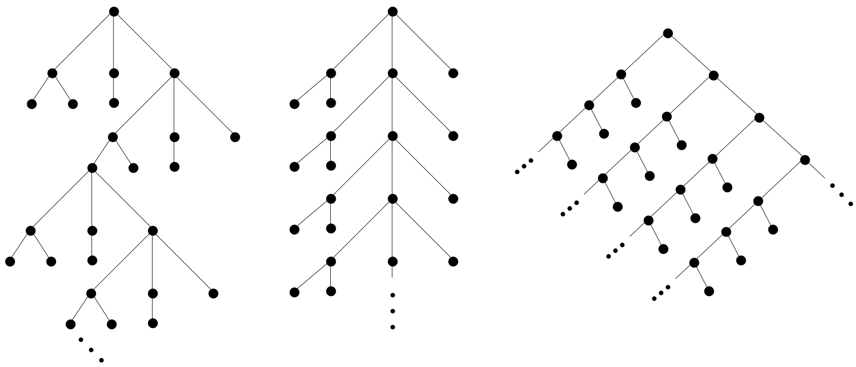


Abbildung 7.3: Unendliche rationale Bäume.

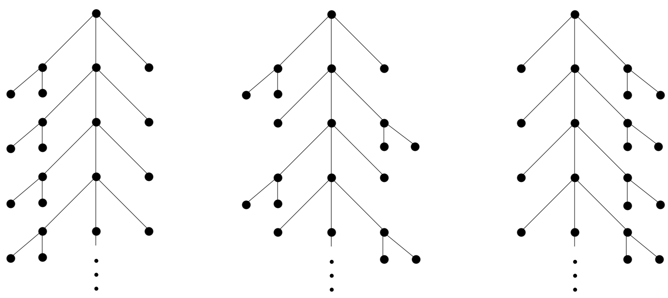
Abbildung 7.3 zeigt drei unendliche rationale Bäume. Ein sehr häufig verwendetes kombinatorisches Prinzip ist das nachfolgende.

Lemma 7.34 (Königs Lemma) *Ein endlich verzweigender Baum $\mathcal{T}$ ist unendlich genau dann, wenn $\mathcal{T}$ einen unendlichen Pfad besitzt.*

Beweis. Falls $\mathcal{T}$ einen unendlichen Pfad besitzt, so ist $\mathcal{T}$ nach Lemma 7.29 unendlich. Es sei umgekehrt $\mathcal{T}$ unendlich. Da die Wurzel n_0 nur endlich viele Kinder hat, muß einer der durch die Kinder bestimmten unmittelbaren Teilbäume unendlich sein. Es sei dies etwa der durch das Kind n_1 bestimmte Teilbaum. Wir setzen $e_1 := \langle n_0, n_1 \rangle$. Durch Iteration dieses Verfahrens ergibt sich ein unendlicher Pfad $e_1, e_2, \ldots$ wo jeweils der durch $n_k := \mathit{ende}(e_k)$ bestimmte Teilbaum von $\mathcal{T}$ unendlich ist. ■

Geordnete Bäume

Bei den bisher betrachteten Bäumen repräsentieren die Kanten die einzige festgelegte Ordnungsbeziehung zwischen den Knoten. Die Abbildungen suggerieren vielleicht fälschlicherweise, daß es noch zusätzlich eine festgelegte links-rechts Abfolge der Knoten gibt. Jedoch ist diese Art der Ordnungsbeziehung lediglich ein ungewollter Nebeneffekt der graphischen Darstellung. Beispielsweise repräsentieren alle drei nachfolgenden Bilder denselben ungeordneten Baum.

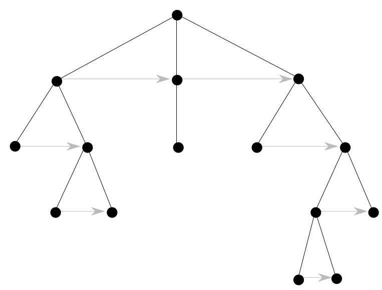


In vielen Zusammenhängen ist es indes durchaus erwünscht, tatsächlich eine Reihenfolge zwischen den verschiedenen Kindern eines Knotens explizit festzulegen. Dies führt auf den Begriff des geordneten Baums.

Definition 7.35 Ein *geordneter Baum* ist ein Tupel der Form $\mathcal{T} = \langle N, E, n_0, < \rangle$, wo gilt

1. $\langle N, E, n_0 \rangle$ ist ein ungeordneter Baum,
2. "$<$" ist eine strikte partielle Ordnung auf N, so daß zwei Knoten n und m in der Beziehung $n < m$ oder $m < n$ stehen genau dann, wenn m und n verschiedene Kinder eines gemeinsamen Elternknotens sind.

Die nachfolgende Abbildung repräsentiert einen geordneten Baum. Die strikte partielle Ordnung „$<$" ist durch graue Pfeile angedeutet.



Bei endlichen geordneten Bäumen ist es möglich[3], den Knoten eine „Standardbenennung" zu geben, so daß die Knotennamen implizit die Struktur

[3]Verallgemeinerung auf bestimmte Arten unendlicher Bäume sind möglich.

des Baums — insbesondere auch die Ordnung „$<$" — widerspiegeln. Dazu führen wir den folgenden Begriff ein.

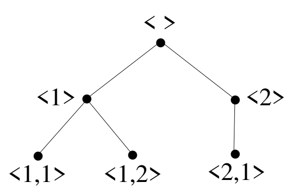
Definition 7.36 Ein *Baumskelett* ist eine endliche Menge $S \neq \emptyset$ endlicher Folgen von positiven natürlichen Zahlen mit den folgenden Eigenschaften:

1. Ist $\langle n_1, \ldots, n_k \rangle \in S$, $k \geq 1$ und $n_k > 1$, so ist auch $\langle n_1, \ldots, n_k - 1 \rangle \in S$.
2. Ist $\langle n_1, \ldots, n_{k-1}, n_k \rangle \in S$ und $k \geq 1$, so ist auch $\langle n_1, \ldots, n_{k-1} \rangle \in S$.

Die *Höhe* von S ist die maximale Länge einer Folge in S.

Es folgt sofort, daß jedes Baumskelett S die leere Folge „$\langle\ \rangle$" enthält, die wir den *Wurzelknoten* von S nennen. Jedem Baumskelett können wir in kanonischer Weise den endlichen geordneten Baum $\mathcal{T}(S) = \langle S, E, \langle\ \rangle, < \rangle$ zuordnen, wo wir von jeder Folge $\langle n_1, \ldots, n_k \rangle$ in S zu jeder Verlängerung der Form $\langle n_1, \ldots, n_k, n_{k+1} \rangle$ in S eine Kante haben, und wo der Wert des letzten Folgenglieds n_{k+1} die links-rechts Anordnung der Kinder eines Knotens $\langle n_1, \ldots, n_{k-1}, n_k \rangle$ bestimmt. Die Elemente von S werden auch die *Positionen* von $\mathcal{T}(S)$ genannt.

Beispiel 7.37 Der dem Baumskelett $S := \{\langle\ \rangle, \langle 1 \rangle, \langle 2 \rangle, \langle 1,1 \rangle, \langle 1,2 \rangle, \langle 2,1 \rangle\}$ zugeordnete Baum $\mathcal{T}(S)$ läßt sich wie folgt repräsentieren:

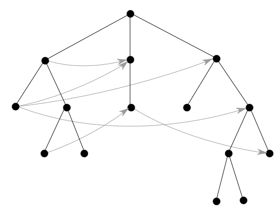


Umgekehrt kann man zu jedem endlichen geordneten Baum $\mathcal{T}$ sofort ein Baumskelett S angeben, so daß $\mathcal{T}$ zu $\mathcal{T}(S)$ isomorph ist.

Auf geordneten Bäumen wird oft eine links-rechts Beziehung der Knoten betrachtet, die nicht nur auf gemeinsame Kinder eines Knotens beschränkt ist. Um sie einzuführen, benötigen wir einen Hilfsbegriff. Wir nennen einen Knoten n' einen *reflexiven Vorgänger* des Knotens n genau dann, wenn entweder $n = n'$ gilt oder wenn n' ein Vorgänger von n ist.

Definition 7.38 Zwei Knoten m und n stehen in der *links-rechts Beziehung* in $\mathcal{T} = \langle N, E, n_0, < \rangle$ genau dann, wenn m und n reflexive Vorgänger m' respektive n' haben, so daß $m' < n'$ gilt.

Wir verwenden das Symbol „$<_{lr}$" für die links-rechts Beziehung. In der nachfolgenden Abbildung sind mit grauen Pfeilen einige Knotenpaare angedeutet, die in der links-rechts Beziehung stehen.



Die folgende Definition führt zwei wichtige *lineare* Ordnungen auf den Knoten eines endlichen geordneten Baums ein. Die Ordnungen sind sehr nützlich, um alle Knoten des Baums in systematischer Weise nacheinander zu besuchen.

Definition 7.39 Es sei $\mathcal{T} = \langle N, E, n_0, < \rangle$ ein endlicher geordneter Baum. Die *Präordnung* auf N ist die strikte lineare Ordnung, die wir erhalten, wenn wir beginnend von der Wurzel jeden Teilbaum von $\mathcal{T}$ so durchlaufen, daß wir

1. als erstes die Wurzel besuchen, dann
2. nacheinander die unmittelbaren Teilbäume in der durch die $<$-Abfolge der Wurzeln bestimmten Reihenfolge. Hierbei wird jeder unmittelbare Teilbaum vollständig durchlaufen, bevor wir den nächsten besuchen.

Bei der *Postordnung* werden zuerst die Knoten der unmittelbaren Teilbäume in der durch die links-rechts Abfolge der Wurzeln bestimmten Reihenfolge durchlaufen, bevor als letztes die Wurzel aufgezählt wird.

Abbildung 7.4 illustriert an einem Beispiel die Präordnungsabfolge (links) und die Postordnungsabfolge (rechts) der Knoten eines geordneten Baums.

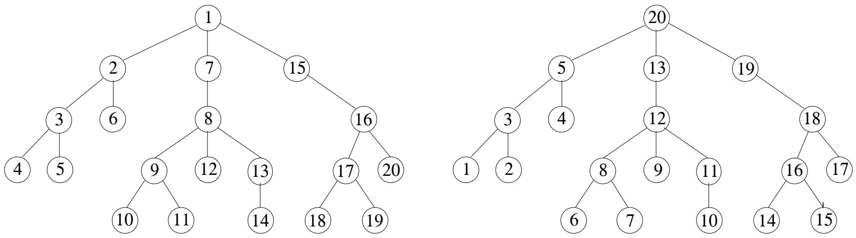


Abbildung 7.4: Präordnung und Postordnung.

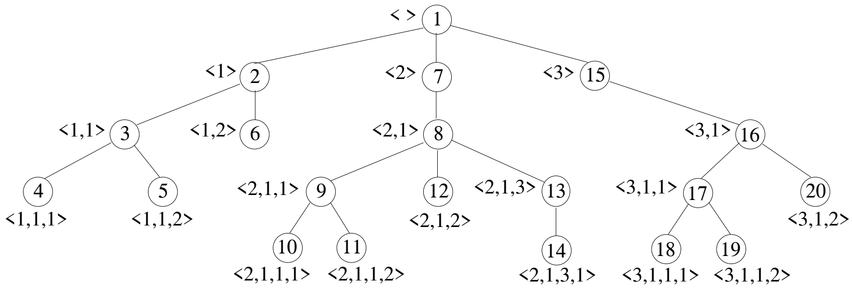


Abbildung 7.5: Präordnung als lexikographische Ordnung des Baumskeletts.

Bemerkung 7.40 Bilden die Knoten des Baums $\mathcal{T}$ ein Baumskelett S, so stimmt die Präordnung genau mit der lexikographischen Ordnung (vgl. Beispiel 4.28 Nr. 5) der Folgen in S überein. Der Zusammenhang ist in Abbildung 7.5 dargestellt.

Lemma 7.41 *Es bezeichne $<_p$ die Präordnung auf den Knoten des endlichen geordneten Baums $\mathcal{T} = \langle N, E, n_0, < \rangle$, mit $<_{lr}$ sei die links-rechts Ordnung der Knoten bezeichnet. Es seien m und n Knoten aus N.*

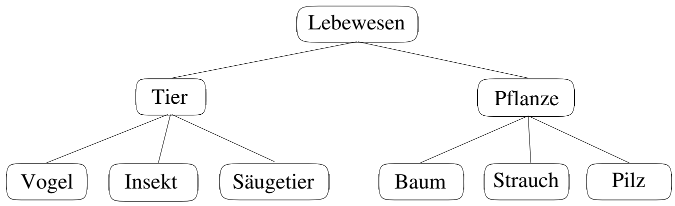
1. *Es gilt $m <_{lr} n$ genau dann, wenn $m <_p n$ und wenn n kein Nachfolger von m ist.*

2. *Falls $m <_{lr} n$ und $n \leq_p k$ so folgt $m <_{lr} k$.*

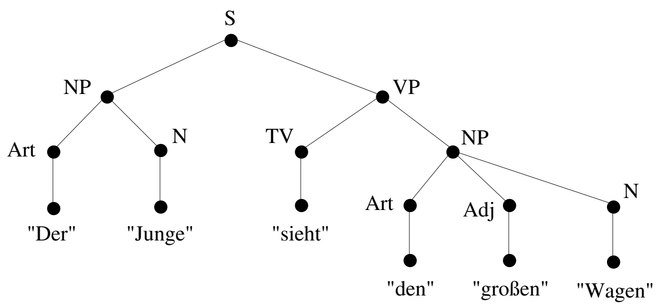
Wie bei allgemeinen Graphen werden auch bei ungeordneten oder geordneten Bäumen häufig Kanten oder Knoten gelabelt.

Gelabelte Bäume

Beispiel 7.42 Begriffshierarchien können adäquat in der Form endlicher ungeordneter Bäume mit gelabelten Knoten dargestellt werden, wie im nachfolgenden Beispiel.



Beispiele 7.43 Die Struktur natürlich-sprachlicher Sätze wird häufig in Gestalt von *Parsebäumen* angegeben. Es handelt sich hierbei um geordnete Bäume mit gelabelten Knoten, die die hierarchische Struktur des Satzes in Teilphrasen wiedergeben. Im nachfolgenden Beispiel steht „S“ für Satz, „NP“ für Nominalphrase, „VP“ für Verbalphrase, „Art“ für Artikel, „N“ für Nomen, „TV“ für transitives Verb und „Adj“ für Adjektiv.



7.3 Algebraische Beschreibung von Bäumen

In diesem Abschnitt betrachten wir endliche geordnete Bäume, deren Knoten mit Funktionssymbolen einer gegebenen Signatur gelabelt sind. Die Stelligkeit des Labels gibt hierbei immer die Zahl der Kinder eines Knotens an. Bäume dieser Art werden zur Repräsentation von Termen verwendet. Auf diesen Sachverhalt gehen wir allerdings erst im nächsten Abschnitt ein.

Nachfolgend sei $\Sigma = \Sigma_\mathcal{F} \cup \Sigma_\mathcal{E}$ stets eine Signatur mit *nichtleerer* Menge $\Sigma_\mathcal{E}$ von Individuenkonstanten und mit der Menge der Funktionssymbole $\Sigma_\mathcal{F}$.

Definition 7.44 Ein *Baum über der Signatur* Σ ist ein Paar $T = \langle S, D \rangle$ wo gilt:

1. S ist ein Baumskelett,

2. D ist eine „Dekoration" für S, das heißt eine Abbildung $S \to \Sigma$ mit den folgenden Eigenschaften:

   (a) Für jeden Knoten $\eta \in S$, der genau $k \geq 1$ Kinder hat, ist $D(\eta)$ ein k-stelliges Funktionssymbol aus $\Sigma_\mathcal{F}$,

   (b) für jedes Blatt $\eta \in S$ ist $D(\eta) \in \Sigma_\mathcal{E}$.

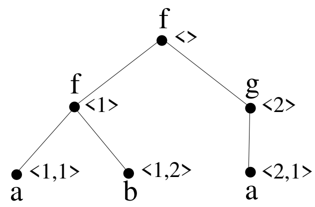
Mit $B(\Sigma)$ bezeichnen wir die Menge aller Bäume über der Signatur Σ.

Bäume der obigen Art werden wir kurz auch als Σ-*Bäume* bezeichnen. Die Dekoration $D(\eta)$ eines Knotens η wird auch als das *Label* von η bezeichnet, anstatt von einer Dekoration spricht man manchmal auch von einem „Labeling". Begriffe wie die *Höhe* oder die Menge der *Positionen* übertragen sich in der offenkundigen Weise vom Baumskelett S auf den Baum $T = \langle S, D \rangle$.

Beispiel 7.45 Die Signatur Σ enthalte ein zweistelliges Funktionssymbol f, ein einstelliges Funktionssymbol g und die Individuenkonstante a und b. Es sei $S := \{\langle\ \rangle, \langle 1 \rangle, \langle 2 \rangle, \langle 1,1 \rangle, \langle 1,2 \rangle, \langle 2,1 \rangle\}$ das Baumskelett aus Beispiel 7.37 und D die Dekoration

$$\begin{bmatrix} \langle\ \rangle & \mapsto & f \\ \langle 1 \rangle & \mapsto & f \\ \langle 2 \rangle & \mapsto & g \\ \langle 1,1 \rangle & \mapsto & a \\ \langle 1,2 \rangle & \mapsto & b \\ \langle 2,1 \rangle & \mapsto & a \end{bmatrix}$$

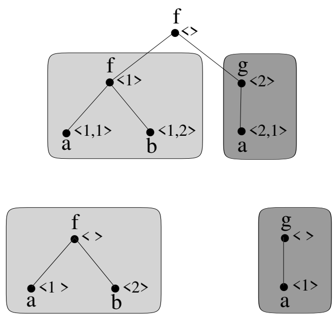
Dann ist $T := \langle S, D \rangle$ ein Baum in $B(\Sigma)$. Hier eine graphische Repräsentation.



In Definition 7.32 hatten wir den Begriff des (unmittelbaren) Teilbaums eines ungeordneten Baums eingeführt. Bei den hier betrachteten geordneten Bäumen kann man den Begriff des Baumskeletts dazu verwenden, um genauer Teilbäume einer festen Position zu definieren. Hierzu bezeichne „$\circ$“ die Konkatenation (Verkettung) endlicher Folgen natürlicher Zahlen, wo also $\langle n_0, \ldots, n_k \rangle \circ \langle m_0, \ldots, m_l \rangle := \langle n_0, \ldots, n_k, m_0, \ldots, m_l \rangle$ gilt.

Definition 7.46 Es sei $T = \langle S, D \rangle \in B(\Sigma)$. Der Baum $\langle S_1, D_1 \rangle \in B(\Sigma)$ heißt *Teilbaum* von T genau dann, wenn eine Position $\eta \in S$ existiert, so daß $S_1 = \{\eta' \mid \eta \circ \eta' \in S\}$ und $D_1(\eta') = D(\eta \circ \eta')$ für alle $\eta' \in S_1$. Es wird dann η die *Position* des Teilbaums $\langle S_1, D_1 \rangle$ in $\langle S, D \rangle$ genannt. Hat $\eta = \langle i \rangle$ die Länge 1, so heißt $\langle S_1, D_1 \rangle$ der *unmittelbare Teilbaum* von T der Position $\langle i \rangle$.

In der nachfolgenden Abbildung sind die zwei unmittelbaren Teilbäume des oberen Baums hervorgehoben. Sie haben die Positionen $\eta_1 := \langle 1 \rangle$ und $\eta_2 := \langle 2 \rangle$. Um eine korrekte Formalisierung dieser Bäume zu erreichen, müssen die Baumskelette wie unten angedeutet die leere Folge als Wurzel haben.



Wir wollen nun auf der Menge $B(\Sigma)$ der Σ-Bäume eine einfache algebraische Struktur einführen. Zur Vorbereitung dient das folgende

Lemma 7.47 *Es sei $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol und $T_1, \ldots, T_n$ seien Σ-Bäume. Dann gibt es einen eindeutig bestimmten Baum $T \in B(\Sigma)$, dessen Wurzelknoten mit f dekoriert ist, und wo $T_1, \ldots, T_n$ die unmittelbaren Teilbäume der Positionen $\langle 1 \rangle, \ldots, \langle n \rangle$ sind.*

Beweis. Wir beweisen zunächst die Existenz. Es habe T_i die Form $\langle S_i, D_i \rangle$, für $i = 1, \ldots, n$. Wir setzen

$$S := \{\langle\ \rangle\} \cup \bigcup_{i=1}^{n} \{\langle i \rangle \circ \eta \mid \eta \in S_i\}.$$

Offenkundig ist S ein Baumskelett und die disjunkte Vereinigung der angeführten $n+1$ Positionsmengen. Setzen wir nun $D(\langle\ \rangle) := f$ und $D(\langle i \rangle \circ \eta) := D_i(\eta)$ für alle $\eta \in S_i$ und $i = 1, \ldots n$, so erhalten wir eine Dekoration von S über Σ. Es ist einfach nachzurechnen, daß $T := \langle S, D \rangle$ die oben genannten Bedingungen erfüllt.

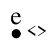
Der Nachweis der Eindeutigkeit bleibt dem Leser überlassen. Man zeigt zunächst, daß das Baumskelett eindeutig bestimmt ist, danach die Eindeutigkeit der Dekoration. ■

Offenkundig gilt auch der folgende einfache Sachverhalt.

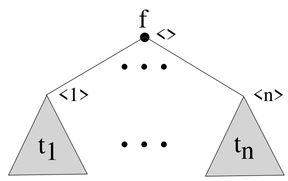
Lemma 7.48 *Jeder Baum $T \in B(\Sigma)$ hat eine eindeutig bestimmte Wurzeldekoration. Diese ist entweder eine Individuenkonstante, wobei T dann Höhe 0 hat, oder sie ist ein Funktionssymbol $f \in \Sigma$ der Stelligkeit $n \geq 1$. In letzterem Fall hat T eine Höhe $k > 0$, die unmittelbaren Teilbäume der Positionen $\langle 1 \rangle, \ldots, \langle n \rangle$ von T sind eindeutig bestimmt und haben kleinere Höhe.*

Definition 7.49 Auf der Menge $B(\Sigma)$ führen wir nun die folgende Interpretation I der Zeichen aus Σ ein:

1. für $e \in \Sigma_{\mathcal{E}}$ sei $I(e)$ der Baum, der nur aus dem mit e dekorierten Wurzelknoten $\langle\ \rangle$ besteht.



2. für $f \in \Sigma_{\mathcal{F}}$ (wo f n-stellig ist) sei $I(f)$ die n-stellige Funktion auf $B(\Sigma)$, die den Bäumen $T_1, \ldots T_n$ den nach Lemma 7.47 eindeutig bestimmten Baum $T \in \mathcal{T}(\Sigma)$ zuordnet, dessen Wurzelknoten mit f dekoriert ist, und wo $T_1, \ldots, T_n$ die unmittelbaren Teilbäume mit Positionen $\langle 1 \rangle, \ldots, \langle n \rangle$ sind.



Dann heißt $\mathcal{B}(\Sigma) := \langle B(\Sigma), I \rangle$ die *Baumalgebra zur Signatur* Σ.

Es ist nicht schwer zu sehen, daß die Baumalgebra $\mathcal{B}(\Sigma)$ bereits von der leeren Teilmenge erzeugt wird (Aufgabe 7.20). Sie hat darüberhinaus folgende bemerkenswerte Eigenschaft.

Theorem 7.50 (Initialität der Baumalgebra) *Ist $\mathcal{A}$ eine beliebige Σ-Algebra, so gibt es stets einen eindeutig bestimmten Homomorphismus von $\mathcal{B}(\Sigma)$ nach $\mathcal{A}$.*

Beweis. Wir zeigen zunächst die *Existenz* eines Homomorphisms. Hierzu definieren wir ein Abbildung $h : \mathcal{B}(\Sigma) \to A$ induktiv über die Höhe des Baumskeletts.

1. Es sei $T \in B(\Sigma)$ ein Baum der Höhe 0. Dann ist T nach Lemma 7.48 ein Blatt, das mit einer Individuenkonstante $e \in \Sigma_{\mathcal{E}}$ gelabelt ist, und es gilt $T = I(e)$, wobei I wie oben die Interpretationsfunktion der Baumalgebra bezeichnet. Wir setzen $h(T) := e_{\mathcal{A}}$ (zur Notation vgl. Bemerkung 6.24).

2. Es sei die Abbildung h bereits für alle Bäume der Höhe $k \geq 0$ definiert. Weiter sei $T \in B(\Sigma)$ ein Baum der Höhe $k+1$. Es sei f das nach Lemma 7.48 eindeutig bestimmte Label des Wurzelknotens von T, weiter seien $T_1, \ldots, T_n$ die eindeutig bestimmten unmittelbaren Teilbäume von T. Wir setzen $h(T) := f_{\mathcal{A}}(h(T_1), \ldots, h(T_n))$, wobei die Argumente $h(T_i)$ von $f_{\mathcal{A}}$ nach Voraussetzung bereits definiert sind.

Durch 1 und 2 wird somit eine Abbildung $h : B(\Sigma) \rightarrow A$ erklärt. Offenkundig ist h aufgrund von Lemma 7.48 nach Konstruktion ein Homomorphismus.

Es bleibt noch die Eindeutigkeit des Homomorphismus nachzuweisen. Es sei h_1 ein Homomorphismus von $\mathcal{B}(\Sigma)$ nach $\mathcal{A}$. Nach Voraussetzung haben h und h_1 denselben Definitionsbereich. Es bleibt nach Lemma 3.39 zu zeigen, daß beide Abbildungen denselben Bäumen T dieselben Elemente aus A zuordnen. Dies zeigen wir per Induktion über die Baumhöhe von T, wobei wir verwenden, daß h und h_1 Homomorphismen sind:

1. Es sei T ein Baum der Höhe 0. Dann ist T nach Lemma 7.48 ein Blatt, das mit einer Individuenkonstante $e \in \Sigma$ gelabelt ist. Es folgt $h_1(T) = h_1(I(e)) = e_{\mathcal{A}} = h(T)$.
2. Es sei T ein Baum der Höhe $k+1 > 0$. Es sei f das nach Lemma 7.48 eindeutig bestimmte Label des Wurzelknotens von T, und $T_1, \ldots, T_n$ seien die eindeutig bestimmten unmittelbaren Teilbäume von T. Nach Induktionsvoraussetzung gilt $h_1(T_i) = h(T_i)$ für $i = 1, \ldots, n$. Damit folgt $h_1(T) = f_{\mathcal{A}}(h_1(T_1), \ldots, h_1(T_n)) = f_{\mathcal{A}}(h(T_1), \ldots, h(T_n)) = h(T)$.

Dies beendet den Beweis. ■

7.4 Die Termalgebra

Die in diesem Abschnitt zu besprechenden Terme stellen syntaktische Beschreibungen von Elementen einer Struktur dar. Terme sind ein wichtiger Bestandteil der Syntax der Prädikatenlogik (vgl. Abschnitt 12.1), sie treten beim Programmieren an vielen Stellen auf, sie bilden darüberhinaus die Grundlage von logischen Programmiersprachen wie Prolog. Terme sind ähnlich zu Bäumen, stellen allerdings flache Zeichenreihen dar, wo die Baumstruktur lediglich implizit, durch Klammern, festgelegt ist. Wir werden sehen, daß sich Terme und Bäume zu gegebener Signatur wechselseitig ineinander übersetzen lassen.

Nachfolgend sei $\Sigma = \Sigma_{\mathcal{F}} \cup \Sigma_{\mathcal{E}}$ wie im vorigen Abschnitt eine Signatur mit nichtleerer Menge $\Sigma_{\mathcal{E}}$ von Individuenkonstanten und mit der Menge der Funktionssymbole $\Sigma_{\mathcal{F}}$. Zusätzlich betrachten wir eine abzählbar unendliche Menge X, die disjunkt zu Σ ist. Die Elemente der Menge X nennen wir *Variablen.*

Definition 7.51 Die Menge der *Terme über* Σ *und* X ist die kleinste Menge von Symbolfolgen, die unter den folgenden Bildungsregeln abgeschlossen ist.

1. Jede Variable $x \in X$ und jede Individienkonstante $e \in \Sigma_{\mathcal{E}}$ ist ein Term über Σ und X.
2. Sind $t_1, \ldots, t_n$ Terme über Σ und X, und ist $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol, so ist $f(t_1, \ldots, t_n)$ ein Term über Σ und X.

Terme in $X \cup \Sigma_{\mathcal{E}}$ werden *atomare Terme* genannt.

Wir schreiben $T(\Sigma, X)$ für die Menge aller Terme über Σ und X.

Satz 7.52 (Eindeutige Lesbarkeit von Termen) *Es sei* $t \in T(\Sigma, X)$ *ein Term über* Σ *und* X. *Dann ist entweder* t *eine Variable, oder eine Individuenkonstante, oder* t *läßt sich in eindeutiger Weise in der Form* $f(t_1, \ldots, t_n)$ *darstellen, wo* $f \in \Sigma_{\mathcal{F}}$ *und die* t_i *Terme sind* $(1 \leq i \leq n)$.

Wir skizzieren den Beweis lediglich: hierzu versehen wir die Klammern eines nichtatomaren Term mit einem Label, das die Klammerungstiefe angibt. Bei jeder öffnenden Klammer zählen wir um eins nach oben, bei jeder schließenden Klammer zählen wir um eins herunter. Für den Term $f(f(g(g(a)), a), g(f(g(b), f(a, b))))$ ergibt sich etwa

$$f({}_1 f({}_2 g({}_3 g({}_4 a)_3)_2, a)_1, g({}_2 f({}_3 g({}_4 b)_3, f({}_4 a, b)_3)_2)_1)_0.$$

Es ist einfach, zu zeigen, daß es in jedem nichtatomaren Term jeweils genau eine öffnende Klammer mit dem Label 1 unmittelbar nach dem Kopffunktionssymbol aus $\Sigma_{\mathcal{F}}$ und genau eine schließende Klammer mit dem Label 0 am Termende gibt. Um die Folge der Teilterme $t_1, \ldots, t_n$ eines Term der Form $f(t_1, \ldots, t_n)$ zu bestimmen, kann man sich im wesentlichen darauf beschränken, Paare von gelabelten Klammern der Form $({}_2 \ldots)_1$ zu ermitteln. Zwischen den beiden Klammern dürfen hierbei weitere auftreten, aber keine Klammer der Form „$)_1$“. Die Klammerpaare führen auf die nichtatomaren Teilterme t_i, die atomaren Terme der Folge $t_1, \ldots, t_n$ sind hieraus sofort zu bestimmen. Die nichtatomaren Teilterme von Teiltermen ergeben sich entsprechend durch Klammerpaare der Form $({}_3 \ldots)_2$ etc. Es ist nicht schwer zu zeigen, daß die derart bestimmte Teiltermstruktur die einzig mögliche Zerlegung in Teilterme angibt. ■

Definition 7.53 Die Menge $TT(t)$ der *Teilterme* des Terms t ist induktiv wie folgt erklärt:

1. Ist t eine Variable $x \in X$ oder eine Individuenkonstante $e \in \Sigma_{\mathcal{E}}$, so ist $TT(t) := \{t\}$.
2. Ist t ein Term der Gestalt $f(t_1, \ldots, t_n)$ wo $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol ist und $t_1, \ldots, t_n$ Terme sind, so ist $TT(t) := \{t\} \cup \bigcup_{i=1}^{n} TT(t_i)$.

In Fall 2 heißen die Terme $t_1, \ldots, t_n$ heißen die *unmittelbaren Teilterme* des Terms t.

Es folgt aus Satz 7.52, daß die Menge der Teilterme von t eindeutig bestimmt ist.

Definition 7.54 Die *Tiefe* $w(t)$ eines Terms t ist induktiv wie folgt erklärt:

1. Ist t eine Variable aus X oder eine Individuenkonstante aus $\Sigma_{\mathcal{E}}$, so gilt $w(t) := 0$.
2. Hat t die Form $f(t_1, \ldots, t_n)$, so ist $w(t)$ das Maximum der Tiefen der unmittelbaren Teilterme $t_1, \ldots, t_n$ plus 1.

Es folgt wiederum aus Satz 7.52, daß die Tiefe eines Terms eindeutig bestimmt ist. Das nachfolgende Beispiel illustriert die nunmehr eingeführten Begriffe.

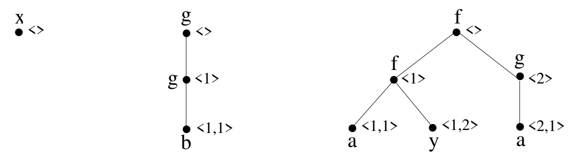
Beispiel 7.55 Die Signatur Σ enthalte ein zweistelliges Funktionssymbol f, ein einstelliges Funktionssymbol g und die Individuenkonstante a und b. Weiter seien x und y Variablen aus X. Dann sind x, a, $g(g(b))$ und $f(f(a,y), g(a))$ Terme über Σ und X. Es gilt $w(x) = w(a) = 0$ und $w(f(f(a,y), g(a))) = 3$. Die unmittelbaren Teilterme von $f(f(a,y), g(a))$ sind $f(a,y)$ und $g(a)$. Weiter gilt $TT(f(f(a,y), g(a))) = \{f(f(a,y), g(a)), f(a,y), g(a), a, y\}$.

Nachfolgend wird von der Baumalgebra $\mathcal{B}(\Sigma \cup X)$ die Rede sein. Hierbei betrachten wir die Menge $\Sigma \cup X$ stets als eine erweiterte Signatur, wo die

Elemente von X die Rolle zusätzlicher Individuenkonstante spielen. Aus diesem Blickwinkel kann man Terme als Bäume über einer erweiterten Signatur betrachten, die in flacher, linearisierter Weise dargestellt sind. Formaler wird dieser Zusammenhang durch die folgende rekursive Definition der Abbildung $\mathit{fl} : B(\Sigma \cup X) \to T(\Sigma, X)$ beschrieben:

1. Ist T ein Baum der Höhe 0 mit Label $x \in X$ (resp. $e \in \Sigma_{\mathcal{E}}$), so ist $\mathit{fl}(T) := x$ (resp. $\mathit{fl}(T) := e$),
2. ist T ein Baum der Höhe $k > 0$ mit Wurzeldekoration f und mit den unmittelbaren Teilbäumen $T_1, \dots, T_n$, so ist $\mathit{fl}(T) := f(\mathit{fl}(T_1), \dots, \mathit{fl}(T_n))$.

Beispiele 7.56 Als Bilder der drei Bäume



unter „fl“ erhält man die Terme x, $g(g(b))$ und $f(f(a, y), g(a))$.

Umgekehrt bildet die nachfolgende Abbildung $\mathit{tree} : T(\Sigma, X) \to B(\Sigma, X)$ jeden Term $t \in T(\Sigma, X)$ auf einen Baum ab:

1. Hat t die Tiefe 0, so sei $\mathit{tree}(t)$ der eindeutig bestimmte Baum der Höhe 0, wo der Wurzelknoten $\langle\,\rangle$ mit dem Label t dekoriert ist.
2. Hat t die Form $f(t_1, \dots, t_n)$, wo $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol ist $(n \geq 1)$, so sei $\mathit{tree}(t)$ der eindeutig bestimmte Baum der Höhe $w(t)$, wo der Wurzelknoten $\langle\,\rangle$ mit dem Label f dekoriert ist, und wo $\mathit{tree}(t_1), \dots, \mathit{tree}(t_n)$ die unmittelbaren Teilbäume sind.

Die zentralen Eigenschaften der Abbildungen fl und tree sind die folgenden.

Lemma 7.57 *Die Abbildung „fl“ ist eine Bijektion von $B(\Sigma \cup X)$ auf $T(\Sigma, X)$. Umgekehrt ist die Abbildung „tree“ ist eine Bijektion von $T(\Sigma \cup X)$*

auf $B(\Sigma, X)$. *Es ist* $\mathit{fl} \circ \mathit{tree}$ *die Identitätsfunktion auf* $B(\Sigma, X)$ *und* $\mathit{tree} \circ \mathit{fl}$ *die Identitätsfunktion auf* $T(\Sigma, X)$.

Beweis. Nach Aufgabe 3.26 reicht es zu zeigen, daß $\mathit{fl} \circ \mathit{tree}$ die Identitätsfunktion auf $B(\Sigma, X)$ und $\mathit{tree} \circ \mathit{fl}$ die Identitätsfunktion auf $T(\Sigma, X)$ ist. Wir zeigen die letztere Eigenschaft, der Beweis der ersten Eigenschaft ist ähnlich. Es sei also $t \in T(\Sigma, X)$. Wir zeigen $t = \mathit{fl}(\mathit{tree}(t))$ mittels vollständiger Induktion über die Tiefe $w(t)$.

Induktionsanfang. Falls $w(t) = 0$ gilt, so ist $\mathit{tree}(t)$ der Baum, der nur aus einem mit t gelabelten Wurzelknoten besteht. Es folgt sofort $t = \mathit{fl}(\mathit{tree}(t))$.

Induktionsannahme. Es sei $k \geq 0$. Für alle Terme s mit $w(s) \leq k$ gelte $t = \mathit{fl}(\mathit{tree}(t))$.

Induktionsschritt. Es sei nun t ein Term mit Tiefe $w(t) = k + 1 > 0$. Dann kann t keine Variable oder Individuenkonstante sein. Demnach hat t die Form $f(t_1, \ldots, t_n)$ wo die unmittelbaren Teilterme t_i Tiefe kleinergleich k haben. Nach Induktionsannahme gilt also $t_i = \mathit{fl}(\mathit{tree}(t_i))$ für $1 \leq i \leq n$. Es ist $T := \mathit{tree}(t)$ der eindeutig bestimmte Baum mit Wurzeldekoration f, wo $\mathit{tree}(t_1), \ldots, \mathit{tree}(t_n)$ die unmittelbaren Teilbäume sind. Folglich hat $\mathit{fl}(T) = \mathit{fl}(\mathit{tree}(t))$ die Form $f(\mathit{fl}(\mathit{tree}(t_1)), \ldots, \mathit{fl}(\mathit{tree}(t_n)))$. Nach Induktionsannahme erhalten wir als Wert den Term $f(t_1, \ldots, t_n)$. ■

Wie vorher im Fall der Bäume können wir nun auf der Menge der Terme über Σ und X eine algebraische Struktur einführen.

Definition 7.58 Die *Termalgebra* zur Signatur Σ über X ist die Algebra $\mathcal{T}(\Sigma, X) := \langle T(\Sigma, X), J \rangle$ wo J die folgende Interpretationsfunktion ist:

1. für $e \in \Sigma_{\mathcal{E}}$ ist $J(e) := e$,
2. ist $f \in \Sigma_F$ ein n-stelliges Funktionssymbol, so ist $J(f)$ diejenige n-stellige Funktion auf $T(\Sigma, X)$, die den Termen $t_1, \ldots, t_n$ den Term $f(t_1, \ldots, t_n)$ zuordnet.

Das folgende Lemma ergibt sich unmittelbar aus Lemma 7.57 und aus der obigen Wahl der Interpretationsfunktion J.

Lemma 7.59 *Die Abbildung „fl" ist ein Isomorphismus zwischen der Baumalgebra $\mathcal{B}(\Sigma \cup X)$ und der Termalgebra $\mathcal{T}(\Sigma, X)$.*

Hieraus ergibt sich die folgende wichtige Eigenschaft der Termalgebra.

Theorem 7.60 (Absolute Freiheitseigenschaft der Termalgebra) *Ist $\mathcal{A}$ eine beliebige Σ-Algebra, so besitzt jede Abbildung von X nach A eine eindeutig bestimmte Erweiterung zu einem Homomorphismus von $\mathcal{T}(\Sigma, X)$ nach $\mathcal{A}$.*

Beweis. Ist eine Abbildung h_0 von X nach $\mathcal{A}$ gegeben, so können wir $\mathcal{A}$ dadurch zu einer Algebra über der erweiterten Signatur $\Sigma \cup X$ machen, daß wir jede Variable $x \in X$ gerade durch $h_0(x)$ interpretieren und ansonsten die Interpretation der Symbole aus Σ unverändert lassen. Nun sagt Theorem 7.50, hier angewendet auf die erweiterte Signatur $\Sigma \cup X$, daß es einen eindeutig bestimmten Homomorphismus f der Baumalgebra $\mathcal{B}(\Sigma \cup X)$ in die $(\Sigma \cup X)$-Algebra $\mathcal{A}$ gibt. Nach Aufgabe 6.18 ist die Komposition $h := \mathit{tree} \circ f$ ein Homomorphismus von $\mathcal{T}(\Sigma, X)$ nach $\mathcal{A}$. Aufgrund der gewählten Interpretation der Symbole aus X in $\mathcal{A}$ gilt, daß dieser Homomorphismus die Abbildung h_0 erweitert.

Es bleibt die Eindeutigkeit zu verifizieren. Ist h' irgendein Homomorphismus von $\mathcal{T}(\Sigma, X)$ nach $\mathcal{A}$, der h_0 erweitert, so ist gemäß Aufgabe 6.18 die Abbildung $\mathit{fl} \circ h'$ ein Homomorphismus zwischen den $(\Sigma \cup X)$-Algebren $\mathcal{B}(\Sigma \cup X)$ und $\mathcal{A}$. Wegen der Eindeutigkeitsaussage aus Theorem 7.50 folgt $\mathit{fl} \circ h' = f$ und mit Lemma 7.57 damit $h' = \mathit{tree} \circ \mathit{fl} \circ h' = \mathit{tree} \circ f = h$. ∎

Man sagt aufgrund der in Theorem 7.60 dargestellten Eigenschaft, daß $\mathcal{T}(\Sigma, X)$ die *absolut freie Algebra der Signatur Σ über der Menge X* ist. Dadurch, so wie durch die Tatsache, daß $\mathcal{T}(\Sigma, X)$ als Σ-Algebra von der Menge der Variablen X erzeugt wird, ist $\mathcal{T}(\Sigma \cup X)$ bis auf Isomorphie gekennzeichnet.

Definition 7.61 Ein *Grundterm* über der Signatur Σ ist ein Term über Σ und X, der kein Vorkommen einer Variablen enthält. Grundterme werden auch als *geschlossene Terme* bezeichnet, Terme mit Variablen als *offene Terme*.

Beispiel 7.62 Das Alphabet Σ enthalte die Funktionssymbole f (zweistellig), g (einstellig) und die Individuenkonstante c. Dann ist $f(f(x, y), g(g(c)))$

ein offener Term, und $g(f(c, g(g(c))))$ ist ein Grundterm.

Die Menge der Grundterme über der Signatur Σ bezeichnen wir mit $T(\Sigma)$. Offenkundig ist $T(\Sigma)$ die Grundmenge einer Teilalgebra $\mathcal{T}(\Sigma)$ von $\mathcal{T}(\Sigma, X)$. Diese Algebra wird als die *Grundtermalgebra* zur Signatur Σ bezeichnet.

Lemma 7.63 *Die Einschränkung der Abbildung „fl“ auf die Menge $\mathcal{T}(\Sigma)$ ist ein Isomorphismus zwischen der Baumalgebra $\mathcal{B}(\Sigma)$ und der Grundtermalgebra $\mathcal{B}(\Sigma)$.*

Aus Theorem 7.50 ergibt sich nun

Theorem 7.64 (Initialität der Grundtermalgebra) *Ist $\mathcal{A}$ eine beliebige Σ-Algebra, so gibt es stets einen eindeutig bestimmten Homomorphismus von $\mathcal{T}(\Sigma)$ nach $\mathcal{A}$.*

7.5 Ergänzungen

In diesem Abschnitt gehen wir auf zwei spezifische Themenbereiche ein, die unmittelbar mit den Darstellungen dieses Kapitels verbunden sind. Die Ausführungen zur Unifikation von Termen im ersten Teil werden wir später aufgreifen. Die Diskussion Eulerscher Kreise im zweiten Abschnitt stellt einen Seitenausflug dar.

7.5.1 Unifikation von Termen

Im letzten Abschnitt hatten wir erwähnt, daß Terme die Grundlage für logische Programmiersprachen und andere Systeme im Bereich der Informatik darstellen. Bei vielen Anwendungen wird ein gegebener Term t hierbei in einem weiteren Sinn verwendet, nämlich um in kompakter Form all diejenigen Terme zu repräsentieren, die sich aus t durch Ersetzung der Variablen durch andere Terme ergeben. Diese Terme werden *Instanzen* von t genannt. Zum Beispiel sind die Terme $f(a, f(a, a))$ und $f(f(a, z), f(a, a))$ Instanzen des Terms $f(x, f(y, y)$. Ein grundlegendes algorithmisches Problem, das als zentraler Rechenschritt bei vielen Anwendungen auftritt, ergibt sich nun daraus, zu zwei gegebenen Termen t_1 und t_2 eine Termrepräsentation all derjenigen Terme zu finden, die *sowohl* von t_1 *als auch* von t_2 repräsentiert werden. Dieses Problem wird durch die sogenannte *Unifikation* von Termen gelöst, die wir an dieser Stelle formal beschreiben wollen. Nachfolgend sei $\mathcal{T}(\Sigma, X)$ die Termalgebra zur Signatur Σ.

Definition 7.65 Eine *Substitution* ist ein Endomorphismus σ der Termalgebra $\mathcal{T}(\Sigma, X)$ in sich derart, daß die Menge $\{x \in X \mid \sigma(x) \neq x\}$ endlich ist.

Die Menge $\{x \in X \mid \sigma(x) \neq x\}$ wird mit $dom(\sigma)$ bezeichnet. Wegen Theorem 7.60 läßt sich jede Substitution σ in eindeutiger Weise durch Angabe der Bilder der Variablen aus $dom(\sigma)$ unter σ beschreiben. Ist etwa $dom(\sigma) = \{x_1, \ldots, x_n\}$, so kann man das Tupel $\langle x_1/\sigma(x_1), \ldots, x_n/\sigma(x_n)\rangle$ zur Darstellung von σ verwenden.

Definition 7.66 Eine Substitution σ heißt *Unifikator* der Terme $t_1, t_2 \in T(\Sigma, X)$, falls $\sigma(t_1) = \sigma(t_2)$ gilt. Eine Substitution μ heißt *allgemeinster Unifikator* von t_1 und t_2, falls μ ein Unifikator von t_1 und t_2 ist, und falls zu jedem Unifikator σ von t_1 und t_2 eine Substitution λ existiert derart, daß $\sigma = \mu \circ \lambda$ gilt.

Beispiel 7.67 Es enthalte Σ die Funktionssymbole f (2-stellig), g (1-stellig) und die Individuenkonstante a und b. Es seien x, y, z Variablen aus X. Wir betrachten die Terme $t_1 := f(x, g(y))$ und $t_2 := f(g(z), z)$. Die Substitutionen

$$\begin{aligned}
\sigma_1 &:= \langle x/g(g(a)), y/a, z/g(a)\rangle \\
\sigma_2 &:= \langle x/g(g(g(b))), y/g(b), z/g(g(b))\rangle \\
\sigma_3 &:= \langle x/g(g(f(a,z))), y/f(a,z), z/g(f(a,z))\rangle
\end{aligned}$$

sind Unifikatoren von t_1 und t_2. In der Tat gilt

$$\begin{aligned}
\sigma_1(t_1) &= f(g(g(a)), g(a)) = \sigma_1(t_2) \\
\sigma_2(t_1) &= f(g(g(g(b))), g(g(b))) = \sigma_2(t_2) \\
\sigma_3(t_1) &= f(g(g(f(a,z))), g(f(a,z))) = \sigma_3(t_2)
\end{aligned}$$

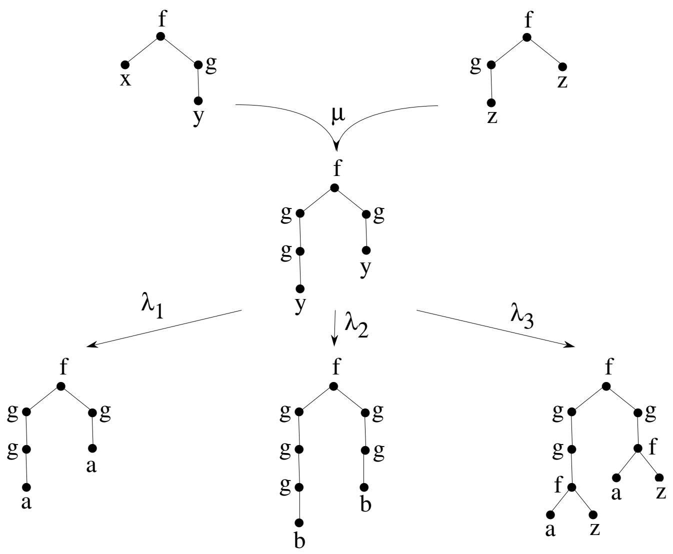
Alle drei genannten Substitutionen sind jedoch keine allgemeinsten Unifikatoren von t_1 und t_2. Ein allgemeinster Unifikator ist zum Beispiel

$\mu := \langle x/g(g(y)), z/g(y)\rangle$. Setzen wir etwa

$$\begin{array}{lcl} \lambda_1 & := & \langle y/a\rangle \\ \lambda_2 & := & \langle y/g(b)\rangle \\ \lambda_3 & := & \langle y/f(a,z)\rangle \end{array}$$

so gilt $\sigma_i = \mu \circ \lambda_i$, für $i = 1, 2, 3$.

Die Rolle von Substitutionen, Unifikatoren und allgemeinsten Unifikatoren wird besonders im Baumbild klar. Die nachfolgende Figur illustriert die Situation im vorausgegangenen Beispiel.



Substitutionen ersetzen also Variablen-Blätter durch Teilbäume. Zu beachten ist, daß für dieselbe Variable überall derselbe Teilbaum eingesetzt wird, und daß alle Ersetzungen *gleichzeitig* durchgeführt werden. Es wird an unserem Beispiel deutlich, daß der allgemeinste Unifikator die „vorsichtigste“ Art und Weise darstellt, wie man zwei gegebene Terme auf dieselbe Form bringen kann. Es bleibt dann immer noch jede andere gemeinsame Form durch eine weitere Instantiierung erreichbar.

Offenkundig wäre im obigen Beispiel auch $\mu' := \langle x/g(g(x)), z/g(x)\rangle$ ein allgemeinster Unifikator. Dies zeigt, daß allgemeinste Unifikatoren nur bis auf „Umbenennung“ von Variablen eindeutig sind.

Natürlich besitzen nicht alle Paare von Termen $t_1, t_2 \in T(\Sigma, X)$ einen Unifikator. Wenn t_1 und t_2 etwa verschiedene oberste Funktionssymbole besitzen, so sind sie sicherlich nicht zu unifizieren. Eine andere Ursache dafür, daß zwei Terme nicht unifizierbar sind, kann darin bestehen, das der Versuch, die beiden Terme zu unifizieren, in einem unendlichen Zyklus landet. Dies tritt etwa dann ein, wenn wir versuchen, die Terme x und $g(x)$ zu unifizieren. Sicherlich können wir x durch $g(x)$ substituieren, dadurch erhalten wir aber lediglich die Terme $g(x)$ und $g(g(x))$, die uns letzlich vor dasselbe Problem stellen. Zwei Fragen stellen sich nun:

1. Gibt es für je zwei unifizierbare Terme $t_1, t_2 \in T(\Sigma, X)$ immer einen allgemeinsten Unifikator?

2. Wie kann man gegebenenfalls einen allgemeinsten Unifikator der Terme t_1 und t_2 berechnen?

Beide Fragen wurden von J.A. Robinson beantwortet, indem er einen Algorithmus angab, der als Eingabe zwei beliebige Terme $t_1, t_2 \in T(\Sigma, X)$ nimmt und

1. mit „fail" endet, falls die Terme t_1 und t_2 nicht unifizierbar sind und

2. andernfalls einen allgemeinsten Unifikator μ von t_1 und t_2 ausgibt.

Mit Hilfe der folgenden drei Beispiele wollen wir eine informelle Darstellung des *Robinsonschen Unifikations-Algorithmus* geben.

Beispiele 7.68 Angenommen, wir wollen die Terme $t_1 := f(x, g(y))$ und $t_2 := f(g(z), z)$ unifizieren. Dazu lesen wir zunächst die Terme von links nach rechts, bis wir zur ersten Stelle kommen, wo wir auf beiden Seiten auf unterschiedliche Symbole treffen. In unserem Beispiel tritt dies dort auf, wo wir links auf x, rechts auf g treffen. Anstelle von g betrachten wir den zugehörigen Teilterm, der mit g beginnt, also $g(z)$. Wir unifizieren nun das „Unterschiedspaar" x und $g(z)$ durch die Substitution $\mu_1 := \langle x/g(z)\rangle$.

Durch Anwendung von μ_1 auf t_1 und t_2 erhalten wir die Terme $t_1' := f(g(z), g(y))$ und $t_2' := f(g(z), z)$. Wir suchen nun das nächste Unterschiedspaar, das in diesem Fall $g(y)$ und z ist. Demgemäß wenden wir die Substitution $\mu_2 := \langle z/g(y)\rangle$ an. Wir erhalten aus t_1' und t_2' nun das identische

Termpaar t_1" $:= f(g(g(y)), g(y))$ und $t_2' := f(g(g(y)), g(y))$. In diesem Fall ist die Komposition $\mu = \mu_1 \circ \mu_2 = \langle x/g(g(y)), z/g(y)\rangle$ ein allgemeinster Unifikator von t_1 und t_2.

Eine erste Variation ergibt sich, wenn wir versuchen, die Terme $s_1 := f(x, g(y))$ und $s_2 := f(g(z), a)$ zu unifizieren. Der erste Schritt bleibt unverändert und führt auf die Terme $s_1' := f(g(z), g(y))$ und $s_2' := f(g(z), a)$. Das nächste Unterschiedspaar ist nun jedoch $g(y)$ und a, und offenkundig sind diese Terme nicht unifizierbar (wir haben einen sogenannten „Clash“). Daher scheitert der Algorithmus mit „fail“.

Eine zweite Variation erhalten wir, wenn wir die Terme $r_1 := f(x, g(y))$ und $r_2 := f(g(z), y)$ betrachten. Wieder bleibt der erste Schritt unverändert. Er führt diesmal auf die Terme $s_1' := f(g(z), g(y))$ und $s_2' := f(g(z), y)$. Das nächste Unterschiedspaar ist $g(y)$ und y. Da y in $g(y)$ vorkommt, können die beiden Terme nicht unifiziert werden, und der Algorithmus scheitert wieder mit „fail“. Das hier auftretende Problem ist unter dem Namen „occur-check failure“ bekannt.

Tatsächlich sind auftretende Clashes und occur-check failure die einzigen Ursachen, die zum Scheitern des Unifikationsversuchs führen können.

7.5.2 Eulersche Kreise

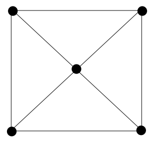
Bei endlichen zusammenhängenden Graphen stellt sich das bekannte Problem, ob es möglich ist, alle Kanten in einem Pfad zusammenzufassen, ohne dabei eine Kante zweimal zu durchlaufen.

Definition 7.69 Sei $\mathcal{G}$ ein zusammenhängender Graph. Ein Pfad $e_1, \ldots, e_k$ heißt *Eulerscher*[4] *Pfad* von $\mathcal{G}$ genau dann, wenn $e_1, \ldots, e_k$ für jede ungerichtete Kante $\{\langle n, m\rangle, \langle n, m\rangle\}$ von $\mathcal{G}$ genau ein Vorkommen einer der Kanten $\langle n, m\rangle$ oder $\langle m, n\rangle$ hat. Der Eulersche Pfad $e_1, \ldots, e_k$ heißt *Eulerscher Kreis* genau dann, wenn $start(e_1) = ende(e_k)$ gilt.

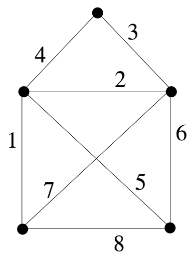
Die Definition beinhaltet insbesondere, daß in einem Eulerschen Pfad nie zwei Kanten $\langle n, m\rangle$ und $\langle m, n\rangle$ auftreten. Es wird damit jede ungerichtete Kante genau einmal und jeweils nur in einer Richtung durchlaufen. Bei

[4]Benannt nach dem Schweizer Mathematiker Leonhard Euler (1707-1783).

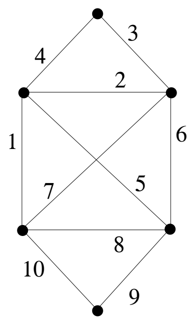
einem Eulerschen Kreis muß man am Ende wieder am Anfangspunkt ankommen. Man kann nicht immer einen Eulerschen Pfad finden, wie das folgende Beispiel zeigt.



Das bekanntest Beispiel eines zusammenhängenden Graphen mit einem Eulerschen Pfad ist das ,,Haus vom Nikolaus“:



Die Kanten 1–8, in der offenkundigen Richtung durchlaufen, bilden einen von mehreren Eulerschen Pfaden. Allerdings gelingt es hier nicht, auch weitergehend einen Eulerschen Kreis anzugeben. Dies ändert sich, wenn wir einen Knoten und zwei Kanten wie folgt hinzufügen:



In der Tat ist bilden nun die Kanten 1–10 einen Eulerschen Kreis. Wir wollen nachfolgend auf den Hintergrund eingehen. Eine erste einfache Beobachtung ist die folgende.

Lemma 7.70 *Der symmetrische Graph $\mathcal{G}$ besitze einen Eulerschen Kreis. Dann hat jeder Knoten von $\mathcal{G}$ geraden Grad.*

Beweis. Wenn wir den Eulerschen Kreis durchlaufen, müssen wir jeden Knoten, den wir besuchen, auch wieder verlassen. Da wir hierzu immer verschiedene Kanten verwenden, so folgt, daß jeder vom Startknoten verschiedene Knoten geraden Grad hat. Dasselbe gilt aber auch für den Startknoten, den wir ganz am Anfang zwar nur verlassen, dafür aber ganz am Ende nur „betreten". ■

Lemma 7.71 *Der symmetrische Graph $\mathcal{G}$ besitze einen Eulerschen Pfad. Dann hat entweder jeder Knoten von $\mathcal{G}$ geraden Grad, oder es gibt genau zwei Knoten mit ungeradem Grad.*

Beweis. Es sei $e_1, \ldots, e_k$ ein Eulerscher Pfad von $\mathcal{G}$. Handelt es sich um einen Eulerschen Kreis, so hat nach Lemma 7.70 jeder Knoten von $\mathcal{G}$ geraden Grad. Im anderen Fall fügen wir zu $\mathcal{G}$ eine neue Kante von $ende(e_k)$ zu $start(e_1)$ hinzu. Offenkundig erhalten wir auf dem neuen Graphen nun einen Eulerschen Kreis. Lemma 7.70 zeigt, daß dort alle Knoten geraden Grad haben. Damit haben in $\mathcal{G}$ aber die Knoten $ende(e_k)$ und $start(e_1)$ ungeraden Grad, alle anderen Knoten geraden Grad. ■

Euler konnte zeigen, daß die Umkehrung von Lemma 7.70 auch richtig ist.

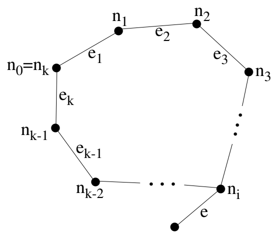
Lemma 7.72 (Satz von Euler) *Es sein $\mathcal{G}$ ein endlicher zusammenhängender Graph. Wenn in $\mathcal{G}$ jeder Knoten geraden Grad hat, so besitzt $\mathcal{G}$ einen Eulerschen Kreis.*

Beweis. Es sei $\mathcal{G} = \langle N, E \rangle$ ein endlicher zusammenhängender Graph, wo jeder Knoten von $\mathcal{G}$ geraden Grad hat. Wir setzen voraus, daß $\mathcal{G}$ zumindest zwei Knoten hat, ansonsten ist das Lemma trivial.

Wir führen zunächst einen Hilfsbegriff ein. Ein Pfad $e_1, \ldots, e_k$ von $\mathcal{G}$ heiße *s-einfach* genau dann, wenn für jede Kante $\langle n, m \rangle \in E$ die Folge $e_1, \ldots, e_k$ höchstens ein Vorkommen einer Kante in $\{\langle n, m \rangle, \langle m, n \rangle\}$ hat. Da $\mathcal{G}$ zusammenhängend ist, besitzt G mindestens eine Kante, damit gibt es zumindest einen s-einfachen Pfad. Wir wählen nun unter allen s-einfachen Pfaden von $\mathcal{G}$ einen von maximaler Länge. Dies sei der Pfad $e_1, \ldots, e_k$. Für $1 \leq i \leq k$ definieren wir $n_{i-1} := \mathit{start}(e_i)$ und $n_i := \mathit{ende}(e_i)$. Der Pfad beginnt also bei n_0, durchläuft $n_1, \ldots, n_{k-1}$ und endet bei n_k.

Wir behaupten zunächst, daß $n_0 = n_k$ gilt. Um dies zu sehen, färben wir die Kanten $e_1, \ldots, e_k$ in der gegebenen Reihenfolge rot ein. Wäre nun $n_0 \neq n_k$, so hätten wir eine ungerade Zahl von Kanten, die von n_k ausgehen, rot eingefärbt. In der Tat, für jedes Vorkommen von n_k in der Teilfolge $n_1, \ldots, n_{k-1}$ färben wir zwei Kanten von n_k rot, eine beim Erreichen, eine beim Verlassen des Knotens. Hinzu kommt eine rote Kante e_k beim letztmaligen Erreichen des Knotens am Ende der Folge. Da n_k jedoch geraden Grad in $\mathcal{G}$ hat, hätte n_k zumindest noch eine ungefärbte Kante e. Damit wäre aber $e_1, \ldots, e_k, e$ ein s-einfacher Pfad der Länge $k+1$. Dies widerspricht der Wahl von $e_1, \ldots, e_k, e$. Folglich gilt $n_0 = n_k$.

Als nächstes zeigen wir, daß jeder Knoten aus N in der Folge $n_0, \ldots, n_k$ zumindest einmal auftritt. Angenommen es gibt einen Knoten $n \in N$, der nicht in der Folge auftritt. Da $\mathcal{G}$ zusammenhängend ist, gibt es einen Pfad von n_0 zu n. Dieser Pfad muß an einer Stelle eine Kante durchlaufen, die nicht zu $\{e_1, \ldots, e_k\}$ gehört, da ansonsten n in der Folge $n_0, \ldots, n_k$ läge. Es sei e die erste solche Kante des neuen Pfads. Wir können nun, von $n_i := \mathit{start}(e)$ beginnend, zunächst einmal den alten Pfad durchlaufen, und danach die Kante e. Auf diese Weise erhalten wir einen s-einfachen Pfad der Länge $k+1$. Dies ist ein Widerspruch. Damit muß also in der Tat jeder Knoten aus N in $n_0, \ldots, n_k$ zumindest einmal auftreten.



Es bleibt zu zeigen, daß in $e_1, \ldots, e_k$ jede Kante aus $\mathcal{G}$ vorkommt. Wäre e eine Kante, die nicht in $e_1, \ldots, e_k$ auftritt, so könnten wir wieder, von $n_i := \mathit{start}(e)$ beginnend, zunächst einmal den alten Pfad durchlaufen, und danach die Kante e. Wir erhalten denselben Widerspruch wie vorher. ■

Aus dem vorigen Satz erhält man leicht die entsprechende Aussage zur Existenz Eulerscher Pfade.

Lemma 7.73 *Es sei $\mathcal{G}$ ein endlicher zusammenhängender Graph. Wenn $\mathcal{G}$ genau zwei Knoten mit ungeradem Grad hat, so besitzt $\mathcal{G}$ einen Eulerschen Pfad.*

Den Beweis lassen wir als Übungsaufgabe offen (vgl. Aufgabe 7.30).

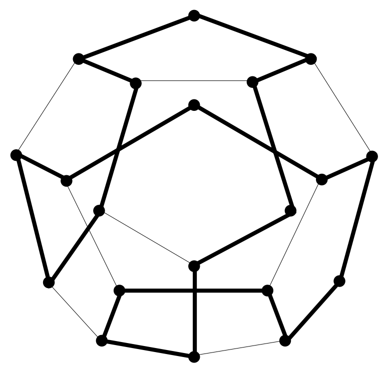
Neben den Eulerschen Kreisen, wo das Traversieren aller Kanten die Hauptbedingung darstellt, gibt es das verwandte Problem, in einem Graphen alle Punkte zu durchlaufen, ohne einen Knoten zweimal zu besuchen.

Definition 7.74 Ein Pfad $e_1, \ldots, e_k$ eines symmetrischen Graphen $\mathcal{G}$ heißt *Hamiltonscher Pfad* genau dann, wenn die Knotenfolge $start(e_1), \ldots, start(e_k), ende(e_k)$ alle Knoten von $\mathcal{G}$ umfaßt und wenn kein Knoten zweimal in der Folge auftritt.

Definition 7.75 Ein Pfad $e_1, \ldots, e_k$ eines symmetrischen Graphen $\mathcal{G}$ heißt *Hamiltonscher Kreis* genau dann, wenn die Knotenfolge $start(e_1), \ldots, start(e_{k-1})$ ein Hamiltonscher Pfad ist und wenn $ende(e_k) = start(e_1)$ gilt.

Das Problem, für einen gegebenen symmetrischen Graphen zu entscheiden, ob es einen Hamiltonsche Kreis gibt, geht auf den W. Hamilton[5] zurück. Die nachfolgende Figur stellt eine Lösung des ursprünglichen Beispielproblems dar.

[5] Sir William Hamilton, irischer Mathematiker, 1805-1865.

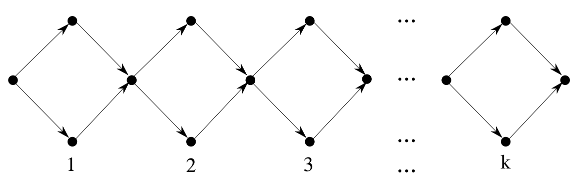


7.6 Aufgaben zu Kapitel 7

Aufgaben zu Teilkapitel 7.1

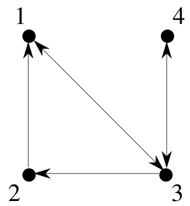
Aufgabe 7.1 Zeigen Sie, daß die Abbildungen 2 und 3 in Bemerkung 7.2 isomorphe gerichtete Graphen repräsentieren. Benennen Sie hierzu die Knoten und geben Sie dann einen Isomorphismus an.

Aufgabe 7.2 Wieviele Pfade der Länge $2k$ besitzt der nachfolgend dargestellte gerichtete Graph?



Stellen Sie sich nun jede Kante mit einem spezifischen Gewicht versehen vor. Das Gewicht eines Pfads soll als die Summe der Gewichte der beteiligten Kanten definiert sein. Geben Sie eine Methode an, den Pfad der Länge $2k$ mit dem geringstem Gewicht zu finden. Vermeiden Sie dabei, für jeden möglichen Pfad der Länge $2k$ sein Gewicht zu berechnen.

Aufgabe 7.3 Geben Sie im nachfolgend abgebildeten Graphen für jedes Element die (direkten) Nachfolger und Vorgänger an. Geben Sie alle (geschlossenen) Pfade an.



Sind alle geschlossenen Pfade einfach?

Aufgabe 7.4 Beweisen Sie Lemma 7.8: Der Knoten n des gerichteten Graphen $\mathcal{G}$ sei ein Nachfolger (resp. Vorgänger) des Knotens m. Dann ist jeder Nachfolger (resp. Vorgänger) von n auch ein Nachfolger (resp. Vorgänger) von m.

Aufgabe 7.5 Stellen Sie die Menge $\{\{\{a, b, c\}, \{\{a, b\}\}\}\}$ als DAG dar, wo die Kanten die Elementbeziehung beschreiben.

Aufgabe 7.6 Kann man zu *jedem* endlichen DAG $\mathcal{G}$ eine Menge M angeben, so daß $\mathcal{G}$ isomorph ist zum Graphen, den wir als DAG-Repräsentation von M enthalten?

Aufgabe 7.7 Sei E^* die reflexiv-transitive Hülle der Kantenrelation E des symmetrischen Graphen $\mathcal{G} = \langle N, E \rangle$. Zeigen Sie: $\mathcal{G}$ ist genau dann zusammenhängend, wenn N/E^* genau ein Element hat.

Aufgaben zu Teilkapitel 7.2

Aufgabe 7.8 Geben Sie bis auf Isomorphie alle ungeordneten Bäume mit drei (vier) Kanten an. Kanten sollen hierbei in Vorzugsrichtung gerichtet sein.

Aufgabe 7.9 Es sei $\mathcal{T}$ ein ungeordneter Baum, wo jeder Knoten maximal Verzweigungsgrad n hat. Die maximale Länge eines Pfads von der Wurzel zu einem Blatt betrage $k \in \mathbb{N}$. Wieviele Blätter kann $\mathcal{T}$ maximal haben? Es sei nun $n = 2$. Wieviele Knoten kann $\mathcal{T}$ maximal haben?

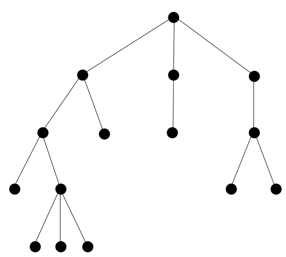
Aufgabe 7.10 Es sei $\mathcal{T}$ ein endlicher ungeordneter Baum mit m Knoten und n Kanten. Beweisen Sie, daß $n = m - 1$ gilt.

Aufgabe 7.11 Geben Sie einen endlich verzweigenden Baum an, der nicht uniform endlichen Verzweigungsgrad hat.

Aufgabe 7.12 Zeigen Sie, daß ein rationaler Baum stets uniform endlichen Verzweigungsgrad hat.

Aufgabe 7.13 Erweitern Sie den Begriff des Baumskeletts dahingehend, daß eine Beschreibung geeigneter unendlicher geordneter Bäume möglich ist. Geben Sie ein Beispiel eines unendlichen Baumskeletts und des zugehörigen Baums. Warum kann man nicht jeden geordneten Baum durch ein Baumskelett beschreiben?

Aufgabe 7.14 Bestimmen Sie für den nachfolgend dargestellten geordneten Baum das zugehörige Baumskelett, die Präordnung und die Postordnung auf der Menge der Knoten.



Aufgaben zu Teilkapitel 7.3

Aufgabe 7.15 Wir hatten in Abschnitt 7.3 vorausgesetzt, daß die Signatur Σ zumindest eine Individuenkonstante enthält. Warum?

Aufgabe 7.16 Es sei Σ eine nichtleere Signatur mit zumindest einer Individuenkonstante. Gibt es immer unendlich viele Σ-Bäume?

Aufgabe 7.17 Es sei Σ eine Signatur mit mit zumindest einer Individuenkonstante. Geben Sie eine induktive Definition der Menge aller Σ-Bäume.

Aufgabe 7.18 Die Signatur Σ enthalte die Individuenkonstante a und b, das einstellige Funktionssymbol g und das zweistellige Funktionssymbol f. Geben Sie alle Σ-Bäume der Höhe 0, 1 und 2 an.

Aufgabe 7.19 Zeigen Sie: ist die Signatur Σ endlich, und enthält Σ zumindest eine Individuenkonstante, so ist die Menge der Σ-Bäume endlich oder abzählbar unendlich. Verwenden Sie die Aussage von Aufgabe 5.3.

Aufgabe 7.20 Zeigen Sie, daß die Baumalgebra $\mathcal{B}(\Sigma)$ bereits von der leeren Teilmenge erzeugt wird.

Aufgabe 7.21 Die Signatur Σ enthalte genau eine Individuenkonstante e sowie die einstelligen Funktionssymbole $f_1, \ldots, f_n$. Es sei $A := \{a_1, \ldots, a_n\}$ ein endliches Alphabet mit genau n Symbolen. Auf A^* definieren wir die n Nachfolgerfunktionen $succ_a : w \mapsto w \circ a$ $(a \in A)$. Zeigen Sie, daß die Baumalgebra $\mathcal{B}(\Sigma)$ zur resultierenden Algebra auf A^* isomorph ist.

Aufgabe 7.22 Die Signatur Σ enthalte die Individuenkonstante *Eins* und das zweistellige Funktionssymbol Add. Beschreiben Sie den nach Theorem 7.50 eindeutig bestimmten Homomorphismus h von $\mathcal{B}(\Sigma)$ in $\langle \mathbb{N}, 1, + \rangle$.

Aufgabe 7.23 Die Signatur Σ enthalte die Individuenkonstante *Null* und das zweistellige Funktionssymbol Add. Es sei A eine nichtleere Menge, $\mathcal{R}(A)$ die Menge aller zweistelligen Relationen auf A, „$\circ$“ die in Definition 3.23 erklärte Komposition von Relationen und Id_A die Identität auf A. Beschreiben Sie den nach Theorem 7.50 eindeutig bestimmten Homomorphismus h von $\mathcal{B}(\Sigma)$ in $\langle \mathcal{R}(A), Id_A, \circ \rangle$.

Aufgaben zu Teilkapitel 7.4

Aufgabe 7.24 Geben Sie die Menge aller Teilterme des Terms $f(f(g(a), f(a, b)), g(g(b)))$ an. Geben Sie für jeden Teilterm seine Tiefe an.

Aufgabe 7.25 Die Signatur Σ enthalte die Individuenkonstante *Null* und das zweistellige Funktionssymbol Add. Es bezeichne $\mathcal{Z}$ die Menge der ganzen Zahlen und $h_0 : X \to \mathcal{Z}$ eine Abbildung, die Variablen auf Zahlen abbildet. Beschreiben Sie den nach Theorem 7.60 eindeutig bestimmten Homomorphismus h von $\mathcal{T}(\Sigma, X)$ in $\langle \mathbb{N}, 0, + \rangle$, der h_0 fortsetzt.

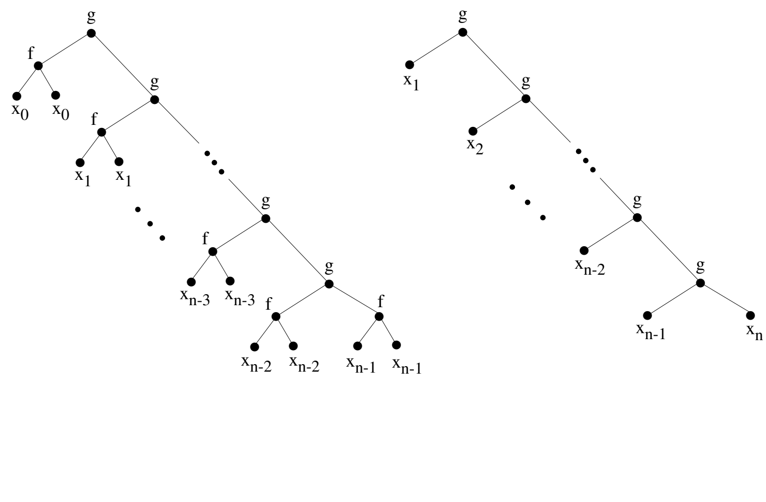
Aufgaben zu Teilkapitel 7.5.1

Aufgabe 7.26 Wenden Sie die Subsitution $\langle x/f(x, y), y/f(x, y) \rangle$ auf die Terme a, $f(x, y)$ und $f(f(a, g(g(y))), g(x))$ an.

Aufgabe 7.27 Welche der nachfolgenden Termpaare sind unifizierbar? Geben Sie gegebenenfalls einen allgemeinsten Unifikator an.

$$\begin{array}{lcl} f(f(a,x),x) & \text{und} & f(y, f(y,a)) \\ f(f(a,x),y) & \text{und} & f(z, f(a,z)) \end{array}$$

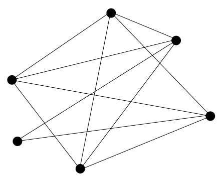
Aufgabe 7.28 Unifizieren Sie nachfolgend in Baumform dargestellten Terme. Es sei μ ein allgemeinster Unifikator.

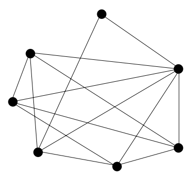


Zeigen Sie, daß die Größe (Zahl der Knoten im Baumbild) des Terms $\mu(x_n)$ mit n exponentiell wächst. Wie könnte man zu einer kompakteren Darstellung des allgemeinsten Unifikators mittels DAGs kommen?

Aufgaben zu Teilkapitel 7.5.2

Aufgabe 7.29 Geben Sie in den nachfolgend abgebildeten zwei Graphen jeweils einen Eulerschen Kreis an.





Aufgabe 7.30 Beweisen Sie Lemma 7.73: Es sei $\mathcal{G}$ ein endlicher zusammenhängender Graph. Wenn $\mathcal{G}$ genau zwei Knoten mit ungeradem Grad hat, so besitzt $\mathcal{G}$ einen Eulerschen Pfad.

7.7 Bibliographische Angaben

Ausführlichere Darstellungen von Graphen und Bäumen mit Beschreibungen einiger wichtiger Algorithmen finden sich in [Big89, RW92], weitergehend in [Jun99]. Eine kompakte Darstellung sogenannter planarer Graphen ist in [Bra88] enthalten. Eine mathematisch sehr detaillierte Beschreibung von Graphen mit anderer Gewichtung bietet [SS93]. Die Theorie der endlichen Automaten ist in allen Büchern über formale Sprachen und Automaten dargestellt, exemplarisch sei auf [HU79, Koz97] verwiesen. Der Inhalt von 7.5.2 lehnt sich an die Darstellung in [RW92] an. Dort finden sich auch mehr Informationen zu Hamiltonschen Kreisen. Der Robinsonsche Unifikations-Algorithmus wurde in [Rob65] eingeführt. Eine formale Analyse der Unifikation von Termen ist in [LMM87] zu finden.