Mathematische Grundlagen der Informatik und Linguistik
Teil 1: von Mengen zur Aussagenlogik

Klaus U. Schulz

4. Oktober 2001

Inhaltsverzeichnis

1

Zum Umgang mit mathematischen Aussagen

Es ist eine ganz wesentliche Grundlage der mathematischen Denkweise, daß alle Behauptungen stets durch entsprechende Beweise zu verifizieren sind. Gerade dieser Aspekt der Mathematik wird jedoch in der Schulmathematik häufig nur am Rande thematisiert. Beim Einstieg in den Mathematikunterricht an der Hochschule bereitet es anfangs fast immer erhebliche Schwierigkeiten, ein Gefühl dafür zu erhalten, was denn eigentlich als ein Beweis zu betrachten sei. Eines unserer Ziele wird sein, diese Unsicherheit soweit wie möglich zu beseitigen. Hierzu gibt es kein Patentrezept, insbesondere es unmöglich, an dieser Stelle eine formale Definition eines Beweises zu geben. Man kann aber doch eine Reihe einfacher Regeln vorab beschreiben, von denen in Beweisen immer wieder—mehr oder weniger stillscheigend—Gebrauch gemacht wird. Aus der Erfahrung, daß sich damit viele Unsicherheiten und Verwirrungen beheben lassen, werden wir in diesem einleitenden Kapitel einige dieser Prinzipien darstellen.

1.1 Aussagen und Wahrheitswerte

Eine *Aussage* im Bereich der natürlichen Sprache ist ein Satz, der wahr oder falsch sein kann. Typisch für die natürliche Sprache sind aber auch Aussagen, wo es fraglich erscheint, ob die bloße Auswahl zwischen Wahrheitswerten „wahr" oder „falsch" eine adäquate Beurteilung liefern kann. Nach der Richtigkeit von Sätzen wie

„München ist schön"
„Der Mond ist rund"

gefragt, werden wir vielleicht dazu tendieren, weitere Charakterisierungen wie „Geschmacksfrage", „vage richtig" oder „ambig" als sinnvolle und mögliche Wahrheitswerte heranzuziehen.

Auch in der Mathematik hat man ständig mit Aussagen zu tun. Um zu einer hinreichend formalen Betrachtungsweise zu gelangen, verlangt man aber, *daß mathematische Aussagen stets entweder wahr oder falsch sind*, weitere Möglichkeiten gibt es nicht. Dieses Prinzip bedeutet natürlich nicht, daß der Wahrheitswert einer mathematischen Aussage bekannt sein muß, es bedeutet nicht einmal, daß der Wahrheitswert prinzipiell ermittelt werden kann. Es ist z.B. kein Verfahren bekannt, mit dem festgestellt werden könnte, ob die Ziffernfolge

$$w = 122333444455555666666$$

in der Dezimaldarstellung der Kreiszahl π vorkommt. Gleichwohl wissen wir, daß die Aussage

„die Ziffernfolge w kommt in der Dezimaldarstellung von π vor"

entweder unzweifelhaft stimmt oder eindeutig falsch ist.

Sicherlich wäre es trotzdem ohne weiteres möglich, mit weiteren Wahrheitswerten zu arbeiten. Insofern stellt die angegebene Grundregel lediglich eine — im folgenden für uns bindende — Konvention dar.

Wir werden später fortlaufend mit konkreten mathematischen Aussagen konfrontiert sein. Um den Umgang mit solchen Aussagen vorab zu erklären, führen wir Symbole

$$\alpha, \beta, \gamma, \ldots, \alpha_0, \beta_0, \gamma_0, \ldots, \alpha_1, \beta_1, \gamma_1, \ldots$$

ein[1], die für beliebige mathematische Aussagen stehen sollen. Als Wahrheitswerte verwenden wir die Symbole „1" und „0" , es bedeutet

$$WW(\alpha) = 1 \text{ (bzw. } WW(\alpha) = 0)$$

[1] α, β, γ liest sich: alpha, beta, gamma.

daß die Aussage α wahr (bzw. falsch) ist.

1.2 Aussagenlogische Junktoren

Während die Frage, ob für eine mathematische Aussage α nun $WW(\alpha) = 1$ oder $WW(\alpha) = 0$ gilt, oft schwierig oder gar nicht zu beantworten ist, ist es einfacher, Regeln festzulegen, die beschreiben, welche Wahrheitswerte man erhält, wenn man Aussagen bestimmter Wahrheitswerte mit Hilfe sogenannter *Junktoren* oder *Bindewörter* zu komplexeren Aussagen kombiniert. Die gebräuchlichsten Junktoren sind, in ihrer umgangsprachlichen Bezeichnung,

und, oder, falls, wenn–dann, genau–dann–wenn, nicht.

Diese Junktoren werden auch in mathematischen Aussagen sehr oft verwendet, manchmal natürlich–sprachlich, manchmal abgekürzt durch folgende Symbole:

$\wedge$	(und, Konjunktion)
$\vee$	(oder, Disjunktion)
$\Rightarrow$	(wenn-dann, Implikation)
$\Leftrightarrow$	(genau-dann-wenn, Biimplikation)
$\neg$	(nicht, Negation)

In der Umgangssprache können diese Junktoren — zumindest zur Verbindung von Teilphrasen — mit unterschiedlichem Sinn verwendet werden:

„An Sonn– und Feiertagen fährt der Bus nicht.“
„An kühlen und regnerischen Tagen verkriecht sich der Dachs.“

In der Mathematik hingegen folgt die Verwendung dieser Junktoren genau festgelegten Spielregeln, um mathematischen Aussagen einen eindeutigen Inhalt zu geben. Diese Regeln wollen wir hier vorab darstellen.[2] Die Verwendung der Negation ist naheliegend. In der Mathematik wie im Alltag

[2]Für Leser mit etwas Hintergrundwissen: wir sollten klarstellen, daß wir an dieser Stelle keine Einführung in die Aussagenlogik im Sinn haben, auch wenn wir im folgenden Teile der Aussagenlogik oberflächlich streifen. Uns geht es hier zunächst nur darum, die Verwen-

empfindet man eine Aussage der Form $\neg\alpha$ genau dann als wahr, wenn α falsch ist. Kürzer:

$$WW(\neg\alpha) = 1 \Leftrightarrow WW(\alpha) = 0.$$

In einer Wahrheitswert-Tabelle kann man diese Regel so festhalten:

α	$\neg\alpha$
0	1
1	0

Ähnlich legt die folgende Tabelle den mathematischen Gebrauch der übrigen Standardjunktoren fest.

α	β	$\alpha \wedge \beta$	$\alpha \vee \beta$	$\alpha \Rightarrow \beta$	$\alpha \Leftrightarrow \beta$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

Hierzu einige Kommentare. Die Festlegung der Konjunktion „$\wedge$“ entspricht der Erwartung: eine Aussage der Form $\alpha \wedge \beta$ ist genau dann wahr, wenn sowohl α als auch β wahr ist. Die Verwendung des Junktors „$\vee$“ (Disjunktion) entspricht dem lateinischen „vel“, also dem nicht–ausschließenden oder. Am willkürlichsten ist die Festlegung des Gebrauchs der Implikation „$\Rightarrow$“. Gemäß der Tabelle ist eine Aussage der Form $\alpha \Rightarrow \beta$ automatisch richtig, wenn das Vorderglied α falsch ist. Dieses Prinzip ist unter dem Stichwort „ex falso quodlibet“ bekannt. Wenn wir α und β mit natürlich-sprachlichen Inhalten füllen, bedeutet dies etwa, daß ein Satz wie

„Wenn Prag am Rhein liegt, dann ist zwei und zwei fünf.“

als wahr eingestuft wird. Dies ist zumindest diskutabel — man könnte auch den Standpunkt vertreten, daß ein derartiger Satz überhaupt keine Aussage macht, da die Prämisse falsch ist. Man kann aber auch die Festlegung kritisieren, daß Aussagen der Form „wenn α, dann β“ stets als wahr eingestuft werden, wann immer α und β wahr sind. Nehmen wir den Satz

dungsweise von Junktoren in mathematischen Aussagen zu klären. Insbesondere führen wir keine Unterscheidung in Objekt- und Metasprache ein, demzufolge ist für uns ein Junktorensymbol wirklich nur eine Abkürzung für den entsprechenden natürlich-sprachlichen Ausdruck in seiner mathematischen Verwendung.

„Wenn morgen Montag ist, dann ist zwei und zwei vier."

Auch an einem Sonntag geäußert, kann man in als inakzeptabel einstufen, da die Prämisse keinerlei *Relevanz* für die Konklusion hat.

Hier wird offensichtlich, daß die Wahrheitswert-Tabellen letztlich eine *Übereinkunft* darstellen, die den Gebrauch der Junktoren bindend festsetzt, und nicht etwa eine „objektive und ewig und überall unumstößliche Wahrheit", die lediglich einmal explizit festgehalten werden muß. Natürlich versucht man solche Übereinkünfte, die an anderer Stelle etwa auch in der Form von Axiomen in die Mathematik eingehen, intuitiv richtig und vernünftig zu treffen. Der Vorteil der gegebenen Formalisierung der Implikation ist ihre Einfachheit. Insbesondere folgt aus unserer Festlegung das folgende

Kompositionalitätsprinzip: *Es sei J einer der Junktoren $\wedge, \vee, \Rightarrow, \Leftrightarrow$. Dann kann der Wahrheitswert einer Aussage der Form $\alpha\ J\ \beta$ allein aus den Wahrheitswerten $WW(\alpha)$ und $WW(\beta)$ sowie aus der Kenntnis von J ermittelt werden. Entsprechendes gilt für Aussagen der Form $\neg\alpha$.*

Entscheidend hierbei ist, daß in die „innere Form" der Bestandteile α und β nicht hineingeschaut zu werden braucht, wenn ihre Wahrheitswerte bekannt sind.

1.3 Aussagenlogische Tautologien

Nachdem die Wahrheitswert-Tabellen den Gebrauch der Junktoren festschreiben, stellt man nun fest, daß es gewisse zusammengesetzte Aussagen gibt, die allein aufgrund ihrer Junktoren-Struktur immer wahr sind, unabhängig davon, wie die Wahrheitswerte der einfachsten Bestandteile aussehen. Wir werden solche Aussagen als *aussagenlogische Tautologien* bezeichnen. Beispielsweise ist jede Aussage der Form $\alpha \vee \neg\alpha$ oder $\neg(\alpha \wedge \neg\alpha)$ eine aussagenlogische Tautologie, wie sich mittels einer Wahrheitswert-Tabelle an folgender Fallunterscheidung ablesen läßt:

α	$\neg\alpha$	$\alpha \vee \neg\alpha$	$\alpha \wedge \neg\alpha$	$\neg(\alpha \wedge \neg\alpha)$
0	1	1	0	1
1	0	1	0	1

Da die Aussage α entweder wahr oder falsch ist, ist die Fallunterscheidung komplett. Der nachfolgende Satz gibt eine Reihe weiterer Tautologien. Die Länge der Liste wird dadurch motiviert, daß uns die meisten der angegebenen Tautologien noch in anderen Zusammenhängen begegnen werden.

Satz 1.1 *Seien α, β und γ Aussagen. Die nachfolgenden Aussagen sind aussagenlogische Tautologien:*

1. $(\alpha \vee \alpha) \Leftrightarrow \alpha$ *(Idempotenz von $\vee$),*
2. $(\alpha \wedge \alpha) \Leftrightarrow \alpha$ *(Idempotenz von $\wedge$),*
3. $(\alpha \wedge \beta) \Leftrightarrow (\beta \wedge \alpha)$ *(Kommutativität von $\wedge$),*
4. $(\alpha \vee \beta) \Leftrightarrow (\beta \vee \alpha)$ *(Kommutativität von $\vee$),*
5. $((\alpha \wedge \beta) \wedge \gamma) \Leftrightarrow (\alpha \wedge (\beta \wedge \gamma))$ *(Assoziativität von $\wedge$),*
6. $((\alpha \vee \beta) \vee \gamma) \Leftrightarrow (\alpha \vee (\beta \vee \gamma))$ *(Assoziativität von $\vee$),*
7. $(\alpha \wedge (\beta \vee \gamma)) \Leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$ *(Distributivität von $\wedge$ und $\vee$),*
8. $(\alpha \vee (\beta \wedge \gamma)) \Leftrightarrow (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$ *(Distributivität von $\vee$ und $\wedge$),*
9. $\neg(\neg\alpha) \Leftrightarrow \alpha$,
10. $\neg(\alpha \wedge \beta) \Leftrightarrow (\neg\alpha \vee \neg\beta)$,
11. $\neg(\alpha \vee \beta) \Leftrightarrow (\neg\alpha \wedge \neg\beta)$,
12. $(\alpha \Rightarrow \beta) \Leftrightarrow ((\alpha \vee \beta) \Leftrightarrow \beta)$,
13. $(\alpha \Rightarrow \beta) \Leftrightarrow ((\alpha \wedge \beta) \Leftrightarrow \alpha)$,
14. $(\alpha \Rightarrow \beta) \Leftrightarrow (\neg\beta \Rightarrow \neg\alpha)$,
15. $\neg(\alpha \Rightarrow \beta) \Leftrightarrow (\alpha \wedge \neg\beta)$,
16. $(\alpha \Leftrightarrow \beta) \Leftrightarrow ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha))$,
17. $(\alpha \wedge (\alpha \Rightarrow \beta)) \Rightarrow \beta$,
18. $((\alpha \Rightarrow \beta) \wedge \neg\beta)) \Rightarrow \neg\alpha$,
19. $((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \gamma)) \Rightarrow (\alpha \Rightarrow \gamma)$ *(Transitivität von $\Rightarrow$),*

20. $((\alpha \Leftrightarrow \beta) \wedge (\beta \Leftrightarrow \gamma)) \Rightarrow (\alpha \Leftrightarrow \gamma)$ *(Transitivität von* $\Leftrightarrow$*).*

Beweis. Ein Beweis ergibt sich jeweils aus einer Fallunterscheidung über die möglichen Wahrheitswerte der Aussagen α, β und γ. Wir führen dies exemplarisch für Tautologie 12 vor:

α	β	$\alpha \Rightarrow \beta$	$\alpha \vee \beta$	$(\alpha \vee \beta) \Leftrightarrow \beta$	11.
0	0	1	0	1	1
0	1	1	1	1	1
1	0	0	1	0	1
1	1	1	1	1	1

Kommen drei Teilaussagen α, β und γ vor, so sind natürlich $8 = 2^3$ Fälle zu unterscheiden. ■

Viele der angegebenen Tautologien können in der Weise gelesen werden, daß sie eine Strategie zur Beweisführungen für eine Behauptung beinhalten. So besagt zum Beispiel Tautologie 19 (Transitivität von „$\Rightarrow$“ folgendes: wenn eine Aussage α eine Zwischenbehauptung β impliziert, und wenn die Zwischenbehauptung β ihrerseits die Aussage γ impliziert, so ist damit auch gezeigt, daß aus α die Aussage γ folgt. Die Beobachtung wird in Aufgabe 1.2 vertieft.

1.4 Äquivalente Aussagen

Man nennt zwei Aussagen α und β *äquivalent* genau dann, wenn $\alpha \Leftrightarrow \beta$ wahr ist. Zum Beispiel sind α und β stets „aussagenlogisch“ äquivalent, wenn $\alpha \Leftrightarrow \beta$ eine aussagenlogische Tautologie ist. In Satz 1.1 geben die ersten vierzehn Tautologien aussagenlogische Äquivalenzen wieder. Die Transitivität von „$\Leftrightarrow$“ (Tautologie 20) macht deutlich, daß der Beweis der Äquivalenz zweier Aussagen α und γ auch mit Hilfe zweier Zwischenschritte erfolgen kann, wo man die Äquivalenz von α und γ zu einer Aussage β verifiziert. Offenkundig können dann auch mehrere solcher Zwischenschritte verwendet werden. Ketten gültiger Äquivalenzen

$$\alpha_1 \Leftrightarrow \alpha_2, \alpha_2 \Leftrightarrow \alpha_3, \ldots, \alpha_{n-1} \Leftrightarrow \alpha_n$$

zeigen damit eine gültige Äquivalenz $\alpha_1 \Leftrightarrow \alpha_n$. Viele Äquivalenzbeweise folgen dieser einfachen Struktur, wobei darüberhinaus häufig zusätzlich von

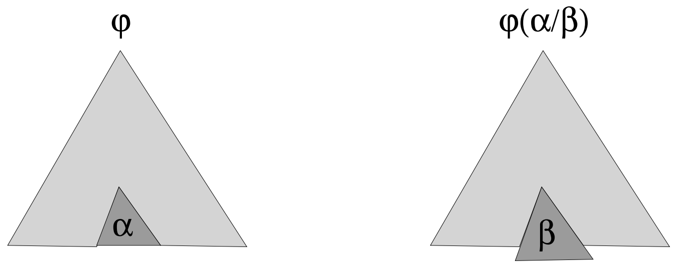


Abbildung 1.1: Sind α und β äquivalent, so auch φ und $\varphi(\alpha/\beta)$.

folgendem Prinzip Gebrauch gemacht wird, das man als eine Verallgemeinerung des Kompositionalitätsprinzips ansehen kann.

Ersetzungsprinzip: *Ersetzt man in einer Aussage φ einige Vorkommen einer Teilaussage α durch eine zu α äquivalente Aussage β, so ist die entstehende Aussage $\varphi(\alpha/\beta)$ äquivalent zu φ. Es ist also $\varphi(\alpha/\beta)$ wahr genau dann, wenn φ wahr ist.*

Abbildung 1.1 bietet eine Illustration. Ein verwandtes Prinzip, das in der Mathematik oft verwendet wird, ist das folgende

Leibniz-Prinzip: *Es seien s und t zwei Terme, das heißt Ausdrücke, die Elemente bezeichnen. Kommt s in einer Aussage φ vor und gilt $s = t$, so ist φ wahr genau dann, wenn auch die Aussage $\varphi(s/t)$ wahr ist, die aus φ entsteht, wenn man einige Vorkommen von s durch t ersetzt.*

1.5 Implikationen

Sehr häufig sind in der Mathematik Implikationen zu verifizieren. Eine Implikation ist eine Aussage der Form

$$(\alpha_1 \wedge \ldots \wedge \alpha_n) \Rightarrow \beta.$$

Hierbei ist $n \geq 1$. Wegen der Assoziativität von „$\wedge$“ (Satz 1.1 Nr. 5) führen alle expliziten Klammerungen der α_i's zu äquivalenten Aussagen.

Bemerkung: *Um die Gültigkeit (Wahrheit) einer solchen Aussage zu beweisen, kann man annehmen, daß alle Prämissen α_i $(i = 1, ..., n)$ wahr* *sind. Kann unter dieser Voraussetzung die Konklusion β gezeigt werden, so ist die Implikation bewiesen.*

Entscheidend ist, daß all diejenigen Fälle, wo eine Prämisse α_i falsch ist, überhaupt nicht betrachtet werden müssen. Ein Rechtfertigung ergibt sich unmittelbar aus den Wahrheitstafeln für „$\wedge$“ und für „$\Rightarrow$“. Wie im Falle der Biimplikation kann man auch Ketten gültiger Implikationen

$$\alpha_1 \Rightarrow \alpha_2, \alpha_2 \Rightarrow \alpha_3, \ldots, \alpha_{n-1} \Rightarrow \alpha_n$$

zu einer gültigen Implikation $\alpha_1 \Rightarrow \alpha_n$ zusammenfassen, wie sich aus Tautologie 19 aus Satz 1.1 ergibt.

1.6 Aussageformen

Neben Aussagen treten in der Mathematik häufig *Aussageformen* auf. Aussageformen sind ähnlich wie Aussagen, enthalten aber Variable, die mit bestimmten Werten belegt werden können. Wenn x eine Variable ist, die als Wert eine natürliche Zahl annehmen kann, so sind

„*x ist eine Primzahl*“ und
„$(x > 14) \Rightarrow (x > 10)$“

Aussageformen. Beim Umgang mit Aussageformen treten in der Mathematik häufig der Allquantor „$\forall$“ (lies: für alle) und der Existenzquantor „$\exists$“ (lies: es existiert ein) auf. Mit diesen Quantoren kann man die Variablen einer Aussageform binden. Wenn alle Variablen gebunden sind, erhält man eine Aussage:

„$\exists x$: *x ist eine Primzahl*“ und
„$\forall x$: $(x > 14) \Rightarrow (x > 10)$“

sind wahre Aussagen über die natürlichen Zahlen, die sich ausführlicher geschrieben wie folgt lesen:

„*es existiert ein x mit der Eigenschaft x ist eine Primzahl*“
„*für alle x gilt die Eigenschaft* $(x > 14) \Rightarrow (x > 10)$“

Allgemein ist eine Aussage der Form $\forall x\colon \alpha(x)$ in einem Bereich M wahr genau dann, wenn all diejenigen Aussagen $\alpha(x/n)$ wahr sind, die man durch Belegung der Variablen x in der Aussageform $\alpha(x)$ mit Werten $n \in M$ erhält. Eine Aussage der Form $\exists x\colon \alpha(x)$ ist wahr genau dann, wenn es zumindest einen erlaubten Wert n für x gibt, so daß $\alpha(x/n)$ wahr ist. Der Bereich M der erlaubten Werte ergibt sich aus dem Kontext, oder er wird explizit angegeben, wie in den Aussagen

„$\exists x \in \mathbb{N}$: x *ist eine Primzahl*“ und
„$\forall x \in \mathbb{N}\colon (x > 14) \Rightarrow (x > 10)$“,

wo $\mathbb{N}$ die Menge der natürlichen Zahlen bezeichnet.

Wir wollen kurz festhalten, wie „universelle“ Aussagen (Aussagen der Form $\forall x\colon \alpha(x)$) und „existentielle“ Aussagen (Aussagen der Form $\exists x\colon \alpha(x)$) typischerweise verifiziert oder falsifiziert werden. Diese Rezepte werden insbesondere dann wichtig, wenn die Variable x unendlich viele Werte annehmen kann, die man nicht mehr einzeln inspizieren kann. Wir werden von der Sprechweise Gebrauch machen, sich „einen beliebigen, (generischen, gedachten) Wert herauszugreifen“. Was damit gemeint ist, läßt sich am besten mit folgendem Beispiel aus der Biologie klarmachen. Wie würde man einem Zweifler plausibel machen, daß alle Pinguine Vögel sind? Man könnte die Argumentation etwa in der folgenden Form beginnen:

> „Nehmen Sie sich irgendeine Pinguindame, nennen wir sie Frieda. Sie werden feststellen, daß Frieda in der Tat Federn hat und Eier legt...“

Natürlich ist nicht von einem bestimmten Pinguin die Rede, sondern „Frieda“ ist ein generisches, nur gedachtes Objekt, das stellvertretend für jeden beliebigen—weiblichen—Pinguin steht. Dieselbe Art der Argumentation wird auch in der Mathematik sehr oft verwendet.

Bemerkung 1.2 Um eine Aussage der Form $\forall x\colon \alpha(x)$ zu *beweisen*, kann man sich einen „generischen“ möglichen Wert g von x herausgreifen. Angenommen, wir machen über g keinerlei spezielle Annahmen. Wenn wir dennoch zeigen können, daß $\alpha(x/g)$ wahr ist, so heißt dies, daß man dieselbe Art der Argumentation auch auf jeden konkreten möglichen Wert n von

x anwenden könnte. Man käme stets zum Ergebnis, daß $\alpha(x/n)$ wahr ist. Deshalb ist $\forall x\colon \alpha(x)$ verifiziert. Um eine Aussage der Form $\forall x\colon \alpha(x)$ zu *widerlegen*, reicht es ein Gegenbeispiel anzugeben, das heißt einen möglichen konkreten Wert n von x, für den $\alpha(x/n)$ falsch ist.

Dual hierzu lassen sich für existentielle Aussagen folgende Regeln angeben.

Bemerkung 1.3 Um eine Aussage der Form $\exists x\colon \alpha(x)$ zu *beweisen*, reicht es ein Beispiel anzugeben, das heißt einen möglichen konkreten Wert n von x, für den $\alpha(x/n)$ wahr ist. Um eine Aussage der Form $\exists x\colon \alpha(x)$ zu *widerlegen*, kann man sich einen beliebigen, generischen Wert g von x herausgreifen. Angenommen, wir machen über g keinerlei spezielle Annahmen. Wenn wir dennoch zeigen können, daß $\alpha(x/g)$ falsch ist, so heißt dies, daß man dieselbe Art der Widerlegung auch auf jeden anderen konkreten Wert n von x anwenden könnte. Deshalb ist $\exists x\colon \alpha(x)$ widerlegt.

Wenn wir eine Aussage der Form $\forall x\colon \alpha(x)$ widerlegen, so ist dies gleichbedeutend damit, die Aussage $\neg\forall x\colon \alpha(x)$ nachzuweisen. Dual gilt: Wenn wir eine Aussage der Form $\exists x\colon \alpha(x)$ widerlegen, so ist dies gleichbedeutend damit, die Aussage $\neg\exists x\colon \alpha(x)$ nachzuweisen. Daher folgt aus Bemerkungen 1.2 und 1.3 zusammen mit Tautologie 9 aus Satz 1.1 folgende Beobachtung.

Bemerkung 1.4 Ist α eine Aussageform mit der Variablen x, so sind

$$\begin{array}{rcl}
\neg\forall x\colon \alpha(x) & \text{und} & \exists x\colon \neg\alpha(x),\\
\neg\exists x\colon \alpha(x) & \text{und} & \forall x\colon \neg\alpha(x),\\
\forall x\colon \alpha(x) & \text{und} & \neg\exists x\colon \neg\alpha(x),\\
\exists x\colon \alpha(x) & \text{und} & \neg\forall x\colon \neg\alpha(x)
\end{array}$$

jeweils äquivalente Aussagen.

1.7 Aufgaben zu Kapitel 1

Aufgabe 1.1 Führen Sie den Beweis von Satz 1.1 für weitere Fälle durch. Geben Sie einige weitere Tautologien an.

Aufgabe 1.2 Beweisen Sie, daß die folgenden Aussagen stets aussagenlogische Tautologien sind. Man kann diesen Tautologien eine inhaltliche Interpretation dergestalt geben, daß sie etwas über die mögliche Architektur eines Beweises aussagen. Geben Sie jeweils eine solche Interpretation.

1. $((\alpha \Rightarrow \beta) \wedge ((\alpha \wedge \beta) \Rightarrow \gamma)) \Rightarrow (\alpha \Rightarrow \gamma)$,
2. $(\alpha \Leftrightarrow \beta) \Rightarrow ((\alpha \Rightarrow \gamma) \Leftrightarrow (\beta \Rightarrow \gamma))$,
3. $((\alpha \wedge \beta) \Rightarrow (\gamma \wedge \neg\gamma)) \Rightarrow (\alpha \Rightarrow \neg\beta)$,
4. $(\alpha \Rightarrow \beta) \Leftrightarrow (\neg\beta \Rightarrow \neg\alpha)$.

Aufgabe 1.3 Man sagt manchmal, daß eine allquantifizierte Aussage eine möglicherweise unendliche „Konjunktion" repräsentiert, und eine existentielle Aussage eine möglicherweise unendliche „Disjunktion". Verdeutlichen Sie diese Betrachtungsweise im Spezialfall, wo die quantifizierten Variablen als Werte die natürlichen Zahlen annehmen können.

Aufgabe 1.4 Was bedeuten die folgenden Aussagen über die natürlichen Zahlen in ausführlicher Schreibweise? Welche Aussagen sind wahr?

1. $\exists x \forall y\colon x \geq y$,
2. $\exists x \forall y\colon x \leq y$,
3. $\forall x\colon (x = 0 \vee \exists y\colon x = y + 1)$,
4. $\forall x \exists y\colon (x < y \wedge \forall z\colon (x \leq z \leq y \Rightarrow (z = x \vee z = y)))$,
5. $\exists x\colon (x \neq 0 \wedge \forall y\colon (x \cdot y < x + y))$,
6. $\forall x \forall y \exists z \forall u\colon (((\exists k\colon x = u \cdot k) \wedge (\exists l\colon y = u \cdot l)) \Rightarrow (\exists m\colon z = u \cdot m))$.

Aufgabe 1.5 Geben Sie für die folgenden Aussagen über die natürlichen Zahlen eine mathematische Kurzdarstellung unter Verwendung der Junktoren „$\exists$" und „$\forall$" sowie der Zeichen „$+$", „$\leq$" und „$\cdot$".

1. Jede Zahl, die höchstens um 2 kleiner ist als ihr Doppeltes, ist kleinergleich 2.

2. Es gibt eine Zahl x, so daß jede Zahl y, die kleinergleich x ist, schon mit x identisch ist.

3. Für alle Zahlen x und y gilt: x und y sind identisch, oder x ist kleiner als y, oder y ist kleiner als x.

Aufgabe 1.6 Für Aussagen α und β sei $\alpha + \beta$ eine Abkürzung für

$$(\alpha \wedge \neg\beta) \vee (\beta \wedge \neg\alpha).$$

Berechnen Sie die Wahrheitswert-Tabelle von des Junktors „+“. Was ist die umgangssprachliche Bedeutung des Junktors „+“? Zeigen Sie, daß stets $\alpha+\beta$ und $(\alpha \vee \beta) \wedge \neg(\alpha \wedge \beta)$ äquivalente Aussagen sind. Zeigen Sie ferner, daß für beliebige Aussagen α, β und γ die folgenden Aussagen stets aussagenlogische Tautologien darstellen:

(1) $(\alpha + \beta) \Leftrightarrow (\beta + \alpha)$

(2) $((\alpha + \beta) + \gamma) \Leftrightarrow (\alpha + (\beta + \gamma))$

(3) $(\alpha \wedge (\beta + \gamma)) \Leftrightarrow ((\alpha \wedge \beta) + (\alpha \wedge \gamma))$

(4) $(\alpha + \alpha) \Leftrightarrow (\beta \wedge \neg\beta)$

(5) $\alpha + \neg\alpha$

(6) $(\alpha + (\alpha + \beta)) \Leftrightarrow \beta$

2

Mengen und Mengenoperationen

In der Mathematik ist man bestrebt, jedes zu verwendende Konzept präzise zu definieren. Gerade für die einfachsten mathematischen Entitäten, wie Zahlen, Punkte, Geraden oder Funktionen fällt es aber schwer, diese Begriffe formal auf noch einfachere Dinge zurückzuführen. Die Bedeutung der Mengenlehre in der Mathematik beruht teilweise darauf, daß man tatsächlich alle mathematischen Begriffe—also auch die oben genannten—definieren (oder vielleicht besser modellieren) kann, wenn erst einmal der Begriff der Menge als Grundlage anerkannt ist. Über diesen Modellierungsaspekt hinaus bilden mengentheoretische Prinzipien aufgrund ihrer Abstraktheit und universellen Anwendbarkeit eine unerläßliche Grundlage aller mathematischen Disziplinen.

In diesem Kapitel geht es um eine minimale Präzisierung des Mengenbegriffs, damit einhergehend werden grundlegende Operationen zwischen Mengen und die dafür geltenden Gesetze besprochen. In Abschnitt 2.4 werden wir darüberhinaus einen kleinen roten Faden beginnen, der später verschiedentlich aufgegriffen wird: es handelt sich um die Beobachtung, daß bestimmte Gesetzmäßigkeiten in ähnlicher Form bei der Behandlung verschiedener mathematischer Objekte immer wiederkehren. Hierzu stoßen wir die Nase auf die Verwandtschaft zwischen aussagenlogischen Tautologien und mengentheoretischen Identitäten, die sich oft wechselseitig ineinander übersetzen lassen. Natürlich kann dies bei einem genaueren Blick auf die Definition der einfachen mengentheoretischen Operationen im Abschnitt 2.3 nicht überraschen. Im weiteren Verlauf gehen wir kurz auf auf die sogenannte Russellsche Antinomie ein, die zeigt, daß der Mengenbegriff bei genauem Hinsehen problematischer ist, als es unsere „naive" Darstellung im ersten Abschnitt erwarten läßt. Das Kapitel schließt mit ergänzenden Bemerkungen zu sogenannten Multimengen umd zum Problem der Fundiertheit von Mengen.

2.1 Mengen und ihre Darstellung

Unsere Beschreibung des Mengenbegriffs basiert auf folgender

„Definition" (G. Cantor)[1]: *Unter einer Menge verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche Elemente von M genannt werden) zu einem Ganzen.*

Natürlich ist dies keine Definition im eigentlichen Sinne! Es bleibt unklar, was „Zusammenfassungen" oder „wohlunterschiedene Objekte unserer Anschauung oder unseres Denkens" sind. Es handelt sich also eher um eine Umschreibung, von der man sich anschaulich leiten lassen kann.

Wir schreiben

$$a \in M$$

falls a Element der Menge M ist und

$$a \notin M$$

im anderen Fall. Für die Darstellung von Mengen stehen uns zu Beginn im wesentlichen zwei Grundmethoden zur Verfügung:

1. Die Elemente einer *endlichen* Menge lassen sich oft explizit aufzählen. Man kann dann die Klammerschreibweise verwenden.

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

etwa besagt, daß M die Menge mit den Elementen 1, 2 und 3 ist. Es ist $\{M, i, s, p\}$ die Menge der Buchstaben des Wortes „Mississippi". Mit „$\emptyset$"

[1] Der deutsche Mathematiker Georg Cantor (1840-1918) begründete im letzten Viertel des 19. Jahrhunderts die moderne Mengenlehre.

oder „{ }“ wird die „leere Menge“ bezeichnet, die kein Element enthält. Wir werden bei Definition 2.2 sehen, daß es genau eine leere Menge gibt.

Mengen treten oft selbst als Elemente von komplexeren Mengen auf. Hier sind einige Beispiele:

- $\{\emptyset\}$ hat genau ein Element, nämlich die leere Menge $\emptyset$.
- $\{\{1,2,3\},\{1,2\}\}$ hat zwei Elemente, $\{1,2,3\}$ und $\{1,2\}$.
- $\{\{\emptyset\},\emptyset\}$ hat zwei Elemente, nämlich $\{\emptyset\}$ und die leere Menge.

2. Man kann die Elemente einer Menge mit Hilfe einer Eigenschaft φ beschreiben, die Objekten x zukommen kann oder nicht. Es bezeichnet

$$\{x \mid \varphi(x)\}$$

die Menge aller Objekte x, die die Eigenschaft φ besitzen. Es stellt also bei dieser Darstellungsweise die auf das Zeichen „|“ folgende Aussageform die Bedingung dar, die auf ein Element x zutreffen muß, damit es zu der beschriebenen Menge gehört. Man spricht hier manchmal von Mengenbildung durch „Abstraktion“. Der verwendete Objektname x ist natürlich irrelevant, die obige Menge ist also identisch zur Menge $\{y \mid \varphi(y)\}$. Verwandt zur Abstraktion ist die Bildung einer Menge durch „Aussonderung“ aus einer bestehenden Menge. So sondert etwa

$$M := \{x \in \mathbb{N} \mid \exists y \in \mathbb{N}\colon x = 2y\}$$

die Menge aller geraden Zahlen aus der Menge $\mathbb{N}$ der natürlichen Zahlen aus.[2] Die hier genannten Darstellungsarten für Mengen werden insbesondere verwendet, um unendliche Mengen zu beschreiben, aber auch in einem Fall wie

$$M := \{n \in \mathbb{N} \mid n \text{ ist die kleinste Primzahl oberhalb } k\}$$

kann man das einzige Element von M nicht explizit angeben, wenn k eine hinreichend große natürliche Zahl ist, sondern ist auf eine beschreibende Eigenschaft angewiesen.

Später werden wir eine weitere wichtige Darstellungsart für Mengen, nämlich die der induktiven Definition, kennenlernen. Auch die im Anschluß

[2] Der Doppelpunkt in „:=“ deutet an, daß M *definiert ist* als die rechtsstehende Menge, die sich ausführlich liest als „Menge aller x aus $\mathbb{N}$, die folgende Bedingung erfüllen: es gibt ein $y \in \mathbb{N}$ mit der Eigenschaft $x = 2y$“.

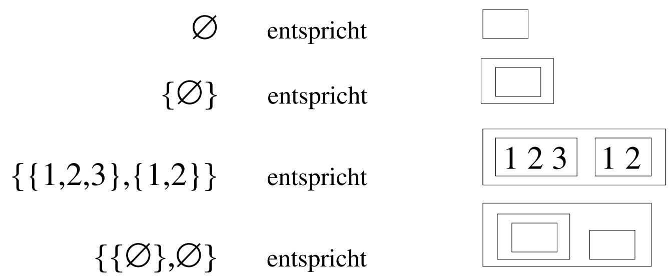


Abbildung 2.1: Schachteldarstellung von Mengen.

zu besprechenden Mengenoperationen lassen sich natürlich zur Definition von Mengen verwenden.

Eine einfache Intuition, die hilft, den Mengenbegriff besser zu verstehen, und etwa den Unterschied zwischen der leeren Menge „$\emptyset$" und der Menge $\{\emptyset\}$, die die leere Menge als einziges Element enthält, zu begreifen, beruht auf dem Bild der Menge als „Schachtel". Eine leere Schachtel ist offenkundig etwas anderes, als eine Schachtel, in der sich als einziger Inhalt eine leere Schachtel befindet. In Abbildung 2.1 sind einige der vorher erwähnten Mengen als Schachteln dargestellt. Die Elemente sind die Objekte, die man innerhalb der äußersten Schachtel *auf oberster Ebene* findet. Die dritte Menge zum Beispiel besitzt genau zwei Elemente, nämlich die Mengen $\{1, 2, 3\}$ und $\{1, 2\}$. Hingegen sind 1, 2 und 3 selbst keine Elemente der Menge.

Bemerkung 2.1 Die folgenden speziellen Mengen werden wir—wie zum Teil bereits geschehen—ohne besondere Erläuterung zu Illustrationszwecken verwenden. Mit $\mathbb{N} = \{0, 1, \ldots\}$ bezeichnen wir die Menge der *natürlichen Zahlen*. Die Zahl 0 wird also als natürliche Zahl behandelt. Gelegentlich referieren wir auf die Menge $\mathcal{Z}$ der *ganzen Zahlen*, die neben den natürlichen Zahlen $0, 1, \ldots$ auch die negativen ganzen Zahlen $-1, -2, \ldots$ als Elemente enthält, sowie auf die Menge $\mathcal{Q}$ der *rationalen Zahlen*, zu der genau diejenigen Zahlen gehören, die sich als Quotient (Bruch) zweier ganzer Zahlen mit von Null verschiedenem Nenner darstellen lassen. Mit $\mathbb{R}$ bezeichnen wir die Menge der *reellen Zahlen*. Diese Menge umfasst neben den rationalen Zahlen auch irrationale Zahlen wie die Kreiszahl Π und $\sqrt{2}$. $\mathbb{R}$ wird oft dargestellt als die Menge aller Punkte auf einer nach beiden Seiten unendlichen Zahlengerade, man redet dann auch vom eindimensionalen euklidischen Raum. Für alle erwähnten Zahlen kann man mengentheoretische Kodierungen angeben, hierzu vergleiche man die Literaturangaben am Kapitelende. Wir werden auf diese Zusammenhänge nicht eingehen, wenn nicht ausdrücklich etwas anderes gesagt wird, so stellen Zahlen für uns grundlegende „Objekte unserer Anschauung" dar, deren Natur wir an dieser Stelle nicht weiter hinterfragen. Neben den Zahlen werden zu Illustrationszwecken auch Buchstaben wie a, b, c als Elemente von Mengen verwendet.

2.2 Gleichheit und Inklusion

Zwei fundamentale Beziehungen zwischen Mengen sind Gleichheit und Inklusion.

Definition 2.2 (Extensionalitätsprinzip für Mengen) Zwei Mengen M und N sind gleich genau dann, wenn sie dieselben Elemente enthalten[3]:

$$M = N \quad :\Leftrightarrow \quad \forall x\colon (x \in M \Leftrightarrow x \in N).$$

Bemerkung 2.3 Beweistechnisch bedeutet das Extensionalitätsprinzip, daß man die Gleichheit zweier Mengen M und N dadurch verifizieren kann, daß man zeigt, daß jedes Element von M auch Element von N ist und umgekehrt. Dieses *Standardrezept* werden wir nachfolgend häufig anwenden.

Inhaltlich besagt das Extensionalitätsprinzip, daß bei dem Prozeß der „Zusammenfassung" davon abstrahiert wird, *welche* „Anschauung" bei der Auswahl der Elemente maßgebend wahr. Demgemäß kann eine Gleichheit

$$\{x \mid \varphi(x)\} = \{x \mid \psi(x)\}$$

gelten, auch wenn die Beschreibungen φ und ψ nichts Erkennbares miteinander zu tun haben. Identitäten wie

$$\{4\} = \{n \in \mathbb{N} \mid 3 < n \wedge n < 5\} = \{\sqrt{16}\}$$

resultieren. Oder, mit einem aus der Semantik bekannten Gegenstand,

$$\{\text{Morgenstern}\} = \{\text{Abendstern}\}.$$

[3]Nachfolgend deutet der Doppelpunkt in „$:\Leftrightarrow$" an, daß es sich um eine definierende Äquivalenz handelt. Die Gleichheit zwischen Mengen M und N ist also *definiert* durch die rechte Seite.

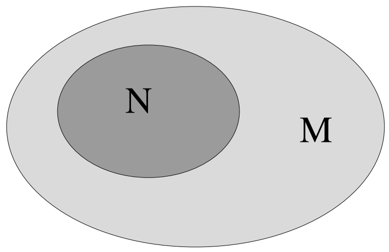


Abbildung 2.2: Teilmengenbeziehung $N \subseteq M$ im Venn-Diagramm.

Natürlich wird in einer Menge auch keine Reihenfolge ausgezeichnet, und es gilt etwa $\{1, 2, 3\} = \{3, 1, 2\}$. Weiterhin ergibt sich aus dem Extensionalitätsprinzip, daß die Mengen $\{1, 1\}$ und $\{1\}$ identisch sind. Es macht also bei Mengen keinen Sinn, nach der Vielfachheit des Vorkommens eines Elements zu fragen. Wir werden bei der expliziten Beschreibung von endlichen Mengen durch Aufzählung der Elemente eine Mehrfachnennung desselben Elements nicht grundsätzlich verbieten, sie jedoch nur in ganz wenigen begründeten Ausnahmefällen verwenden.

Definition 2.4 M und N seien Mengen. N heißt *Teilmenge* von M, im Zeichen $N \subseteq M$, genau dann, wenn jedes Element von N auch Element von M ist:

$$M \subseteq N \quad :\Leftrightarrow \quad \forall x\colon (x \in M \Rightarrow x \in N).$$

N heißt *echte* Teilmenge von M, $N \subset M$, genau dann, wenn $N \subseteq M$ und $N \neq M$.

Man schreibt oft $M \not\subseteq N$ (resp. $M \not\subset M$) um auszudrücken, daß M *keine* Teilmenge (resp. keine echte Teilmenge) von N ist. Abbildung 2.2 gibt eine Darstellung der Teilmengenbeziehung in Form eines sogenannten *Venn-Diagramms.*[4]

Beispiele 2.5 Die Menge $\{1, 4\}$ ist eine echte Teilmenge der Menge $\{1, 4, 7\}$. Die Menge $\mathbb{N}$ der natürlichen Zahlen ist eine echte Teilmenge der Menge $\mathbb{R}$ der reellen Zahlen.

[4]Beim Arbeiten mit Venn-Diagrammen sollte man sich stets vergegenwärtigen, daß alle Teilflächen auch leere Flächen repräsentieren können.

Lemma 2.6 *Für beliebige Mengen L, M und N gilt:*

1. *wenn $L \subseteq M$ und $M \subseteq N$, so $L \subseteq N$,*
2. *wenn $M \subseteq N$ und $N \subseteq M$, so $M = N$,*
3. *$\emptyset \subseteq M$ und $M \subseteq M$.*

Beweis. Wir zeigen Teilaussage 1, die Beweise der Teile 2 und 3 bleiben als Übung offen. Sei $x \in L$. Aus $L \subseteq M$ folgt $x \in M$, aus $M \subseteq N$ folgt $x \in N$. Also gilt $L \subseteq N$. ■

Während wir später den Beweis der Teilaussage 1 als „trivial“ bezeichnen und weglassen würden, wollen wir an dieser Stelle die verwendete Argumentation einmal mit der Lupe betrachten, um zu zeigen, wie bisher beschriebenen Beweisprinzipien bei einer genauen Betrachtung eingehen.

Die zu beweisende Aussage redet zunächst einmal über beliebige Mengen. Bemerkung 1.2 folgend nehmen wir zunächst „generische“ Mengen L, M und N, für die wir nun die Behauptung verifizieren. Die Behauptung stellt eine Implikation mit den zwei Prämissen $L \subseteq M$ und $M \subseteq N$ dar. Nach Abschnitt 1.5 reicht es, die Situation zu betrachten, wo beide Prämissen wahr sind. Die Prämissen enthalten die definierte Teilmengenbeziehung „$\subseteq$“. Indem wir die definierten Beziehungen jeweils durch die gemäß Definition 2.4 äquivalenten Bedingungen ersetzen, erhalten wir die wahren Aussagen

$$\begin{array}{ll} (1) & \forall x\colon (x \in L \Rightarrow x \in M) \\ (2) & \forall x\colon (x \in M \Rightarrow x \in N). \end{array}$$

Zu zeigen ist die Konklusion $L \subseteq N$. Nach Definition 2.4 müssen wir also

$$(3) \qquad \forall x\colon (x \in L \Rightarrow x \in M).$$

zeigen. Hierzu nehmen wir an, daß g ein beliebiges Objekt ist. Wir wollen zeigen, daß $g \in L$ auch $g \in N$ impliziert. Es reicht wiederum den Fall zu betrachten, wo tatsächlich $g \in L$ gilt. Durch Einsetzen von g für x in Voraussetzung (1) erhalten wir die wahre Aussage $(g \in L \Rightarrow g \in M)$. Wie in Tautologie 17 von Satz 1.1 sichtbar gemacht wird, implizieren die beiden letztgenannten Formeln nun $g \in M$. Durch Spezialisierung der Voraussetzung (2) erhalten wir die wahre Aussage $(g \in M \Rightarrow g \in N)$. Durch eine zweite Anwendung von Tautologie 17 erhalten wir $g \in N$. Damit ist die

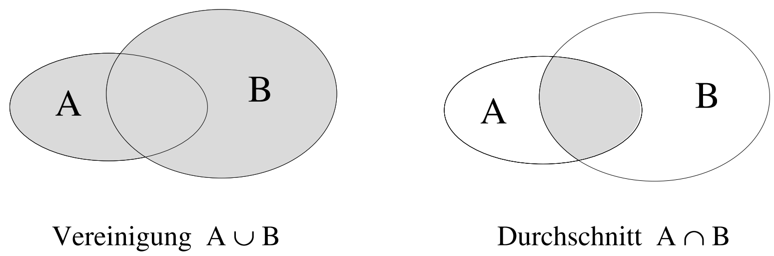


Abbildung 2.3: Vereinigung $A \cup B$ (links) und Durchschnitt $A \cap B$ (rechts).

Implikation $g \in L \Rightarrow g \in N$ gezeigt. Da g beliebig war, ist (3) gezeigt. Man sollte bemerken, daß wir ohne besondere Erwähnung auch das durch Tautologie 1 aus Aufgabe 1.2 (Kapitel 1) beschriebene Prinzip verwendet haben, nach dem alles, was sich aus den Voraussetzungen herleiten läßt, dann zum Herleiten der Konklusion weiterverwendet werden kann.

Viele der in den ersten Kapiteln auftretenden Beweise lassen sich ähnlich auf die in Kapitel 1 vorgestellten Prinzipien—oder einfache Varianten—zurückführen. Nachfolgend werden wir natürlich stets eine verkürzte Darstellungsart wählen. Es sei jedoch empfohlen, sich anfangs ab und zu zu vergegenwärtigen, auf welchen Tautologien oder Regeln die verwendeten Schlüsse beruhen.

2.3 Mengenoperationen und Gesetze

Wir geben nun einige grundlegende Operationen zwischen Mengen an. Im folgenden stehen Ausdrücke wie A, B, C, I, J und A_i (für $i \in I$) stets für Mengen.

Definition 2.7 Die Menge $A \cup B := \{x \mid x \in A \vee x \in B\}$ heißt die *Vereinigung* von A und B. Die Menge $A \cap B := \{x \mid x \in A \wedge x \in B\}$ heißt der *Durchschnitt* von A und B.

Abbildung 2.3 gibt eine Darstellung von Vereinigung und Durchschnitt in Form eines Venn-Diagramms. Die Mengen A und B heißen *disjunkt* genau dann, wenn $A \cap B = \emptyset$ gilt.

Beispiele 2.8 Die Vereinigung der Mengen $\{1,3,4\}$ und $\{1,2,4,9\}$ ist $\{1,2,3,4,9\}$. Die Vereinigung der Mengen $\{\{1,3,4\}\}$ und $\{\{1,2,4,9\}\}$ ist $\{\{1,3,4\},\{1,2,4,9\}\}$. Im „Schachtelbild" vereinigen wir jeweils den Inhalt der beiden Ausgangsschachteln in einer neuen Schachtel, wobei doppelte Vorkommen eliminiert werden. Der Durchschnitt der Mengen $\{1,3,4\}$ und $\{1,2,4,9\}$ ist $\{1,4\}$. Der Durchschnitt von $\{\{1,3,4\}\}$ und $\{\{1,2,4,9\}\}$ ist die leere Menge.

$$\boxed{1\ 3\ 4} \cup \boxed{1\ 2\ 4\ 9} = \boxed{1\ 2\ 3\ 4\ 9}$$

$$\boxed{\boxed{1\ 3\ 4}} \cup \boxed{\boxed{1\ 2\ 4\ 9}} = \boxed{\boxed{1\ 3\ 4}\ \boxed{1\ 2\ 4\ 9}}$$

$$\boxed{1\ 3\ 4} \cap \boxed{1\ 2\ 4\ 9} = \boxed{1\ 4}$$

$$\boxed{\boxed{1\ 3\ 4}} \cap \boxed{\boxed{1\ 2\ 4\ 9}} = \boxed{}$$

Lemma 2.9 *Für beliebige Mengen A, B, C gelten die folgenden Identitäten:*

1. $A \cup A = A$, $A \cap A = A$ *(Idempotenz von $\cup$ und $\cap$),*
2. $A \cup B = B \cup A$, $A \cap B = B \cap A$ *(Kommutativität von $\cup$ und $\cap$),*
3. $A \cup (B \cup C) = (A \cup B) \cup C$,
   $A \cap (B \cap C) = (A \cap B) \cap C$ *(Assoziativität von $\cup$ und $\cap$),*
4. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *(Distributivität von $\cup$ und $\cap$),*
5. $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$.

Beweis. Wir beweisen die erste Identität der Teilaussage 2, der Rest wird als Übung offengelassen. Um die Gleichheit der Mengen $A \cup B$ und $B \cup A$ zu beweisen, verwenden wir das in Bemerkung 2.3 dargestellte Standardrezept: Wir zeigen, daß jedes Element von $A \cup B$ auch Element von $B \cup A$ ist und umgekehrt. Hierzu betrachten wir die folgenden Aussagen:

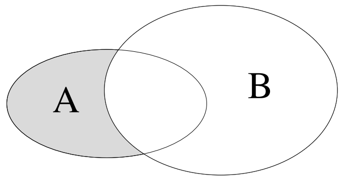
$$\begin{array}{ll} (1) & x \in A \cup B \\ (2) & (x \in A) \vee (x \in B) \end{array}$$

(3) $(x \in B) \vee (x \in A)$
(4) $x \in B \cup A$.

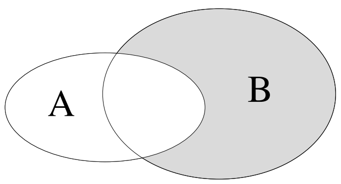
Beim Übergang zwischen den Aussagen (1) und (2)—analog bei (3) und (4)—haben wir nur die Vereinigung durch ihre Definition ersetzt haben. Die Aussagen (2) und (3) sind nach Satz 1.1 (Tautologie 4) äquivalent. Damit sind auch die Aussagen (1) und (4) äquivalent. Gemäß dem Extensionalitätsprinzip folgt nun $A \cup B = B \cup A$. ■

Es ist bemerkenswert, daß wir viele der Gesetze aus Lemma 2.9 in ähnlicher Form vom Rechnen mit Zahlen schon kennen. Die Vereinigung verhält sich ähnlich zur Addition, die Durchschnittsbildung ähnlich zur Multiplikation. Die leere Menge spielt in diesem Bild die Rolle der Zahl 0, sie ist „Neutralelement“ bezüglich der Vereinigung und „absorbierendes“ Element bezüglich des Durchschnitts.

Definition 2.10 Die Menge $A \setminus B := \{x \mid x \in A \wedge x \notin B\}$ heißt *Differenz* der Mengen A und B.



Differenz A \ B



Differenz B \ A

Man sollte sich merken, daß in einer Differenz $A \setminus B$ die Menge B keinesfalls Teilmenge von A sein muß!

Beispiele 2.11 Es gilt $\{4, 1, 2\} \setminus \{1, 5, 6\} = \{4, 2\}$ und $\{1, 5, 6\} \setminus \{4, 1, 2\} = \{5, 6\}$. Es ist $\{\{1, 3, 4\}\} \setminus \{\{1, 2, 4, 9\}\} = \{\{1, 3, 4\}\}$.

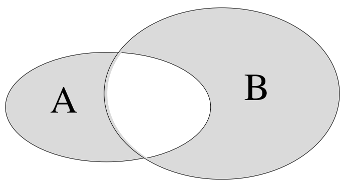
Das nachfolgende Lemma, das wir ohne Beweis geben, stellt zwei einfache Beobachtungen zur Differenz von Mengen dar.

Lemma 2.12 *Für beliebige Mengen A und B gilt stets:*

1. *$A \setminus \emptyset = A$ und $A \setminus A = \emptyset$.*

2. *$A \cap B = \emptyset \Leftrightarrow A \setminus B = A \Leftrightarrow B \setminus A = B$.*

Neben der Differenz tritt seltener auch die sogenannte *symmetrische Differenz* $A \bowtie B := (A \setminus B) \cup (B \setminus A)$ auf.



Symmetrische Differenz von A und B

Oft wird „+“ anstelle von „$\bowtie$“ als Symbol verwendet.

Beispiel 2.13 Es gilt $\{4,1,2\} \bowtie \{1,5,6\} = \{4,2,5,6\}$ und $\{4,1,5,6\} \bowtie \{1,5,6\} = \{4\}$.

Lemma 2.14 *Für beliebige Mengen A, B und C gilt stets*

1. *$A \bowtie B = (A \cup B) \setminus (A \cap B)$,*

2. *$A \bowtie B = B \bowtie A$ (Kommutativität von $\bowtie$),*

3. *$A \bowtie (B \bowtie C) = (A \bowtie B) \bowtie C$ (Assoziativität von $\bowtie$),*

4. *$A \bowtie A = \emptyset$,*

5. *$A \bowtie (A \bowtie B) = B$,*

6. *$A \cap (B \bowtie C) = (A \cap B) \bowtie A \cap C)$,*

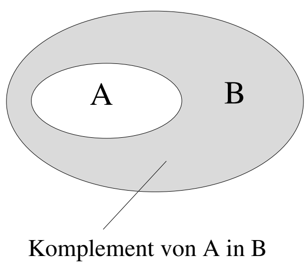
7. *$A \cap B = \emptyset \iff A \bowtie B = A \cup B$.*

Beweis. Wir beweisen Teil 1, die übrigen Teile bleiben als Übungsaufgabe (vgl. Aufgabe 2.15). Da wir die Gleichheit zweier Mengen zu beweisen haben, verwenden wir wieder die Methode aus Bemerkung 2.3. Wir betrachten die Aussagen

$$
\begin{aligned}
&(1) \quad x \in A \bowtie B, \\
&(2) \quad x \in (A \setminus B) \cup (B \setminus A), \\
&(3) \quad (x \in A \wedge \neg x \in B) \vee (x \in B \wedge \neg x \in A), \\
&(4) \quad (x \in A \vee x \in B) \wedge \neg(x \in B \wedge x \in A), \\
&(5) \quad x \in (A \cup B) \setminus (A \cap B).
\end{aligned}
$$

Nach der Definition von „$\bowtie$“ sind (1) und (2) äquivalent. Die Äquivalenz der Aussagen (2) und (3), analog die von (4) und (5), folgt aus der Definition von Vereinigung, Durchschnitt und Mengendifferenz. Die Äquivalenz von (3) und (4) ergibt sich aus Aufgabe 1.6. ■

Definition 2.15 Ist A eine Teilmenge von B, so wird die Differenz $B \setminus A$ auch *Komplement von A in B* genannt.



Für die Bezeichnung des Komplements von A in B gibt es unterschiedliche notationelle Konventionen. Manche Autoren verwenden die Schreibweise $\mathcal{C}_B A$, wo das Symbol $\mathcal{C}$ für „complement“ steht. Wenn klar ist, in welcher Menge B die Komplemente gebildet werden, dann wird allerdings zumeist einfach $\bar{A}$ oder $-A$ geschrieben. Die zuletztgenannte Schreibweise werden wir im nachfolgenden verwenden.

Lemma 2.16 (De Morgansche Regeln[5], einfache Form) *Es seien A und B beliebige Teilmengen einer Menge M. Das Symbol „$-$“ bezeichne*

[5]Benannt nach dem englischen Mathematiker Augustus De Morgan (1806-1871).

Komplementbildung in M. Dann gilt

$$\begin{aligned} -(A \cap B) &= (-A) \cup (-B), \\ -(A \cup B) &= (-A) \cap (-B). \end{aligned}$$

Beweis. Wir verwenden das Standardrezept aus Bemerkung 2.3 und betrachten die Aussagen

$$\begin{array}{ll} (1) & x \in -(A \cap B) \\ (2) & (x \in M) \wedge \neg((x \in A) \wedge (x \in B)) \\ (3) & (x \in M) \wedge (\neg(x \in A) \vee \neg(x \in B)) \\ (4) & ((x \in M) \wedge \neg(x \in A)) \vee ((x \in M) \wedge \neg(x \in B)) \\ (5) & x \in (-A) \cup (-B). \end{array}$$

Die Aussagen (1) und (2) sowie (4) und (5) sind gemäß der Definition des Komplements äquivalent. Nach Tautologie 10 aus Satz 1.1 sind die Aussagen (2) und (3) äquivalent, nach Tautologie 7 desselben Satzes sind auch (3) und (4) äquivalent. Damit sind (1) und (5) äquivalent, es enthalten die beiden Mengen $-(A \cap B)$ und $(-A) \cup (-B)$ dieselben Elemente und sind nach dem Extensionalitätsprinzip 2.2 gleich. ■

Lemma 2.17 *Es seien A, B beliebige Teilmengen einer Menge M. Das Symbol „$-$" bezeichne Komplementbildung in M. Dann gilt*

1. $-(-A) = A$,
2. $A \subseteq B$ *genau dann, wenn* $-B \subseteq -A$,
3. $-\emptyset = M$ *und* $-M = \emptyset$,
4. $A \subseteq B$ *genau dann, wenn* $-A \cup B = M$.

Beweis. Wir zeigen Teil 2, der restliche Beweis bleibt dem Leser als Übung überlassen. Es gelte $A \subseteq B$. Dann folgt aus $x \in A$ stets $x \in B$, umgekehrt daher aus $x \notin B$ stets $x \notin A$ (hierzu vergleiche man Tautologie 14 aus Satz 1.1). Daher folgt aus $x \in -B$ beziehungsweise $x \in M \wedge x \notin B$, auch $x \in M \wedge x \notin A$ beziehungsweise $x \in -A$. Dies zeigt $-B \subseteq -A$.

Es gelte nun umgekehrt $-B \subseteq -A$. Für jedes $x \in M$ folgt dann aus $x \notin B$ stets $x \notin A$, umgekehrt also aus $x \in A$ stets $x \in B$. Da nach Voraussetzung aber jedes Element von A in M liegt, folgt $A \subseteq B$. ■

2.4 Aussagenlogische Tautologien und mengentheoretische Identitäten

Die *mengentheoretischen Identitäten* des vorangegangenen Teilkapitels haben wir jeweils dadurch beweisen können, daß wir sie auf entsprechende *aussagenlogische Tautologien* zurückgeführt haben. Dies ist nicht verwunderlich, da wir die mengentheoretischen Operationen „$\cup$“, „$\cap$“ und „$-$“ mit Hilfe der aussagenlogische Junktoren „$\vee$“, „$\wedge$“ und „$\neg$“ definiert hatten. Dies führt zu einer allgemeineren Beobachtung:

Satz 2.18 *Es seien A, B und C beliebige Teilmengen der Menge M. Dann sind alle mengentheoretischen Aussagen wahr, die wir aus den Tautologien 1 - 11 in Satz 1.1 dadurch erhalten, daß wir jedes Vorkommen von α, β und γ durch A respektive B bzw. C ersetzen, die aussagenlogischen Junktoren „$\wedge$“ durch „$\cap$“, „$\vee$“ durch „$\cup$“, „$\neg$“ durch „$-$“ (wobei „$-$“ als Komplementbildung in M zu interpretieren ist) und „$\Leftrightarrow$“ durch „$=$“.*

Beweis. Durch Inspektion. ■

Ganz allgemein läßt sich *jede* aussagenlogische Tautologie der Form $\alpha \Leftrightarrow \beta$, wo die Teilaussagen α und β nur die Junktoren „$\vee$“, „$\wedge$“ und „$\neg$“ enthalten, sofort auf dieselbe Weise in eine mengentheoretische Identität übersetzen. Wir belassen es jedoch bei der Feststellung und verzichten auf einen Beweis.

Auch einige aussagenlogischen Tautologien aus Satz 1.1 einer anderen Form lassen sich in allgemeingültige Aussagen über Mengen übersetzen. Aus den Tautologien 12 und 13 von Satz 1.1 etwa erhält man folgende Gesetzmäßigkeit:

Lemma 2.19 *Für beliebige Mengen A und B gilt:*

$$\begin{aligned}(A \subseteq B) \quad &\Leftrightarrow \quad (A \cup B = B),\\ (A \subseteq B) \quad &\Leftrightarrow \quad (A \cap B = A).\end{aligned}$$

Es stellt sich natürlich an dieser Stelle die Frage, ob man auch umgekehrt aus jeder allgemeingültigen mengentheoretischen Identität, die nur Mengensymbole wie $A, B, C \ldots$ und die Operatoren „$\cap$“, „$\cup$“ und „$-$“ verwendet, durch Rückübersetzung stets eine aussagenlogische Tautologie erhält. Die Antwort ist positiv, aber dies werden wir erst zu einem späteren Zeitpunkt zeigen können.

2.5 Operationen für Mengenfamilien

Eine Menge, deren sämtliche Elemente selbst wiederum Mengen sind, werden wir als *Mengenfamilie* bezeichnen. In diesem Abschnitt gehen wir auf Operationen ein, wo die Eingabe oder das Resultat eine Mengenfamilie darstellt. Häufig treten Mengenfamilien in der Form auf, daß die Elemente in geeigneter Weise mit den Elementen eines Indexbereichs I durchnummeriert sind. Als Indexbereich I werden in der Praxis meist Anfangsabschnitte von $\mathbb{N}$ oder $\mathbb{N}$ selbst, aber auch andere Mengen wie $\{a, b, c\}$ verwendet. Die Mengenfamilie kann dann in der Form $\{A_i \mid i \in I\}$ dargestellt werden. Es wird dabei nicht immer vorausgesetzt, daß je zwei Mengen mit unterschiedlichem Index verschieden sind.

Beispiel 2.20 Es sei $A_1 := \{4\}$, $A_2 := \{2, 5\}$ und $A_3 := \{8, 9\}$. Dann ist $\{\{4\}, \{2, 5\}, \{8, 9\}\}$ eine Mengenfamilie mit den Elementen A_1, A_2 und A_3. Sie kann in der Form $\{A_i \mid i \in \{1, 2, 3\}\}$ dargestellt werden.

Beispiel 2.21 Für jedes $n \in \mathbb{N}$ sei $A_n := \{k \in \mathbb{N} \mid k < n\}$. Dann ist $\{A_n \mid n \in \mathbb{N}\}$ eine Mengenfamilie. Sie enthält alle Anfangsabschnitte $\emptyset, \{0\}, \{0, 1\}, \{0, 1, 2\}, \ldots$ als Elemente.

Die ersten der nun zu besprechenden Mengenoperationen verallgemeinern die Vereinigung und Durchschnittsbildung von je zwei Mengen in der Weise, daß nun alle Mengen einer vorgegebenen Mengenfamilie vereinigt beziehungsweise geschnitten werden. Zur Definition dieser Operationen werden die Quantoren „$\exists$“ und „$\forall$“ verwendet. Dementsprechend werden sich nachfolgende Umformungs- und Rechenregeln nicht mehr ausschließlich auf *aussagenlogische* Tautologien zurückführen lassen.

Definition 2.22 Sei A eine Mengenfamilie. Die Menge

$$\bigcup A := \{x \mid \exists B \in A \colon x \in B\}$$

heißt die *Vereinigung* von A. Hat A die Form $A = \{A_i \mid i \in I\}$, so wird $\bigcup A$ auch in der Form $\bigcup_{i \in I} A_i$ notiert. Es sei nun $A \neq \emptyset$. Die Menge

$$\bigcap A := \{x \mid \forall B \in A\colon x \in A\}$$

heißt der *Durchschnitt* von A. Hat A die Form $A = \{A_i \mid i \in I\}$, so wird $\bigcap A$ auch in der Form $\bigcap_{i \in I} A_i$ notiert.

Es enthält also $\bigcup A$ (resp. $\bigcap A$) genau diejenigen Elemente x, die in irgendeinem Element (resp. in allen Elementen) von A als Element auftreten. Wenn wir im nachfolgenden eine Menge der Form $\bigcap A$ verwenden, so sei stets vorausgesetzt, daß A nichtleer ist. In $\bigcup A$ liegen im allgemeinen keine Elemente von A, sondern Elemente *von Elementen* aus A. Entsprechendes gilt für den Durchschnitt. Hier sind einige Beispiele:

- Sei $A = \{A_1, A_2\}$ wo $A_1 = \{1,3,4\}$, $A_2 = \{1,2,4,9\}$. Dann ist $\bigcup A = \{1,2,3,4,9\}$ und $\bigcap A = \{1,4\}$. Im „Schachtelbild" erhalten wir also die Vereinigung einer Mengenfamilie A dadurch, daß wir die Inhalte aller obersten Schachteln in A in einer gemeinsamen Schachtel zusammenführen. Doppelte Vorkommen werden wie immer eliminiert.

$$\bigcup \boxed{\boxed{1\ 3\ 4} \qquad \boxed{1\ 2\ 4\ 9}} = \boxed{1\ 2\ 3\ 4\ 9}$$

- $\bigcup\{\{\{\emptyset\}\}\} = \{\{\emptyset\}\}$.
- $\bigcup\{\emptyset, \{\emptyset\}\} = \{\emptyset\}$.
- Für $n \in \mathbb{N}$ sei $A_n = \{0, 1, n\}$. Dann ist $\bigcup_{n \in \mathbb{N}} A_n = \mathbb{N}$ und $\bigcap_{n \in \mathbb{N}} A_n = \{0, 1\}$.

Die Beispiele machen deutlich, daß im Fall einer *endlichen* Mengenfamilie $A = \{A_1, \ldots, A_n\}$ die Vereinigung $\bigcup A$ identisch zur Vereinigung der Elemente $A_1 \cup \ldots \cup A_n$ ist, dasselbe gilt für den Durchschnitt. Wir hätten für diesen Fall keinen neuen Begriff einzuführen brauchen. Die Bedeutung von Definition 2.22 beruht vor allem darauf, daß diese auch eine Vereinigung beziehungsweise Durchschnittsbildung über *unendlich* viele Mengen möglich macht und erklärt.

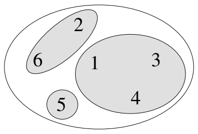
Wir führen nun noch die Begriffe der „Partition" und der „Potenzmenge" einer gegebenen Menge ein. Partitionierung und Potenzmengenbildung führen ausgehend von einer beliebigen Menge jeweils zu einer eng verwandten Mengenfamilie.

Definition 2.23 Sei M eine Menge. Die Mengenfamilie P heißt eine *Partition* oder *Zerlegung* von M genau dann, wenn gilt:

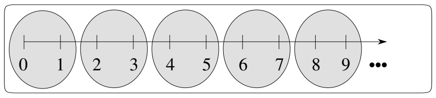
1. jedes Element von P ist eine *nicht-leere* Teilmenge von M,
2. $\bigcup P = M$,
3. je zwei verschiedene Elemente von P sind disjunkt.

Für jede nichtleere Menge M sind beispielweise $\{M\}$ und $\{\{m\} \mid m \in M\}$ stets Partitionen von M. Diese „Extremfälle“ sind jedoch meist von geringem Interesse. Häufiger werden Partitionen betrachtet, die zu mehreren Partitionsklassen mit mehreren Elementen führen.

Beispiel 2.24 Es ist $\{\{1,3,4\},\{2,6\},\{5\}\}$ eine Partition von $\{1,2,3,4,5,6\}$.



Beispiel 2.25 Für jedes $n \in \mathbb{N}$ bezeichne A_n die Menge $\{n, n+1\}$. Dann ist $Q := \{A_n \mid n \in \mathbb{N}\}$ keine Zerlegung von $\mathbb{N}$. Zwar gelten die Bedingungen 1 und 2 aus Definition 2.23, aber Bedingung 3 ist verletzt. Hingegen ist $P := \{A_{2n} \mid n \in \mathbb{N}\}$ eine Zerlegung von $\mathbb{N}$.



Definition 2.26 Die Menge $\mathcal{P}(A) := \{M \mid M \subseteq A\}$ heißt die *Potenzmenge* von A.

In der Potenzmenge einer Menge A liegen also alle *Teilmengen* von A. Beim letzten Beispiel der nachfolgenden Liste sieht man, daß nicht auszuschließen ist, daß $\mathcal{P}(A)$ auch Elemente von A enthält — wenn A Elemente enthält, die gleichzeitig Teilmenge sind.

- $\mathcal{P}(\emptyset) = \{\emptyset\}$,
- $\mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}$,
- $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$,
- $\mathcal{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$,
- $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$.

Für die Potenzmenge einer Menge A wird oft auch das Symbol 2^A verwendet. Eine Erklärung dieser Schreibweise werden wir später (im Abschnitt 3.6.2) geben. Ein erster Hinweis ergibt sich aus dem folgendem Lemma, für dessen Beweis wir das im Anschluß näher erläuterte Beweisprinzip der vollständigen Induktion verwenden.

Lemma 2.27 *Es sei A eine endliche Menge mit n Elementen. Dann hat $\mathcal{P}(A)$ genau 2^n Elemente.*

Beweis. Induktionsanfang: $n = 0$. In diesem Fall gilt $A = \emptyset$ und es folgt $\mathcal{P}(A) = \{\emptyset\}$. Wie behauptet hat also $\mathcal{P}(A)$ genau $1 = 2^0 = 2^n$ Elemente.

Induktionshypothese: Die Behauptung sei richtig für jede Menge A mit k Elementen, wo $k \geq 0$.

Induktionsschritt: Es sei nun A eine Menge mit $k+1 \geq 1$ Elementen. Zu zeigen ist, daß $\mathcal{P}(A)$ genau 2^{k+1} Elemente hat. Da $k+1 \geq 1$ können wir ein Element a aus A herausnehmen, die Menge $A_0 := A - \{a\}$ hat dann k Elemente. Wir zerlegen nun $\mathcal{P}(A)$ in die folgenden zwei Teile:

$$\begin{aligned} \mathcal{P}_1 &= \{M \subseteq A \mid a \notin M\} \\ \mathcal{P}_2 &= \{M \subseteq A \mid a \in M\}. \end{aligned}$$

Offenkundig sind diese beiden Mengen disjunkt und es gilt $\mathcal{P}(A) = \mathcal{P}_1 \cup \mathcal{P}_2$, daher ist die Anzahl der Elemente von $\mathcal{P}(A)$ genau die Summe der Anzahl der Elemente von $\mathcal{P}_1$ respektive $\mathcal{P}_2$. Man sieht leicht, daß $\mathcal{P}_1$ und $\mathcal{P}_2$ dieselbe

Anzahl von Elementen haben: in der Tat entspricht jedem Element M von $\mathcal{P}_1$ genau ein Element der Form $M \cup \{a\}$ von $\mathcal{P}_2$ und umgekehrt (eine genauere Argumentation erfolgt in Beispiel 3.52). Nun gilt aber $\mathcal{P}_1 = \mathcal{P}(A_0)$, nach Induktionsvoraussetzung hat also $\mathcal{P}_1$ genau 2^k Elemente. Daraus ergibt sich, daß $\mathcal{P}(A)$ genau $2^k + 2^k = 2 \times 2^k = 2^{k+1}$ Elemente hat. ■

Das Beweisprinzip der vollständigen Induktion. Es ist sicher angebracht zu erklären, warum die obige Argumentation einen Beweis darstellt. Wenn wir von der konkreten Aussage etwas abstrahieren, dann besagt das Lemma, daß eine bestimmte Eigenschaft φ für jede beliebige Zahl $n \in \mathbb{N}$ gilt. Um dies zu beweisen, reicht es, zwei Aussagen zu verifizieren:

1. Die Eigenschaft φ trifft auf die Zahl 0 zu.
2. Falls die Eigenschaft φ auf eine beliebige natürliche Zahl $k \geq 0$ zutrifft, so auch auf $k + 1$.

Warum können wir aus 1 und 2 schließen, daß alle natürlichen Zahlen die Eigenschaft φ haben? Wir wissen, daß wir jede natürliche Zahl n ausgehend von der Null durch eine endliche Folge von Nachfolgerschritten

$$0 \to 1, 1 \to 2, \ldots, n-1 \to n$$

erreichen können. Indem wir die Regel 2 auf $k = 0$ anwenden, können wir mit Hilfe der wahren Aussage 1 zunächst schließen, daß die Eigenschaft φ auch auf die Zahl 1 zutrifft. Daraus folgt durch eine erneute Anwendung der Schlußregel 2, daß φ auch auf die Zahl 2 zutrifft. In dieser Weise fortfahrend erreichen wir schließlich nach endlich vielen Schritten die Zahl n und sehen somit, daß φ auch auf die Zahl n zutrifft. Dieses explizite „Hochhangeln“ zu immer größeren Zahlen k bis hin zu n ist jedoch offenkundig völlig überflüssig, da wir in Form von Aussage 2 ja ein für alle mal eingesehen haben, daß es prinzipiell immer klappt. Daher ist die Eigenschaft φ tatsächlich für jedes $n \in \mathbb{N}$ bewiesen.

Beim eigentlichen Beweis der Schlußregel 2 verfahren wir wieder nach bekanntem Rezept. Wir nehmen an, daß die Prämisse gilt, wobei wir dies explizit als *Induktionshypothese* formulieren, und zeigen im *Induktionsschritt*, daß unter dieser Voraussetzung auch die Konklusion gilt. Manchmal wird die Idee des *Induktionsbeweises* auch als „Leiterprinzip“ veranschaulicht: wenn man die unterste Stufe einer Leiter erreichen kann, und der Sprossenabstand so eingerichtet ist, daß man von einer beliebigen Sprosse stets die nächsthöhere erreichen kann, so ist klar, daß man jede Sprosse erreichen kann. Entscheidend hierbei ist freilich die Voraussetzung, daß man vom Boden jede einzelne Sprosse in endlich vielen Schritten erreichen kann.

Ehrlichkeitshalber sollte man hinzufügen, daß wir auch mit der obigen Argumentation letztlich nicht „bewiesen“ haben, daß das Prinzip der vollständigen Induktion ein gültiges Beweisprinzip ist. Bei einer genaueren Beschreibung der natürlichen Zahlen wird üblicherweise die Gültigkeit des Prinzips der vollständigen Induktion als ein *Teil der definitorischen Beschreibung* dieses Zahlbereichs eingeführt.

2.6 Die Russellsche Antinomie

An dieser Stelle wollen wir noch kurz zeigen, daß der Mengenbegriff doch nicht völlig naiv gehandhabt werden kann, will man sich nicht in unauflösbare Widersprüche verwickeln. Das nachfolgende Problem ist unter dem Namen *Russellsche*[6] *Antinomie* bekannt.

Wenn man annimmt, daß man wirklich in beliebiger Weise wohlunterschiedene Objekte zu einer Menge zusammenfassen kann, so ist auch M eine Menge, wo

$$M := \{A \mid A \text{ ist Menge}, A \notin A\}.$$

Dies ist die Menge aller Mengen, die sich nicht selbst als Element enthalten. Nun führt jedoch die Annahme $M \in M$ sofort zum Widerspruch $M \notin M$, und umgekehrt die Annahme $M \notin M$ sofort zum Widerspruch $M \in M$. Da eine der Annahmen richtig sein muß, wenn M eine Menge ist, kann M nicht wirklich eine Menge sein.

Das Russellsche Paradoxon machte erstmals offenkundig, daß man nicht davon ausgehen kann, daß jede Eigenschaft genau die Elemente einer Menge beschreibt, will man sich nicht in Widersprüche verwickeln. Da es erklärtes Ziel der Mengenlehre war, eine formale Grundlage für die Mathematik zu bilden, wurden zu Beginn des 20. Jahrhunderts Wege der Einschränkungen bei der Bildung von Mengen entwickelt, die Antinomien wie die oben genannte vermeiden. Im wesentlichen haben sich zwei Methoden herauskristallisiert, zum einen die sogenannte Typentheorie, zum andern die axiomatische Be-

[6]Der Mathematiker und Philosoph Bertrand A. W. Russell (1872-1970) stieß 1901 auf dieses Problem.

schreibung der Konstruktion möglicher Mengen (siehe etwa [Lev79]). In diesen Formalisierungen wird dann nicht jede Sammlung von Objekten, die wir sprachlich beschreiben können, automatisch eine Menge, und man unterscheidet zwischen Mengen und sogenannten echten „Klassen". Wir wollen hier jedoch die Diskussion abbrechen.

2.7 Ergänzungen

In diesem Abschnitt gehen wir auf drei Themen ein, die man beim ersten Lesen überschlagen kann. Im ersten Teil stellen wir einige Umformungsregeln für Mengenausdrücke zusammen, in denen Mengenfamilien vorkommen. Im zweiten Teil geben wir eine informelle Definition sogenannter Multimengen, die in der Informatik häufig als abstrakter Datentyp verwendet werden. Im dritten Abschnitt gehen wir kurz auf das Problem der Fundiertheit von Mengen ein.

2.7.1 Gesetze für Operationen zwischen Mengenfamilien

Die nachfolgenden mengentheoretischen Identitäten für Operationen zwischen Mengenfamilien sind nur der Vollständigkeit halber angeführt, um sie bei Bedarf nachschlagen zu können. Wie auch bislang setzen wir nachfolgend stets voraus, daß für alle Ausdrücke der Form $\bigcap_{i\in I} B_i$ stets I nichtleer ist.

Lemma 2.28 *Für beliebige Mengen gelten die folgenden Identitäten:*

1. $A \cap \bigcup_{i\in I} B_i = \bigcup_{i\in I}(A \cap B_i)$,
2. $A \cup \bigcap_{i\in I} B_i = \bigcap_{i\in I}(A \cup B_i)$,
3. $(\bigcup_{i\in I} A_i) \cup (\bigcup_{j\in J} B_j) = \bigcup_{i\in I}(\bigcup_{j\in J}(A_i \cup B_j)) = \bigcup_{j\in J}(\bigcup_{i\in I}(A_i \cup B_j))$,
4. $(\bigcap_{i\in I} A_i) \cap (\bigcap_{j\in J} B_j) = \bigcap_{i\in I}(\bigcap_{j\in J}(A_i \cap B_j)) = \bigcap_{j\in J}(\bigcap_{i\in I}(A_i \cap B_j))$.

Beweis. Wir zeigen Teil 1, die weiteren Identitäten folgen ähnlich.

$$x \in A \cap \bigcup_{i\in I} B_i \quad\Leftrightarrow\quad x \in A \wedge (\exists i \in I\colon x \in B_i)$$

$$\begin{aligned}
&\Leftrightarrow \quad \exists i \in I\colon (x \in A \wedge x \in B_i)\\
&\Leftrightarrow \quad \exists i \in I\colon x \in A \cap B_i\\
&\Leftrightarrow \quad x \in \bigcup_{i \in I} (A \cap B_i). \blacksquare
\end{aligned}$$

Lemma 2.29 *Für beliebige Mengen gelten die folgenden Identitäten:*

1. $(\bigcup_{i\in I} A_i) \cap (\bigcup_{j\in J} B_j) = \bigcup_{i\in I}(\bigcup_{j\in J}(A_i \cap B_j))$,
2. $(\bigcup_{i\in I} A_i) \cap (\bigcap_{j\in J} B_j) = \bigcup_{i\in I}(\bigcap_{j\in J}(A_i \cap B_j)) = \bigcap_{j\in J}(\bigcup_{i\in I}(A_i \cap B_j))$.

Beweis. Wir zeigen Teil 2, Teil 1 folgt ähnlich. Setze $A^* = \bigcup_{i\in I} A_i, B^* = \bigcap_{j\in J} B_j$. Mit Lemma 2.28, Teil 1, folgt

$$\begin{aligned}
A^* \cap B^* &= (\bigcup_{i\in I} A_i) \cap B^*\\
&= \bigcup_{i\in I} (A_i \cap B^*)\\
&= \bigcup_{i\in I} (\bigcap_{j\in J} (A_i \cap B_j)).\\
A^* \cap B^* &= A^* \cap (\bigcap_{j\in J} B_j)\\
&= \bigcap_{j\in J} (A^* \cap B_j)\\
&= \bigcap_{j\in J} (\bigcup_{i\in I} (A_i \cap B_j)). \blacksquare
\end{aligned}$$

Lemma 2.30 (De Morgansche Regeln, allgemeine Form) *Es sei $A_i \subseteq B$ für alle $i \in I$. Es bezeichne „$-$“ die Komplementbildung in B. Dann gilt*

$$\begin{aligned}
-(\bigcap_{i\in I} A_i) &= \bigcup_{i\in I} -A_i,\\
-(\bigcup_{i\in I} A_i) &= \bigcap_{i\in I} -A_i.
\end{aligned}$$

Beweis. Für den Beweis der ersten Identität betrachten wir folgende Äquivalenzen:

$$x \in -(\bigcap_{i\in I} A_i) \quad \Leftrightarrow \quad x \in B \wedge x \notin \bigcap_{i\in I} A_i$$

$$\begin{aligned}
&\Leftrightarrow \quad x \in B \wedge (\exists j \in I\colon x \notin A_j) \\
&\Leftrightarrow \quad \exists j \in I\colon (x \in B \wedge x \notin A_j) \\
&\Leftrightarrow \quad \exists j \in I\colon x \in -A_j \\
&\Leftrightarrow \quad x \in \bigcup_{i \in I} -A_i.
\end{aligned}$$

Der Beweis der zweiten Identität ist ähnlich. ■

2.7.2 Multimengen

Während wir bei einer Menge nur danach fragen, ob ein gegebenes Elemente dazugehört oder nicht, sind in manchen Bereichen der Informatik Verallgemeinerungen dieses Konzepts interessant, wo jedes Element *endlich* oft auftreten kann, und wo wir Mengen nach der Vielfachheit des Auftretens eines gegebenen Elements unterscheiden können. Diese Mengen sind unter dem Begriff *Multimengen* bekannt. Eine formale Definition von Multimengen auf der Grundlage des normalen Mengenbegriffs wird in Kapitel 3 erfolgen (vgl. Beispiele 3.34 Nr. 13), setzt jedoch den Begriff der Funktion voraus, den wir erst im nächsten Kapitel einführen werden. Wie bei der Cantorschen „Definition" einer Menge sollte jedoch auch hier klar sein, was wir unter einer Multimenge verstehen.

Wir verwenden für die Darstellung von Multimengen die Klammern „[" und „]". Beispielsweise ist $[a, a, b, e, e, g, m, m, O, r, r, u]$ die Multimenge der Buchstaben des Wortes „Oberammergau", und $[2, 2, 3, 3, 5, 11]$ ist die Multimenge der Primfaktoren der Zahl 1980.

Definition 2.31 Zwei Multimengen sind gleich genau dann, wenn sie dieselben Elemente in der selben Vielfachheit enthalten. Eine Multimenge A ist *Teilmultimenge* einer Multimenge B, $A \subseteq_m B$, wenn jedes Element von A auch in B mit zumindest gleich großer Vielfachheit auftritt.

Definition 2.32 Die *Vereinigung* zweier Multimengen A und B ist diejenige Multimenge $A \cup_m B$, die genau diejenigen Elemente umfaßt, die in A oder B vorkommen, und wo ein Element x mit der Vielfachheit $\max\{n, m\}$ in $A \cup_m B$ auftritt, wo $n \geq 0$ ($m \geq 0$) die Vielfachheit von x in A (resp. B) angibt. Der *Durchschnitt* $A \cap_m B$ zweier Multimengen läßt sich analog definieren, wenn wir $\max\{n, m\}$ durch $\min\{n, m\}$ ersetzen. Bei der *Summe*

$A + B$ zweier Multimengen ersetzen wir $\max\{n, m\}$ durch $n + m$. Weiter erhalten wir $A \setminus_m B$ dadurch, daß wir aus A die Elemente von B entsprechend ihrer Vielfachheit in B herausnehmen. Sobald wir hierbei die Vielfachheit 0 erreichen, tritt das Element natürlich in der Differenz nicht mehr auf.

Es gilt also zum Beispiel

$$
\begin{aligned}
&[1, 1, 2, 2, 2, 5] \cup_m [1, 3, 3, 3, 5, 5] = [1, 1, 2, 2, 2, 3, 3, 3, 5, 5],\\
&[1, 1, 2, 2, 2, 5] \cap_m [1, 3, 3, 3, 5, 5] = [1, 5],\\
&[1, 1, 2, 2, 2, 5] + [1, 3, 3, 3, 5, 5] = [1, 1, 1, 2, 2, 2, 3, 3, 3, 5, 5, 5].
\end{aligned}
$$

Ist A (B) die Multimenge der Primfaktoren der Zahl m (n), so ergibt $A \cup_m B$ (resp. $A \cap_m B$) gerade die Multimenge der Primfaktoren des kleinsten gemeinsamen Vielfachen (des größten gemeinsamen Teilers) von m und n, und $A + B$ ergibt die Multimenge der Primfaktoren von $m \cdot n$.

2.7.3 Fundierte und nichtfundierte Mengen

Bei einer genaueren Formalisierung des Mengenkonzepts wird in aller Regel das folgende *Prinzip der Wohlfundiertheit* gefordert:

> *Jede nichtleere Menge M enthält ein Element A, so daß $M \cap A = \emptyset$ gilt.*

Falls $A \in M$ und $M \cap A = \emptyset$ gilt, so wird A auch ein *$\in$-minimales Element* von M genannt. Das Prinzip der Wohlfundiertheit fordert also, daß jede nichtleere Menge ein $\in$-minimales Element enthält. Das Prinzip hat zur Folge, daß es keine unendlich absteigenden Ketten von Mengen gibt, die in einer Elementschaftsbeziehung

$$\cdots \in M_{i+1} \in M_i \in \cdots \in M_2 \in M_1 \in M_0$$

stehen. Allerdings wurde in jüngerer Zeit von verschiedener Seite bemerkt, daß es für die Modellierung gewisser Probleme in der Informatik und Linguistik angebrachter ist, auf das oben genannte Prinzip zu verzichten und auch nichtfundierte Mengen zuzulassen. Nichtfundierte Menge können hilfreich sein für die Formalisierung zyklischer Datenstrukturen, oder auch in der Linguistik, etwa für die logische Modellierung von Situationen, bei denen Phänomene der Selbstreflexion relevant sind. Die im nachfolgenden zu besprechenden Eigenschaften von Mengen und den daraus abgeleiteten Konzepten bleiben von der Frage der Fundiertheit der Mengen allerdings unbetroffen, wir werden darum nicht näher auf nichtfundierte Mengen eingehen. Dem interessierten Leser bleibt der Verweis auf die Literaturangaben am Kapitelende.

2.8 Aufgaben zu Kapitel 2

Aufgaben zu Teilkapitel 2.1

Aufgabe 2.1 Geben Sie eine kompakte Darstellung

1. der Menge aller natürlichen Zahlen, die sich als Summe zweier Quadratzahlen darstellen lassen,
2. der Menge aller natürlichen Zahlen, deren sämtliche Teiler kleiner als 1000 sind.

Eine natürliche Zahl k heißt *Teiler* der natürlichen Zahl n genau dann, wenn es eine Zahl m gibt, so daß $n = k \cdot m$.

Aufgabe 2.2 Für $n \in \mathbb{N}$ sei $A_n := \{n\}$. Wie sieht die Menge $\{A_n \mid n \in \mathbb{N}\}$ aus?

Aufgabe 2.3 Geben Sie eine Schachteldarstellung der folgenden Mengen.

(a) $\{1, 2, \{1, \{1, 2\}\}, \{2, \{1, 2\}\}\}$ (b) $\{\{1\}, \{\{1\}\}, \{\{\{1\}\}\}\}$
(c) $\{\{1\}, \{\{1\}, \{\{1\}, 1\}\}\}$ (d) $\{\emptyset, \{\{\{\emptyset, \{\{\{\emptyset, \{\{\emptyset\}\}\}\}\}\}\}\}\}$.

Geben Sie jeweils die Zahl der Elemente an.

Aufgaben zu Teilkapitel 2.2

Aufgabe 2.4 Welche der nachfolgenden Aussagen sind richtig, welche sind falsch, was stellt aus anderen Gründen gar keine wohlgeformte Aussage dar? Zahlen sollen hierbei *nicht* als Mengen aufgefaßt werden.

(a)	$1 \subset \{1\}$	(b)	$1 \in \{5, 1, 3\}$
(c)	$4 \notin \{5, 1, 3\}$	(d)	$\{2, 4, 6\} \subset \mathbb{N}$
(e)	$\{\emptyset\} \subset \mathbb{N}$	(f)	$\{\emptyset\} \subseteq \{\emptyset, \{\emptyset\}\}$
(g)	$\{\{\emptyset\}\} \subset \{\emptyset, \{\{\emptyset\}\}\}$	(h)	$\mathbb{N} \in \{\mathbb{N}\}$
(i)	$\mathbb{N} = \mathbb{N} \cup \{0, 1, 2\}$	(j)	$\mathbb{N} \subset \mathbb{N} \cup \{\mathbb{N}\}$
(k)	$5 \in \{5, 1, 3\} \setminus \{5\}$	(l)	$\{5, 3\} \in \{5, 1, 3\} \setminus \{1\}$.

Aufgabe 2.5 Es sei $a = b$. Wie kann man die Menge

$$\{\{\{a\}, \{\{a\}, \{b\}\}\}, \{\{\{b, a\}\}, \{\{b\}, \{a\}\}, \{b\}\}\}$$

möglichst einfach darstellen?

Aufgabe 2.6 Beweisen Sie Lemma 2.6, Teil 2.

Aufgaben zu Teilkapitel 2.3

Aufgabe 2.7 Geben Sie alle Mengen B an, so daß $\{1, 2, 3\} \cup B = \{1, 2, 3, 4, 5, 6\}$.

Aufgabe 2.8 Für welche der folgenden Gleichungen gibt es keine, genau eine, endlich viele, für welche gibt es unendlich viele Lösungen B? Wie lassen sich jeweils die möglichen Mengen B charakterisieren?

$$\begin{aligned}
\{1, 2, 3\} \cap B &= \{1, 2, 3, 4\}, \\
\{1, 2, 3\} \cap B &= \{1, 2\}, \\
\{1, 2, 3, 4, 5\} \setminus B &= \{1, 2\}, \\
B \setminus \{1, 2, 3\} &= \{4, 5\}, \\
B \bowtie \{1, 2, 3\} &= \{1\}.
\end{aligned}$$

Aufgabe 2.9 Gegeben sei eine der folgenden Mengengleichungen.

$$\begin{aligned} A \cap B &= C, \\ A \cup B &= C, \\ A \setminus B &= C, \\ A \bowtie B &= C. \end{aligned}$$

Zwei Spieler 1 und 2 fügen nun bei $A := B := C := \emptyset$ beginnend abwechselnd Elemente zu einer der drei Mengen. Es dürfen auch Elemente zu einer Menge hinzugefügt werden, die bereits in der Menge sind. Spieler 1 beginnt, es folgt Spieler 2, Spieler 1, Spieler 2. Nach diesen vier Zügen hat Spieler 2 gewonnen, falls die Mengengleichung erfüllt ist, Spieler 2 andernfalls. Für welche der Gleichungen gewinnt Spieler 1, für welche Spieler 2, wenn man intelligentes Spielen voraussetzt?

Aufgabe 2.10 Gegeben seien folgenden Mengenpaare:

$$\begin{array}{rcl} \{1,2,3\} & \text{und} & \{2,3,4\}, \\ \{1,2,3\} & \text{und} & \{2\}, \\ \{0,2,3,5\} & \text{und} & \{5,0,1,7\}, \\ \{\emptyset,\{\emptyset\}\} & \text{und} & \{\emptyset\}, \\ \{\{\emptyset,\{\emptyset\}\}\} & \text{und} & \{\{\emptyset\}\}. \end{array}$$

Berechnen Sie jeweils Vereinigung, Durchschnitt, symmetrische Differenz und (beide) Differenzen.

Aufgabe 2.11 Gegeben seien folgenden Mengenpaare:

$$\begin{array}{rcl} \{n \in \mathbb{N} \mid \exists k \in \mathbb{N}\colon n = 4k\} & \text{und} & \{n \in \mathbb{N} \mid \exists k \in \mathbb{N}\colon n = 6k\}, \\ \{n \in \mathbb{N} \mid n \geq 17\} & \text{und} & \{n \in \mathbb{N} \mid n \geq 23\}, \\ \{n \in \mathbb{N} \mid n \geq 17\} & \text{und} & \{n \in \mathbb{N} \mid n \leq 23\}, \\ \mathbb{N} & \text{und} & \mathbb{R}, \\ \mathbb{N} & \text{und} & \emptyset. \end{array}$$

Berechnen Sie jeweils Vereinigung, Durchschnitt und zumindest eine Differenz.

Aufgabe 2.12 Gibt es Mengen A und B, für die $A \setminus B = B \setminus A$ gilt? Falls ja, was kann man über A und B sagen?

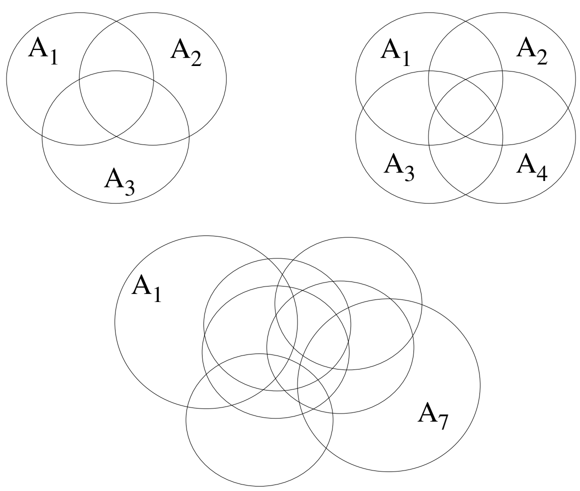
Aufgabe 2.13 Es seien A und B Mengen. Unter welcher Voraussetzung gilt $(A \setminus (B \setminus A)) = ((A \setminus B) \setminus A)$?

Aufgabe 2.14 Beweisen Sie Lemma 2.9, Teile 1, 3, 4 und 5.

Aufgabe 2.15 Beweisen sie Lemma 2.14, Teile 2 bis 7. Tip: verwenden Sie Aufgabe 1.6.

Aufgabe 2.16 Geben Sie eine Darstellung aller Mengen A und B mit der Eigenschaft $A \bowtie B = \{1\}$.

Aufgabe 2.17 Nach Teil 3 von Lemma 2.14 ist die symmetrische Differenz endlich vieler Mengen unabängig von der Reihenfolge. Man könnte vorschnell annehmen, daß in der n-fachen symmetrischen Differenz genau diejenigen Elemente liegen, die in genau einer der beteiligten Mengen liegen. Überprüfen Sie dies: in jedem der drei folgenden Venn-Diagrammen sind verschiedene Mengen $A_1, \ldots, A_n$ dargestellt. Geben sie jeweils eine geometrische Interpretation der symmetrischen Differenz $A_1 \bowtie \cdots \bowtie A_n$ aller überlagerter Mengen, indem Sie genau diejenigen Bereiche grau färben, die zur symmetrischen Differenz dazugehören.



Was zeigen die Beobachtungen über das in Aufgabe 1.6 definierte ausschließende oder „+“?

Aufgabe 2.18 Beweisen Sie die zweite Identität von Lemma 2.16.

Aufgabe 2.19 Verdeutlichen Sie die Aussagen von Lemma 2.16 mittels eines geeigneten Venn-Diagramms.

Aufgabe 2.20 Beweisen Sie die Teile 1, 3 und 4 von Lemma 2.17.

Aufgaben zu Teilkapitel 2.4

Aufgabe 2.21 Beweisen Sie Lemma 2.19.

Aufgaben zu Teilkapitel 2.5

Aufgabe 2.22 Beweisen Sie Lemma 2.28, Teile 2–4.

Aufgabe 2.23 Zeigen Sie: für je zwei Mengen M und N gilt $M = N$ genau dann, wenn $\mathcal{P}(M) = \mathcal{P}(N)$ gilt.

Aufgabe 2.24 Es sei $M = \{1, 2\}$. Berechnen Sie $\mathcal{P}(\mathcal{P}(M))$.

Aufgabe 2.25 Es sei $A := \{\{1, 2, 3\}, \{2, 4\}, \{2, 3, 5\}\}$. Berechnen Sie $\bigcup A$ und $\bigcap A$.

Aufgabe 2.26 Für jedes $i \in \mathbb{N}$ sei $A_i := \{n \in \mathbb{N} \mid n \geq i\}$. Berechnen Sie $\bigcup_{i \in \mathbb{N}} A_i$ und $\bigcap_{i \in \mathbb{N}} A_i$.

Aufgabe 2.27 Es sei

$$A := \{\{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 4\}\}, \{\{1, 2, 3\}, \{1, 2, 5\}, \{1, 4\}\}\}.$$

Berechnen Sie $\bigcup(\bigcup A)$, $\bigcup(\bigcap A)$, $\bigcap(\bigcup A)$ und $\bigcap(\bigcap A)$.

Aufgabe 2.28 Geben Sie eine nichtleere Mengenfamilie A an, wo $\bigcup A = \bigcap A$ gilt. Versuchen sie dann, ein einfaches Kriterium zu finden, wann für eine nichtleere Mengenfamilie A die Gleichung $\bigcup A = \bigcap A$ gilt und wann nicht.

Aufgabe 2.29 Geben Sie alle Partitionen der Menge $\{1, 2, 3, 4\}$ an.

Aufgabe 2.30 Es sei M eine endliche Menge mit $n \geq 1$ Elementen. Es nimmt nun zunächst Spieler 1 eine Teilmenge mit $1, 2$ oder 3 Elementen heraus. Danach nimmt Spieler 2 aus der verbleibenden Menge 1, 2 oder 3 Elemente. Abwechselnd wird weiter so verfahren. Spieler 2 gewinnt, wenn er das letzte Elemente aus der Menge nimmt. Für welche Zahlen n gewinnt Spieler 1, wenn er sich nur geschickt verhält, für welche Zahlen gewinnt Spieler 2?

Aufgabe 2.31 Es lassen sich beliebig viele Mengenfamilien A einer ganz speziellen Form angeben, für die $\mathcal{P}(\bigcup A) = A$ gilt. Welche Form ist das?

Aufgabe 2.32 Es sei P eine Partition der Menge M. Wie sieht $\bigcap P$ aus?

Aufgabe 2.33 Wir hatten bemerkt, daß für jede Menge M stets $\{M\}$ und $\{\{m\} \mid m \in M\}$ Partitionen von M sind. Bedeutet das, daß jede Menge zumindest zwei Partitionen hat? Welche Mengen haben genau zwei Partitionen? Kann es Mengen geben, die genau drei Partitionen besitzen?

Aufgabe 2.34 Betrachtet werden Aussagen der Form

$$\alpha \Rightarrow \alpha \tag{2.1}$$
$$(\alpha \Rightarrow \alpha) \Rightarrow \alpha \tag{2.2}$$
$$((\alpha \Rightarrow \alpha) \Rightarrow \alpha) \Rightarrow \alpha \tag{2.3}$$
$$(((\alpha \Rightarrow \alpha) \Rightarrow \alpha) \Rightarrow \alpha) \Rightarrow \alpha \tag{2.4}$$

Welche dieser Aussagen sind stets (das heißt für jede mathematische Aussage α) aussagenlogische Tautologien? Verallgemeinern Sie Ihre Beobachtung auf Aussagen der entsprechenden Form mit $n > 0$ Vorkommen von α. Verwenden Sie dazu vollständige Induktion.

Aufgaben zu Teilkapitel 2.6

Aufgabe 2.35 Kann es einen Frisör geben, der genau diejenigen Münchener rasiert, die sich nicht selbst rasieren, und sonst keine weitere Person? Wenn ja, kann dieser Frisör ein Münchener sein?

Aufgabe 2.36 Wir nennen ein Adjektiv w der deutschen Sprache *autologisch* genau dann, wenn auf w selbst die durch w ausgedrückte Eigenschaft zutrifft. Das Adjektiv „kurz" ist etwa autologisch, da es nur vier Buchstaben besitzt. Umgekehrt nennen wir ein Adjektiv w *heterologisch* genau dann, wenn w die durch w ausgedrückte Eigenschaft *nicht* besitzt. Das Adjektiv „lang" ist heterologisch. Ist das Adjektiv „heterologisch" autologisch oder heterologisch?

Aufgaben zu Teilkapitel 2.7

Aufgabe 2.37 Kann es eine Mengenfamilie A geben, wo in $\bigcap A$ ein Element von A liegt? Von welcher Frage hängt dies ab?

2.9 Bibliographische Angaben

Kurzdarstellungen zu mengentheoretischen Prinzipien, die zum hier dargestellten Inhalt vergleichbar sind, finden sich in fast allen Büchern zur „diskreten" Mathematik, wie etwa [Big89, Bra88, FS91, Ger72, RW92, uLW74]. Viele dieser Bücher eignen sich auch zur ergänzenden Lektüre bei den nachfolgenden Kapiteln. Die Darstellungen in [Big89, Bra88] bieten mehr Hintergrund zu den Zahlbereichen. Eine kurze Darstellung der mengentheoretischen Kodierung von natürlichen, ganzen, rationalen und reellen Zahlen ist in [PtMW90] zu finden. Einen ausführlichen Hintergrund zu diesem Thema mit vielen Bemerkungen zur historischen Entwicklung der Mengenlehre bietet das dreibändige Werk [Fel79]. Es gibt zahlreiche Bücher, die ausführlichere Einführungen in die Mengenlehre anbieten. Darunter fallen etwa [Hal87, Roi90]. Unter den anspruchsvolleren weiterführenden Werken sei [Lev79] genannt. Cantors Originalbeiträge sind in [Can55] wiedergegeben. Die Entwicklung der nichtfundierten Mengenlehre geht wesentlich auf [Acz88] zurück.

3

Relationen und Funktionen

Die in diesem Kapitel zu besprechenden Konzepte der Relation und der Funktion erlauben es, bestimmte *strukturelle Eigenschaften* eines Bereichs formal zu erfassen. Während Relationen allgemeine Beziehungen zwischen Elementen beschreiben, sind Funktionen Zuordnungen, die jedem Element einer Menge A genau ein Element einer Menge B zuweisen. Um diese Begriffe zu präzisieren, führen wir im ersten Abschnitt ausgehend vom Mengenbegriff zunächst den Begriff des geordneten Paars zweier Elemente und den Begriff des kartesischen Produkts von Mengen ein. Allgemeiner werden n-Tupel und Wörter über einer gegebenen Menge charakterisiert. Aufbauend auf dem Begriff des kartesischen Produkts werden dann in Abschnitt 3.2 zunächst Relationen formal definiert und in 3.3 Operationen auf Relationen betrachtet. Im Anschluß stellen wir in Abschnitt 3.4 den Funktionsbegriff dar und gehen in 3.5 auf wichtige Eigenschaften von Funktionen ein. Das Kapitel schließt mit einer Darstellung des wichtigen Prinzips der induktiven Definition von Mengen und mit Bemerkungen zu charakteristischen Funktionen, höheren Funktionen und unendlichen kartesischen Produkten.

3.1 Tupel, kartesische Produkte und Wörter

Der naive Begriff einer „Zuordnung“ beinhaltet eine Reihenfolge oder Richtung, anders als der Mengenbegriff. Kartesische Produkte, die wir nun besprechen werden, erlauben es, eine Reihenfolge zwischen Elementen festzuhalten. Um kartesische Produkte zu definieren, benötigt man den Begriff des „geordneten Paars“ $\langle a, b\rangle$ zweier Elemente. Wir könnten das geordnete Paar als primitiven Begriff einführen. Es ist jedoch auch eine Rückführung auf schon vorhandene Begriffe möglich. Wer sich nicht für Feinheiten einer mengentheoretischen Kodierung interessiert, kann die nachfolgende Definition nebst zugehörigem Lemma ohne Schaden überlesen. Schreibweisen wie $\{a, b\}$ sollen dort nicht implizieren, daß die Elemente a und b notwendigerweise verschieden sind. Im Fall $a = b$ folgt natürlich $\{a, b\} = \{a\} = \{b\}$ sowie $\{\{a\}, \{a, b\}\} = \{\{a\}\}$.

Definition 3.1 Seien a und b Elemente einer Menge M. Das *geordnete Paar* von a und b, notiert $\langle a, b\rangle$, ist die Menge $\{\{a\}, \{a, b\}\}$.

Lemma 3.2 *Für beliebige Elemente a, b, c und d einer Menge M gilt:*

(i) *Falls $\{a, c\} = \{b, c\}$, so gilt $a = b$.*

(ii) *Falls $\langle a, b\rangle = \langle c, d\rangle$, so gilt $a = c$ und $b = d$, und umgekehrt.*

Beweis. (i) Es gelte $\{a, c\} = \{b, c\}$. Wir unterscheiden nun zwei Fälle. Falls $a = c$ gilt, so folgt $b \in \{b, c\} = \{a, c\} = \{a\}$, also $a = b$. Falls $a \neq c$ gilt, so folgt aus $a \in \{a, c\} = \{b, c\}$ ebenfalls $a = b$.

(ii) Es gelte $\langle a, b\rangle = \langle c, d\rangle$. Aus $\{a\} \in \langle a, b\rangle = \langle c, d\rangle = \{\{c\}, \{c, d\}\}$ folgt $\{a\} = \{c\}$ oder $\{a\} = \{c, d\}$. Im ersten Fall folgt $a = c$, im zweiten sogar $a = c = d$. Stets erhalten wir $\{a\} = \{c\}$ und $a = c$. Nach Voraussetzung gilt $\{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$. Aufgrund dieser Gleichung folgt wegen $\{a\} = \{c\}$ nach Teil (i) nun $\{a, b\} = \{c, d\}$, hieraus durch nochmalige Anwendung von (i) aber $b = d$. Bei der Umkehrung ist nichts zu zeigen. ■

Wegen Lemma 3.2 (ii) können wir vom vorderen Eintrag a und vom hinteren Eintrag b des geordneten Paars $\langle a, b\rangle$ sprechen. Beide Einträge können identisch sein. Derselbe Effekt ließe sich allerdings auch durch viele Varianten von Definition 3.1 erreichen.

Definition 3.3 Seien A und B Mengen. Die Menge

$$A \times B := \{\langle a, b\rangle \mid a \in A, b \in B\}$$

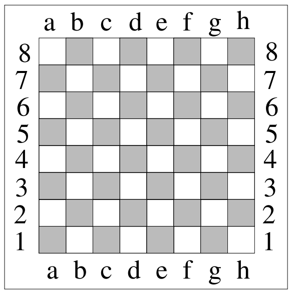
heißt *kartesisches Produkt* der Mengen A und B.

Beispiele 3.4 Zur Verdeutlichung dieses Begriffs hier einige Beispiele.

1. Es sei sei $A = \{1, 2\}$, und $B = \{k, l, m\}$. Dann ist $A \times B = \{\langle 1, k\rangle, \langle 1, l\rangle, \langle 1, m\rangle, \langle 2, k\rangle, \langle 2, l\rangle, \langle 2, m\rangle\}$.

2. Es sei $A = \{0, 1\}$. Dann ist $A \times A = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 0\rangle, \langle 1, 1\rangle\}$. Die Wahrheitswert-Tabelle aus dem Beweis von Satz 1.1 unterscheidet also je einen Fall für jedes Element von $\{0, 1\} \times \{0, 1\}$.

3. Die übliche Schreibkonvention für Schachfelder basiert darauf, daß das Schachbrett



   als kartesisches Produkt der Mengen $\{a, b, c, d, e, f, g, h\}$ und $\{1, 2, 3, 4, 5, 6, 7, 8\}$ betrachtet wird. Felderbezeichnungen wie „$f5$“ stellen nur eine Kurznotation für das entsprechende geordnete Paar $\langle f, 5\rangle$ dar.

Sind A und B endlich, mit m respektive n Elementen, so hat des kartesische Produkt $A \times B$ gerade $m \cdot n$ viele Elemente. Dies folgt sofort aus Teil (ii) von Lemma 3.2. Die zweite der nun folgenden Identitäten ergibt eine weitere formale Ähnlichkeit zum Rechnen mit Zahlen. Man beachte, daß dort das erste Argument A festgehalten wird.

Lemma 3.5 *Die folgenden Identitäten gelten für beliebige Mengen:*

$$\begin{array}{rrcl} (i) & (A \times B) \cap (C \times D) & = & (A \cap C) \times (B \cap D), \\ (ii) & (A \times B) \cup (A \times C) & = & A \times (B \cup C). \end{array}$$

Beweis. (i) Da wir hier die Gleichheit zweier Mengen zu zeigen haben, folgen wir dem Standardrezept aus Bemerkung 2.3. Wir zeigen, daß jedes Element der linken Seite auch ein Element der rechten Seite ist. Die Umkehrung folgt analog. Sei $x \in (A \times B) \cap (C \times D)$. Nach Definition des Durchschnitts gilt dann

$$\begin{aligned} &(1) \quad x \in A \times B, \\ &(2) \quad x \in C \times D. \end{aligned}$$

Nach der Definition des kartesischen Produkts (Def. 3.3) folgt aus (1) und (2) nun

$$\begin{aligned} &(3) \quad \text{es existieren } a \in A, b \in B\text{: } x = \langle a, b \rangle, \\ &(4) \quad \text{es existieren } c \in C, d \in D\text{: } x = \langle c, d \rangle. \end{aligned}$$

Da somit $\langle a, b \rangle = \langle c, d \rangle$ gilt, folgt mit Lemma 3.2 (ii) nun $a = c$ und $b = d$. Somit ist $a = c \in A \cap C$ und $b = d \in B \cap D$, das heißt $x \in (A \cap C) \times (B \cap D)$.

(ii) Wir zeigen zuerst $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$. Sei d ein beliebiges Element von $(A \times B) \cup (A \times C)$. Nach Definition der Vereinigung ist dann $d \in A \times B$ oder $d \in A \times C$. Wir führen eine Fallunterscheidung durch.

Fall 1, $d \in A \times B$. Nach Definition 3.3 existieren dann Elemente $a \in A$ und $b \in B$ mit $d = \langle a, b \rangle$. Damit existieren Elemente $a \in A$ und $b \in B \cup C$ mit $d = \langle a, b \rangle$. Nach Definition 3.3 gilt dann $d \in A \times (B \cup C)$.

Fall 2, $d \in A \times C$. Nach Definition 3.3 existieren dann Elemente $a \in A$ und $c \in C$ mit $d = \langle a, c \rangle$. Damit existieren Elemente $a \in A$ und $c \in B \cup C$ mit $d = \langle a, c \rangle$. Nach Definition 3.3 gilt dann $d \in A \times (B \cup C)$.

In beiden Fällen folgt also $d \in A \times (B \cup C)$, wodurch $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$ bewiesen ist.

Wir zeigen nun umgekehrt, daß $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$ gilt. Es sei d ein beliebiges Element von $A \times (B \cup C)$. Nach Definition 3.3 existieren dann Elemente $a \in A, e \in B \cup C$ mit $d = \langle a, e \rangle$. Wir führen eine Fallunterscheidung durch.

Fall 1, $e \in B$. Nach Definition 3.3 ist dann $d = \langle a, e \rangle \in A \times B$. Damit ist auch $d \in (A \times B) \cup (A \times C)$.

Fall 2, $e \in C$. Nach Definition 3.3 ist dann $d = \langle a, e \rangle \in A \times C$. Damit ist auch $d \in (A \times B) \cup (A \times C)$.

In beiden Fällen ergibt sich $d \in (A \times B) \cup (A \times C)$. Damit ist $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$ bewiesen. Da wir bereits $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$

gezeigt haben, folgt $(A \times B) \cup (A \times C) = A \times (B \cup C)$ nach Lemma 2.6 Teil 2.

Wenn man etwas Vertrautheit beim Umgang mit dem Existenzquantor voraussetzt, kann man die Argumentation wesentlich kompakter darstellen:

$$\begin{aligned}
& d \in A \times (B \cup C) \\
\Leftrightarrow\ & \exists a \in A, \exists e \in B \cup C\colon d = \langle a, e\rangle \\
\Leftrightarrow\ & \exists a \in A, \exists e\colon ((e \in B \vee e \in C) \wedge d = \langle a, e\rangle) \\
\Leftrightarrow\ & \exists a \in A, \exists e\colon ((e \in B \wedge d = \langle a, e\rangle) \vee (e \in C \wedge d = \langle a, e\rangle)) \\
\Leftrightarrow\ & (\exists a \in A, \exists e \in B\colon d = \langle a, e\rangle) \ \vee \ (\exists a \in A, \exists e \in C\colon d = \langle a, e\rangle) \\
\Leftrightarrow\ & (d \in A \times B) \vee (d \in A \times C) \\
\Leftrightarrow\ & d \in (A \times B) \cup (A \times C).
\end{aligned}$$

■

Hinweis 3.6 Wir haben in Teil (i) des vorangegangenen Beweises eine „Vorsichtsmaßnahme“ verwendet, die sich generell empfiehlt, um Fehlschlüsse zu vermeiden: als wir in den Zeilen (3) und (4) die beiden Informationen $x \in A \times B$ und $x \in C \times D$ aufschlüsselten, hatten wir bewußt *vier* Variablen a, b, c, d verwendet. Aus Definition 3.3 alleine folgen keineswegs die Identitäten $a = c$ und $b = d$, wir mußten wirklich erst Lemma 3.2 (ii) verwenden. Ganz generell sollte man *besser diesselbe Variable nur dann zweimal verwenden, wenn sicher ist, daß damit stets dasselbe Objekt bezeichnet wird.*

Wir wollen nun noch kartesische Produkte mit n Faktoren definieren, wobei $n \geq 0$ eine beliebige natürliche Zahl ist. Hierzu führen wir das Symbol „$\langle\ \rangle$“ ein, das nachfolgend das *leere Tupel* repräsentiert (wir verzichten auf eine Kodierung als Menge). Weiter betrachten wir die Elemente einer Menge A als Eintupel, die wir wahlweise in der Form a oder $\langle a\rangle$ schreiben. Für beliebige Elemente $a_1, \ldots, a_{n+1}$ mit $n \geq 2$ definieren wir nun induktiv

$$\langle a_1, \ldots, a_n, a_{n+1}\rangle := \langle\langle a_1, \ldots, a_n\rangle, a_{n+1}\rangle$$

als das $n+1$-*Tupel* mit den Elementen $a_1, \ldots, a_n, a_{n+1}$.

Die hier verwendete Technik der induktiven Definition beruht auf demselben „Leiterprinzip“, das wir bereits im Zusammenhang mit dem Beweisprinzip der vollständigen Induktion im vorigen Kapitel kennengelernt hatten. Wenn wir wissen, wie geordnete Tupel der Stelligkeit 0, 1 oder 2 definiert sind, und wenn wir angeben, wie sich die Definition des geordneten $(n+1)$-Tupels für $n \geq 2$ auf die des geordneten n-Tupels zurückführen läßt, so haben wir tatsächlich Tupel beliebiger (endlicher) Stelligkeit definiert. Am Ende des Kapitels werden wir ausführlicher auf die eng verwandte induktive Definition von Mengen eingehen.

Definition 3.7 Es sei $n \in \mathbb{N}$ und $A_1, \ldots, A_n$ Mengen. Die Menge

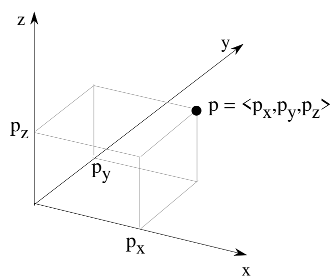
$$\prod_{i=1}^{n} A_i := \{\langle a_1, \ldots, a_n \rangle \mid a_i \in A_i \quad \text{für} \quad i = 1, \ldots, n\}$$

heißt *kartesisches Produkt* der Mengen $A_1, ..., A_n$.

Das kartesische Produkt wird oft auch in der Form $A_1 \times A_2 \times \ldots \times A_n$ notiert. Gilt $A_1 = \ldots = A_n =: A$, so schreibt man für $\prod_{i=1}^{n} A_i$ auch kürzer A^n.

Beispiele 3.8 Auch hierzu einige Beispiele.

1. Stets gilt $A^0 = \{\langle\ \rangle\}$ und $A^1 = \{\langle a \rangle \mid a \in A\} = \{a \mid a \in A\} = A$.
2. Es sei sei $A = \{1, 2\}$, $B = \{b\}$ und $C = \{c\}$. Dann ist $A \times B \times C = \{\langle 1, b, c \rangle, \langle 2, b, c \rangle\}$.
3. Bezeichnet $\mathbb{R}$ die Menge der reellen Zahlen, so ist $\mathbb{R}^3$ der dreidimensionale euklidische Raum. Jedes Element $p = \langle p_x, p_y, p_z \rangle \in \mathbb{R}^3$ bezeichnet also einen Punkt im Raum, die drei Komponenten geben die Koordinaten bezüglich des üblichen Koordinatensystems mit den Achsen x, y, z wieder.



Die folgende Definition motiviert, warum wir auch geordnete 1-Tupel und das leere Tupel $\langle\ \rangle$ als Spezialfälle geordneter Tupel eingeführt haben.

Definition 3.9 Es sei A eine Menge. Dann heißt

$$A^* := \bigcup_{n \in \mathbb{N}} A^n$$

die Menge der *Wörter über der Menge* A. Jedes Element $w \in A^n$ heißt ein *Wort der Länge* n über A.

Man sollte sich nicht dadurch verwirren lassen, daß wir an dieser Stelle zwei unterschiedliche Bezeichnungen für dieselbe Art von Objekt verwenden: formal sind Wörter der Länge n schlicht n-Tupel mit Elementen aus A. Von Wörtern spricht man in der Regel dann, wenn man Tupel unterschiedlicher Länge betrachtet, und wenn man betonen möchte, daß man die Tupel einfach als Zeichenfolgen auffassen möchte. Die Basismenge A wird in diesem Kontext meist als das „Alphabet" bezeichnet. Wörter über dem Alphabet A werden häufig in der vereinfachten Form $a_1 \cdots a_n$ statt $\langle a_1, \ldots, a_n \rangle$ notiert. Man sagt deshalb auch, ein Wort über A sei eine *endliche Folge* von Symbolen aus A.

Beispiel 3.10 Es sei A das Alphabet mit allen Buchstaben der deutschen Sprache. Dann sind

$$\mathit{Montag}, \mathit{Dienstag}, \ldots, \mathit{Sonntag}$$

Wörter über A. Es ist *Montag* eine notationelle Variante für $\langle M, o, n, t, a, g \rangle$ und analog für die anderen Wörter.

Definition 3.11 Es sei A eine Menge. Dann heißt jede Teilmenge L von A^* eine *formale Sprache* über dem Alphabet A. Mit $\mathcal{L}(A)$ bezeichnen wir die Menge aller formalen Sprachen über A.

Solange keine Konfusion zu befürchten ist, werden wir formale Sprachen auch einfach „Sprachen" nennen.

Beispiele 3.12 Die nachfolgenden Beispiele illustrieren den Begriff und werden zum Teil später wieder aufgegriffen.

1. Es sei A das Alphabet mit den Ziffern $0, 1, \ldots, 9$. Dann ist die Menge der natürlichen Zahlen in der üblichen Dezimaldarstellung, wo die erste Ziffer einer von Null verschiedenen Zahl stets ungleich 0 ist, eine Sprache über A.

2. Es sei A das Alphabet mit allen Buchstaben der deutschen Sprache und mit dem Leerzeichen „␣“, dem Punkt „.“ und den Ziffern aus $\{0, 1, \ldots, 9\}$. Dann sind

$$\begin{aligned} L_1 &:= \{\mathit{Montag}\_, \mathit{Dienstag}\_, \ldots, \mathit{Sonntag}\_\}, \\ L_2 &:= \{\mathit{den}\_\}, \\ L_3 &:= \{1.\_, 2.\_, \ldots, 31.\_\}, \\ L_4 &:= \{\mathit{Januar}\_, \mathit{Februar}\_, \ldots, \mathit{Dezember}\_\}, \\ L_5 &:= \{2000, 2001, \ldots, 3000\}, \end{aligned}$$

formale Sprachen über A. Wir werden später sehen, wie wir durch eine einfache Form der Komposition dieser Sprachen eine Sprache mit vollständigen Datumsangaben der Art

$$\mathit{Dienstag\_den\_}11.\_\mathit{Januar}\_2000$$

erreichen können (vgl. Beispiel 3.34 Nr. 7).

3. Eine Menge von *Wörtern* L kann selbst als Alphabet dienen. Ist etwa $L := \{\mathit{Max}, \mathit{raucht}, \mathit{lacht}\}$, so ist die Menge mit den Sätzen (formal: Wörtern über L) „*Max raucht*“ und „*Max lacht*“ eine Sprache über L. In diesem Sinn kann jede Menge von Sätzen der deutschen Sprache als eine formale Sprache über einem Wortalphabet beschrieben werden.

3.2 Relationen

Mit Relationen lassen sich Beziehungen zwischen Elementen von Mengen formalisieren. Von einer Relation spricht man genauer, wenn immer dieselbe Anzahl von Elementen in Beziehung steht.

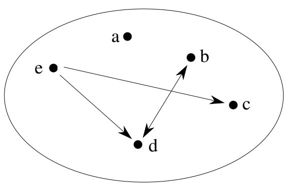
Definition 3.13 Eine Menge R wird n–stellige *Relation* genannt genau dann, wenn es Mengen $A_1, \ldots, A_n$ gibt ($n \geq 1$), wo $R \subseteq \prod_{i=1}^{n} A_i$. Statt $\langle a_1, \ldots, a_n \rangle \in R$ schreiben wir kurz $R(a_1, \ldots, a_n)$. Gilt $A_1 = \ldots = A_n =: A$, so heißt R eine *n–stellige Relation auf A*.

Beispiel 3.14 Tabellen der Art

Name	Geschlecht	Dienstalter	Geburtsdatum	Gehaltsklasse
Maier	m	7	11.2.63	I
Kurz	w	13	11.3.67	III
Müller	w	14	2.12.73	I
Santorini	m	9	5.5.64	IV
...	...	...	...	...

stellen Relationen dar, die Zahl der Tabellenspalten bestimmt die Stelligkeit. Die Beispieltabelle enthält 5-Tupel aus dem kartesischen Produkt einer Menge A_1 von Namen, der Menge $A_2 := \{m, w\}$, einer Menge A_3 von Geburtsdaten und der Menge A_4 der Gehaltsklassen. In der Informatik werden Tabellen dieser Art in sogenannten relationalen Datenbanken gespeichert und abfragbar gemacht.

Viele der in der Mathematik auftretenden Relationen sind zweistellig. Zur Darstellung von zweistelligen Relationen auf endlichen Mengen verwendet man häufig Schaubilder mit Pfeilen. Ein Pfeil von einem Element x zu einem Element y deutet an, daß $R(x, y)$ gilt.

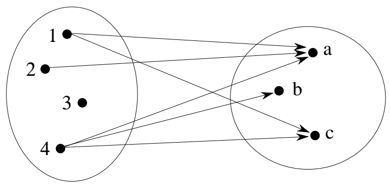


Alternativ können Relation auf endlichen Mengen als *Matrix* von Wahrheitswerten dargestellt werden. Wir schreiben $M[i, j]$ für den Eintrag in Zeile i und Spalte j der Matrix M. Ein Eintrag $M[i, j] = 1$ (resp. 0) steht für $R(i, j)$ (resp. $\neg R(i, j)$). Für die oben abgebildete Relation R ergibt sich folgende Matrix M.

R	a	b	c	d	e
a	0	0	0	0	0
b	0	0	0	1	0
c	0	0	0	0	0
d	0	1	0	0	0
e	0	0	1	0	1

Diagramm- und Matrixdarstellung können natürlich auch für beliebige zweistellige Relationen $R \subseteq A \times B$ verwendet werden.

Beispiel 3.15 Die Relation



enthält genau die Paare $\langle 1, a\rangle, \langle 1, c\rangle, \langle 2, a\rangle, \langle 4, a\rangle, \langle 4, b\rangle$ und $\langle 4, c\rangle$. In Matrixform ergibt sich

R	a	b	c
1	1	0	1
2	1	0	0
3	0	0	0
4	1	1	1

Charakteristisch für zweistellige Relationen—im Gegensatz zu den nachfolgend zu besprechenden Funktionen—ist es, daß ein gegebenes Element zu *mehreren* anderen Elementen in Beziehung stehen kann, wie auch die beiden vorausgegangenen Beispiele verdeutlichen.

Beispiele 3.16 Hier eine kleine Auswahl weiterer Relationen.

1. Es sein $A = \{a, b, c\}$. Wie in Definition 3.9 bezeichne A^* die Menge aller Wörter über dem Alphabet A. Die Relation R, die auf ein Paar $\langle v, w\rangle \in A^* \times A^*$ zutrifft genau dann, wenn v ein Präfix (Wortanfang) von w ist, ist eine zweistellige Relation auf A^*. Die Präfixe von $aabc$ sind beispielsweise $\langle\ \rangle$ (leeres Wort, oft als „ϵ“ notiert) sowie die Wörter a, aa, aab und $aabc$ (eine formale Definition des Begriffs „Präfix“ werden wir in Beispiel 3.34 Nr. 5 geben). Verwandt ist die Suffixrelation, die auf ein Paar $\langle v, w\rangle \in A^* \times A^*$ zutrifft genau dann, wenn v ein Suffix (Wortende) von w ist. Die Suffixe von $aabc$ sind $\langle\ \rangle$, c, bc, abc und $aabc$.

2. Die übliche „kleinergleich“–Relation „$\leq$“ ist eine zweistellige Relation auf der Menge $\mathbb{R}$ der reellen Zahlen. Statt „$\leq (a, b)$“ schreibt man in Infixnotation „$a \leq b$“. Auch die „kleinergleich“–Relation auf $\mathbb{N}$ ist eine

zweistellige Relation. Obwohl beide Relationen formal unterschiedliche Mengen sind, wird dasselbe Zeichen „$\leq$" verwendet.

3. Für natürlich Zahlen n und m ist die *Teilbarkeitsrelation* T definiert durch $T(n, m) :\Leftrightarrow n$ teilt $m \Leftrightarrow \exists k \in \mathbb{N}$: $n \cdot k = m$. Falls n die Zahl m teilt, schreibt man kurz $n|m$.

4. Jede Teilmenge A einer Menge M ist eine einstellige Relation auf M und umgekehrt, wobei M wie in Beispiel 3.8 Nr. 1 als Produkt mit nur einem Faktor aufgefaßt wird. In einer Menge von Personen M definiert die Eigenschaft, weiblich zu sein, eine einstellige Relation W auf M. In der Linguistik wird oft die Bedeutung von intransitiven Verben, Nomen, und (unter Einschränkungen) Adjektiven durch einstellige Relationen auf einer Menge M, die durch den Sprechkontext gegeben ist, formalisiert.

5. Einfach transitive Verben wie „liebt" oder „kennt" entsprechen zweistelligen Relationen auf der durch den Sprechkontext gegebenen Menge M.

6. Ist A eine Menge von Personen, so lassen sich viele Verwandschaftsbeziehungen als binäre Relationen formalisieren. Wir können etwa die „Elternbeziehung" E definieren durch $E(a, b) :\Leftrightarrow a$ ist Elternteil von b. Analog definieren wir die Kinderbeziehung K durch $K(a, b) :\Leftrightarrow a$ ist Kind von b. In ähnlicher Weise können wir Vater-, Mutter-, Schwester-, Bruder-, Vettern- und andere Beziehungen definieren.

7. Es sei $\mathcal{B} := \{a, b, c, d, e, f, g, h\} \times \{1, 2, 3, 4, 5, 6, 7, 8\}$ die Menge der Felder auf dem Schachbrett, $\mathcal{F}$ sei eine Figur, etwa ein Springer oder ein Läufer. Dann bildet die Menge $R_{\mathcal{F}}$ aller Felder-Paare $\langle\langle x, n\rangle, \langle y, m\rangle\rangle \in \mathcal{B} \times \mathcal{B}$, so daß man (auf dem ansonsten leeren Schachbrett) mit der Figur $\mathcal{F}$ von $\langle x, n\rangle$ nach $\langle y, m\rangle$ ziehen kann, eine zweistellige Relation auf der Menge $\mathcal{B}$.

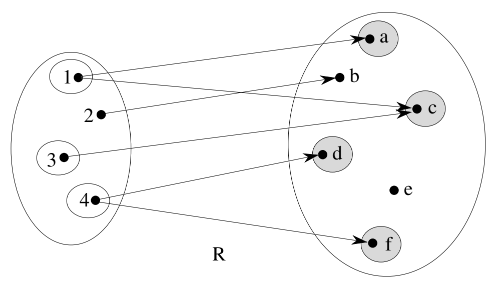
8. Wenn wir die Menge der Figuren bei einer bestimmten Stellung auf einem Schachbrett nehmen, so bilden die geordneten Paare von Figuren derselben Farbe, die sich bei der aktuellen Stellung gegenseitig decken, eine zweistellige Relation. Eine andere Relation bilden die geordneten Paare $\langle x, y\rangle$ von Figuren unterschiedlicher Farbe, wo Figur x die Figur y bedroht.

Definition 3.17 Sei $R \subseteq A \times B$ eine zweistellige Relation, $X \subseteq A$ und $Y \subseteq B$. Die Menge $R(X) := \{y \in B \mid \exists x \in X\colon R(x,y)\}$ heißt das *Bild* von X unter R, Die Menge $R\lceil X := \{\langle x,y\rangle \mid R(x,y), x \in X\}$ heißt die *Einschränkung* von R auf X.

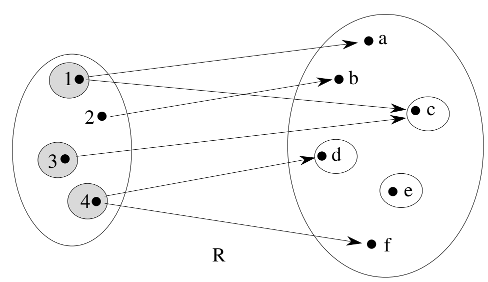
Die Mengen $R\lceil X$ und $R(X)$ sollten nicht verwechselt werden. Man beachte, daß die Einschränkung $R\lceil X$ eine Relation—damit eine Menge geordneter Paare—ist, hingegen das Bild $R(X)$ eine einfache Teilmenge von B repräsentiert.

Beispiele 3.18 Hierzu einige Illustrationen.

1. In der nachfolgenden Abbildung ist $\{a,c,d,f\}$ das Bild von $\{1,3,4\}$ unter der durch Pfeile symbolisierten Relation R:



   Um diese Menge zu erhalten, muß man die hinteren Einträge y all derjenigen geordneten Paare $\langle x,y\rangle$ aus R zusammenfassen, wo der vordere Eintrag x aus $\{1,3,4\}$ ist. Entsprechend wäre $\{b,d,f\}$ das Bild von $\{2,4\}$. Das Urbild von $\{c,d,e\}$ unter R ist $\{1,3,4\}$:

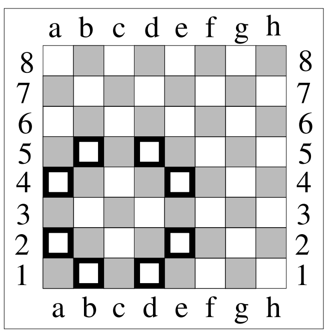


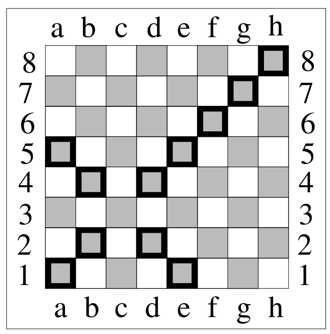
2. Für die Relation R, die in obigen Abbildungen dargestellt ist, ist die Einschränkung von R auf $\{1,2,3\}$ die Menge

$\{\langle 1,a\rangle,\langle 1,c\rangle,\langle 2,b\rangle,\langle 3,c\rangle\}$. Die Paare $\langle 4,d\rangle$ und $\langle 4,f\rangle$ aus R gehören nicht zu $R\lceil\{1,2,3\}$, da $4\notin\{1,2,3\}$.

3. Ist A eine Menge von Personen, K wie in Beispiel 3.16 Nr. 6 die Kinderbeziehung, $a\in A$, so ist $K(\{a\})$ die Menge der Eltern der Person a. Ist auch $b\in A$, so bilden die (nicht notwendig gemeinsamen) Eltern von a und b die Menge $K(\{a,b\})$.

4. Es sei $\mathcal{B}:=\{a,b,c,d,e,f,g,h\}\times\{1,2,3,4,5,6,7,8\}$ die Menge der Felder auf dem Schachbrett, S sei ein Springer, L ein Läufer. Die Relationen R_S und R_D auf $\mathcal{B}$ seien wie in Beispiel 3.16 Nr. 7 erklärt. Dann bilden die umrandeten Felder das Bild der Feldermenge $\{C3\}$ unter R_S bzw. R_L.





Lemma 3.19 *Es sei $R\subseteq A\times B$ und $X_1\subseteq X_2\subseteq A$. Dann ist $R(X_1)\subseteq R(X_2)$. Ist $Y_1\subseteq Y_2\subseteq B$, so ist $R^{-1}(Y_1)\subseteq R^{-1}(Y_2)$.*

Beweis. Es sei $b\in R(X_1)$. Dann existiert ein $x\in X_1$ mit $\langle x,b\rangle\in R$. Wegen $X_1\subseteq X_2$ existiert auch ein $x\in X_2$ mit $\langle x,b\rangle\in R$. Damit gilt $b\in R(X_2)$. Die zweite Behauptung ist eine Instanz der ersten. ∎

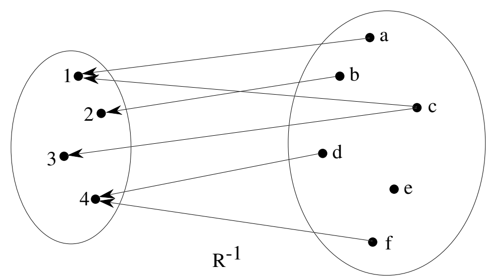
3.3 Umkehrrelation und Komposition von Relationen

Das Vorgehen ist nun dasselbe wie im vorhergehenden Kapitel: nachdem wir Relationen eingeführt haben, betrachten wir nun Operationen, die es erlauben, aus bestehenden Relationen neue zu bilden. Zwei Operationen, Komposition und Umkehrung, sind hierbei von vorrangiger Bedeutung.

Definition 3.20 Sei $R \subseteq A \times B$ eine zweistellige Relation. Die Menge $R^{-1} := \{\langle y, x\rangle \mid R(x,y)\}$ heißt die zu R *inverse* Relation oder die *Umkehrrelation* von R. Ist $Y \subseteq B$, so heißt $R^{-1}(Y)$ das *Urbild* von Y unter R.

Beispiele 3.21 Einige Beispiele zur Illustration der Umkehrrelation:

1. Ist die Relation R durch Pfeile dargestellt, so ergibt sich die inverse Relation R^{-1} durch Umkehrung aller Pfeile. Die Umkehrrelation der in Beispiel 3.18 Nr. 1 abgebildeten Relation R hat damit die folgende Form:



2. Ist $R \subseteq A \times A$ in Matrixform dargestellt, so ergibt sich R^{-1} durch Spiegelung an der Diagonalen.

R	1	2	3	4	5
1	0	0	1	0	1
2	0	0	0	0	0
3	0	0	0	1	1
4	0	0	0	0	0
5	0	0	1	0	0

R^{-1}	1	2	3	4	5
1	0	0	0	0	0
2	0	0	0	0	0
3	1	0	0	0	1
4	0	0	1	0	0
5	1	0	1	0	0

3. Formalisiert die Relation R die Bedeutung eines transitiven Verbs, so ergibt sich die Relation R^{-1} durch Passivbildung. Die inverse Relation zu „liebt“ ist also „wird-geliebt-von“.

4. Es sei A eine Menge von Personen, K wie in Beispiel 3.16 Nr. 6 die Kinderbeziehung. Die zu K inverse Relation K^{-1} ist die „Elternbeziehung“ E, und $E^{-1} = K$.

Lemma 3.22 *Es seien $R \subseteq A \times B$ und $S \subseteq A \times B$ zweistellige Relationen. Dann gilt stets*

1. $(R^{-1})^{-1} = R$,
2. $(R \cup S)^{-1} = R^{-1} \cup S^{-1}$,
3. $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$.

Beweis. Wir zeigen Teil 2, die anderen Beweise bleiben dem Leser überlassen (vgl. Aufgabe 3.16). Es sei $\langle b, a\rangle \in (R \cup S)^{-1}$. Dann gilt $\langle a, b\rangle \in (R \cup S)$ und somit $\langle a, b\rangle \in R$ oder $\langle a, b\rangle \in S$. Im ersten Fall folgt $\langle b, a\rangle \in R^{-1}$, im zweiten Fall $\langle b, a\rangle \in S^{-1}$, somit gilt stets $\langle b, a\rangle \in R^{-1} \cup S^{-1}$. Wir haben damit die Inklusion $(R \cap S)^{-1} \subseteq R^{-1} \cap S^{-1}$ bewiesen. Die Umkehrung folgt analog. ■

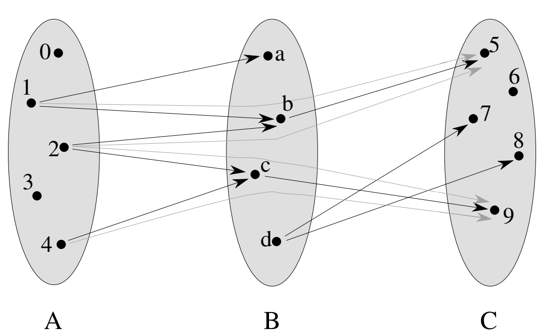
Definition 3.23 Es seien $R \subseteq A \times B$ und $S \subseteq B \times C$ zwei Relationen. Ihr *Produkt* (oder ihre *Komposition*) $R \circ S$ ist erklärt durch

$$R \circ S := \{\langle a, c\rangle \in A \times C \mid \exists b \in B\colon R(a, b) \text{ und } S(b, c)\}.$$

Für das n-fache Produkt $R \circ \cdots \circ R$ einer Relation mit sich selbst schreiben wir kurz R^n $(n \geq 1)$.[1]

Hinweis 3.24 Die Notation für das Produkt zweier Relationen ist in der Literatur nicht einheitlich. In manchen Büchern wird für die oben definierte Relation $R \circ S$ genau umgekehrt $S \circ R$ geschrieben. Man sollte sich also gegebenenfalls vergewissern, welche Konvention der Notation zugrundeliegt.

Offensichlich ist mit $R \subseteq A \times B$ und $S \subseteq B \times C$ auch $R \circ S$ wieder eine Relation, und zwar eine Teilmenge von $A \times C$. Das folgende Bild bietet ein illustrierendes Beispiel. Die Relation $R \circ S$ ist durch graue Pfeile symbolisiert.



[1] Die Reihenfolge der Klammerung ist hier irrelevant, vgl. 3.27 unten.

Sind A, B, C endliche Mengen mit m, n beziehungsweise l Elementen, und sind die Relationen $R \subseteq A \times B$ und $S \subseteq B \times C$ durch die Matrizen M_R und M_S repräsentiert, so erhält man die Matrix M von $R \circ S$ durch eine spezielle Art der Matrixmultiplikation aus M_R und M_S. M hat m Zeilen und l Spalten. Der Eintrag $M[i, j]$ ergibt sich als Disjunktion (i.e., Maximum) über alle Produkte $M_R[i, k] \cdot M_S[k, j]$ für $1 \leq k \leq n$. Im Fall der oben abgebildeten Relationen R und S erhält man

R	a	b	c	d
0	0	0	0	0
1	1	1	0	0
2	0	1	1	0
3	0	0	0	0
4	0	0	1	0

$\times$

S	5	6	7	8	9
a	0	0	0	0	0
b	1	0	0	0	0
c	0	0	0	0	1
d	0	0	1	1	0

$=$

$R \circ S$	5	6	7	8	9
0	0	0	0	0	0
1	1	0	0	0	0
2	1	0	0	0	1
3	0	0	0	0	0
4	0	0	0	0	1

Beispiel 3.25 Wenn wir wie in Beispiel 3.16 Nr. 6 auf einer Menge von Personen die üblichen Verwandtschaftsbeziehungen einführen, so ist die Komposition der Vaterbeziehung mit der Elternbeziehung die Großvater-Relation. Die Komposition der Kinderbeziehung mit sich selbst ergibt die Enkelbeziehung.

Der einfache Beweis des nachfolgenden Lemmas wird als Übung dem Leser überlassen (vgl. Aufgabe 3.19).

Lemma 3.26 *Es seien A, B und C Mengen. Für jeden Index $i \in I \cup \{1, 2\}$ sei stets $R_i \subseteq A \times B$ und $S_i \subseteq B \times C$.*

1. *Ist $R_1 \subseteq R_2$ und $S_1 \subseteq S_2$, so auch $R_1 \circ S_1 \subseteq R_2 \circ S_2$.*
2. *Es sei $R \subseteq A \times B$. Dann ist $R \circ (\bigcup_{i \in I} S_i) = \bigcup_{i \in I}(R \circ S_i)$.*
3. *Es sei $S \subseteq B \times C$. Dann ist $(\bigcup_{i \in I} R_i) \circ S = \bigcup_{i \in I}(R_i \circ S)$.*

Lemma 3.27 *Die Komposition zweistelliger Relationen ist assoziativ: Sind $R \subseteq A \times B$, $S \subseteq B \times C$ und $T \subseteq C \times D$ Relationen, so gilt $R \circ (S \circ T) = (R \circ S) \circ T$.*

Beweis. Zunächst folgt aus Definition 3.23 unmittelbar, daß $R \circ (S \circ T) \subseteq A \times D$ und $(R \circ S) \circ T \subseteq A \times D$ gilt. Nun gilt für alle $a \in A$ und $d \in D$

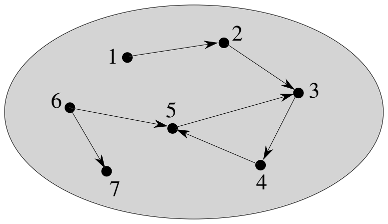
$$\begin{aligned}
\langle a,d\rangle \in R \circ (S \circ T) \quad &\Leftrightarrow \quad \exists b \in B\colon (\langle a,b\rangle \in R \wedge \langle b,d\rangle \in S \circ T) \\
&\Leftrightarrow \quad \exists b \in B, c \in C\colon (\langle a,b\rangle \in R \wedge \langle b,c\rangle \in S \wedge \langle c,d\rangle \in T) \\
&\Leftrightarrow \quad \exists c \in C\colon (\langle a,c\rangle \in R \circ S \wedge \langle c,d\rangle \in T) \\
&\Leftrightarrow \quad \langle a,d\rangle \in (R \circ S) \circ T
\end{aligned}$$

wodurch die Behauptung folgt. ■

Die folgende Begriffsbildung wird später verschiedentlich nützlich.

Definition 3.28 Es sei $R \subseteq A \times A$. Für $n \geq 0$ heißt eine Folge $a_0, \ldots, a_n$ von (nicht notwendigerweise verschiedenen) Elementen aus A eine *R-Kette* der Länge n von a_0 nach a_n genau dann, wenn $R(a_i, a_{i+1})$ für $i = 0, \ldots, n-1$ gilt. Eine R–Kette $a_0, \ldots, a_n$ heißt *R–Zyklus* genau dann, wenn $a_0 = a_n$ gilt. Ein R–Zyklus heißt *nicht-trivial* genau dann, wenn er zumindest zwei verschiedene Elemente enthält.

In der nachfolgenden Abbildung ist $1,2,3,4,5,3,4$ eine R–Kette der Länge 7 der durch Pfeile dargestellten Relation R. Es ist $3,4,5,3$ ein nichttrivialer R–Zyklus.



Lemma 3.29 *Es sei R eine zweistellige Relation auf A und $n \geq 1$. Dann gilt $R^n = \{\langle a,b\rangle \in A \times A \mid$ es ex. eine R–Kette der Länge n von a nach $b\}$.*

Der Beweis ist offenkundig und wird ausgelassen.

3.4 Funktionen

Funktionen sind spezielle Relationen. Während ein Element a im allgemeinen zu sehr vielen anderen Elementen $b_1, b_2 \ldots$ in einer vorgegebenen Relation R stehen kann, fordert man bei Funktionen, daß es höchstens ein solches Element b gibt, das man dann als *Bild* von a unter der Funktion bezeichnet. Fast immer betrachtet man Funktionen *von einer Menge A in eine Menge B*. Diese Sprechweise beinhaltet dann, daß auch *jedes* $a \in A$ ein Bild in B besitzt, welches dann eindeutig bestimmt ist.[2]

Definition 3.30 Eine zweistellige Relation $R \subseteq A \times B$ heißt *Funktion* oder *Abbildung* (engl. „function" oder „mapping") genau dann, wenn jedes Element von A höchstens ein Bild unter R hat, das heißt, wenn gilt

$$\forall x, y, z\colon ((R(x,y) \wedge R(x,z)) \Rightarrow y = z).$$

gilt. Die Menge $Def(R) := \{a \in A \mid \exists b \in B\colon R(a,b)\}$ wird dann der *Definitionsbereich* (engl. „domain"), B der *Bildbereich* (engl. „codomain") der Funktion R genannt.

Um Funktionen zu bezeichnen, verwendet man typischerweise Buchstaben wie f, g und ähnliche.

Definition 3.31 Ist $f \subseteq A \times B$ eine Funktion mit Definitionsbereich A, so heißt f *Funktion von A nach B*, man schreibt kurz $f\colon A \to B$. Für jedes $a \in A$ ist dann das *Bild* $b \in B$ mit $f(a,b)$ eindeutig bestimmt, wir schreiben $b = f(a)$. Die Elemente von A heißen auch die *Argumente* der Funktion, $f(a)$ heißt auch der *Wert* von a unter f. Hat speziell A die Form B^n, so heißt f eine *n–stellige Funktion auf B*.

Abbildung 3.1 illustriert diese Konzepte.

Bemerkung 3.32 Auf drei Punkte, die gelegentlich zu Unsicherheiten führen, sei besonders hingewiesen:

[2] Manchmal spricht man dann genauer von einer *totalen* Funktion, um zu betonen, daß *jedes* $a \in A$ ein Bild hat. Wenn man explizit erlauben will, daß manche Elemente kein Bild haben, so spricht man auch von einer *partiellen* Funktion.

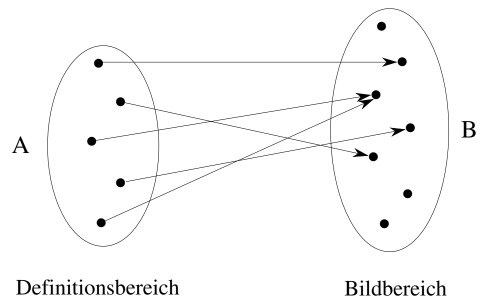


Abbildung 3.1: Beispiel einer Funktion $f\colon A \to B$.

1. Vorsicht: eine n-stellige Funktion auf B ist — als Relation betrachtet — $(n+1)$-stellig.

2. Es sei an dieser Stelle nochmals ausdrücklich darauf hingewiesen, daß jede Funktion f eine Menge geordneter Paare darstellt, genauer gilt stets $f = \{\langle x, f(x)\rangle \mid x \in \mathit{Def}(f)\}$.

3. Man beachte, daß Definition 3.31 ein gewisses Ungleichgewicht beinhaltet. Ist $f\colon A \to B$, so hat jedes $a \in A$ ein Bild unter f in B. Es braucht aber *nicht* jedes $b \in B$ als Bild aufzutreten, wie auch Abbildung 3.1 deutlich macht. Demzufolge ist der Bildbereich einer gegebenen Funktion f mit Definitionsbereich A leider nicht eindeutig. Jede Menge M, die die Menge $\{f(a) \mid a \in A\}$ umfaßt, ist im Prinzip ein möglicher Bildbereich. Bei konkreten Beispielen bietet sich jedoch oft in natürlicher Weise ein ganz bestimmter Bildbereich an.

Mit
$$B^A := \{f \mid f\colon A \to B\}$$
bezeichnen wir nachfolgend die Menge aller Funktionen von A nach B. Da die Notation oft verwendet wird, sollte man sie gut in Erinnerung behalten. Eine Motivation für die Schreibweise ergibt sich aus folgendem Lemma. Den Beweis, der vollständige Induktion verwendet, überlassen wir als Übung dem Leser (vgl. Aufgabe 3.22).

Lemma 3.33 *Sind A und B endliche Mengen mit n resp. m Elementen, so hat B^A genau m^n Elemente.*

Zur Darstellung von Funktionen gibt es zahlreiche Methoden. Funktionen auf kleinen endlichen Mengen werden oft in Form von Diagrammen wie

$$\begin{bmatrix} a & \mapsto & 1 \\ b & \mapsto & 3 \\ c & \mapsto & 2 \\ d & \mapsto & 1 \end{bmatrix}$$

repräsentiert, wobei allerdings der volle Bildbereich im allgemeinen nicht zu rekonstruieren ist. Eng verwandt ist die tabellarische Darstellung der Funktionswerte zu gegebenem Argument, die bei etwas umfangreicherem endlichen Definitionsbereich üblich ist. Natürlich ist für Funktionen mit endlichem Definitions- und Bildbereich auch eine Matrixdarstellung wie vorher bei den Relationen möglich (vergleiche Aufgabe 3.20). Oft werden Funktionen auch in Form einer Vorschrift spezifiziert, die angibt, wie das Bild eines Elements aussieht, wie im Beispiel

$$f\colon \mathbb{N} \to \mathbb{N}; x \mapsto 2x$$

wo $x \mapsto 2x$ besagt, daß unter f jedes Element auf sein Zweifaches abgebildet wird.

Beispiele 3.34 Nachfolgend nun eine längere Liste von Funktionen, wobei die Beispiele aus dem Leben zum Teil streng genommen eine striktere Formalisierung voraussetzen. Die ersten beiden Beispiele sollen dazu dienen, zwei unterschiedliche Intuitionen, die mit dem Funktionsbegriff verbunden sind, zu verdeutlichen.

1. Eine Klasse von Funktionen läßt sich auf der intuitiven Ebene dadurch beschreiben, daß man für eine Eingabe einen bestimmten Wert berechnen möchte. Wenn zum Beispiel danach gefragt wird, wieviel Geld man bei einem Anfangskapital von 1000 DM und einem Zinssatz von 2% nach $n = 0, 1, 2, \ldots$ Jahren angespart hat, so wird die Antwort durch eine Funktion von $\mathbb{N}$ in die rationalen Zahlen beschrieben. Sehr viele weitere numerische Funktionen, etwa zur Inhalts- oder Flächenberechnung, sind ihrer Natur und Verwendung nach ähnlich.

2. Bei einer anderen Klasse von Funktionen steht der Gedanke im Vordergrund, in Form der Funktion eine Art *Korrespondenz* zwischen Argumenten und ihren Bildern zu formalisieren, und es macht kaum Sinn, von der „Berechnung eines Werts“ zu sprechen. In einer Schließanlage

können wir jedem Schlüssel das zugehörige Schließfach zuweisen. Dies definiert eine Funktion von der Menge der Schlüssel in die Menge der Schließfächer. Verwandt ist die Festlegung einer Reihenfolge, wo wir eine Korrespondenz zwischen den Elementen einer Menge M und einem Anfangsintervall $A = [1, 2, \ldots, n]$ von $\mathbb{N}$ erhalten, die formal durch eine Funktion $M \to A$ erfaßt wird. Für Funktionen dieser Art ist der (formal synonyme!) Begriff „Abbildung" näher an der Intuition.

3. Jede Klassifikationen, bei der jedes Objekt einer gegebenen Menge M genau einer Klasse einer Menge von Klassen K zugeordnet wird, repräsentiert eine Funktion von M in K.

4. Ist M eine Menge, so sind Vereinigung „$\cup$" und Durchschnitt „$\cap$" zweistellige Funktionen auf der Potenzmenge $\mathcal{P}(M)$. Komplementbildung bezüglich M (vgl. Def. 2.15) ist eine einstellige Funktion auf $\mathcal{P}(M)$.

5. Ist A^* die in Definition 3.9 beschriebene Menge der Wörter über dem Alphabet A, so ist die *Konkatenation*„$\circ$" von Wörtern eine zweistellige Funktion auf A^*. Die Konkatenation wird in Infixschreibweise notiert, und ist durch $a_1 \cdots a_n \circ b_1 \cdots b_m := a_1 \cdots a_n b_1 \cdots b_m$ definiert. Zum Beispiel ergibt die Konkatenation von „*Computer*" mit „*linguistik*" das Wort „*Computerlinguistik*". Mittels des Konkatenationsbegriffs könnnen wir die Präfixrelation und die Suffixrelation (vgl. Beispiel 3.16 Nr. 1) formal definieren: Das Wort $v \in A^*$ ist ein *Präfix* (resp. *Suffix*) des Wortes $w \in A^*$ genau dann, wenn ein Wort $u \in A^*$ existiert, so daß $w = v \circ u$ (resp. $w = u \circ v$) gilt.

6. Es sei A eine Menge und $\mathcal{R}(A)$ die Menge aller zweistelligen Relationen auf A. Die in Definition 3.23 eingeführte Komposition „$\circ$" von Relationen ist eine zweistellige Funktion auf $\mathcal{R}(A)$.

7. Es bezeichne $\mathcal{L}(A)$ die in Definition 3.11 eingeführte Menge aller Sprachen über dem Alphabet A. Für Sprachen $L_1, L_2 \in \mathcal{L}(A)$ heißt

$$L_1 \odot L_2 := \{u \circ v \mid u \in L_1, v \in L_2\}$$

die *Komposition* von L_1 und L_2. Dabei bezeichnet „$\circ$" die in Beispiel 5 erklärte Konkatenation. Die Komposition „$\odot$" ist eine zweistellige Funktion auf $\mathcal{L}(A)$. Wir haben nur der Verständlichkeit halber ein neues Symbol verwendet, oft wird auch für die Komposition von Sprachen das Symbol „$\circ$" verwendet. Es ist leicht zu sehen, daß die

Komposition von Sprachen assoziativ ist, das heißt es gilt für beliebige Sprachen L_1, L_2, L_3 über A stets $(L_1 \odot L_2) \odot L_3 = L_1 \odot (L_2 \odot L_3)$. Die Komposition $L_1 \odot L_2 \odot \cdots \odot L_5$ der in Beispiel 3.12 Nr. 2 erwähnten Sprachen $L_1, \ldots, L_5$ ergibt die dort beschriebene Menge aller Datumsangaben des dritten Jahrtausends.

8. Die Alters-Relation, zu der ein Paar $\langle p, n \rangle$ gehört genau dann, wenn n das Alter von p ist, hat funktionalen Charakter. Ähnlich werden in behördlichen Formularen sehr viele Werte weiterer „Funktionen" abgefragt: *Beruf*, *Adresse*, *Religion*, *Steuerklasse*, *Zahl-der-Kinder* etc. Sind im Gegensatz dazu mehrere Antworten möglich, wie etwa bei *Kinder* oder *Erziehungsberechtigt*, so hat die Abfrage relationalen Charakter.

9. Für das n-stellige kartesische Produkt $A_1 \times \cdots \times A_n$ definieren wir die n *Projektionsfunktionen*

$$\pi_i \colon A_1 \times \cdots \times A_n \to A_i \colon \langle a_1, \ldots, a_n \rangle \mapsto a_i \ (1 \leq i \leq n).$$

   Kartesische Produkte und zugehörige Projektionsfunktionen zeichnen sich durch die folgende Universalitätseigenschaft aus: ist M eine Menge, und sind Funktionen $f_i \colon M \to A_i$ für $1 \leq i \leq n$ vorgegeben, so gibt es eine eindeutig bestimmte Funktion $g \colon M \to \Pi_{i=1}^n A_i$, so daß $f_i = g \circ \pi_i$ für $1 \leq i \leq n$ gilt (vgl. Aufgabe 3.7).

10. Addition und Multiplikation (Sinus und Cosinus) sind zweistellige (einstellige) Funktionen auf der Menge $\mathbb{R}$ der reellen Zahlen. Die Einschränkung der Addition von $\mathbb{R}$ auf die Menge $\mathcal{Q}$ der rationalen Zahlen ist eine zweistellige Funktion auf $\mathcal{Q}$, nämlich die übliche Addition auf $\mathcal{Q}$. Beide genannten Additionsfunktionen sind als Mengen unterschiedlich, obwohl in der Regel dasselbe Symbol „+" verwendet wird.

11. Die in der Informatik verwendeten „Arrays" weisen (im eindimensionalen Fall) jedem Wert i aus einem Anfangsintervall $[1, n] = \{1, 2, \ldots, n\}$ der natürlichen Zahlen einen Eintrag aus einer Menge B zu. Damit kann man ein Array auffassen als Modellierung einer Funktion $f \colon [1, n] \to B$. Man beachte auch die Verwandtschaft zu den oben angegebenen Diagrammdarstellungen von endlichen Funktionen.

12. Bei geeigneter Betrachtungsweise definiert jede binäre Relation $R \subseteq A \times B$ eine Funktion von $\mathcal{P}(A)$ nach $\mathcal{P}(B)$: in der Tat, auf der Basis von Definition 3.17 ordnet R jeder Teilmenge X von A genau eine Teilmenge $R(X)$ von B zu. Analog ist auch $R^{-1} \colon \mathcal{P}(B) \to \mathcal{P}(A)$ eine Funktion.

13. Mittels des Begriffs der Funktion können wir nun *Multimengen* (vgl. Abschnitt 2.7.2) formalisieren. Eine *Multimenge* ist eine Abbildung $m_A \colon A \to \mathbb{N}$. Intuitiv gibt das Bild eines Elements $a \in A$ unter der Abbildung m_A die Zahl seiner Vorkommen in der Multimenge an.

Hier noch ein weiteres Beispiel, das später verschiedentlich auftritt.

Definition 3.35 Es sei A eine Menge. Die Funktion $Id_A := \{\langle a, a\rangle \mid a \in A\}$ wird *Identität auf* A (auch *Identitätsrelation* oder *Identitätsfunktion*) genannt.

Beispiel 3.36 Es sei $A := \{1, 2, 3, 4, 5\}$. Dann kann Id_A in Matrixform wie folgt dargestellt werden.

R	1	2	3	4	5
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	0	0	1	0
4	0	0	0	0	1

Dies verdeutlicht, warum Id_A manchmal die *Diagonale* von A genannt wird.

Mit den nun eingeführten Schreibkonventionen können wir die in den Definitionen 3.17 und 3.20 eingeführten Begriffe wie folgt für Funktionen darstellen.

Definition 3.37 Sei $f \colon A \to B$ eine Funktion, $X \subseteq A, Y \subseteq B$.

(a) Die Menge $f(X) := \{f(x) \mid x \in X\}$ heißt das *Bild* von X unter der Funktion f.

(b) Die Menge $f^{-1} := \{\langle y, x\rangle \mid x \in A, f(x) = y\} \subseteq B \times A$ heißt *Umkehrrelation* zur Funktion f. Die Menge $f^{-1}(Y)$ heißt das *Urbild* von Y unter f.

(c) Die Menge $f\lceil X := \{\langle x, y\rangle \mid x \in X, f(x) = y\}$ heißt die *Einschränkung* von f auf X. Wir fassen $f\lceil X$ auf als eine Funktion mit Definitionsbereich X und Bildbereich B.

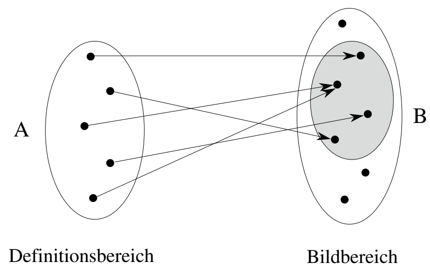


Abbildung 3.2: Unterschied zwischen Bildbereich und Bild des Definitionsbereichs.

Man beachte, daß das in (a) erklärte Bild von A unter f vom „Bildbereich“ B von f durchaus verschieden sein kann, wie Abbildung 3.2 illustriert. Die Wortwahl in Teil (b) macht deutlich, daß die Umkehr*relation* f^{-1} einer Funktion f im allgemeinen keine Funktion ist! Auch dieser Sachverhalt läßt sich aus Abbildung 3.2 leicht ersehen. Wir werden in Kapitel 3.5 spezielle Bedingungen dafür angeben, daß auch f^{-1} wieder eine Funktion ist.

Definition 3.38 Eine einstellige Funktion $f\colon A \to A$ heißt

1. *nilpotent* genau dann, wenn $f \circ f = Id_A$ gilt,
2. *idempotent* genau dann, wenn $f \circ f = f$ gilt.

Offenkundig läßt sich nun Id_A dadurch charakterisieren, daß es die einzige Funktion von A in A ist, die gleichzeitig nilpotent und idempotent ist.

Ähnlich wie bei Mengen können auch Funktionen, die in unterschiedlicher Weise definiert wurden, gleich sein. Das folgende Lemma beinhaltet ein Standardverfahren, wie man die Gleichheit zweier Funktionen nachweisen kann.

Lemma 3.39 (Extensionalitätsprinzip für Funktionen) *Zwei Funktionen f und g sind genau dann gleich, $f = g$, wenn sie denselben Definitionsbereich haben und für dieselben Argumente dieselben Bilder ergeben.*

Beweis. Wie in Bemerkung 3.32 Teil 2 festgehalten, gilt stets

$$f \quad = \quad \{\langle x, f(x)\rangle \mid x \in Def(f)\}$$

$$g \quad = \quad \{\langle x, g(x)\rangle \mid x \in \mathit{Def}(g)\}.$$

Es sei nun $f = g$. Ist $a \in \mathit{Def}(f)$, so folgt aus $\langle a, f(a)\rangle \in f$ auch $\langle a, f(a)\rangle \in g$, das heißt, es gilt $a \in \mathit{Def}(g)$. Damit folgt $\mathit{Def}(f) \subseteq \mathit{Def}(g)$. Umgekehrt folgt analog $\mathit{Def}(g) \subseteq \mathit{Def}(f)$, das heißt, es gilt $\mathit{Def}(f) = \mathit{Def}(g)$. Aus $\langle a, f(a)\rangle \in f = g$ und $\langle a, g(a)\rangle \in g$ folgt $f(a) = g(a)$, da g Funktion ist.

Es gelte umgekehrt, daß f und g denselben Definitionsbereich haben und für dieselben Argumente dieselben Bilder ergeben. Sei $\langle a, b\rangle \in f$. Dann ist $a \in \mathit{Def}(f) = \mathit{Def}(g)$. Da $b = f(a)$ folgt nach Voraussetzung $b = g(a)$. Also gilt $\langle a, b\rangle \in g$. Wir haben gezeigt, daß $f \subseteq g$. Umgekehrt folgt analog $g \subseteq f$, und damit mit Lemma 2.6 (ii) wie gewünscht $f = g$. ■

Das nachfolgende Lemma macht Aussagen über die Bilder von Teilmengen unter einer Funktion. Wir greifen auf die in Definition 3.37, Teil (a), entwickelte Notation zurück.

Lemma 3.40 *Sei $f\colon A \to B$. Dann gilt stets*

$$\begin{array}{rl}
(i) & \forall X, X' \subseteq A\colon f(X \cup X') = f(X) \cup f(X'), \\
(ii) & \forall X, X' \subseteq A\colon f(X \cap X') \subseteq f(X) \cap f(X'), \\
(iii) & \forall X \subseteq A\colon X \subseteq f^{-1}(f(X)), \\
(iv) & \forall Y \subseteq B\colon f(f^{-1}(Y)) \subseteq Y.
\end{array}$$

Beweis. Teil (i) wird als Übung offengelassen (vgl. Aufgabe 3.24). Zu Teil (ii): Es sei $X, X' \subseteq A$. Dann gilt für beliebiges b

$$\begin{array}{rcl}
b \in f(X \cap X') & \Leftrightarrow & \exists a \in X \cap X'\colon f(a) = b \\
& \Leftrightarrow & \exists a\colon a \in X \wedge a \in X' \wedge f(a) = b \\
& \Rightarrow & \exists a \in X\colon f(a) = b \ \wedge \ \exists a' \in X'\colon f(a') = b \\
& \Leftrightarrow & b \in f(X) \cap f(X').
\end{array}$$

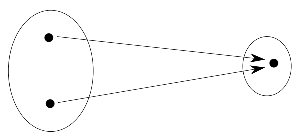
Hieraus folgt die Behauptung.

Zu Teil (iii): Sei $X \subseteq A$ und $x \in X$. Dann ist $f(x) \in f(X)$. Somit gilt $\{f(x)\} \subseteq f(X)$ und nach Lemma 3.19 auch $f^{-1}(\{f(x)\}) \subseteq f^{-1}(f(X))$. Trivialerweise gilt $x \in f^{-1}(\{f(x)\})$. Damit folgt auch $x \in f^{-1}(f(X))$.

Zu Teil (iv): Es sei $Y \subseteq B$ und $b \in f(f^{-1}(Y))$. Dann existiert gemäß Definition 3.37, Teil (a), ein Element $a \in f^{-1}(Y)$ mit $f(a) = b$. Da $a \in f^{-1}(Y)$

existiert nach Definition 3.17, Teil (a), ein Element $y \in Y$ mit $\langle y, a\rangle \in f^{-1}$. Aus $\langle y, a\rangle \in f^{-1}$ ergibt sich nach Definition 3.37, Teil (d), daß $f(a) = y$. Da aber f Funktion ist, folgt aus $f(a) = b$ und $f(a) = y$ nun $b = y \in Y$. ■

In der vorletzten Zeile im Beweis von (ii) hatten wir wieder sicherheitshalber mit zwei Variablennamen gearbeitet, damit klarer wird, daß man die Richtung der zugehörigen Implikation nicht notwendig umkehren kann. In der Tat gilt im allgemeinen *nicht* die Identität $f(X \cap X') = f(X) \cap f(X')$, wie folgendes einfache Beispiel zeigt (warum?):



Der Leser möge sich ähnliche Beispiele ausdenken, die zeigen, daß man auch in den Aussagen (iii) und (iv) im allgemeinen keine Gleichheit hat.

Da Funktionen spezielle Relationen sind, können wir auch die Komposition von Funktionen betrachten. Das folgende Lemma zeigt, daß die Komposition zweier Abbildungen gerade die Hintereinanderausführung der Abbildungen darstellt.

Lemma 3.41 *Es seien $f\colon A \to B$ und $g\colon B \to C$ Funktionen. Dann ist auch $f \circ g\colon A \to C$ eine Funktion, deren Werte durch $(f \circ g)(a) = g(f(a))$ gegeben sind $(a \in A)$.*

Beweis. Zunächst folgt aus Definition 3.23 sofort, daß $(f \circ g) \subseteq A \times C$ gilt. Um zu zeigen, daß $f \circ g$ eine Funktion ist, müssen wir nachweisen, daß f jedem Element aus A genau ein Bild aus C zuordnet.

Wir zeigen zuerst („Existenz“), daß f jedem Element $a \in A$ zumindest ein Bild aus C zuordnet. In der Tat, mit $\langle a, f(a)\rangle \in f$ und $\langle f(a), g(f(a))\rangle \in g$ folgt nach Definition 3.23 $\langle a, g(f(a))\rangle \in (f \circ g)$.

Es bleibt noch die „Eindeutigkeit“ der Zuordnung zu zeigen. Seien $c, c' \in C$, und es gelte $\langle a, c\rangle \in (f \circ g)$ sowie $\langle a, c'\rangle \in (f \circ g)$. Dann existieren Elemente $b, b' \in B$ mit $\langle a, b\rangle \in f$ und $\langle b, c\rangle \in g$ sowie $\langle a, b'\rangle \in f$ und $\langle b', c'\rangle \in g$. Da f Funktion ist, folgt $b = b'$. Da g Funktion ist, folgt $c = c'$. Wir haben auch die Eindeutigkeit des Bildes nachgewiesen. Zusammenfassend haben wir gezeigt, daß $f \circ g\colon A \to C$ Funktion ist.

Alle Paare $\langle a, g(f(a))\rangle$ (für $a \in A$) gehören zu $f \circ g$, wie wir im „Existenzteil" gesehen haben. Damit ist $g(f(a))$ das eindeutig bestimmte Bild von a unter der Funktion $f \circ g$. ■

Aus Lemma 3.27 folgt sofort als Spezialfall

Lemma 3.42 *Die Komposition von Funktionen ist assoziativ, für* $f\colon A \to B$, $g\colon B \to C$ *und* $h\colon C \to D$ *gilt stets* $f \circ (g \circ h) = (f \circ g) \circ h$.

3.5 Injektivität, Surjektivität und Bijektivität

Wir kommen zu einigen wichtigen Eigenschaften von Funktionen.

Definition 3.43 Eine Funktion $f\colon A \to B$ heißt *injektiv* (oder *eineindeutig*) genau dann, falls verschiedene Argumente stets verschiedene Werte unter f haben, das heißt, falls gilt

$$\forall x, y \in A\colon (f(x) = f(y) \Rightarrow x = y).$$

Eine injektive Funktion wird auch *Injektion* genannt. Wenn es eine injektive Funktion f von A nach B gibt, so bedeutet dies, daß B „bis auf Namensgebung" eine Kopie der Elemente von A enthält. Das Bild $f(A)$ hat dann nämlich dieselbe Anzahl von Elementen wie A selbst, wie in Abbildung 3.3 veranschaulicht wird. Der in Beispiel 3.34 Nr. 2 erwähnte „Korrespondenzgedanke" tritt hier offen zutage. Bezeichnen wir für eine endliche Menge M mit $|M|$ die Anzahl der Elemente von M (auch *Kardinalität* von M genannt), so gilt also:

Lemma 3.44 *Es seien* A *und* B *endliche Mengen. Gibt es eine injektive Funktion* $f\colon A \to B$, *so folgt* $|A| \leq |B|$. *Die Umkehrung gilt ebenfalls.*

Eine Umformulierung diese Lemmas ist in der Literatur unter der Bezeichung „pigeonhole principle" (Taubenloch-Prinzip) bekannt: gilt $|A| > |B|$, so kann es keine injektive Funktion von A nach B geben. Oder suggestiver: will man m Tauben auf n Löcher verteilen, und gilt $m > n$, so muß man zumindest in ein Loch zwei oder mehr Tauben stecken. Diese Prinzip hat als Beweismittel viele Anwendungen. Hierzu vergleiche man Aufgabe 3.27 am Kapitelende.

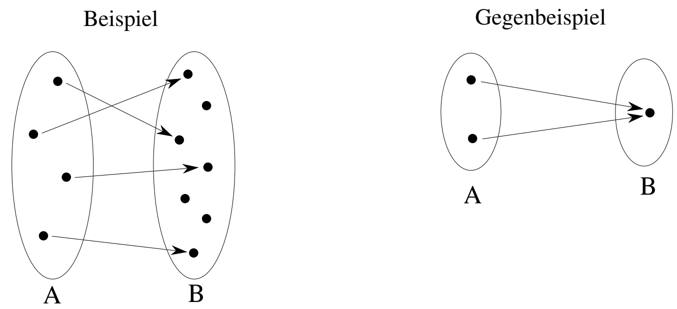
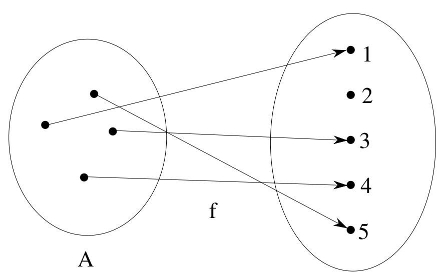


Abbildung 3.3: Injektivität von Funktionen.

Beispiele 3.45 Die Sitzordnung im Hörsaal definiert eine injektive Funktion von der Menge der Zuhörer auf die Menge der Plätze im Hörsaal. In einem Konzertsaal definiert die Platznummer eine injektive Funktion der Menge der Plätze in $\mathbb{N}$.

Beispiel 3.46 Injektive Funktionen $f\colon A \to B$ werden häufig in Situationen verwendet, wo wir auf Elemente aus A referieren wollen, sprachlich aber besser auf die Elemente von B zugreifen können. Betrachten wir als Beispiel das folgende Bild einer injektiven Funktion $f\colon A \to \{1, 2, 3, 4, 5\}$.



Obwohl wir kaum in direkter Weise über die Elemente aus A reden können, geht dies problemlos unter Zuhilfenahme der Funktion f. Das „auf 1 abgebildete Element“ beispielweise ist durch diese Beschreibung eindeutig charakterisiert. Beispiele dieser Art begegnen uns in vielen Zusammenhängen. Alle Definitionen, Beispiele, Lemmata und Sätze in diesem Buch sind mit einer Nummer versehen. Die Nummerierung liefert eine injektive Abbildung f von der Menge A der Definitionen, Beispiele, Lemmata und Sätze in die

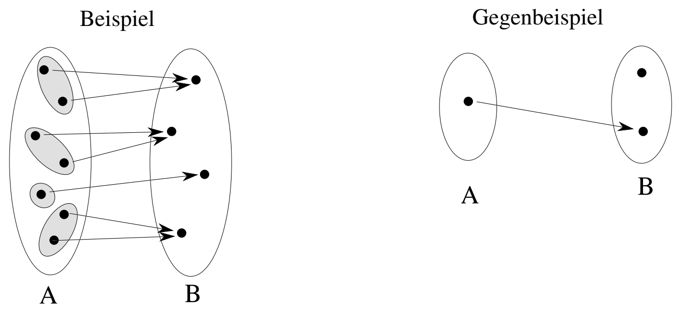


Abbildung 3.4: Surjektivität von Funktionen.

Menge B der in Frage kommenden Nummern. Wenn wir auf ein Element aus A referieren, etwa auf Lemma 3.22, so verwenden wir hierzu nicht das Element, sondern sein Bild unter f in B. Ein Nachteil ist, daß dasselbe Lemma in unterschiedlichen Büchern in aller Regel verschiedene Nummern trägt. Aus diesem Grund werden wirklich wichtige Prinzipien oft noch in anderer Weise benannt. Beispielsweise hatten wir den — allgemein üblichen — Namen „Extensionalitätsprinzip für Funktionen“ für den in Lemma 3.39 dargestellten Sachverhalt eingeführt. Viele wichtige Sätze in der Mathematik sind nach ihren Entdecker benannt.

Definition 3.47 Eine Funktion $f\colon A \to B$ heißt *surjektiv* (oder Funktion von A *auf* B) genau dann, wenn jedes $b \in B$ als Bild unter f auftritt, das heißt, falls gilt

$$\forall b \in B\ \exists a \in A\colon f(a) = b.$$

Eine surjektive Funktion wird auch *Surjektion* genannt. Wenn es eine surjektive Funktion f von A auf B gibt, so bedeutet dies, daß man aus A mittels einer geeigneten „Verschmelzung“ von Elementen eine Menge erhält, die ebensoviele Elemente wie B enthält. Hierzu identifiziert man genau diejenigen Elemente von A, die dasselbe Bild unter f haben. Abbildung 3.4 verdeutlicht diesen Zusammenhang. Damit zusammenhängend wird deutlich, daß die Menge

$$\{f^{-1}(\{b\}) \mid b \in B\}$$

eine Partition von A ist. Bezüglich der Zahl der Elemente von A und B gilt die folgende Aussage.

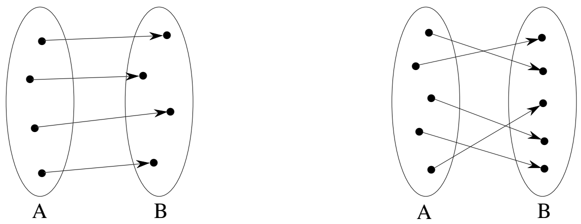


Abbildung 3.5: Beispiele bijektiver Funktionen.

Lemma 3.48 *Es seien A und B endliche Mengen. Gibt es eine surjektive Funktion $f\colon A \to B$, so folgt $|A| \geq |B|$. Die Umkehrung gilt ebenfalls.*

Definition 3.49 Eine Funktion $f\colon A \to B$ heißt *bijektiv* genau dann, wenn f sowohl injektiv als auch surjektiv ist.

Eine bijektive Funktion wird auch *Bijektion* genannt. Eine Bijektion einer Menge M auf sich selbst wird auch *Permutation* von M genannt. Abbildung 3.5 verdeutlicht das Konzept der Bijektion. Der in Beispiel 3.34 Nr. 2 erwähnte „Korrespondenzgedanke“ begegnet uns hier in der reinsten Form, da bei bijektiven Funktionen die Korrespondenz *alle* Elemente des Bildbereichs miteinschließt.

Lemma 3.50 *Es seien A und B endliche Mengen. Gibt es eine bijektive Funktion $f\colon A \to B$, so folgt $|A| = |B|$. Die Umkehrung gilt ebenfalls.*

Beispiel 3.51 Sind bei einer Vorlesung alle Plätze besetzt (ohne daß sich zwei Zuhörer einen Platz teilen), so definiert die Sitzordnung eine Bijektion zwischen der Menge der Zuhörer und der Menge der Plätze.

Beispiel 3.52 Dieses Beispiel präzisiert die Argumentation im Beweis von Lemma 2.27. Es sei A eine nichtleere Menge und $a \in A$. Definieren wir $\mathcal{P}_1 := \{M \subseteq A \mid a \notin M\}$ und $\mathcal{P}_2 := \{M \subseteq A \mid a \in M\}$, so stellt die Abbildung $f\colon M \mapsto M \cup \{a\}$ eine Bijektion zwischen $\mathcal{P}_1$ und $\mathcal{P}_2$ dar. In der Tat, aus $f(M_1) = f(M_2) = N$ folgt $M_1 = M_2 = N \setminus \{a\}$, woraus sich die Injektivität ergibt. Ist $N \in \mathcal{P}_2$ so gilt $N \setminus \{a\} \in \mathcal{P}_1$ und $f(N \setminus \{a\}) = N$, wodurch sich die Surjektivität ergibt. Aus Lemma 3.50 folgt nun im Fall, wo A endlich ist, daß $|\mathcal{P}_1| = |\mathcal{P}_2|$.

Die folgenden Lemmata stellen eine Verbindung her zwischen Injektivität und Surjektivität einerseits und den Aussagen (ii) und (iv) aus Lemma 3.40 andererseits:

Lemma 3.53 *Sei $A \neq \emptyset$ und $f\colon A \to B$. Die folgenden Aussagen sind äquivalent:*

(*i*) *f ist injektiv,*

(*ii*) $\forall X, X' \subseteq A\colon f(X \cap X') = f(X) \cap f(X')$,

(*iii*) *f besitzt ein Rechtsinverses, das heißt, es gibt eine Funktion $g\colon B \to A$ so daß $f \circ g = Id_A$.*

Beweis. „(i) $\Rightarrow$ (ii)“: Sei f injektiv. Sind nun $X, X' \subseteq A$ so gilt nach Lemma 3.40 (ii) stets $f(X \cap X') \subseteq f(X) \cap f(X')$. Es bleibt daher die umgekehrte Inklusion $f(X) \cap f(X') \subseteq f(X \cap X')$ zu zeigen. Sei $b \in f(X) \cap f(X')$. Dann gilt

$$\exists x \in X\colon b = f(x) \ \wedge \ \exists x' \in X'\colon b = f(x')$$

Da f injektiv ist, folgt $x = x' \in X \cap X'$ und somit $b \in f(X \cap X')$.

„(ii) $\Rightarrow$ (iii)“: Es gelte Eigenschaft (ii). Wir zeigen zunächst, daß f^{-1} eine Funktion ist. Dazu seien $\langle y, x_1\rangle$ und $\langle y, x_2\rangle$ in f^{-1}. Wir müssen $x_1 = x_2$ zeigen. Es sind $\langle x_1, y\rangle$ und $\langle x_2, y\rangle$ in f, nach Definition von f^{-1}. Wäre nun $x_1 \neq x_2$, so wäre $\{x_1\} \cap \{x_2\} = \emptyset$ und wegen (ii) somit

$$\{y\} = f(\{x_1\}) \cap f(\{x_2\}) = f(\{x_1\} \cap \{x_2\}) = f(\emptyset) = \emptyset,$$

was unmöglich ist. Daher ist tatsächlich $x_1 = x_2$. Somit ist f^{-1} eine Funktion. Es ist klar, daß $Def(f^{-1}) = f(A)$. Wir wählen nun irgendein $a \in A$, was geht, da $A \neq \emptyset$ nach Voraussetzung. Durch die Definition

$$g(y) = \begin{cases} f^{-1}(y) & \text{für } y \in f(A) \\ a & \text{für } y \in B \setminus f(A) \end{cases}$$

erhalten wir die Funktion $g\colon B \to A$. Es sei $x \in A$ gegeben. Es gilt

$$(f \circ g)(x) = g(f(x)) = f^{-1}(f(x)) = x,$$

damit ist (iii) gezeigt.

„(iii) $\Rightarrow$ (i)“: Wir nehmen an, daß es eine Funktion $g\colon B \to A$ mit $(f \circ g)(x) = x$ für alle $x \in A$ gibt . Es seien $x_1, x_2 \in A$ und $f(x_1) = f(x_2)$. Damit folgt $g(f(x_1)) = g(f(x_2)) = x_1 = x_2$. Somit ist f injektiv. ■

Der obige Beweis verwendet einen sogenannten *Ringschluß*: obwohl wir nur drei Richtungen gezeigt haben, folgt nun sofort, daß tatsächlich alle Aussagen äquivalent sind. Alle nicht explizit gezeigten Richtungen folgen wegen der Transitivität der Implikation.

Lemma 3.54 *Sei $f\colon A \to B$. Es ist f^{-1} eine Funktion genau dann, wenn f injektiv ist. In diesem Fall ist auch f^{-1} wieder injektiv, und jedes $a \in A$ tritt als Bild unter f^{-1} auf. Ist f außerdem surjektiv, so ist $f^{-1}\colon B \to A$ eine Bijektion und es gilt $f \circ f^{-1} = Id_A$ und $f^{-1} \circ f = Id_B$.*

Beweis. Übung. ■

Lemma 3.55 *Es sei $f\colon A \to B$. Die folgenden Aussagen sind äquivalent:*

(i) f *ist surjektiv,*

(ii) $\forall Y \subseteq B : f(f^{-1}(Y)) = Y$,

(iii) f *hat ein Linksinverses, das heißt, es gibt eine Funktion* $g\colon B \to A$ *mit* $g \circ f = Id_B$.

Der Nachweis der Äquivalenz von (i) und (ii) wird als Übung empfohlen. Die Äquivalenz zu (iii) ist schwieriger[3], daher verzichten wir auf einen Beweis.

Lemma 3.56 *Es seien $f\colon A \to B$ und $g\colon B \to C$ Funktionen.*

(i) Sind f und g injektiv, so ist auch $f \circ g$ injektiv.

(ii) Sind f und g surjektiv, so ist auch $f \circ g$ surjektiv.

(iii) Sind f und g bijektiv, so ist auch $f \circ g$ bijektiv und es gilt

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1},$$

(iv) Ist $f \circ g$ bijektiv, so ist g surjektiv und f injektiv.

Beweis. Übung. ■

[3] Sie benötigt das sogenannte „Auswahlaxiom“.

3.6 Ergänzungen

Zum Schluß dieses Kapitels werden drei Themenbereiche angerissen, von denen die letzten beiden beim ersten Lesen auch überschlagen werden können. Im ersten Teil, der ungleich wichtiger ist, erläutern wir das Prinzip der (simultanen) induktiven Definition von Mengen, das wir bereits in Beispielen verwendet hatten. Da diese Form der Definition von Mengen gerade in der Informatik und Linguistik häufig auftritt, ist sie von großer Bedeutung. Im zweiten Teil zeigen wir, wie man mit Hilfe von Funktionen auch Teilmengen und Relationen beschreiben kann. Danach verwenden wir den Funktionsbegriff, um auch unendliche kartesische Produkte einzuführen.

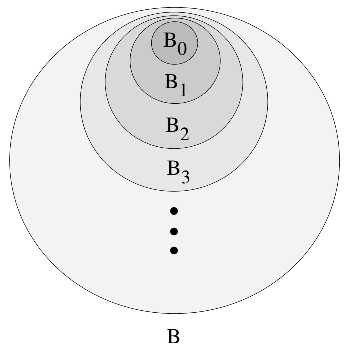
3.6.1 Induktive Definition von Mengen

Eine sogenannte *induktive Definition* einer Menge B besteht aus einer endlichen Menge von Regeln, wobei jede Regel entweder eine *Basisregel* ist oder eine *induktive Regel*.

- Eine Basisregel gibt an, daß bestimmte Elemente zu B gehören, oder daß eine bereits definierte und bekannte Menge Teilmenge von B ist.

- Eine induktive Regel hat als *Prämisse* eine endliche Zahl von Elementschaftsbedingungen, die besagen, daß bestimmte Elemente $b_1, \ldots, b_n$ zu B oder zu bekannten und vorgegebenen Mengen gehören sollen. Die *Konklusion* besagt, daß dann ein mit Hilfe der Elemente $b_1, \ldots, b_n$ zu definierendes Element b zu B gehört.

Die auf diese Weise definierte Menge B enthält genau diejenigen Elemente, deren Zugehörigkeit zu B durch endlich viele Anwendungen der Regeln verifiziert werden kann. Anschaulich kann man sich die Menge B mit Hilfe der folgenden Figur vorstellen:



Die Basisregeln definieren eine erste Teilmenge B_0 von B. Durch Anwenden der induktiven Regeln auf die bereits erhaltenen Elemente in B_0 und auf die Elemente der bekannten und vorgegebenen Mengen erhalten wir (in der Regel) neue Elemente. Durch Zusammenfassen von B_0 mit diesen neuen Elementen erhalten wir die Obermenge B_1. Durch Anwenden der induktiven Regeln auf die bereits erhaltenen Elemente in B_1 und auf die Elemente der bekannten und vorgegebenen Mengen erhalten wir (in der Regel) neue Elemente, die wir mit B_1 zur Obermenge B_2 zusammenfassen etc. Wenn wir diese Erweiterungsschritte unendlich oft wiederholen, erhalten wir schließlich die gesuchte Menge B als Zusammenfassung aller Mengen B_i $(i \in \mathbb{N})$.

Beispiel 3.57 Die Menge G der geraden natürlichen Zahlen kann induktiv etwa wie folgt definiert werden. Die einzige Basisregel sei

$$0 \in G.$$

Die einzige induktive Regel besagt

$$\text{Ist } n \in G, \text{ so ist auch } n + 2 \in G.$$

Eine Kurzschreibweise, die beide Regeln zusammenfaßt, ist

$$G \ ::= \ 0 \mid G + 2.$$

Im obigen Bild erhalten wir zunächst die „Basismenge“ $G_0 = \{0\}$. Im nächsten Schritt ergibt sich $G_1 = G_0 \cup \{g + 2 \mid g \in G_0\} = \{0, 2\}$. Hierbei repräsentiert G_0 die zu diesem Zeitpunkt bereits erhaltene Teilmenge von G, die Menge $\{g+2 \mid g \in G_0\}$ stellt die Menge derjenigen Elemente dar, die sich durch Anwenden der induktiven Regel auf die bisherigen Elemente ergeben. Im dritten Schritt erhalten wir analog $G_2 = G_1 \cup \{g+2 \mid g \in G_1\} = \{0, 2, 4\}$ etc. Durch Vereinigen aller so konstruierbaren Mengen $G_0, G_1, G_2, \ldots$ erhalten wir die gesuchte Menge G.

Beispiel 3.58 Es sei A eine Menge. Dann kann die in Definition 3.9 eingeführte Menge A^* aller Wörter über A wie folgt induktiv definiert werden.

(1) Das leere Wort „ϵ“ gehört zu A^*.

(2) Ist $w \in A^*$ und $a \in A$, so ist auch $w \circ a \in A^*$

Hierbei ist (1) Basisregel, (2) eine induktive Regel, „$\circ$“ steht für die in Beispiel 3.34 Nr. 5 definierte Konkatenation. Dieselbe Definition in Kurzschreibweise wäre

$$A^* \ ::= \ \epsilon \mid A^* \circ A.$$

Im obigen Bild erhalten wir hier zunächst die Basismenge $\{\epsilon\} = A^0$. Im nächsten Schritt ergibt sich $A^0 \cup \{w \circ a \mid w \in \{\epsilon\}, a \in A\} = A^0 \cup A^1$. Hierbei repräsentiert $\{\epsilon\} = A^0$ die zu diesem Zeitpunkt bereits erhaltene Teilmenge von A^*, die Menge $\{w \circ a \mid w \in \{\epsilon\}, a \in A\}$ stellt die Menge derjenigen Elemente dar, die sich durch Anwenden der induktiven Regel auf die bisherigen Elemente ergeben. Im dritten Schritt erhalten wir analog $A^0 \cup A^1 \cup A^2$ etc. Durch Vereinigen aller so konstruierbaren Mengen erhalten wir die gesuchte Menge A^*.

Die Idee der induktiven Definition von Mengen läßt sich noch etwas verallgemeinern. Eine *simultane induktive Definition* einer endlichen Familie $\mathcal{A} = \{A_1, \ldots, A_n\}$ von Mengen besteht aus einer endlichen Menge von Regeln. Jede Regel ist entweder eine *Basisregel* oder eine *induktive Regel*. Eine Basisregel gibt an, daß bestimmte Elemente zu einer Menge A_i gehören. Eine induktive Regel hat als *Prämisse* eine endliche Zahl von Elementschaftsbedingungen, die besagen, daß bestimmte Elemente zu vorgegebenen Mengen aus $\mathcal{A}$ gehören sollen. Die *Konklusion* besagt, daß dann ein daraus abzuleitendes Element zu einer bestimmten Menge A_i gehört. Ein Element gehört zu einer der auf diese Weise definierten Mengen A_i genau dann, falls man dies durch endlich viele Anwendungen der Regeln verifizieren kann.

Beispiel 3.59 Die folgenden Regeln stellen eine simultane induktive Definition der Menge G der geraden natürlichen Zahlen und der Menge U der ungeraden natürlichen Zahlen dar:

(1) $0 \in G$,

(2) $1 \in U$,

(3) $n + m \in G$ falls $n \in G$ und $m \in G$,

(4) $n + m \in G$ falls $n \in U$ und $m \in U$,

(5) $n + m \in U$ falls $n \in U$ und $m \in G$.

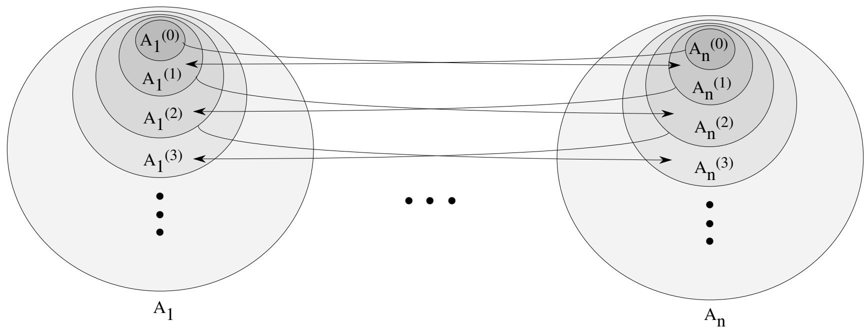


Abbildung 3.6: Illustration zur simultanen induktiven Definition von Mengen $A_1, \ldots, A_n$.

Die Kurzschreibweise ist hier

$$\begin{aligned} G &::= 0 \mid G+G \mid U+U, \\ U &::= 1 \mid G+U. \end{aligned}$$

Man kann die durch eine simultane induktive Definition beschriebenen Mengen $A_1, \ldots, A_n$ wie im Fall der einfachen induktiven Definition durch eine *Turm-* oder *Limeskonstruktion* expliziter beschreiben. Dazu setzt man in einem ersten Schritt $A_i^{(0)}$ als Menge derjenigen Elemente fest, die aufgrund der gegebenen Basisregeln zu A_i gehören müssen ($1 \leq i \leq n$). Sind nun für ein $k \in \mathbb{N}$ für $1 \leq i \leq n$ die Mengen $A_i^{(k)}$ bereits definiert, so setzen wir $A_i^{(k+1)}$ als die Vereinigung von $A_i^{(k)}$ mit der Menge derjenigen Elemente fest, die sich durch Anwendung der induktiven Regeln auf Elemente in den Mengen $A_1^{(k)}, \ldots, A_n^{(k)}$ ergeben. Dies ist in Abbildung 3.6 illustriert. Offenkundig ist für jedes $1 \leq i \leq n$ die Folge

$$A_i^{(0)} \subseteq A_i^{(1)} \subseteq \cdots \subseteq A_i^{(k)} \subseteq A_i^{(k+1)} \subseteq \cdots$$

in der Tat ein aufsteigender Turm von Mengen. Es ist dann $\bigcup_{k \in \mathbb{N}} A_i^{(k)}$ die gesuchte Menge A_i, für $1 \leq i \leq n$.

Beispiel 3.60 Die Limeskonstruktion zu der in Beispiel 3.59 gegebenen simultanen induktiven Definition der Mengen G und U der geraden und ungeraden natürlichen Zahlen stellt sich wie folgt dar.

$$G_0 = \{0\},$$

$$
\begin{array}{lcl}
U_0 & = & \{1\}, \\
G_1 & = & G_0 \cup \{n+m \mid n,m \in G_0\} \cup \{n+m \mid n,m \in U_0\} = \{0,2\}, \\
U_1 & = & U_0 \cup \{n+m \mid n \in U_0, m \in G_0\} = \{1\}, \\
G_2 & = & G_1 \cup \{n+m \mid n,m \in G_1\} \cup \{n+m \mid n,m \in U_1\} = \{0,2,4\}, \\
U_2 & = & U_1 \cup \{n+m \mid n \in U_1, m \in G_1\} = \{1,3\}, \\
G_3 & = & G_2 \cup \{n+m \mid n,m \in G_2\} \cup \{n+m \mid n,m \in U_2\} = \{0,2,\ldots,8\}, \\
U_3 & = & U_2 \cup \{n+m \mid n \in U_2, m \in G_2\} = \{1,3,\ldots,7\}, \\
\ldots & & \ldots
\end{array}
$$

Offenkundig ist dann $\bigcup_{i \in \mathbb{N}} G_i = G$ und $\bigcup_{i \in \mathbb{N}} U_i = U$.

3.6.2 Charakteristische Funktionen und höhere Funktionen

An dieser Stelle wollen wir nochmal zur Schreibweise 2^M für die Potenzmenge von M zurückkehren. Die Potenzmenge $\mathcal{P}(M)$ enthält genau alle Teilmengen von M als Elemente. Nun läßt sich eine Teilmenge N von M eindeutig dadurch charakterisieren, daß man für jedes $m \in M$ angibt, ob $m \in N$ gilt (1) oder nicht (0). Eine solche Beschreibung liefert die *charakteristische Funktion* $\chi_N \colon M \to \{0,1\}$ mit

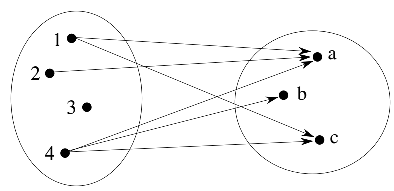
$$
\chi_N(m) = \begin{cases} 1 & \text{falls } m \in N, \\ 0 & \text{sonst.} \end{cases}
$$

Man sieht sofort, daß die Abbildung $N \mapsto \chi_N$ eine Bijektion von $\mathcal{P}(M)$ auf die Menge $\{0,1\}^M$ aller Abbildungen von M nach $\{0,1\}$ ist. Die Natürlichkeit dieser Bijektion erlaubt es, die beiden Mengen zu „identifizieren“ , was wir durch die symbolische Gleichung $\mathcal{P}(M) = \{0,1\}^M$ andeuten wollen. Machen wir nun von der in der Mengenlehre üblichen Konvention Gebrauch, die Zahl $n \in \mathbb{N}$ durch $\{0,1,\ldots,n-1\}$ zu kodieren, so können wir für $\mathcal{P}(M) = \{0,1\}^M$ auch 2^M schreiben.

Ganz allgemein wird jede Funktion $\chi\colon M \to \{0,1\}$ als *charakteristische Funktion* auf M bezeichnet. Mit Hilfe charakteristischer Funktionen lassen sich beliebige Relationen adäquat durch komplexe Funktionen repräsentieren. Nehmen wir zur Illustration die Relation aus Beispiel 3.15:

$$\begin{bmatrix} 1 & \longmapsto & \begin{bmatrix} a & \mapsto & 1 \\ b & \mapsto & 0 \\ c & \mapsto & 1 \end{bmatrix} \\ \\ 2 & \longmapsto & \begin{bmatrix} a & \mapsto & 1 \\ b & \mapsto & 0 \\ c & \mapsto & 0 \end{bmatrix} \\ \\ 3 & \longmapsto & \begin{bmatrix} a & \mapsto & 0 \\ b & \mapsto & 0 \\ c & \mapsto & 0 \end{bmatrix} \\ \\ 4 & \longmapsto & \begin{bmatrix} a & \mapsto & 1 \\ b & \mapsto & 1 \\ c & \mapsto & 1 \end{bmatrix} \end{bmatrix}$$

Abbildung 3.7: Funktionale Darstellung von Relationen.



Für jedes $x \in \{1,2,3,4\}$ ist $R(\{x\})$, also das Bild von $\{x\}$ unter R, eine Teilmenge von $\{a,b,c\}$. Wenn wir für jedes $x \in \{1,2,3,4\}$ diese Teilmenge $R(\{x\})$ kennen, so kennen wir umgekehrt auch R. Indem wir nun wieder Teilmengen durch charakteristische Funktionen ersetzen, erhalten wir die in Abbildung 3.7 dargestellte Repräsentation von R, die nur auf Funktionen basiert: Man kann natürlich diese funktionale Darstellung als Variante der in Beispiel 3.15 angegebenen Matrixdarstellung auffassen.

Ganz analog läßt sich jede Relation $R \subseteq A \times B$ in eine Funktion

$$F_R\colon A \to (2^B)\colon F_R(x) = \begin{cases} \text{die Funktion } \chi_x\colon B \to \{0,1\} \\ \text{mit } \chi_x(b) = 1 \text{ gdw. } R(x,b). \end{cases}$$

übersetzen. Die Funktionen F_R zeigen, daß der Bildbereich einer Funktion wiederum selbst aus Funktionen bestehen kann. Diese „Verschachtelung“

von Funktionen kann beliebig weiter getrieben werden. Man kann zum Beispiel auch jede n-stellige Funktion auf einer Menge B (vgl. Definition 3.31) durch eine Verschachtelung von „einfachen" Funktionen (wo alle Argumente einfache Elemente sind) repräsentieren. Auch dieser technische Trick, der manchmal als „Currying[4]" bezeichnet wird, sei anhand eines kleinen Beispiels illustriert. Es sei f die folgende zweistellige Funktion auf $\{a, b\}$:

$$f := \begin{bmatrix} \langle a,a\rangle & \mapsto & a \\ \langle a,b\rangle & \mapsto & a \\ \langle b,a\rangle & \mapsto & a \\ \langle b,b\rangle & \mapsto & b \end{bmatrix}$$

Dann kann f auch in der folgenden Weise kodiert werden:

$$\begin{bmatrix} a & \longmapsto & \begin{bmatrix} a & \mapsto & a \\ b & \mapsto & a \end{bmatrix} \\ & & \\ b & \longmapsto & \begin{bmatrix} a & \mapsto & a \\ b & \mapsto & b \end{bmatrix} \end{bmatrix}$$

3.6.3 Unendliche kartesische Produkte

Erinnern wir uns, daß es uns bei der Definition des geordneten n-Tupels im wesentlichen darum ging, eine Reihenfolge zwischen Elementen fixieren zu können. Anstatt n-Tupel zu verwenden, kann man die Reihenfolge der Elemente auch explizit mit Hilfe von Funktionen festhalten, die auf „Anfangsabschnitten" von $\mathbb{N}$ der Form $\{0, \ldots, n-1\}$ definiert sind. Die Funktion f der Form

$$\begin{bmatrix} 0 \mapsto a \\ 1 \mapsto d \\ 2 \mapsto b \\ 3 \mapsto a \end{bmatrix}$$

beispielsweise entspricht gerade dem Viertupel $\langle a, d, b, a\rangle$ beziehungsweise dem Wort *adba*. Ein kartesisches Produkt $A \times B \times C \times D$ kann auch durch die Menge

$$\{f \in (A \cup B \cup C \cup D)^{\{0,1,2,3\}} \mid f(0) \in A, f(1) \in B, f(2) \in C, f(3) \in D\}$$

[4] Nach dem Logiker H.B. Curry.

beschrieben werden. Im folgenden werden wir in beiden Fällen nur noch von kartesischen Produkten reden. Der springende Punkt ist, daß Produkte, die auf Funktionen beruhen, leicht auf Produkte mit unendlich vielen Faktoren verallgemeinert werden können. So beschreibt etwa

$$\prod_{i \in \mathbb{N}} A_i := \{f\colon \mathbb{N} \to \bigcup_{i \in \mathbb{N}} A_i \mid f(i) \in A_i \;\text{ für alle } i \in \mathbb{N}\}$$

die Menge aller unendlichen Folgen, deren i–tes Element aus A_i kommt ($i \in \mathbb{N}$). Für $A_i = A$ ($i \in \mathbb{N}$) erhalten wir $\prod_{i \in \mathbb{N}} A_i = A^{\mathbb{N}}$. Allgemeiner kann man für beliebigen nichtleeren Indexbereich I und Mengen A_i das Produkt

$$\prod_{i \in I} A_i := \{f\colon I \to \bigcup_{i \in I} A_i \mid f(i) \in A_i \;\text{ für alle } i \in I\}$$

einführen.

3.7 Aufgaben zu Kapitel 3

Aufgaben zu Teilkapitel 3.1

Aufgabe 3.1 Würde Lemma 3.2 (ii) richtig bleiben, wenn man das geordnete Paar $\langle a, b\rangle$ in der Form $\langle a, b\rangle := \{a, \{b\}\}$ definieren würde? Geben Sie einen Beweis oder ein Gegenbeispiel.

Aufgabe 3.2 Es sei $A_i = \emptyset$ für ein $i \in \{1, \ldots, n\}$. Wie sieht $\prod_{i=1,\ldots,n} A_i$ aus?

Aufgabe 3.3 Es sei $A := \{a_1, a_2\}$, $B := \{0, 1\}$ und $C := \{0, a_2, c\}$.

1. Berechnen Sie B^3.
2. Geben Sie $A \times B \times C$ an.

Aufgabe 3.4 Aus dem kartesischen Produkt $A \times B$ haben wir vier Elemente herausgenommen und erhalten $\{\langle a, a\rangle, \langle b, a\rangle, \langle c, d\rangle, \langle c, c\rangle, \langle a, c\rangle\}$. Welche Elemente haben wir herausgenommen?

Aufgabe 3.5 Es seien A und B endliche Mengen. Wieviele Elemente hat die kleinste Teilmenge von $A \times B$, aus der wir A, B und damit $A \times B$ rekonstruieren können?

Aufgabe 3.6 Gilt für beliebige Mengen stets $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$? Geben Sie einen Beweis oder ein Gegenbeispiel.

Aufgabe 3.7 Beweisen Sie, daß für beliebige Mengen A, B und C stets

$$A \times (B \setminus C) = (A \times B) \setminus (A \times C)$$

gilt.

Aufgabe 3.8 Es sei A das Alphabet mit den Symbolen a, b, c. Geben Sie fünf Beispiele für unendliche formale Sprachen über A an.

Aufgaben zu Teilkapitel 3.2

Aufgabe 3.9 Es sei A die Menge $\{1, 2\}$. Geben Sie alle zweistellige Relationen auf A an.

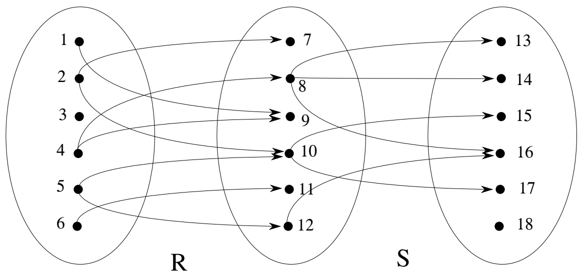
Aufgabe 3.10 Es sei A eine Menge mit n Elementen und $m \geq 1$. Wieviel m-stellige Relationen auf A gibt es?

Aufgabe 3.11 Es sei R die in Beispiel 3.18 Nr. 1 abgebildete Relation. Geben Sie $R(\{2, 3, 4\})$, $R(\{1, 4\})$, $R(\{1, 3\})$ sowie $R\lceil\{1, 2, 4\}$ und $R\lceil\{1, 3\}$ an.

Aufgabe 3.12 Es seien A, B endlich und $R, S \subseteq A \times B$. Wie ergibt sich die Matrixdarstellung von $R \cup S$ (resp. $R \cap S$) aus der Matrixdarstellung von R und S?

Aufgaben zu Teilkapitel 3.3

Aufgabe 3.13 Es seien R und S die nachfolgend abgebildeten Relationen. Berechnen Sie $R \circ S$, $(R \circ S)(\{2, 5, 6\})$, $(R \circ S)^{-1}(\{16, 17\})$.



Aufgabe 3.14 Gegeben seien die Relationen $R = \{\langle 0,2\rangle, \langle 2,4\rangle, \langle 4,6\rangle\}$ und $S := \{\langle 1,3\rangle, \langle 3,5\rangle, \langle 3,4\rangle, \langle 3,0\rangle\}$ auf $\mathbb{N}$. Berechnen Sie R^{-1}, S^{-1}, $R \circ S$, $S \circ R$ und $R^{-1} \circ S^{-1}$. Was fällt beim Vergleich $S \circ R$ und $R^{-1} \circ S^{-1}$ auf?

Aufgabe 3.15 Es sei R eine zweistellige Relation auf der Menge A. Unter welcher Bedingung gilt dann $R = R^{-1}$? Was kann man über A sagen, wenn jede zweistellige Relation R auf A stets $R = R^{-1}$ erfüllt?

Aufgabe 3.16 Beweisen Sie die Teile 1 und 3 aus Lemma 3.22.

Aufgabe 3.17 Es seien $R \subseteq A \times B$ und $S \subseteq B \times C$ Relationen. Was gilt: $(R \circ S)^{-1} = R^{-1} \circ S^{-1}$ oder $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$? Beweisen Sie die richtige Identität.

Aufgabe 3.18 Es sei A eine Menge mit k Elementen und $R, S \subseteq A \times A$. Es gelte $R \circ S = \emptyset$. Wieviel Elemente kann $R \cup S$ maximal haben?

Aufgabe 3.19 Beweisen Sie Lemma 3.26.

Aufgaben zu Teilkapitel 3.4

Aufgabe 3.20 Es seien A, B endliche Mengen und $f \subseteq A \times B$. Wie kann man an der Matrixdarstellung von f ablesen, ob f eine Funktion ist?

Aufgabe 3.21 Zeigen Sie, daß kartesische Produkte und zugehörige Projektionsfunktionen (vgl. Bsp. 3.34 Nr. 9) sich durch die folgende Universalitätseigenschaft auszeichnen: ist M eine Menge, und sind Funktionen

f_i: $M \to A_i$ für $1 \leq i \leq n$ vorgegeben, so gibt es eine eindeutig bestimmte Funktion g: $M \to \Pi_{i=1}^{n} A_i$, so daß $f_i = g \circ \pi_i$ für $1 \leq i \leq n$ gilt.

Aufgabe 3.22 Beweisen Sie Lemma 3.33 mittels vollständiger Induktion über die Zahl n der Elemente von A. Hinweis: es gibt genau eine Abbildung der leeren Menge in die Menge B, nämlich die „leere" Abbildung $\emptyset$.

Aufgabe 3.23 Es sei $A := \{1, 2, 3\}, B := \{2, 4, 6, 8, 10\}$ und f: $A \to B$ die Abbildung $x \mapsto 2x$. Berechnen Sie die Mengen $f(\emptyset)$, $f(\{1, 2\})$, $f^{-1}(\{2, 6, 8\})$, $f\lceil\{1, 3\}$, $(f\lceil\{1, 3\})^{-1}(\{2, 4, 8\})$.

Aufgabe 3.24 Beweisen Sie Lemma 3.40 (i). Hilfe: Sie können hier die Struktur des Beweises von Lemma 3.5 (ii) übernehmen.

Aufgabe 3.25 Geben Sie passende Beispiele an, die zeigen, daß in Lemma 3.40 (iii) und (iv) im allgemeinen die umgekehrte Inklusion nicht richtig ist.

Aufgaben zu Teilkapitel 3.5

Aufgabe 3.26 A und B seien Mengen und f: $A \to B$ sowie g: $B \to A$ Funktionen. Zeigen Sie: falls $f \circ g = Id_A$ und $g \circ f = Id_B$, so sind f und g bijektiv und es gilt $f = g^{-1}$ sowie $g = f^{-1}$.

Aufgabe 3.27 Verwenden Sie das Taubenloch-Prinzip (vgl. Lemma 3.44), um folgendes zu zeigen: es sei P eine Menge von mindestens 2 Personen. Dann gibt es zumindest zwei Personen in P, die dieselbe Zahl von anderen Personen aus P kennen. Hierbei sei vorausgesetzt, daß stets Person b Person a kennt, falls Person a Person b kennt $(a, b \in P)$.

Aufgabe 3.28 Es sei f: $A \to B$ eine surjektive Funktion. Zeigen Sie, daß die Menge $\{f^{-1}(\{b\}) \mid b \in B\}$ eine Partition der Menge A bildet.

Aufgabe 3.29 Beweisen Sie Lemma 3.54.

Aufgabe 3.30 Es seien A, B endlich und f: $A \to B$. Wie kann man an der Matrixdarstellung von f ablesen, ob f injektiv (resp. surjektiv, bijektiv) ist?

Aufgaben zu Teilkapitel 3.6

Aufgabe 3.31 Für $i = 0, 1, 2$ sei M_i die Menge aller natürlichen Zahlen, die beim Teilen durch 3 den Rest i lassen. Zum Beispiel enthält M_1 die Zahlen $1, 4, 7, 10$ etc. Geben Sie eine simultane induktive Definition der Mengen M_1, M_2, M_3, wobei Sie nur die Zahlen 0 und 1 und das Additionszeichen „+“ verwenden.

Aufgabe 3.32 Wenn man geordnete n-Tupel und kartesische Produkte so elegant mit Funktionen beschreiben kann (vgl. Abschnitt 3.6.3), warum haben wir überhaupt erst den Begriff des kartesischen Produktes eingeführt?

3.8 Bibliographische Angaben

Wie im vorigen Kapitel verweisen wir auf [Big89, Bra88, FS91, Ger72, RW92, uLW74] zur begleitenden Lektüre. Eine mathematisch anspruchsvolle Diskussion von Relationen und ihren Eigenschaften bietet [SS93]. Der Inhalt dieses Buches ist auch für Kapitel 7 interessant.

4

Äquivalenzrelationen, Ordnungsrelationen und Hüllenbildungen

Im vorausgegangenen Kapitel hatten wir das Konzept der Relation eingeführt und erläutert. In diesem Abschnitt werden nun einige wichtige Klassen von Relationen vorgestellt. Hierzu gehen wir in Abschnitt 4.1 zunächst auf einige grundlegende Eigenschaften ein, mit denen sich binäre Relationen charakterisieren lassen. Darauf aufbauend führen wir nachfolgend in 4.2 und 4.3 Äquivalenzrelationen und Ordnungsrelationen ein. Dieses sind die mit Abstand wichtigsten Typen mathematischer Relationen. Während Äquivalenzrelationen der Erfassung von Ähnlichkeiten dienen, besteht die Funktion der Ordnungsrelationen vor allem darin, daß sie einen systematischen Vergleich von Elementen ermöglichen. Beide Begriffe bilden die Grundlage für viele Betrachtungen in den nachfolgenden Kapiteln. Dies ist durchaus symptomatisch—man kann sich kaum ein Gebiet vorstellen, das einer mathematischen Beschreibung zugänglich ist, und wo nicht Äquivalenzrelationen oder Ordnungsrelationen in natürlicher Weise auftreten. In Abschnitt 4.4 stellen wir schließlich einige wichtige Formen der Erweiterung einer gegebenen Relation dar, insbesondere das Konzept der reflexiv-transitiven Hülle einer Relation, das im Zusammenhang mit dem Begriff der Ableitung bei Grammatiken und verwandten formalen Systemen große Bedeutung besitzt.

4.1 Charakteristische Eigenschaften zweistelliger Relationen

Die meisten mathematisch relevanten zweistellingen Relationen auf einer gegebenen Menge besitzen charakteristische Eigenschaften wie die im folgenden aufgelisteten.

Definition 4.1 Sei $R \subseteq A \times A$ eine Relation;

(i) R heißt *reflexiv* genau dann, wenn gilt
$\forall a \in A\colon R(a,a)$,

(ii) R heißt *transitiv* genau dann, wenn gilt
$\forall a,b,c \in A\colon (R(a,b) \wedge R(b,c)) \Rightarrow R(a,c))$,

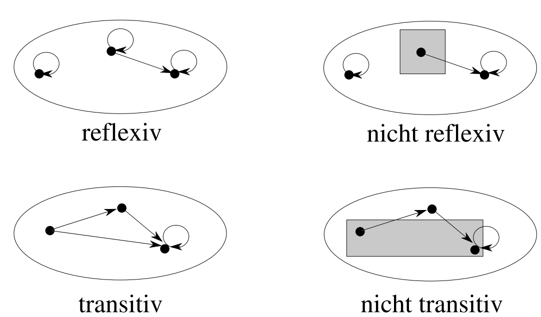
(iii) R heißt *symmetrisch* genau dann, wenn gilt
$\forall a,b \in A\colon (R(a,b) \Rightarrow R(b,a))$,

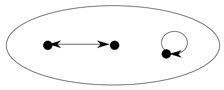
(iv) R heißt *antisymmetrisch* genau dann, wenn gilt
$\forall a,b \in A\colon ((R(a,b) \wedge R(b,a)) \Rightarrow a = b)$.

(v) R heißt *irreflexiv* genau dann, wenn gilt
$\forall a \in A\colon \neg R(a,a)$.

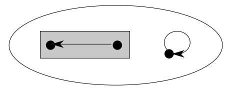
Zu diesen Eigenschaften nun einige Beispiele und Gegenbeispiele. Bei den Gegenbeispielen sind die relevanten Konstellationen unterlegt.

Die folgenden Relationen sind

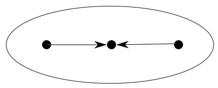




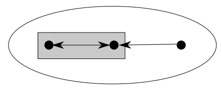
symmetrisch



nicht symmetrisch

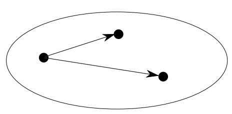


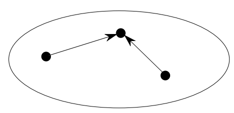
antisymmetrisch



nicht antisymmetrisch

Bemerkung 4.2 Die Bedingungen, die die Transitivität, Symmetrie und Antisymmetrie charakterisieren, beinhalten jeweils spezielle Implikationen, die für beliebige Elemente des Grundbereichs A gelten müssen. Es ist an dieser Stelle wichtig, sich die Bemerkungen aus Abschnitt 1.5 in Erinnerung zu rufen: eine Implikation ist inbesondere dann wahr, wenn bereits die Prämisse falsch ist. Im hier vorliegenden Fall bedeutet dies zum Beispiel, daß die nachfolgend dargestellten Relationen beide transitiv sind.





In der Tat gibt es in beiden Fällen keine Elemente $a, b, c \in A$, die in der Beziehung $R(a, b)$ und $R(b, c)$ stehen. Daher ist die Implikation $(R(a, b) \wedge R(b, c)) \to R(a, c)$ stets in trivialer Weise erfüllt. In beiden Fällen gibt es auch keine Elemente $a, b \in A$, die die Bedingung $R(a, b) \wedge R(b, a)$ erfüllen, daher sind die abgebildeten Relationen auch antisymmetrisch. Aus ähnlichen Gründen ist die leere Relation „$\emptyset$" *immer* transitiv, symmetrisch und antisymmetrisch, die Identitätsrelation Id_A (vgl. Def. 3.35) auf einer Menge A ist stets reflexiv, transitiv, symmetrisch und antisymmetrisch (Aufgabe 4.8).

Reflexivität, Symmetrie und Transitivität können mit Hilfe der Identitätsrelation, der Umkehrrelation und der Komposition wie folgt charakterisiert werden.

Lemma 4.3 *Es sei R eine zweistellige Relation auf der Menge A. Dann gilt:*

1. *R ist reflexiv genau dann, wenn $\mathrm{Id}_A \subseteq R$ gilt.*

R	1	2	3	4	5
1	1	0	0	0	1
2	1	1	0	0	1
3	0	1	1	0	0
4	0	0	0	1	0
5	0	0	1	0	1

S	1	2	3	4	5
1	1	1	0	0	1
2	1	0	1	0	0
3	0	1	0	0	1
4	0	0	0	1	1
5	1	0	1	1	0

Abbildung 4.1: Matrixdarstellung einer reflexiven (R) und einer symmetrischen Relation (S).

2. *R ist symmetrisch genau dann, wenn $R^{-1} \subseteq R$ gilt.*
3. *R ist transitiv genau dann, wenn $R \circ R \subseteq R$ gilt.*
4. *R ist antisymmetrisch genau dann, wenn $R \cap R^{-1} \subseteq Id_A$ gilt.*

Beweis. Übung (vgl. Aufgabe 4.7). ■

Wie Abbildung 4.1 illustriert, ist eine zweistellige Relation R auf einer endlichen Menge A reflexiv genau dann, wenn in der Matrixdarstellung die Diagonale überall mit 1 besetzt ist. Eine Relation ist symmetrisch, wenn die zugehörige Matrix spiegelsymmetrisch zur Diagonalen ist. Das folgende Lemma zeigt, daß reflexive, symmetrische und transitive Relationen auf einer Menge unter Durchschnittsbildung abgeschlossen sind.

Lemma 4.4 *Es seien $\{R_i \mid i \in I\}$ eine nichtleere Menge zweistelliger Relationen auf derselben Menge A. Dann gilt:*

1. *Sind alle Relationen R_i $(i \in I)$ reflexiv, so ist auch $\bigcap_{i \in I} R_i$ reflexiv.*
2. *Sind alle Relationen R_i $(i \in I)$ symmetrisch, so ist auch $\bigcap_{i \in I} R_i$ symmetrisch.*
3. *Sind alle Relationen R_i $(i \in I)$ transitiv, so ist auch $\bigcap_{i \in I} R_i$ transitiv.*

Der Beweis von Lemma 4.4 ist einfach und bleibt als Übungsaufgabe (vgl. Aufgabe 4.2).

Der Vollständigkeit halber seien noch die folgenden Eigenschaften zweistelliger Relationen erwähnt, die gelegentlich in der Literatur auftreten.

Definition 4.5 Eine zweistellige Relation $R \subseteq A \times A$ heißt *euklidisch* genau dann, wenn gilt:

$$\forall a, b, c \in A\colon ((R(a, b) \wedge R(a, c)) \Rightarrow R(b, c)).$$

R heißt *anti–euklidisch* genau dann, wenn gilt

$$\forall a, b, c \in A\colon ((R(a, b) \wedge R(c, b)) \Rightarrow R(a, c)).$$

Die in diesem Abschnitt eingeführten Eigenschaften sind natürlich nicht völlig unabhängig voneinander. Beispielsweise ist jede Relation R, die euklidisch und reflexiv ist, stets auch symmetrisch. Dazu vergleiche man Aufgabe 4.9 am Kapitelende.

4.2 Äquivalenzrelationen

Der Begriff der Äquivalenz verallgemeinert den Begriff der Gleichheit und beinhaltet in einem zu präzisierenden Sinn Gleichheit unter einem bestimmten Aspekt.

Definition 4.6 Eine Relation $R \subseteq A \times A$ heißt *Äquivalenzrelation* auf A, genau dann, wenn R reflexiv, transitiv und symmetrisch ist.

Es ist banal sich klarzumachen, daß für jede Menge A die Gleichheitsrelation Id_A in der Tat eine Äquivalenzrelation ist. Statt $R(a, b)$ schreibt man für Äquivalenzrelationen meistens in Infixnotation $a \sim_R b$, oder einfach $a \sim b$, wenn R aus dem Zusammenhang klar ist.

Beispiele 4.7 Nachfolgend eine kleine Liste von Äquivalenzrelationen. Die Beispiele „aus dem täglichen Leben“ setzen streng genommen eine striktere Formalisierung voraus.

1. Es sei $f\colon A \to B$ eine beliebige Funktion. Dann wird durch

   $$a_1 \sim_f a_2 \;:\Leftrightarrow\; f(a_1) = f(a_2)$$

   eine Äquivalenzrelation auf A definiert. Die Äquivalenzrelation „$\sim_f$“ heißt die durch die Abbildung f *induzierte Äquivalenzrelation.* Wir

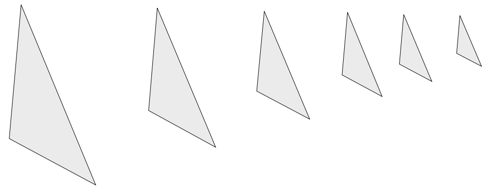


Abbildung 4.2: Ähnlichkeit von Dreiecken als Beispiel für Äquivalenz.

werden unten sehen, daß jede Äquivalenzrelationen bei geeigneter Betrachtungsweise durch eine passende Abbildung induziert ist. Man möge sich dies auch bei den nachfolgenden Beispielen verdeutlichen (vgl. hierzu Aufgabe 4.12).

2. Es sei $0 \neq n$ eine natürliche Zahl. Die Relation R, die auf zwei natürliche Zahlen k und l zutrifft genau dann, wenn beide beim Teilen durch n denselben Rest lassen, ist eine Äquivalenzrelation.

3. Es sei M eine Menge von berufstätigen Personen. Die Relation R, die besagt, daß zwei Personen dieselbe Steuerklasse haben, ist eine Äquivalenzrelation.

4. Es sei M die Menge der Schüler/innen einer Schule. Die Relation R, die besagt, daß zwei Elemente von M in dieselbe Klasse gehen, ist eine Äquivalenzrelation.

5. Die leere Relation $\emptyset$ ist eine Äquivalenzrelation auf der leeren Menge $\emptyset$. Es ist die einzige Äquivalenzrelation auf dieser Menge. Für jede Menge M ist stets $M \times M$ eine Äquivalenzrelation auf M.

6. Die sogenannte „Ähnlichkeit“[1] von Figuren im zweidimensionalen euklidischen Raum definiert eine Äquivalenzrelation auf der Menge aller Dreiecke. In Abbildung 4.2 sind einige ähnliche Dreiecke dargestellt.

Es ist oft zweckmäßig, die Mengen paarweise äquivalenter Objekte unter einem neuen Begriff zusammenzufassen. Beim Beispiel der Schüler/innen sind das die Schulklassen. Für sehr viele praktische Aufgaben, wie das Erstellen eines Stundenplans, kann nämlich von den einzelnen Mitgliedern der Klassen abstrahiert werden, man kann sich stattdessen auf wenige Klassennamen

[1] Informell besagt die Ähnlichkeit zweier Figuren, daß man zwischen den Strecken beider Figuren eine Bijektion herstellen kann, die alle Proportionen und Winkel erhält.

beschränken. Dasselbe prinzipielle Vorgehen ist auch in der Mathematik von großer Bedeutung. Dies rechtfertigt die folgenden Begriffe:

Definition 4.8 Sei $R \subseteq A \times A$ eine Äquivalenzrelation. Für jedes $a \in A$ wird die Menge $\{b \in A \mid b \sim_R a\}$ die *Äquivalenzklasse* von a bezüglich R (oder modulo R) genannt und kurz als $[a]_R$ notiert. Die Menge $\{[a]_R \mid a \in A\}$ aller Äquivalenzklassen modulo R wird mit A/R (lies: Quotientenmenge von A modulo R) oder $A/{\sim_R}$ bezeichnet.

Wenn die Relation R aus dem Zusammenhang klar ist, schreiben wir einfacher $[a]$ und $A/{\sim}$ anstatt $[a]_R$ beziehungsweise $A/{\sim_R}$.

Definition 4.9 Sei $R \subseteq A \times A$ eine Äquivalenzrelation. Jedes Element $b \in [a]$ heißt *Vertreter* oder *Repräsentant* der Äquivalenzklasse $[a]$. Eine Menge $V \subseteq A$ heißt *Vertretersystem* oder *Repräsentantensystem* bezüglich R genau dann, falls V aus jeder Äquivalenzklasse genau ein Element enthält, das heißt, falls gilt

1. $\forall a \in A \; \exists v \in V\colon a \sim v$,
2. $\forall w, v \in V\colon (v \neq w \Rightarrow [v] \neq [w])$.

Lemma 4.10 *Es sei $R \subseteq A \times A$ eine Äquivalenzrelation und $a, b \in A$. Dann sind die folgenden Aussagen äquivalent:*

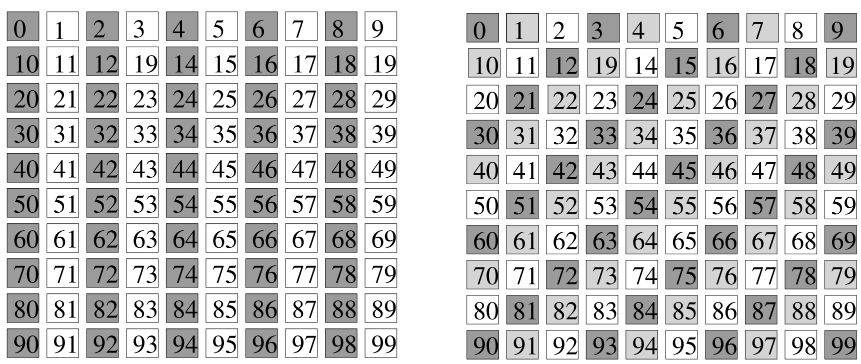
1. $[a] = [b]$,
2. $a \sim b$,
3. $b \sim a$,
4. $a \in [b]$,
5. $b \in [a]$.

Beweis. Wir zeigen die Äquivalenz von 1 und 2, der Rest bleibt als Übung. Falls $[a] = [b]$ gilt, so folgt $a \in [a] = [b]$ wegen der Reflexivität von R. Aus $a \in [b]$ folgt $a \sim b$. Es gelte nun umgekehrt $a \sim b$. Da R symmetrisch ist, gilt auch $b \sim a$. Es sei $c \in [a]$. Dann gilt $c \sim a$, wegen der Transitivität von R folgt $c \sim b$, das heißt $c \in [b]$. Demnach gilt $[a] \subseteq [b]$. Symmetrisch folgt $[b] \subseteq [a]$, insgesamt also $[a] = [b]$. ■

Definition 4.11 Die Funktion $\kappa\colon A \to A/\!\sim;\ a \mapsto [a]$ wird als die *kanonische Abbildung* von A auf die Quotientenmenge $A/\!\sim$ bezeichnet.

Lemma 4.10 zeigt, daß $a \sim b$ gilt genau dann, wenn $\kappa(a) = \kappa(b)$. Jede Äquivalenzrelation ist damit induziert (im Sinn von Beispiele 4.7 Nr. 1) durch die zugehörige kanonische Abbildung κ. Zu den nun eingeführten Begriffen eine Illustration.

Beispiel 4.12 Es sei $0 \neq n$ eine natürliche Zahl. Wir betrachten die Äquivalenzrelation R, die auf zwei natürliche Zahlen k und l zutrifft genau dann, wenn beide beim Teilen durch n denselben Rest lassen. Die nachfolgende Abbildung gibt die zwei (beziehungsweise drei) existierenden Äquivalenzklassen $[0]_R$ und $[1]_R$ ($[0]_R, [1]_R$ und $[2]_R$) für den Fall $n = 2$ (respektive $n = 3$) wieder.



Für den links dargestellten Fall $n = 2$ sind genau die geraden Zahlen die möglichen Vertreter von $[0]_R$, die ungeraden Zahlen sind die möglichen Vertreter von $[1]_R$. Eine zweielementige Teilmenge V von $\mathbb{N}$ ist ein Vertretersystem für R genau dann, wenn V eine gerade und eine ungerade Zahl enthält.

Wir zeigen nun, daß Äquivalenzrelationen und Partitionen (vgl. Definition 2.23) nur zwei verschiedene Formalisierungen derselben Situation darstellen.

Lemma 4.13 *Es sei R eine Äquivalenzrelation auf der Menge A. Dann ist $\Pi := A/\!\sim_R$ eine Partition von A. Ist umgekehrt Π eine Partition der Menge A, dann wird durch*

$$a \sim_\Pi b \quad :\Leftrightarrow \quad \exists B \in \Pi\colon a, b \in B$$

eine Äquivalenzrelation auf A definiert und es gilt $A/{\sim_\Pi} = \Pi$.

Beweis. Um die erste Teilaussage zu beweisen, zeigen wir, daß Π die drei Eigenschaften besitzt, die gemäß Definition 2.23 eine Partition einer Menge A charakterisieren.

1. Offensichtlich sind alle Mengen $[a] \in A/{\sim_R}$ nichtleer für jedes $a \in A$, da stets $a \in [a]$ aufgrund der Reflexivität von „$\sim$“ gilt. Aus der Definition der Menge $[a]$ ergibt sich sofort, daß stets $[a]$ eine Teilmenge von A ist.

2. Aus der verifizierten Eigenschaft 1 folgt leicht

$$\bigcup A/{\sim_R} = \bigcup\{[a] \mid a \in A\} \subseteq A.$$

Ist $a \in A$, so gilt wegen der Reflexivität von „$\sim$“ stets $a \in [a] \in A/{\sim_R}$, damit aber gemäß der Definition der Vereinigung auch $a \in \bigcup A/{\sim_R}$. Damit haben wir die Gleichheit der Mengen $\bigcup A/{\sim_R}$ und A gezeigt.

3. Für $[a] \neq [b]$ sind die Mengen $[a]$ und $[b]$ disjunkt: falls $c \in [a] \cap [b]$, so folgt $c \sim a$ und $c \sim b$. Da „$\sim$“ symmetrisch und transitiv ist, folgt $a \sim c$ und $a \sim b$. Somit $[a] = [b]$ nach Lemma 4.10.

Der Beweis der zweiten Teilaussage bleibt als Übung offen. ■

Zur Erläuterung von Lemma 4.13 und zur graphischen Repräsentation von Äquivalenzrelationen hier noch ein weiteres Beispiel.

Beispiel 4.14 Die Menge $A = \{1, 2, 3, 4\}$ besitzt (unter anderem) die Partition $\Pi = \{\{1, 2\}, \{3, 4\}\}$. Die korrespondierende Äquivalenzrelation $\sim_\Pi$ ist in Abbildung 4.3 in der Mitte durch Pfeile angedeutet. Es fällt auf, daß innerhalb einer Äquivalenzklasse je zwei Elemente immer durch Pfeile verbunden sind, Elemente verschiedener Äquivalenzklassen hingegen nie. Die Äquivalenzklassen von $\sim_\Pi$ sind genau die Elemente (Partitionsmengen) von Π, nämlich $[1] = [2] = \{1, 2\}$ (mit den möglichen Vertretern 1 und 2) sowie $[3] = \{3, 4\}$. Es gibt vier mögliche Vertretersysteme, nämlich $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$ und $\{2, 4\}$. Die Quotientenmenge $A/{\sim}$ besitzt lediglich zwei Elemente, nämlich $[1]$ und $[3]$. Die kanonische Abbildung κ ist durch graue Pfeile angedeutet.

Wie im vorausgegangenen Beispiel gilt bei der Darstellung einer Äquivalenzrelation durch Pfeile allgemein, daß stets alle Elemente innerhalb einer

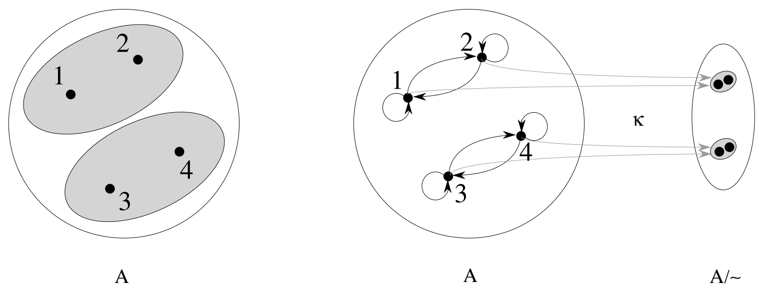
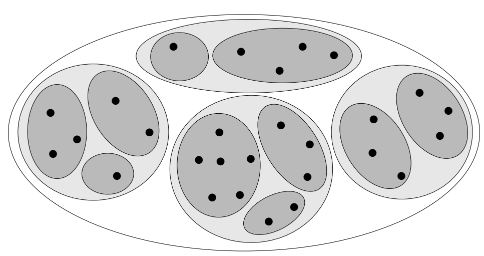


Abbildung 4.3: Partition, korrespondierende Äquivalenzrelation, kanonische Abbildung und Quotientenmenge.

Äquivalenzklasse wechselseitig durch Pfeile verbunden sind, Elemente verschiedener Äquivalenzklassen hingegen nie.

Definition 4.15 Es seien R und S zwei Äquivalenzrelationen auf A. Dann heißt S eine *Verfeinerung* von R genau dann, wenn jede Äquivalenzklasse bezüglich S enthalten ist in der entsprechenden Äquivalenzklasse bezüglich R, das heißt, wenn gilt: $\forall a \in A$: $[a]_S \subseteq [a]_R$.

In nachfolgender Illustration ist die durch dunkle Schattierungen dargestellte Äquivalenzrelation (Partition) eine Verfeinerung der durch die hellen Schattierungen angedeuteten Äquivalenzrelation.



Weitere Beispiele für Verfeinerungen ergeben sich bei Äquivalenzrelationen, die durch komponierte Funktionen induziert sind.

Beispiel 4.16 Seien f: $A \to B$ und g: $B \to C$ Funktionen. Die Äquivalenz-

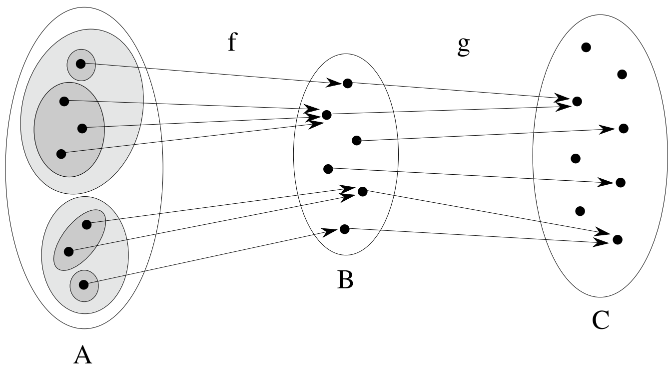


Abbildung 4.4: Verfeinerungen und induzierte Äquivalenzrelationen.

relationen $\sim_f$ und $\sim_{f\circ g}$ auf A seien vermöge

$$\begin{aligned} a_1 \sim_f a_2 \quad &:\Leftrightarrow \quad f(a_1) = f(a_2) \\ a_1 \sim_{f\circ g} a_2 \quad &:\Leftrightarrow \quad g(f(a_1)) = g(f(a_2)) \end{aligned}$$

definiert. Dann ist $\sim_f$ eine Verfeinerung von $\sim_{f\circ g}$. Dies ist in Abbildung 4.4 illustriert.

Das nachfolgende Lemma weist auf eine weitere Situation hin, wo Verfeinerungen von Äquivalenzrelationen in natürlicher Weise auftreten.

Lemma 4.17 *Für jeden Index i der nichtleeren Indexmenge I sei R_i eine Äquivalenzrelation auf der Menge A. Dann ist auch der Durchschnitt $\bigcap_{i\in I} R_i$ wieder eine Äquivalenzrelation auf A. Für jedes $j \in I$ ist $\bigcap_{i\in I} R_i$ eine Verfeinerung von R_j.*

Beweis. Jede der Relationen R_i ist nach Voraussetzung reflexiv, transitiv und symmetrisch. Damit ist nach Lemma 4.4 auch $\bigcap_{i\in I} R_i$ reflexiv, transitiv und symmetrisch. Also ist $\bigcap_{i\in I} R_i$ eine Äquivalenzrelation. Sei nun $a \in A$ und $j \in I$. Aus $a \sim_{\bigcap_{i\in I} R_i} b$ folgt $a \sim_{R_j} b$. Damit gilt $[a]_{\bigcap_{i\in I} R_i} \subseteq [a]_{R_j}$, und $\bigcap_{i\in I} R_i$ ist eine Verfeinerung von R_j. ■

Beispiel 4.18 Abbildung 4.5 repräsentiert zwei Äquivalenzrelationen auf der Menge $\{a, b, c, \ldots, s, t\}$ und ihren Durchschnitt. Wenn sich die Äquivalenzklassen beider Äquivalenzrelationen geometrisch durch Durchführen

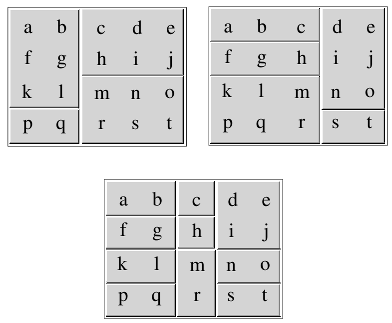


Abbildung 4.5: Äquivalenzklassen des Durchschnitts zweier Äquivalenzrelationen.

bestimmter „Schnitte“ ergeben, so ergibt sich der Durchschnitt der Äquivalenzrelationen durch Durchführen aller Schnitte beider Teilrelationen. Das vorangegangene Beispiel läßt sich verallgemeinern.

Lemma 4.19 *Für jeden Index i der nichtleeren Indexmenge I sei R_i eine Äquivalenzrelation auf der Menge A. Dann gilt für alle $a \in A$ stets* $[a]_{\bigcap_{i\in I} R_i} = \bigcap_{i\in I}[a]_{R_i}$.

Der Beweis wird als Übung offengelassen (vgl. Aufgabe 4.16).

Lemma 4.20 *Die Vereinigung zweier Äquivalenzrelationen R_1 und R_2 auf derselben Menge A ist im allgemeinen keine Äquivalenzrelation, aber stets symmetrisch und reflexiv.*

Beweis. Wir betrachten die Äquivalenzrelation S_1, die durch die Partition $\{\{1,2\},\{3\}\}$ der Menge $\{1,2,3\}$ bestimmt ist, analog sei S_2 durch $\{\{1\},\{2,3\}\}$ bestimmt. $S_1 \cup S_2$ enthält genau die Paare $\langle 1,1\rangle, \langle 2,2\rangle, \langle 3,3\rangle, \langle 1,2\rangle, \langle 2,1\rangle, \langle 2,3\rangle, \langle 3,2\rangle$, nicht jedoch $\langle 1,3\rangle$. Somit ist $S_1 \cup S_2$ nicht transitiv. Sind nun R_1 und R_2 beliebige Äquivalenzrelationen auf A, so mit ihnen offenkundig auch $R_1 \cup R_2$ symmetrisch und reflexiv. ■

4.3 Ordnungsrelationen

Ordnungsrelationen spielen immer dann eine Rolle, wenn man alle Elemente einer Menge in einer systematischen Reihenfolge behandeln will, oder wenn man Elemente unter einem bestimmten Gesichtpunkt vergleichen möchte. Wenn man diese zunächst recht grobe Intuition präzisiert, kommt man zu einer Reihe unterschiedlicher Begriffe.

Definition 4.21 Es sei $R \subseteq A \times A$ eine binäre Relation auf der Menge A.

1. R heißt *Quasi-Ordnung* auf A genau dann, wenn R reflexiv und transitiv ist.
2. R heißt *partielle Ordnung* auf A genau dann, wenn R transitiv, reflexiv und antisymmetrisch ist.
3. Eine partielle Ordnung R auf A heißt *lineare Ordnung* oder *Totalordnung* genau dann, wenn zusätzlich gilt:

$$\forall a, b \in A\colon\ (R(a,b) \ \text{ oder } R(b,a)).$$

Bevor wir Beispiele geben und die Begriffe näher erläutern, gehen wir kurz auf notationelle Konventionen und Varianten ein.

Ist R eine Ordnung eines Typs 1–3, so schreibt man meist $a \leq_R b$ oder $a \leq b$ anstatt $R(a,b)$, außerdem $a < b$ falls $a \leq b$ und nicht $a \leq b$. Wie die Schreibweise andeutet, werden anstelle der oben definierten partiellen und linearen Ordnungen manchmal auch „strikte“ Versionen betrachtet, wo ein Element nie zu sich selbst in Relation steht.

Definition 4.22 Eine *strikte partielle Ordnung* auf der Menge A ist eine Relation $R \subseteq A \times A$, die irreflexiv und transitiv ist. R heißt *strikte lineare Ordnung* genau dann, wenn außerdem für alle $a \neq b \in A$ stets $R(a,b)$ oder $R(b,a)$ gilt.

Üblicherweise werden strikte Ordnungen mit dem Symbol „$<$“ bezeichnet. Man beachte, daß jede strikte (partielle oder lineare) Ordnung „$<$“ automatisch „vollständig unsymmetrisch“ im folgenden Sinn ist: für $a \neq b \in A$ gilt nie zugleich $a < b$ und $b < a$. In der Tat, aufgrund der Transitivität hätte man sofort $a < a$, was der Irreflexivität widerspricht.

Man kann aus jeder partiellen oder linearen Ordnung „$\leq$“ durch Weglassen der Identität eine strikte Version „$<$“ erhalten, und umgekehrt aus jeder strikten partiellen oder linearen Ordnung durch Vereinigung mit der Identitätsrelation eine normale (d.h. reflexive) partielle Ordnung erhalten. Deswegen ist es bei partiellen und linearen Ordnungen gleichgültig, ob wir strikte oder nicht-strikte Versionen näher betrachten. Wir werden auf strikte Ordnungen nicht weiter eingehen.

Definition 4.23 Es sei „$\leq$“ eine partielle Ordnung auf der Menge A.

1. Das Element $a \in A$ heißt *minimal* (*maximal*) bezüglich „$\leq$“ genau dann, wenn für alle $b \in A$ aus $b \leq a$ (resp. $a \leq b$) stets $b = a$ folgt.

2. Das Element $a \in A$ heißt das *kleinste* (*größte*) Element bezüglich „$\leq$“ genau dann, wenn für alle $b \in A$ stets $a \leq b$ (resp. $b \leq a$) gilt.

In den unten aufgeführten Beispielen wird deutlich werden, daß es nicht immer minimale, maximale, kleinste oder größte Elemente bezüglich einer vorgegebenen partiellen Ordnung gibt. Falls das kleinste (größte) Element existiert, ist es eindeutig und stets auch minimal (maximal), die Umkehrung gilt aber in der Regel nicht, da es viele minimale Elemente geben kann.

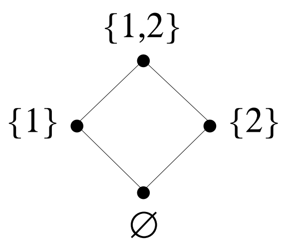
Die ersten drei der nun folgenden Beispiele sollen dazu dienen, die charakteristischen Eigenschaften der drei Ordnungstypen und die Unterschiede zu verdeutlichen. Zunächst mag der Hinweis nützlich sein, daß jede lineare Ordnung stets eine partielle Ordnung ist, jede partielle Ordnung immer eine Quasi-Ordnung. Es handelt sich also um eine hierarchische Begriffsbildung. Man spricht von einer „echten“ partiellen Ordnung (Quasi-Ordnung), wenn eine partielle Ordnung (resp. Quasi-Ordnung) vorliegt, die nicht weitergehend eine lineare (partielle) Ordnung darstellt.

Beispiele 4.24 Die natürliche Ordnung „$\leq$“ auf der Menge $\mathbb{N}$ der natürlichen Zahlen ist eine lineare Ordnung. Es ist 0 das eindeutig bestimmte kleinste Element von $\mathbb{N}$ bezüglich „$\leq$“. Es gibt kein maximales Element bezüglich „$\leq$“. Die natürlichen Ordnungen auf den ganzen, rationalen, oder reellen Zahlen sind gleichfalls lineare Ordnungen. Sie besitzen jeweils keine minimalen oder kleinsten Elemente.

Charakteristisch für alle lineare Ordnungen ist es, daß man je zwei verschiedene Elemente immer bezüglich der Ordnung vergleichen und in eine eindeutige Reihenfolge bringen kann.

Beispiel 4.25 Es sei M eine beliebige Menge. Dann ist die Inklusionsbeziehung „$\subseteq$“ ist eine partielle Ordnung auf der Potenzmenge $\mathcal{P}(M)$ (vgl. Definition 2.26). Hat M zumindest zwei Elemente, so ist diese Ordnung aber keine lineare Ordnung. Es ist „$\emptyset$“ das kleinste Element, M das größte Element bezüglich „$\subseteq$“ .

Bei echten partiellen Ordnungen gibt es Elemente, bei denen ein Größenvergleich fehlschlägt. Hat M zum Beispiel die Form $\{1,2\}$, so sind die Mengen $\{1\}$ und $\{2\}$ bezüglich der Inklusionsbeziehung unvergleichbar[2].



Erhalten bleibt bei partiellen Ordnungen allerdings die folgende Eigenschaft: wenn zwei verschiedene Elemente bezüglich „$\leq$“ vergleichbar sind, so legt die Ordnung stets eine eindeutige Reihenfolge zwischen den Elementen fest.

Beispiel 4.26 Es sei A eine Menge von Personen. Für $a \in A$ sei $f(a)$ die Größe der Person a, gemessen in cm. Mit „$\leq$“ bezeichnen wir die übliche Ordnung auf $\mathbb{N}$. Für $a, b \in A$ definieren wir

$$a \preceq b \quad :\Leftrightarrow \quad f(a) \leq f(b).$$

Dann ist im allgemeinen „$\preceq$“ eine Quasi-Ordnung, aber keine partielle Ordnung. Abbildung 4.6 illustriert den Zusammenhang. Die Funktion f repräsentiert die Messung der Körpergröße der Personen $1, \ldots, 4$. Die Personen 1 und 2 (resp. 3 und 4) haben dieselbe Körpergröße. Die Ordnung „$\preceq$“ auf A ist durch Pfeile angedeutet. Sie wird die durch f und „$\leq$“ *induzierte* Quasi-Ordnung auf A genannt. Offenkundig gilt $1 \preceq 2$ und $2 \preceq 1$, obwohl 1 und 2 verschieden sind. Daher verletzt „$\preceq$“ die Antisymmetrie und ist keine partielle Ordnung.

[2]Wie immer fassen wir hier Zahlen nicht als Mengen auf.

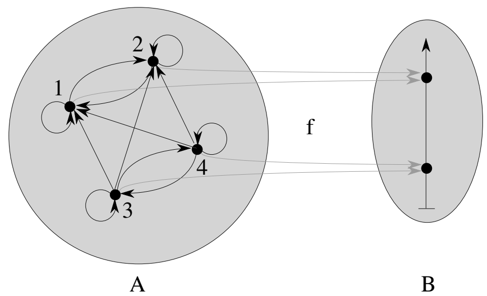


Abbildung 4.6: Durch Messung f induzierte Quasi-Ordnung.

Bei Quasi-Ordnungen kann es also vorkommen, daß wir zwei verschiedene Elemente zwar mittels der Ordnung vergleichen können, aber daraus trotzdem keine eindeutige Reihenfolge der Elemente erhalten. Die meisten in der Praxis auftretenden Quasi-Ordnungen (nicht alle) haben die Eigenschaft, daß man sogar stets zwei Elemente bezüglich der Ordnung vergleichen kann. Alle Arten von Messungen mit einer linearen Meßskala definieren Quasi-Ordnungen diesen Typs auf der Menge der zu messenden Objekte. Viele steigerbaren Adjektive (schneller, heißer, dicker, besser,...) lassen sich mit Quasi-Ordnungen dieses Typs assoziieren.

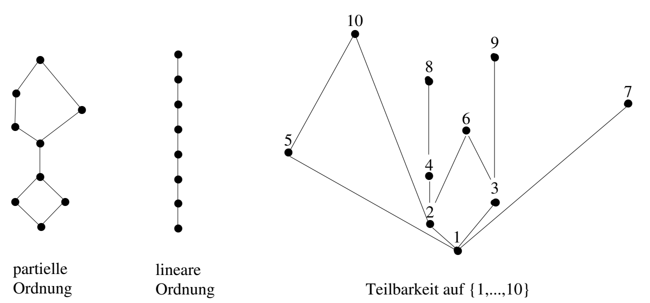
Definition 4.27 Es sei „$\leq$" eine partielle Ordnung auf der Menge A. Das Element $b \in A$ heißt ein *Nachfolger* von $a \in A$ genau dann, wenn $a < b$. In dieser Situation heißt umgekehrt a ein *Vorgänger* von b. Gilt außerdem, daß es kein Element $c \in A$ gibt mit $a < c < b$, so heißt b ein *direkter Nachfolger* von a, und umgekehrt heißt a ein *direkter Vorgänger* von b.

Beispiele 4.28 Die nachfolgende umfangreiche Liste soll dazu dienen, die Konzepte zu vertiefen. Viele der nachfolgenden Beispiele sind aber auch von großer praktischer Bedeutung.

1. Bezeichnet „$\leq$" die übliche Ordnung auf $\mathbb{N}$, so ist 0 das kleinste Element, für jedes $n \in \mathbb{N}$ ist stets $n+1$ der direkte Nachfolger von n. Bezeichnet „$\leq$" hingegen die übliche Ordnung auf den reellen Zahlen, so hat kein Element einen direkten Nachfolger oder einen direkten Vorgänger bezüglich „$\leq$" und es gibt kein minimales oder kleinstes Element.

2. Die Teilbarkeitsrelation T auf $\mathbb{N}$ (vgl. Beispiele 3.16 Nr. 3) beschreibt eine (echte) partielle Ordnung auf $\mathbb{N}$.

3. Endliche partielle Ordnungen lassen sich durch sogenannte *Hasse-Diagramme* darstellen. Ein solches Diagramm besteht aus einer endlichen Menge von Punkten, von denen einige durch Kanten verbunden sind. Man liest alle Kanten einheitlich von unten nach oben (horizontale Kanten sind nicht erlaubt), damit sind die Kanten implizit gerichtet. Die Punkte des Diagramms repräsentieren die Elemente einer Menge. Die Kanten definieren eine partielle Ordnung „$\leq$" wie folgt: Für beliebige Elemente x und y gilt $x \leq y$ genau dann, wenn es einen aufsteigenden Kantenzug der Länge $n \geq 0$ von x nach y gibt. Einzelne Kanten stehen damit für die direkte Nachfolgerbeziehung bezüglich „$\leq$".

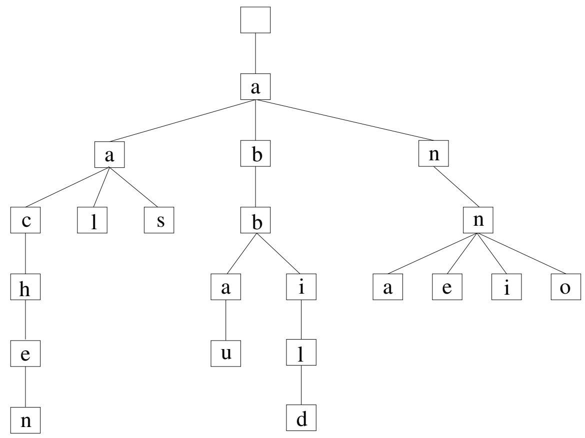


Wir werden unten zeigen, daß jede endliche partielle Ordnung „$\leq$" durch ein Hasse-Diagramm dargestellt werden kann.

4. Es sei $A = \{a_1, \ldots, a_n\}$ ein Alphabet, A^* bezeichne die in Definition 3.9 eingeführte Menge aller Wörter A. Für jede Teilmenge $W \subseteq A^*$ definiert

$$u \prec v \quad :\Leftrightarrow \quad u \text{ ist ein Präfix von } v$$

eine partielle Ordnung auf W, die *Präfixordnung* genannt wird. Wir betrachten den Fall genauer, wo W endlich und unter Präfixbildung abgeschlossen ist. Das Hasse-Diagramm hat hier die Form eines Baums. Der Baum, der sich für die Menge mit den Wörtern *a, aa, aac, aach, aache, aachen, aal, aar, aas, ab, abb, abbau, abbi, abbil, abbild, an, anna, anne, anno, anni* ergibt, ist in der nachfolgenden Figur (hier von oben nach unten gerichtet) dargestellt, wobei sich das einem Knoten η entsprechende Wort durch die Labels aller Knoten von der Wurzel bis zu η ergibt.



Ähnliche Bäume, sogenannte „Tries“, werden in der Informatik und Computerlinguistik zur internen Repräsentation von Wörterbüchern verwendet.

5. Die nun zu besprechende partielle Ordnung stellt eine Verfeinerung der Präfixordnung dar. Es sei $A = \{a_1, \ldots, a_n\}$ ein Alphabet, das durch $a_1 <_A a_2 <_A \ldots <_A a_n$ geordnet sei. Es bezeichne A^* die in Definition 3.9 definierte Menge aller Wörter über dem Alphabet A. Dann wird durch

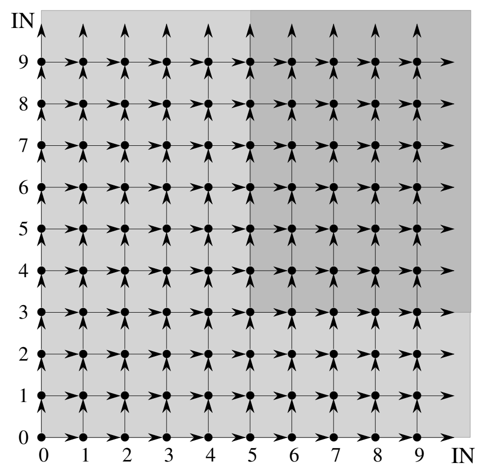
$$x_1 \ldots x_k \leq_L y_1 ... y_l :\Leftrightarrow \begin{cases} k \leq l \quad \text{und } x_1 = y_1, \ldots, x_k = y_k \\ \text{oder} \\ \text{es ex. } i,\ 0 \leq i < min(k,l) \text{ wo} \\ x_1 = y_1, \ldots, x_i = y_i \text{ und } x_{i+1} <_A y_{i+1} \end{cases}$$

eine lineare Ordnung auf A^* definiert, die *lexikographische Ordnung* von A^* genannt wird. Es handelt sich um die übliche Wortreihenfolge in Lexika, Telefonbüchern etc. Die Wortliste *a, aa, aac, aach, aache, aachen, aal, aar, aas, ab, abb, abbau, abbi, abbil, abbild, an, anna, anne, anno, anni* ist lexikographisch geordnet.

6. Ist „$\leq$“ eine partielle (lineare) Ordnung auf der Menge M, und definiert man $a \geq b :\Leftrightarrow b \leq a$, so wird durch „$\geq$“ eine partielle (lineare) Ordnung auf M definiert, die die zu „$\leq$“ *duale Ordnung* heißt.

7. Es sei $M = A_1 \times \ldots \times A_n$ ein kartesisches Produkt. Wenn wir auf jeder der „Faktormengen“ A_i eine partielle Ordnung „$\leq_i$” haben, so können wir auf M wie folgt eine partielle Ordnung $\leq_M$ definieren. Für $\langle a_1, \ldots, a_n \rangle$ und $\langle a'_1, \ldots, a'_n \rangle \in M$ setzen wir $\langle a_1, \ldots, a_n \rangle \leq_M$

$\langle a_1', \ldots, a_n' \rangle$ genau dann, wenn $a_i \leq_i a_i'$ für alle $1 \leq i \leq n$. Die Ordnung $\leq_M$ wird oft als *Produktordnung* der Ordnungen $\leq_1, \ldots, \leq_n$ bezeichnet. Die nachfolgende Abbildung gibt die Produktordnung auf $\mathbb{N} \times \mathbb{N}$ an, wobei die natürliche Ordnung auf $\mathbb{N}$ die jeweilige Komponentenordnung darstellt. Pfeile deuten die zwei direkten Nachfolger eines Elements bezüglich der Produktordnung an, die Menge aller Nachfolger von $\langle 5, 3 \rangle$ zusammen mit $\langle 5, 3 \rangle$ ist dunkel hinterlegt.



8. Sind A, B Mengen und ist „$\leq_B$" eine partielle Ordnung auf B, so definiert

$$f \trianglelefteq g \quad :\Leftrightarrow \quad (\forall a \in A\colon f(a) \leq_B g(a))$$

eine partielle Ordnung auf der Menge B^A aller Abbildungen von A nach B. Der Leser möge sich überlegen, warum dieses Beispiel das vorausgegangene verallgemeinert.

9. Wir können Beispiel 4.26 wie folgt verallgemeinern. Es sei „$\leq$" eine partielle Ordnung auf der Menge B und $f\colon A \to B$ eine Funktion. Auf A definieren wir die Relation „$\preceq$" durch

$$a_1 \preceq a_2 \quad :\Leftrightarrow \quad f(a_1) \leq f(a_2).$$

Dann ist „$\preceq$" eine Quasi-Ordnung, aber im allgemeinen keine partielle Ordnung. Es heißt „$\preceq$" die durch f und „$\leq$" *induzierte* Quasi-Ordnung auf A.

10. Ist A eine Menge von Personen und bezeichnen wir mit E^+ die „Vorfahrbeziehung“ auf A, so ist E^+ eine strikte partielle Ordnung, aber im allgemeinen keine strikte lineare Ordnung.

11. Die Menge aller Äquivalenzrelationen auf einer Menge A wird durch die Verfeinerungsrelation partiell geordnet.

Weiter oben hatten wir dargestellt, daß der prinzipielle Unterschied zwischen echten Quasi-Ordnungen und partiellen Ordnungen darin besteht, daß bei Quasi-Ordnungen ein Vergleich zweier vergleichbarer Elemente möglicherweise keine eindeutige Reihenfolge der Elemente ergibt. Wir stellen nun die Frage, ob es möglich ist, ausgehend von einer beliebigen Quasi-Ordnung „$\preceq$“ auf einer Menge A zu einer „vereinfachten Repräsentation“ dieser Ordnung zu kommen, wo sich dieser unschöne Effekt in Luft auflöst. Dies führt uns auf eine nützliche Verwendungsweise von Äquivalenzrelationen, die uns auch später noch verschiedentlich begegnen wird. Die Idee besteht darin, daß daß man eine geeignete Äquivalenzrelation „$\sim$“ auf A einführt und zur Quotientenmenge $A/\sim$ (vgl. Def. 4.8) übergeht. Diese Menge kann man als eine Vereinfachung der Menge A ansehen, da wir sie durch „Identifizieren“ von Elementen aus A erhalten. Im Anschluß wird dann eine Ordnung „$\leq$“ auf $A/\sim$ eingeführt, die ihrerseits eine Vereinfachung von „$\preceq$“ darstellt und nun tatsächlich eine partielle Ordnung ist.

Beispiel 4.29 Wir wählen das Beispiel aus Abbildung 4.6 zum Ausgangspunkt. Bei Betrachten der Menge A fällt auf, daß alle Elemente der Teilmenge $\{1, 2\}$ wechselseitig in der Relation „$\preceq$“ stehen. Dasselbe gilt für die Elemente der Teilmenge $\{3, 4\}$. Wenn wir die nicht umkehrbaren Pfeile ignorieren, erhalten wir also eine Äquivalenzrelation, man rufe sich hierzu die Diskussion in Beispiel 4.14 und Figur 4.3 in Erinnerung! Zu beachten ist, daß gerade der Vergleich der Elemente 1 und 2 (analog 3 und 4) problematisch ist, da er zu keiner eindeutigen Reihenfolge zwischen 1 und 3 (analog 3 und 4) führt. Wenn wir nun wie in Beispiel 4.14 zur Quotientenmenge $A/\sim$ übergehen, so werden durch die kanonische Abbildung κ die problematischen Paare jeweils identifiziert. Es bleibt noch die Aufgabe zu lösen, auf $A/\sim$ eine Ordnung „$\leq$ einzuführen, die „$\preceq$“ in geeigneter Weise widerspiegelt und eine partielle Ordnung ist. Es gibt nur eine natürliche Weise: da *jedes* Element von $[3] = \{3, 4\}$ bezüglich „$\preceq$“ kleiner ist als *jedes* Element von $[1] = \{1, 2\}$ definieren wir $[3] < [1]$. Das Vorgehen ist in Abbildung 4.7 illustriert.

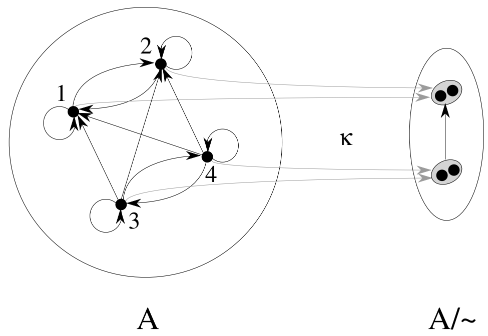


Abbildung 4.7: Übergang von einer Quasi-Ordnung zu vereinfachter partieller Ordnung durch Quotientenbildung.

Im nachfolgenden Lemma wird dieses Beispiel verallgemeinert. Das Lemma zeigt auch, daß bei geeigneter Betrachtungsweise jede Quasi-Ordnung durch eine Abbildung auf eine partiell geordnete Menge induziert ist im Sinn von Beispiel 4.28 Nr. 9. Ein wichtiges technisches Detail beim Beweis ist das folgende. Wir führen unten eine Relation zwischen Äquivalenzklassen dadurch ein, daß wir geeignete Bedingung an „die" Vertreter der Äquivalenzklassen stellen. Nun sind im allgemeinen die Vertreter gar nicht eindeutig bestimmt, da es ja viele Elemente in den Äquivalenzklassen gibt. In einem solchen Fall muß man sicherstellen, daß die Auswertung der Bedingung *nicht von der Wahl der Vertreter abhängt*. Wenn man für jede Auswahl von Vertretern stets dasselbe Resultat erhält, sagt man, daß die so eingeführte Relation *wohldefiniert* ist.

Lemma 4.30 *Es sei „$\preceq$" eine Quasi-Ordnung auf der Menge A. Dann wird durch*

$$a \sim b \quad :\Leftrightarrow \quad (a \preceq b \ \wedge \ b \preceq a)$$

eine Äquivalenzrelation auf A definiert. Auf der Quotientenmenge $A/\sim$ sei die Relation „$\leq$" durch

$$[a] \leq [b] \quad :\Leftrightarrow \quad a \preceq b$$

erklärt. Dann ist „$\leq$" wohldefiniert und eine partielle Ordnung.

Beweis. Wir gehen nur auf die Wohldefiniertheit ein, der restliche Beweis bleibt als Übung. Es seien $[a_1] = [a_2]$ und $[b_1] = [b_2]$ Äquivalenzklassen in $A/\sim$. Wir müssen zeigen, daß wir dasselbe Resultat erhalten, wenn wir zur Definition von „$\leq$“ einerseits $[a_1]$ und $[b_1]$, andererseits $[a_2]$ und $[b_2]$ verwenden. In anderen Worten: es ist hier zu zeigen, daß die Bedingungen $a_1 \preceq b_1$ und $a_2 \preceq b_2$ äquivalent sind. Wegen der Symmetrie der Behauptung reicht es zu zeigen, daß aus $a_1 \preceq b_1$ auch $a_2 \preceq b_2$ folgt. Es gelte also $a_1 \preceq b_1$. Wegen $[a_1] = [a_2]$ und $[b_1] = [b_2]$ gelten auch die folgenden Beziehungen: $a_2 \preceq a_1$ and $b_1 \preceq b_2$. Wegen der Transitivität von „$\preceq$“ folgt nun $a_2 \preceq b_2$. ■

Wenn wir wie in Definition 4.11 mit κ die kanonische Abbildung von A in die Quotientenmenge $A/\sim$ bezeichnen, so ist in der obigen Situation „$\preceq$“ genau die durch κ und „$\leq$“ induzierte Quasi-Ordnung auf A. Das nachfolgende Lemma verdeutlicht nochmals den Unterschied zwischen Quasi-Ordnungen und partiellen Ordnungen aus einer anderen Perspektive.

Lemma 4.31 *Es sei „$\leq$“ eine Quasi-Ordnung auf der Menge A. Dann ist „$\leq$“ eine partielle Ordnung genau dann, wenn A keine nicht-trivialen $\leq$–Zyklen (vgl. Definition 3.28) enthält.*

Beweis. Es sei „$\leq$“ eine partielle Ordung, somit antisymmetrisch. Es sei $a_0, \ldots, a_n$ ein $\leq$–Zyklus. Somit gilt $a_0 = a_n$. Wäre der Zyklus nicht-trivial, so enthielte er ein Element $a_i \neq a_0$. Aus der Transitivität von „$\leq$“ folgte einerseits $a_0 \leq a_i$, andererseits auch $a_i \leq a_n = a_0$. Nunmehr ergibt sich aber ein Widerspruch zur Antisymmetrie: wir haben $a_0 \leq a_i$, $a_i \leq a_0$ und $a_i \neq a_0$. Der Widerspruch zeigt, daß der Zyklus $a_0, \ldots, a_n$ trivial sein muß.

Nun nehmen wir umgekehrt an, daß A keine nicht-trivialen $\leq$–Zyklen enthält. Dann muß „$\leq$“ antisymmetrisch sein: andernfalls gäbe es Elemente $a_0, a_1 \in A$ mit $a_0 \leq a_1$ und $a_1 \leq a_0$ sowie $a_0 \neq a_1$. Dann wäre die Folge a_0, a_1, a_0 ein nicht-trivialer $\leq$–Zyklus. ■

Im ersten Teil des Beweises haben wir die Trivialität des betrachteten Zyklus dadurch bewiesen, daß wir aus der *gegenteiligen Annahme* einen *Widerspruch zur Voraussetzung*, daß $\leq$ eine partielle Ordnung ist, hergeleitet haben. Man nennt diese Technik „Beweis durch Kontraposition“. Aus Lemma 4.31 ergibt sich sofort

Lemma 4.32 *Es sei „$\leq$“ eine partielle Ordnung auf der nichtleeren endlichen Menge A. Dann besitzt A ein bezüglich „$\leq$“ minimales Element und ein bezüglich „$\leq$“ maximales Element.*

Beweis. Da $A \neq \emptyset$ können wir ein Element $a_0 \in A$ wählen. Wenn a_0 minimal ist, sind wir fertig. Wenn a_0 nicht minimal ist, so gibt es ein Element $a_1 < a_0$. Solange wir kein minimales Element gefunden haben, fahren wir auf diese Weise fort. Da A endlich ist und keine nicht-trivialen „$\leq$“–Zyklen enthält, müssen wir nach endlich viele Schritten bei einem minimalen Element angelangen. Wenn wir dieselbe Methode auf die duale Ordnung „$\geq$“ anwenden, erreichen wir ein minimales Element bezüglich „$\geq$“, also ein maximales Element bezüglich „$\leq$“. ∎

Lemma 4.33 *Jede partielle Ordnung „$\leq$“ auf einer endlichen Menge A läßt sich als Hasse-Diagramm darstellen.*

Beweis. Induktion über die Zahl n der Elemente von A.

Induktionsanfang, $n = 0$. In diesem Fall ist „$\leq$“ die leere Menge $\emptyset$, wir nehmen das leere Diagramm zur Repräsentation.

Induktionshypothese. Sein $n \geq 0$. Die Behauptung sei richtig für jede Menge A mit n Elementen.

Induktionsschritt. Es sei A eine Menge mit $n + 1$ Elementen $(n \geq 0)$ und „$\leq$“ eine partielle Ordnung auf A. Wir wählen gemäß Lemma 4.32 ein „$\leq$“-minimales Element a. Die Einschränkung von „$\leq$“ auf $A \setminus \{a\}$ ist offenkundig eine partielle Ordnung. Nach Induktionsvoraussetzung besitzt diese ein Hasse-Diagramm. Wir fügen zu diesem einen Punkt hinzu, der das Element a repräsentiert, desweiteren Kanten von a zu jedem direkten Nachfolger von a. Offenkundig erhalten wir damit ein Hasse-Diagramm für „$\leq$“. ∎

Ähnlich läßt sich zeigen:

Lemma 4.34 *Es sei A eine endliche Menge. Jede partielle Ordnung „$\leq_p$“ auf A läßt sich zu einer linearen Ordnung „$\leq$“ mit „$\leq_p$“ $\subseteq$ „$\leq$“ ergänzen.*

Beweis. Übung. ∎

4.4 Hüllenbildungen bei Relationen

Im vorausgegangenen Abschnitt hatten wir in Lemma 4.30 und in der zugehörigen Abbildung 4.7 an einem Beispiel gesehen, daß man manchmal durch den Übergang zur Quotientenmenge die Eigenschaften von Relationen angenehmer machen kann. In diesem Abschnitt geht es auch darum, wie man von einer Relation, die bestimmte wünschenswerte Eigenschaften nicht besitzt, zu einer verwandten Relation übergehen kann, die besser geartet ist. Die Idee ist diesmal, die Relation solange systematisch zu erweitern, bis die gewünschte Eigenschaft erreicht ist. Formal erreicht wird dies durch sogenannte Hüllenabbildungen. Von großer Bedeutung ist vor allem der unten eingeführte Begriff der (reflexiv-) transitiven Hülle einer Relation, auf den wir ausführlich eingehen. Da das Konzept der Hüllenbildung in sehr vielen Zusammenhängen nützlich ist, stellen wir zu Beginn kurz den allgemeinen Begriff vor.

Definition 4.35 Als *Hüllenabbildung* werden Abbildungen $H\colon \mathcal{P}(M) \to \mathcal{P}(M)$ bezeichnet, die einer Menge $A \subseteq M$ eine Menge $H(A) \subseteq M$ (ihre „Hülle") zuordnen derart, daß drei Bedingungen erfüllt sind:

(i) Es gilt stets $A \subseteq H(A)$ (Erweiterungseigenschaft),

(ii) $A \subseteq B$ impliziert stets $H(A) \subseteq H(B)$ (Monotonie),

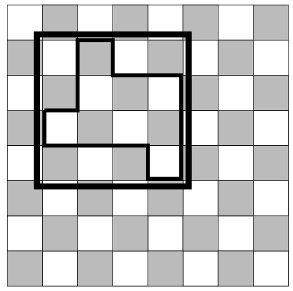
(iii) es gilt stets $H(H(A)) = H(A)$ (Idempotenz).

Hierbei ist M eine beliebige Menge. Ist $A \subseteq M$ und gilt $H(A) = A$, so wird A ein *Fixpunkt* der Hüllenabbildung H genannt.

Beispiel 4.36 Bezeichnet $\mathbb{R}$ die Menge der reellen Zahlen, und ist M_1 die Potenzmenge von $\mathbb{R}$, so ist für $A \in M_1$ die Menge

$$I_1(A) := \{x \in \mathbb{R} \mid \exists a_1, a_2 \in A\colon a_1 \leq x \leq a_2\}$$

das kleinste A umfassende „Intervall" von $\mathbb{R}$. Die Abbildung $I_1\colon M_1 \to M; A \mapsto I_1(A)$ ist eine Hüllenabbildung. Intervallbasierte Hüllenabbildungen gibt es auch in höheren Dimensionen, wie die nachfolgende Abbildung andeutet, wo wir einer gegebenen Fläche das kleinste sie umfassende Planrechteck als Hülle zuordnen. In der Tat kann man Planrechtecke als zweidimensionale Intervalle betrachten.



Es ist trivial zu zeigen, daß für jede Hüllenabbildung $H\colon \mathcal{P}(M) \to \mathcal{P}(M)$ stets die Grundmenge M selbst ein Fixpunkt von H ist. Interessanter ist der folgende Zusammenhang, der zeigt, daß Fixpunkte unter Durchschnittsbildung abgeschlossen sind.

Lemma 4.37 *Es sei $H\colon \mathcal{P}(M) \to \mathcal{P}(M)$ eine Hüllenabbildung und I eine nichtleere Indexmenge. Ist $A_i \subseteq M$ für jeden Index $i \in I$ ein Fixpunkt von H, so ist auch $\bigcap_{i\in I} A_i$ ein Fixpunkt von H.*

Beweis. Es sei A_i für jeden Index $i \in I$ ein Fixpunkt von H, es gelte also $H(A_i) = A_i$. Aufgrund der Erweiterungseigenschaft von H folgt

$$\bigcap_{i\in I} A_i \subseteq H(\bigcap_{i\in I} A_i).$$

Sei nun $j \in I$. Aus $\bigcap_{i\in I} A_i \subseteq A_j$ ergibt sich aufgrund der Monotonie $H(\bigcap_{i\in I} A_i) \subseteq H(A_j) = A_j$. Da diese Inklusion für jeden Index j gilt, folgt

$$H(\bigcap_{i\in I} A_i) \subseteq \bigcap_{i\in I} A_i.$$

Mit Lemma 2.6, Teil 2, erhält man $H(\bigcap_{i\in I} A_i) = \bigcap_{i\in I} A_i$. Damit ist $\bigcap_{i\in I} A_i$ ein Fixpunkt von H. ■

Man kann bei Vorliegen einer Durchschnittseigenschaft ähnlich zu der in Lemma 4.37 genannten auch umgekehrt hieraus eine geeignete Hüllenabbildung definieren. Hierzu vergleiche man Aufgabe 4.28.

Die Diskussion der Hüllenbildungen bei Relationen beginnen wir mit einer Vorbemerkung: Es sei $S \subseteq A \times A$ eine zweistellige Relation auf der Menge A. Es ist $A \times A$ eine S umfassende reflexive, symmetrische und transitive Relation. Aus Lemma 4.4 ergibt sich sofort, daß der Durchschnitt aller

S umfassender reflexiver (resp. symmetrischer, transitiver) Relationen auf der Menge A stets selbst eine S umfassende reflexive (resp. symmetrische, transitive) Relation auf A ist. Daher macht die nachfolgende Definition Sinn.

Definition 4.38 Es sei R eine zweistellige Relation auf der Menge A.

1. Die *reflexive Hülle* von von R ist die Relation
$$\bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ reflexiv}\}.$$
2. Die *symmetrische Hülle* von von R ist die Relation
$$\bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ symmetrisch}\}.$$
3. Die *transitive Hülle* von von R ist die Relation
$$R^+ := \bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ transitiv}\}.$$
4. Die *reflexiv-transitive Hülle* von von R ist die Relation
$$R^* := \bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ reflexiv und transitiv}\}.$$

Offenkundig ist die reflexive (resp. symmetrische, transitive, reflexiv-transitive) Hülle von R genau die kleinste R umfassende reflexive (resp. symmetrische, transitive, reflexive und transitive) Relation auf A. Es ist leicht zu zeigen, daß die Abbildung, die einer gegebenen Relation $R \subseteq A \times A$ ihre reflexive (resp. symmetrische, transitive, reflexiv-transitive) Hülle zuordnet, in der Tat eine Hüllenabbildung $H\colon \mathcal{P}(A \times A) \to \mathcal{P}(A \times A)$ im Sinn von Definition 4.35 ist, hierzu vergleiche Aufgabe 4.25.

Bemerkung 4.39 Aus der Tatsache, daß man den Übergang von einer Relation $R \subseteq A \times A$ zu ihrer reflexiven (resp. symmetrischen, transitiven) Hülle jeweils als Hüllenabbildung interpretieren kann, kann man ersehen, daß die drei Aussagen aus Lemma 4.4 lediglich Spezialfälle des in Lemma 4.37 ausgedrückten allgemeinen Sachverhalts sind. Hierzu muß man nur bemerken, daß jede reflexive (resp. symmetrische, transitive) Relation R ein Fixpunkt der entsprechenden Hüllenabbildung ist. Man vergleiche Aufgabe 4.27. Auch der Übergang von einer Relation $R \subseteq A \times A$ zur kleinsten R umfassenden Äquivalenzrelation ist eine Hüllenabbildung. Damit entpuppt sich auch Lemma 4.17 ein Spezialfall von Lemma 4.37.

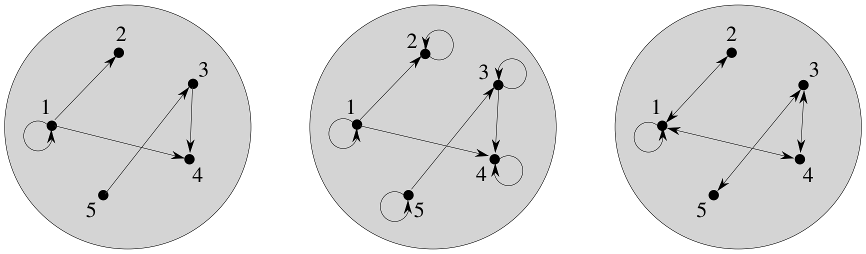


Abbildung 4.8: Reflexive (Mitte) und symmetrische Hülle (rechts) einer binären Relation (links).

Eine explizite Beschreibung der reflexiven und der symmetrischen Hülle liefert das folgende Lemma.

Lemma 4.40 *Es sei $R \subseteq A \times A$ wie oben. Dann ist $\mathrm{Id}_A \cup R$ die reflexive Hülle von R und $R \cup R^{-1}$ die symmetrische Hülle von R.*

Beweis. Es ist unmittelbar klar, daß $Id_A \cup R$ eine R umfassende reflexive Relation ist, und daß jede Relation mit diesen beiden Eigenschaften $Id_A \cup R$ als Teilmenge enthält. Damit gilt $Id_A \cup R = \bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ reflexiv}\}$. Analog ist die Argumentation im Fall der symmetrischen Hülle. ■

Für einen Beweis unter Verwendung von Lemma 4.3 vergleiche man Aufgabe 4.29. Abbildung 4.8 illustriert die reflexive und transitive Hülle einer Relation einer Relation R. Im Matrixbild ergibt sich die reflexive Hülle von R durch Auffüllen der Diagonalen mit Einträgen 1 (Mitte), die symmetrische Hülle durch Spiegeln aller Einträge 1 an der Diagonalen unter Beibehaltung der bisherigen Einträge 1 (rechts).

R	1	2	3	4	5
1	1	1	0	1	0
2	0	0	0	0	0
3	0	0	0	1	0
4	0	0	0	0	0
5	0	0	1	0	0

	1	2	3	4	5
1	1	1	0	1	0
2	0	1	0	0	0
3	0	0	1	1	0
4	0	0	0	1	0
5	0	0	1	0	1

	1	2	3	4	5
1	1	1	0	1	0
2	1	0	0	0	0
3	0	0	0	1	1
4	1	0	1	0	0
5	0	0	1	0	0

Definition 4.41 Es sei R eine binäre Relation auf der Menge A. Eine Relation $S \subseteq A \times A$ heißt *abgeschlossen unter Komposition mit R* (kurz: abgeschlossen unter R) genau dann, wenn $S \circ R \subseteq S$ gilt.

Transitive Hülle und reflexiv-transitive Hülle liefern zwei wichtige Beispiele einer Relation, die unter Komposition mit R abgeschlossen ist. Die nachfolgende Bemerkung gibt explizite Beschreibungen der transitiven Hülle und der reflexiv-transitiven Hülle mittels der in Abschnitt 3.6.1 besprochenen induktiven Definition von Mengen.

Bemerkung 4.42 Es sei R wie oben. Dann können wir die transitive Hülle R^+ induktiv wie folgt definieren. Als Basisregel nehmen wir $R \subseteq R^+$. Die einzige induktive Regel besagt:

Gilt $\langle a, b\rangle \in R^+$ und $\langle b, c\rangle \in R$, so ist $\langle a, c\rangle \in R^+$.

Kürzer können wir schreiben

$$R^+ \ ::= \ R \mid R^+ \circ R.$$

Wenn wir in die Abschnitt 3.6.1 erläuterte schrittweise Turmkonstruktion der definierten Menge durchführen, erhalten wir zunächst R, danach $R \cup R^2$, im dritten Schritt $R \cup R^2 \cup R^3$ etc. Daraus ergibt sich

$$R^+ = \bigcup_{n\geq 1} R^n.$$

Ein formaler Beweis der Gleichung, daß die induktiv definierte Menge $\bigcup_{n\geq 1} R^n$ tatrsächlich die transitive Hülle R^+ von R ist, ergibt sich wie folgt: gemäß Definition der transitiven Hülle gilt $R \subseteq R^+$. Da R^+ überdies transitiv ist, folgt mit Teil 3 aus Lemma 4.3 (wo R^+ an die Stelle von R tritt) $R^+ \circ R^+ \subseteq R^+$. Mit Teil 1 aus Lemma 3.26 ergibt sich hieraus $R \circ R \subseteq R^+$. Nun folgt analog

$$R \circ R \circ R \subseteq R^+ \circ R \subseteq R^+ \circ R^+ \subseteq R^+.$$

Eine einfache Induktion zeigt, daß $R^n \subseteq R^+$ für alle $n \geq 1$ und somit

$$\bigcup_{n\geq 1} R^n \subseteq R^+.$$

Andererseits folgt mit den Teilen 2 und 3 aus Lemma 3.26

$$\begin{aligned}
(\bigcup_{n\geq 1} R^n) \circ (\bigcup_{m\geq 1} R^m) &= \bigcup_{m\geq 1} ((\bigcup_{n\geq 1} R^n) \circ R^m) \\
&= \bigcup_{m\geq 1} (\bigcup_{n\geq 1} (R^n \circ R^m)) \\
&\subseteq \bigcup_{n\geq 1} R^n.
\end{aligned}$$

Daraus folgt mit Teil 3 aus Lemma 4.3, daß $\bigcup_{n\geq 1} R^n$ bereits transitiv ist. Offenkundig gilt $R = R^1 \subseteq \bigcup_{n\geq 1} R^n$. Da somit $\bigcup_{n\geq 1} R^n$ eine R umfassende transitive Relation ist, folgt

$$R^+ \subseteq \bigcup_{n\geq 1} R^n$$

und zusammen mit $\bigcup_{n\geq 1} R^n \subseteq R^+$ nun $R^+ = \bigcup_{n\geq 1} R^n$. ∎

Bemerkung 4.43 Die induktive Definition der reflexiv-transitiven Hülle ist ähnlich. Als Basisregel nehmen wir $Id_A \subseteq R^*$. Die einzige induktive Regel besagt:

Gilt $\langle a, b\rangle \in R^*$ und $\langle b, c\rangle \in R$, so ist $\langle a, c\rangle \in R^*$.

Kürzer können wir schreiben

$$R^* \ ::= \ Id_A \mid R^* \circ R.$$

Wenn wir wieder die schrittweise Konstruktion durchführen, erhalten wir nacheinander Id_A, danach $Id_A \circ \cup R = R = R^1$, daraufhin $Id_A \cup R^1 \cup R^2$ etc. Setzen wir $R^0 := Id_A$ so gilt also

$$R^* = \bigcup_{n\geq 0} R^n.$$

Der Beweis, daß die Menge $\bigcup_{n\geq 0} R^n$ in der Tat die reflexiv-transitive Hülle R^* ist, kann auf ähnliche Weise wie für die transitive Hülle geführt werden (Aufgabe 4.35). ∎

Eine weitere Beschreibung von transitiver und reflexiv-transitiver Hülle folgt unmittelbar aus Lemma 3.29:

Lemma 4.44 *Es sei R eine zweistellige Relation auf der Menge A. Dann gilt $R^+ = \{\langle a, b\rangle \in A \times A \mid \exists$ R-Kette der Länge $n \geq 1$ von a nach $b\}$ und $R^* = \{\langle a, b\rangle \in A \times A \mid \exists$ R-Kette der Länge $n \geq 0$ von a nach $b\}$.*

Beispiele 4.45 Einige Illustrationen sollen das Konzept der transitiven und der reflexiv-transitiven Hülle verdeutlichen.

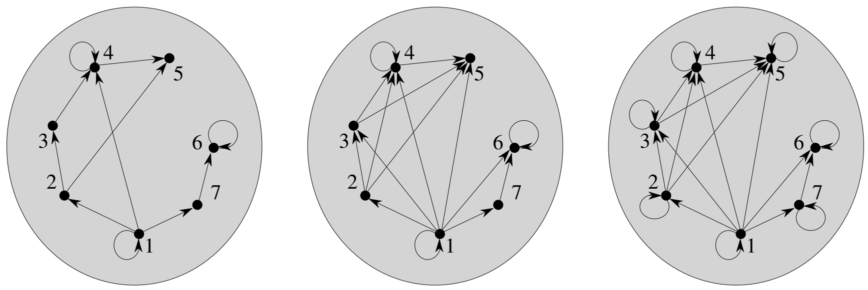


Abbildung 4.9: Transitive und reflexiv-transitive Hülle einer Relation.

1. Wenn die Relation $R \subseteq A \times A$ durch gerichtete Pfeile symbolisiert ist, so gehört ein Paar—nicht notwendig verschiedener—Punkte $\langle a, b\rangle$ genau dann zu R^*, wenn man von a nach b auf einem Weg gelangen kann, wo man immer geeigneten Pfeilen in der angegebenen Richtung folgt. Die Weglänge 0 ist hierbei möglich, alle Paare $\langle a, a\rangle$ gehören stets zu R^*. Bei der transitiven Hülle R^+ muß man zumindest einen R-Pfeil entlang laufen, wobei dieser allerdings durchaus wieder beim Ausgangselement enden kann. In Figur 4.9 ist links die Relation R, in der Mitte ihre transitive Hülle und rechts die reflexiv-transitive Hülle abgebildet.

2. Ist A eine Menge von Personen, und gibt E (resp. K) wie in Beispiel 3.16 Nr. 6 die Elternbeziehung und $K = E^{-1}$ die „ist-Kind-von" Beziehung an, so ist E^+ die Vorfahrbeziehung und K^+ die Nachfahrbeziehung.

3. Wenn R_L wie in den Beispielen 3.16 und 3.18 diejenige Relation ist, die die möglichen Zielfelder eines Läufers auf dem ansonsten leeren Schachbrett bei einem einzelnen Zug beschreibt, so umfaßt R_L^* alle Felderpaare, die zwei nicht notwendig verschiedene Felder derselben Farbe (schwarz oder weiß) enthalten. Etwas weniger trivial ist das Springerbeispiel. Ist R_S diejenige Relation, die die möglichen Zielfelder eines Springers auf dem ansonsten leeren Schachbrett bei einem einzelnen Zug beschreibt, so umfaßt R_S^* alle Felderpaare (vgl. Aufgabe 4.33).

4. Es sei A eine Menge von Punkten in einer Stadt, die Relation R gelte für $\langle a, b\rangle \in A \times A$ genau dann, wenn es eine direkte Busverbindung von a nach b gibt. Dann beschreibt die Relation R^* diejenigen Paare $\langle a, x\rangle$, wo es irgendeine Busverbindung — mit oder ohne Umsteigen, gegebe-

nenfalls ohne überhaupt zu fahren — von a nach x gibt. Insbesondere gehört die „Diagonale“ $\{\langle a, a\rangle \mid a \in A\}$ von $A \times A$ zu R^*.

5. Es sei $R := \{\langle n, n+1\rangle \in \mathbb{N} \times \mathbb{N} \mid n \geq 0\}$ die Nachfolgerrelation auf den natürlichen Zahlen. Dann beschreibt R^* gerade die übliche Ordnung „$\leq$“ auf $\mathbb{N}$.

6. Als Variante des vorausgegangenen Beispiels gilt folgendes: bezeichnet R die direkte Nachfolgerbeziehung der partiellen Ordnung „$\leq$“ auf der endlichen Menge M, so stimmen R^* (resp. R^+) und „$\leq$“ (resp. „$<$“) überein (Aufgabe 4.36).

Unser letztes Beispiel ist etwas komplizierter, verdeutlicht aber die grundsätzliche Bedeutung des Begriffs der reflexiv-transitiven Hülle im Zusammenhang mit sogenannten *Ableitungen* bei formalen Systemen.

Beispiel 4.46 Es sei M das Alphabet mit den Symbolen $\{S, A, B, a, b\}$. Das Symbol S bezeichnen wir als *Startsymbol*, die Symbole S, A und B nennen wir *nichtterminale Symbole*, die Buchstaben a und b nennen wir *terminale Symbole*. Wie üblich bezeichne M^* die Menge aller Wörter über M (vgl. Definition 3.9). Wir betrachten die „Ein-Schritt-Ableitungsrelation“ R auf M^*, die alle Paare der Form $\langle uSv, uASBv\rangle$, $\langle uSv, uv\rangle$, $\langle uAv, uav\rangle$ und $\langle uBv, ubv\rangle$ enthält. Hierbei bezeichnen u und v jeweils beliebige Wörter aus M^*. R kann als eine einfache formale *Grammatik* aufgefaßt werden. Ist $\langle S, u\rangle \in R^*$, so sagen wir, daß das Wort u (in endlich vielen Schritten) aus S *ableitbar* ist. Beispiele für Ableitungen sind

$$\begin{aligned}
S &\rightarrow \epsilon, \\
S &\rightarrow ASB \rightarrow AB \rightarrow aB \rightarrow ab, \\
S &\rightarrow ASB \rightarrow AASBB \rightarrow AABB \rightarrow aABB \rightarrow aAbB \rightarrow aabB \rightarrow aabb
\end{aligned}$$

wo „ϵ“ das leere Wort bezeichnet. Man beachte, daß jeweils jedes Paar aufeinanderfolgender Wörter zu R gehört. Damit gilt in der Tat $\langle S, \epsilon\rangle, \langle S, ab\rangle, \langle S, aabb\rangle \in R^*$. Die Menge $\{u \in \{a, b\}^* \mid \langle S, u\rangle \in R^*\}$ heißt die durch die Grammatik erzeugte *Sprache*. Sie umfaßt bei unseren Beispiel genau alle Wörter der Form $a^n b^n$ $(n \in \mathbb{N})$. Ähnliche Grammatiken werden zur Beschreibung von natürlichen Sprachen und Programmiersprachen verwendet. Das Beispiel zeigt, wie es mittels des Begriffs der reflexiv-transitiven Hülle möglich ist, trotz der endlichen Darstellung der Relation R eine unendliche Menge möglicher Ableitungen und damit eine unendliche Sprache zu charakterisieren. Ähnliche Formen der Ableitung treten bei vielen verwandten Grammatiken und verwandten formalen Systemen auf und werden in aller Regel mit Hilfe des Begriffs der reflexiv-transitiven Hülle einer „Ein-Schritt-Ableitungsrelation“ charakterisiert.

4.5 Ergänzung: Fundierte Ordnungen, wohlfundierte Induktion

Eine spezielle Klasse von Ordnungen, die viele Anwendungen besitzt, sind die sogenannten fundierten (auch: wohlfundierten) Ordnungen. Der Fundiertheitsbegriff für Quasi-Ordnungen ist etwas unübersichtlich, wir beschränken uns einfachheitshalber auf partielle und lineare Ordnungen.

Definition 4.47 Eine partielle Ordnung „$\leq$“ auf einer Menge A heißt *fundiert* genau dann, wenn jede nichtleere Teilmenge B von A ein bezüglich „$\leq$“ minimales Element enthält. Eine fundierte lineare Ordnung wird als *Wohlordnung* bezeichnet.

Die Fundiertheit einer partiellen Ordnung „$\leq$“ bedeutet, daß es keine unendlich absteigenden Ketten von Elementen $\cdots < a_{n+1} < a_n < \cdots a_2 < a_1 < a_0$ gibt (Übungsaufgabe 4.38). Der Begriff der fundierten partiellen Ordnung erlaubt eine Verallgemeinerung des Induktionsprinzips als Beweisprinzip, und parallel hierzu eine Verallgemeinerung des Prinzips der induktiven Definition von Mengen. Wir deuten kurz die Idee des verallgemeinerten Induktionsbeweises an.

Lemma 4.48 (Prinzip der wohlfundierten Induktion) *Es sei „$\leq$“ eine fundierte partielle Ordnung auf der Menge A. Es beschreibe φ eine Eigenschaft, die auf Elemente von A zutrifft oder nicht. Gilt*

(†) *jedes Element $a \in A$ besitzt die Eigenschaft φ, sobald alle Elemente aus der Menge der Vorgänger $\{b \in A \mid b < a\}$ die Eigenschaft φ besitzen,*

so haben alle Elemente aus A die Eigenschaft φ.

Beweis. Angenommen es gibt ein Element $a \in A$, das die Eigenschaft φ nicht besitzt. Wir betrachten die Teilmenge E von A, die all diejenigen Elemente enthält, die nicht die Eigenschaft φ besitzen. E ist nichtleer, da a nach Annahme zu E gehört. Wegen der Fundiertheit muß es ein „$\leq$"-minimales Element e_0 in E geben. Alle Elemente der Vorgängermenge $\{b \in A \mid b < e_0\}$ erfüllen damit φ nach Wahl von e_0. Aus der Voraussetzung (†) folgt jetzt aber, daß e_0 selbst die Eigenschaft φ besitzt. Dies stellt einen Widerspruch dar, damit müssen alle Elemente $a \in A$ die Eigenschaft φ besitzen. ■

Das (hypothetische) Element e_0 aus dem vorangegangenen Beweis wird oft als der „kleinste Verbrecher" bezeichnet, der durch Herleitung des Widerspruchs „gehängt" wird. Man spricht dann von der „Methode des kleinsten Verbrechers". Das Prinzip der vollständigen Induktion ergibt sich als der Spezialfall, wo $A = \mathbb{N}$ ist und „$\leq$" die übliche Wohlordnung auf $\mathbb{N}$.

In der Informatik spielt die Fundiertheit von Ordnungen eine große Rolle bei Terminierungsbeweisen, wenn es also darum geht, zu zeigen, daß ein Algorithmus für jede erlaubte Eingabe nach endlich vielen Schritten stoppt. Dabei bildet man die Zustände, die sich durch die Eingabe oder als Ergebnisse der Rechenschritte ergeben, auf passende Elemente einer Menge mit einer fundierten partiellen Ordnung ab. Wenn man nun zeigen kann, daß jeder Rechenschritt zu einem Zustand führt, der durch ein kleineres Element der partiellen Ordnung charakterisiert ist, so folgt aus dem Fundiertheitsprinzip sofort, daß es bei beliebiger Eingabe nur endlich viele Rechenschritte geben kann. Die Schwierigkeit ist im allgemeinen, eine geeignete fundierte partielle Ordnung zu finden. Die nachfolgend beschriebene Ordnung auf Multimengen ist hier häufig nützlich.

Bemerkung 4.49 Es sei „$<$" eine strikte partielle Ordnung auf der Menge A. Mit $\mathbb{N}^A$ bezeichnen wir die Menge aller Multimengen mit Elementen aus A (vgl. Abschnitt 2.7.2 sowie Beispiele 3.34 Nr. 7). Für $M, N \in \mathbb{N}^A$ setzen wir $N <_m M$ genau dann, wenn es Multimengen $\emptyset \neq K \subseteq_m M$ und $L \subseteq_m N$ gibt derart, daß $N = (M \setminus_m K) \cup_m L$ wo gilt $\forall l \in L\ \exists k \in K\colon k > l$. Hierbei stehen „$\subseteq_m$", „$\cup_m$" und „$\setminus_m$" für die Multimengen-Beziehungen und -Operationen, vgl. Abschnitt 2.7.2. Die Ordnung „$<_m$" wird *Multimengen-Ordnung* über „$<$" genannt. Intuitiv wird eine Multimenge M kleiner bezüglich „$<_m$", wenn wir ein Element herausnehmen und beliebig—aber endlich—viele kleinere (in der Ordnung „$<$") hinzufügen, wobei dieser Vorgang iteriert werden darf. Der wichtige Zusammenhang, den wir hier nicht beweisen ist der folgende: Es ist die Multimengen-Ordnung „$<_m$" stets fundiert, wenn die Ausgangsordnung „$<$" fundiert ist.

Bemerkung 4.50 Das Prinzip der wohlfundierten Induktion läßt sich auch Relationen anwenden, die nicht notwendigerweise transitiv sind. Wir nennen eine Relation $R \subseteq A \times A$ *wohlfundiert* genau dann, wenn es keine unendliche absteigende R-Kette

$$\ldots, R(a_{n+1}, a_n), R(a_n, a_{n-1}) \ldots, R(a_2, a_1), R(a_1, a_0)$$

gibt.[3] Sei im nachfolgenden R wohlfundiert. Es beschreibe φ eine Eigenschaft, die auf Elemente von A zutrifft oder nicht. Gilt

(††) jedes Element $a \in A$ besitzt die Eigenschaft φ, sobald alle Elemente aus der Menge der R-Vorgänger $\{b \in A \mid R^+(b, a)\}$ die Eigenschaft φ besitzen,

so haben alle Elemente aus A die Eigenschaft φ. Gäbe es nämlich ein Element a_0, das die Eigenschaft a_1 nicht besitzt, so würde man wegen (††) unter den R^+-Vorgängern von a_0 ein Element a_1 finden, auf das die Eigenschaft φ ebenfalls nicht zutrifft. Durch Iteration ergäbe sich eine unendliche R^+-Kette

$$\ldots, R^+(a_{n+1}, a_n), R^+(a_n, a_{n-1}) \ldots, R^+(a_2, a_1), R^+(a_1, a_0)$$

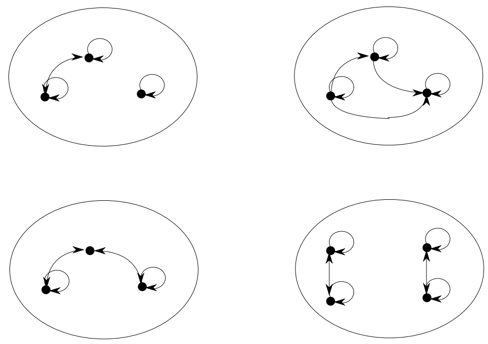
wo die Eigenschaft φ auf keines der Elemente a_i zutrifft ($i \in \mathbb{N}$). Eine solche Kette kann aber nicht existieren, da es ja keine unendliche absteigende R-Kette gibt. Ein anderer Beweis ergibt sich aus der Beobachtung, daß R als wohlfundierte Relation notwendig irreflexiv ist. Damit ist R^+ eine strikte partielle Ordnung. Man kann nun Lemma 4.48 zum Beweis der Behauptung verwenden. Die Einzelheiten bleiben dem Leser überlassen.

4.6 Aufgaben zu Kapitel 4

Aufgaben zu Teilkapitel 4.1

Aufgabe 4.1 Welche der in den nachfolgenden vier Abbildungen dargestellten Relationen sind reflexiv, welche transitiv, welche symmetrisch?

[3] Andere Autoren definieren wohlfundierte Relationen in der Form, daß es keine unendlichen *aufsteigenden* R-Ketten gibt. Man hat also die Richtung der verbotenen Kette zu beachten! Unsere Definition ist konsistent mit der obigen Definition der Fundiertheit einer partiellen Ordnung „$\leq$“.



Aufgabe 4.2 Beweisen Sie Lemma 4.4.

Aufgabe 4.3 Es sei $f\colon A \to A$ eine Funktion. Zeigen Sie:

- f ist nilpotent genau dann, wenn f als Relation betrachtet symmetrisch ist.
- f is idempotent genau dann, wenn f als Relation betrachtet transitiv ist.

Weisen Sie nun nach, daß die folgenden Eigenschaften äquivalent sind.

1. f ist als Relation reflexiv,
2. f ist als Relation idempotent und nilpotent,
3. $f = Id_A$.

Aufgabe 4.4 Es seien R und S zweistellige Relationen auf der Menge A. Zeigen Sie: Ist R oder S antisymmetrisch, so auch $R \cap S$.

Aufgabe 4.5 Ist die Vereinigung zweier reflexiver (resp. symmetrischer, antisymmetrischer, transitiver) Relationen auf einer Menge A stets selbst wieder reflexiv (resp. symmetrisch, antisymmetrisch, transitiv)? Geben Sie jeweils einen Beweis oder ein einfaches Gegenbeispiel.

Aufgabe 4.6 Ist die Komposition zweier reflexiver Relationen auf der Menge A stets selbst wieder reflexiv? Ist die Komposition zweier symmetrischer Relationen auf der Menge A stets selbst wieder symmetrisch? Ist die Komposition zweier antisymmetrischer Relationen auf der Menge A stets selbst wieder antisymmetrisch? Ist die Komposition zweier transitiver Relationen auf der Menge A stets selbst wieder transitiv? Geben Sie jeweils einen Beweis oder ein einfaches Gegenbeispiel.

Aufgabe 4.7 Es sei $R \subseteq A \times A$ eine Relation. Zeigen Sie:

1. R ist reflexiv genau dann, wenn $Id_A \subseteq R$ gilt.
2. R ist symmetrisch genau dann, wenn $R^{-1} \subseteq R$ gilt.
3. R ist transitiv genau dann, wenn $R \circ R \subseteq R$ gilt.
4. R ist antisymmetrisch genau dann, wenn $R \cap R^{-1} \subseteq Id_A$ gilt.

Aufgabe 4.8 Es sei A eine beliebige Menge und Id_A die Identitätsrelation auf A (vgl. Def. 3.35). Zeigen Sie: Id_A ist reflexiv, symmetrisch, antisymmetrisch und transitiv. Es ist Id_A die einzige Relation auf A mit diesen vier Eigenschaften.

Aufgabe 4.9 Es sei R eine binäre Relation auf der Menge A. Zeigen Sie

1. ist R reflexiv und euklidisch, so ist R auch symmetrisch und transitiv,
2. ist R reflexiv und anti-euklidisch, so ist R auch symmetrisch und transitiv.

Aufgabe 4.10 Es sei $M = \{a, b\}$. Geben Sie eine binäre Relation R auf M an, die weder reflexiv, noch transitiv, anti-symmetrisch, euklidisch oder anti-euklidisch ist.

Aufgabe 4.11 Es sei $M = \{a, b, c\}$. Geben Sie eine binäre Relation R auf M an, die keine der Eigenschaften (i)–(v) aus Definition 4.1 hat, und die nicht euklidisch ist.

Aufgaben zu Teilkapitel 4.2

Aufgabe 4.12 Durch welche Abbildungen sind die in den Beispielen 4.7 erwähnten Äquivalenzrelationen induziert (im Sinn von Beispiele 4.7 Nr. 1)?

Aufgabe 4.13 Beweisen Sie die zweite Teilaussage von Lemma 4.13.

Aufgabe 4.14 Es sei A eine nichtleere Menge und A^* die Menge aller Wörter über A (vgl. Definition 3.9). Die Konkatenation „$\circ$" sei wie in Beispiele 3.34 Nr. 5 erklärt. Sei $L \subseteq A^*$ eine formale Sprache über A. Für $v, v' \in A^*$ definieren wir

$$\begin{aligned} v \sim_L^r v' &:\Leftrightarrow (\forall w \in A^*\colon v \circ w \in L \Leftrightarrow v' \circ w \in L), \\ v \sim_L^l v' &:\Leftrightarrow (\forall u \in A^*\colon u \circ v \in L \Leftrightarrow u \circ v' \in L). \\ v \sim_L v' &:\Leftrightarrow (\forall u, w \in A^*\colon u \circ v \circ w \in L \Leftrightarrow u \circ v' \circ w \in L). \end{aligned}$$

Zeigen Sie:

1. die Relationen „$\sim_L^r$", „$\sim_L^l$" und „$\sim_L$" sind Äquivalenzrelationen auf A^*.
2. Gilt $v \sim_L^r v'$, so auch $v \circ w \sim_L^r v' \circ w$ für alle $w \in A^*$.
3. Gilt $v \sim_L^l v'$, so auch $u \circ v \sim_L^r u \circ v'$ für alle $u \in A^*$.
4. Gilt $v \sim_L v'$, so auch $u \circ v \circ w \sim_L u \circ v' \circ w$ für alle $u, w \in A^*$.
5. Gilt $v_1 \sim_L v_1'$ und $v_2 \sim_L v_2'$, so auch $v_1 \circ v_2 \sim_L v_1' \circ v_2'$.

Aufgabe 4.15 Es sei $A := \{a, b\}$ und $L := \{a^m b^n \mid m, n \in \mathbb{N}\}$. Wie sehen die Äquivalenzklassen der drei in Aufgabe 4.14 definierten Äquivalenzrelationen „$\sim_L^r$", „$\sim_L^l$" und „$\sim_L$" aus?

Aufgabe 4.16 Beweisen Sie Lemma 4.19.

Aufgabe 4.17 Es seien R und S zwei Äquivalenzrelationen auf der Menge A. Zeigen Sie: R verfeinert S genau dann, wenn $R \subseteq S$ gilt.

Aufgaben zu Teilkapitel 4.3

Aufgabe 4.18 Wir hatten zahlreiche partielle Ordnungen eingeführt, ohne den Nachweis zu führen, daß es sich um partielle Ordnungen handelt. Beweisen Sie dies für Beispiel 4.25 sowie die Beispiele 4.28 Nr. 8 und Nr. 2.

Aufgabe 4.19 Die Teilbarkeitsrelation T (vgl. 4.28 Nr. 2) definiert eine partielle Ordnung auf der Menge $\{1, 2, \ldots, 30\}$. Wie sehen die maximalen Elemente aus, wie die minimalen?

Aufgabe 4.20 Es bezeichne T die Teilbarkeitsrelation auf $\mathbb{N}$ (vgl. 4.28 Nr. 2). Wie sehen die direkten Nachfolger, wie die direkten Vorgänger einer Zahl n aus?

Aufgabe 4.21 Es sei $A_1 = A_2 := \{1, 2, 3\}$. Auf A_1 und A_2 betrachten wir die natürliche Ordnung der Elemente. Wie sieht die Produktordnung (vgl. 4.28 Nr. 7) auf $A_1 \times A_2$ aus? Zeichnen Sie ein Hasse-Diagramm.

Aufgabe 4.22 Beweisen Sie Lemma 4.30.

Aufgabe 4.23 Es seien „$\leq_1$" und „$\leq_2$" zwei lineare Ordnungen auf M und „$\leq_{1,2}$" ihr Durchschnitt. Zeigen Sie: es ist „$\leq_{1,2}$" eine lineare Ordnung genau dann, wenn „$\leq_1$" und „$\leq_2$" übereinstimmen.

Aufgabe 4.24 Es seien R_1 und R_2 partielle Ordnungen auf der Menge A. Ist auch $R_1 \cap R_2$ (resp. $R_1 \cup R_2$) eine partielle Ordnung auf A? Geben Sie einen Beweis oder ein einfaches Gegenbeispiel. Sie können gegebenenfalls Aufgabe 4.5 verwenden.

Aufgaben zu Teilkapitel 4.4

Aufgabe 4.25 Es sei A eine Menge. Zeigen Sie: die Abbildung, die einer binären Relation $R \subseteq A \times A$ ihre reflexive (resp. symmetrische, transitive, reflexiv-transitive) Hülle zuordnet, ist eine Hüllenabbildung im Sinn von Definition 4.35. Die Fixpunkte dieser Abbildung sind genau die reflexiven (resp. symmetrischen, transitiven, reflexiven und transitiven) Relationen auf A.

Aufgabe 4.26 Es seien R_1, R_2 und S Äquivalenzrelationen auf A. Zeigen Sie:

1. falls S sowohl R_1 als auch R_2 verfeinert, so auch $R_1 \cap R_2$.
2. falls R_1 und R_2 beide S verfeinern, so tut dies auch $(R_1 \cup R_2)^*$.

Aufgabe 4.27 Verwenden Sie Aufgabe 4.25 und Lemma 4.37, um einen einfachen Beweis von Lemma 4.4 zu erhalten.

Aufgabe 4.28 Es sei M eine Menge und E eine Eigenschaft, die auf Teilmengen von M zutrifft oder nicht. Wir sagen, daß E die *Durchschnittseigenschaft* hat, genau dann, falls folgende Bedingungen gelten:

1. M hat die Eigenschaft E.
2. Ist $\{A_i \mid i \in I\}$ eine nichtleere Familie von Teilmengen von M, wo jedes Element A_i die Eigenschaft E hat, so hat auch der Durchschnitt $\bigcap_{i\in I} A_i$ die Eigenschaft E.

Es habe nun E die Durchschnittseigenschaft. Zeigen Sie: die Abbildung H_E, die jeder Teilmenge $B \subseteq M$ den Durchschnitt aller B umfassenden Teilmengen von M mit der Eigenschaft E zuordnet, ist eine Hüllenabbildung. Die Fixpunkte sind genau die Teilmengen mit der Eigenschaft E.

Aufgabe 4.29 Beweisen Sie Lemma 4.40 mit Hilfe von Lemma 4.3 und Lemma 3.22.

Aufgabe 4.30 Beweisen Sie: ist $R \subseteq A \times A$ eine endliche Relation mit n geordneten Paaren, so ist auch R^+ endlich und hat höchstens n^2 geordnete Paare. Ist mit R auch stets R^* endlich?

Aufgabe 4.31 Zeigen Sie: ist $R \subseteq A \times A$ eine symmetrische Relation, so sind auch die transitive Hülle R^+ und die reflexiv-transitive Hülle R^* symmetrisch. Gilt auch jeweils die Umkehrung?

Aufgabe 4.32 Es sei $R \subset \mathbb{N} \times \mathbb{N}$ definiert durch

$$R(n,m) \Leftrightarrow m = n + 2.$$

Beschreiben Sie die Relationen R^i $(i \in \mathbb{N})$ und R^*.

Aufgabe 4.33 Zeigen Sie: ist R_S diejenige Relation, die die möglichen Zielfelder eines Springers auf dem ansonsten leeren Schachbrett bei einem einzelnen Zug beschreibt, so umfaßt R_S^* alle Felderpaare. (Man kann sogar mit dem Springer alle Felder des Schachbretts durchlaufen, ohne daß eines der Felder zweimal besucht wird.)

Aufgabe 4.34 Geben sie eine möglichst kleine Relation $R \subseteq \mathbb{N} \times \mathbb{N}$ an, so daß $R^* = \mathbb{N} \times \mathbb{N}$ gilt.

Aufgabe 4.35 Es sei $R \subseteq A \times A$. Beweisen Sie, daß $\bigcup_{n \geq 0} R^n$ die reflexiv-transitive Hülle von R ist.

Aufgabe 4.36 Zeigen Sie: Bezeichnet R die direkte Nachfolgerbeziehung der partiellen Ordnung „$\leq$“ auf der endlichen Menge M, so stimmen R^* (resp. R^+) und „$\leq$“ (resp. „$<$“) überein. Gilt dies auch stets für partielle Ordnungen auf unendlichen Mengen?

Aufgabe 4.37 Es sei $R \subseteq A \times A$. Gilt stets $(R^{-1})^* = (R^*)^{-1}$? Geben Sie einen Beweis oder ein Gegenbeispiel.

Aufgaben zu Teilkapitel 4.5

Aufgabe 4.38 Es sei „$\leq$“ eine partielle Ordnung auf der Menge A. Zeigen Sie: „$\leq$“ ist fundiert genau dann, wenn es in A keine unendlich absteigende Kette von Elementen $\cdots < a_{n+1} < a_n < \cdots a_2 < a_1 < a_0$ gibt.

4.7 Bibliographische Angaben

Die im letzten Abschnitt betrachteten fundierten partiellen Ordnungen sind Spezialfälle sogenannter *abstrakter Reduktionssysteme*, die in verschiedenen

Gebieten der Informatik, vor allem im Bereich der Termersetzung, betrachtet werden. Mehr Hintergrund zu diesem Thema bietet Kapitel 1 aus [BN98]. Die Multimengen-Ordnung (vgl. Bemerkung 4.49) wurde in [DM79] eingeführt.

5

Abzählbare und überabzählbare Mengen

In diesem Kapitel werden wir zeigen, daß es auch im Bereich der unendlichen Mengen interessante Größenunterschiede gibt. Die „kleinsten" unendlichen Mengen sind die in Abschnitt 5.2 definierten „abzählbar unendlichen" Mengen, für die die Menge der natürlichen Zahlen das Standardbeispiel liefert. Wir werden sehen, daß die Klasse der abzählbar unendlichen Mengen unter vielen mengentheoretischen Operationen abgeschlossen ist. Beispiele für die in Abschnitt 5.3 betrachteten sogenannten „überabzählbaren" Mengen bilden die Menge der reellen Zahlen und die Menge aller Abbildungen von der Menge der natürlichen Zahlen in sich selbst. Es wird schließlich gezeigt, daß in einem präzisen Sinn die Potenzmenge einer Menge stets echt größer ist als die Menge selbst. Insbesondere gilt dies auch für unendliche Mengen. Die Betrachtungen dieses Kapitels werden uns mit dem auf G. Cantor zurückgehenden „Diagonalverfahren" bekannt machen. Hinter diesem Begriff verbirgt sich ein extrem wichtiges Beweisprinzip, das aufgrund seiner Rolle bei vielen fundamentalen Erkenntnissen zu den immanenten Grenzen der Mathematik und Informatik eine fast populärwissenschaftliche Berühmtheit gefunden hat. Am Kapitelende beschreiben wir in informellem Stil zwei Anwendungen. Es wird gezeigt, daß man viele der Sprachen über einem gegebenen Alphabet nicht durch Grammatiken beschreiben kann, und daß die wenigsten Funktionen auf den natürlichen Zahlen durch Computerprogramme berechenbar sind.

5.1 Größenbegriff bei unendlichen Mengen

Wenn man nach der Größe von unendlichen Mengen fragt, muß man sich zunächst von einigen gewohnten Denkweisen freimachen. Vom Bereich des Endlichen sind wir gewohnt, eine echte Teilmenge B einer Menge A auch als echt kleiner einzustufen. Nun ist zum Beispiel die Menge $\mathbb{N}$ der natürlichen Zahlen eine echte Teilmenge der Menge $\mathcal{Z}$ der ganzen Zahlen. Die folgende Überlegung zeigt jedoch, daß beide Mengen in einem klar festgelegten Sinn gleich groß sind: wir können leicht eine Bijektion von $\mathbb{N}$ auf $\mathcal{Z}$ angeben, indem wir allen geraden natürlichen Zahlen $2n$ die Zahl n zuordnen, allen ungeraden natürlichen Zahlen $2n+1$ die negative Zahl $-(n+1)$. Dies zeigt, daß wir den natürlichen Zahlen nur „andere Namen“ zu geben brauchen, um alle ganzen Zahlen zu erhalten. Nun haben die Namen der Elemente offenkundig keinen Einfluß auf die Größe einer Menge. Die Mengen $\mathbb{N}$ und $\mathcal{Z}$ sind also gleichgroß. Da man jede Bijektion als eine ähnliche Umbenennung verstehen kann, ist die folgende Definition gerechtfertigt:

Definition 5.1

1. Zwei Mengen A und B heißen *gleichmächtig* genau dann, wenn es eine Bijektion von A nach B gibt. Wir schreiben $|A| = |B|$.
2. Eine Menge B heißt *echt mächtiger* als eine Menge A, wenn es zwar eine injektive Abbildung $f\colon A \to B$ gibt, aber keine Bijektion. Wir schreiben $|A| < |B|$.
3. Eine endliche Menge hat *Mächtigkeit* (oder *Kardinalität*) $n \in \mathbb{N}$, symbolisch $|A| = n$, genau dann, wenn es eine Bijektion von $\{1, 2, \dots, n\}$ auf A gibt.

Wie wir es von Ordnungen gewohnt sind, können wir die in den Fällen 1 und 2 dargestellten Situationen notationell zusammenfassen. Wir schreiben nachfolgend $|A| \leq |B|$ genau dann, wenn es eine injektive Funktion von A nach B gibt. Dies ist allerdings zunächst lediglich eine Notation. Es stellt sich die Frage, ob die hier dargestellte Beziehung zwischen (Mächtigkeiten von) Mengen die üblichen Gesetzmäßigkeiten erfüllt, die wir von Ordnungen her kennen. Eine Teilanwort ergibt sich aus dem folgenden Lemma.

Lemma 5.2 *Für jede Menge A gilt $|A| \leq |A|$ (Reflexivität). Für beliebige Mengen A, B und C gilt: aus $|A| \leq |B|$ und $|B| \leq |C|$ folgt $|A| \leq |C|$ (Transitivität).*

Die erste Aussage ergibt sich sofort daraus, daß die Identitätsfunktion Id_A (vgl. Def. 3.35) eine injektive Abbildung von A nach A ist. Es bleibt dem Leser überlassen nachzuweisen, daß auch das im Lemma beschriebene Gesetz der Transitivität gilt (vgl. Aufgabe 5.1). Die Regeln, die Quasi-Ordnungen charakterisieren, sind damit erfüllt. Weitergehend gilt auch die Antisymmetrie. Dies ist der Inhalt des Satzes von Cantor-Bernstein.

Satz 5.3 (Satz von Cantor-Bernstein) *Es seien A und B Mengen. Wenn es zwei injektive Abbildungen $f\colon A \to B$ und $g\colon B \to A$ gibt, so sind A und B gleichmächtig. Aus $|A| \leq |B|$ und $|B| \leq |A|$ folgt also $|A| = |B|$.*

Beweis. Es seien $f\colon A \to B$ und $g\colon B \to A$ zwei injektive Abbildungen. Nach Lemma 3.53 gibt es zu f und g rechtsinverse Abbildungen h und k, für die also $f \circ h = Id_A$ und $g \circ k = Id_B$ gilt. Es stimmt h (resp. k) auf $f(A)$ (resp. $g(B)$) mit f^{-1} (resp. g^{-1}) überein. Als *unmittelbaren Vorgänger* eines Elements $z \in f(A) \cup g(B)$ bezeichnen wir das Urbild $f^{-1}(z)$ resp. $g^{-1}(z)$. Für jede natürliche Zahl n definieren wir nun induktiv n-Vorgänger eines Elements $z \in f(A) \cup g(B)$. Hierbei hat z für jedes n maximal einen n-Vorgänger.

Als 0-*Vorgänger* eines Elements $z \in f(A) \cup g(B)$ bezeichnen wir z selbst. Hat z einen n-Vorgänger y, und hat y selbst einen unmittelbaren Vorgänger x, so nennen wir x den (eindeutig bestimmten) $(n+1)$-Vorgänger von z.

Um eine Bijektion F von A auf B anzugeben, teilen wir die Elemente von $a \in A$ in drei Klassen ein:

1. a hat *Typ 1* genau dann, wenn a für jedes $n \in \mathbb{N}$ einen n-Vorgänger hat.
2. a hat *Typ 2* genau dann, wenn a einen letzten Vorgänger hat, und wenn dieser in A liegt.
3. a hat *Typ 3* genau dann, wenn a einen letzten Vorgänger hat, und wenn dieser in B liegt.

In Fall 3 hat a zumindest einen 1-Vorgänger und liegt damit in $g(B)$. Abbildung 5.1 illustriert diese Begriffe. Die Abbildung $F\colon A \to B$ sei wie folgt erklärt:

$$F(y) = \begin{cases} f(a) & \text{falls } a \text{ Typ 1 oder 2 hat,} \\ g^{-1}(a) & \text{falls } a \text{ Typ 3 hat.} \end{cases}$$

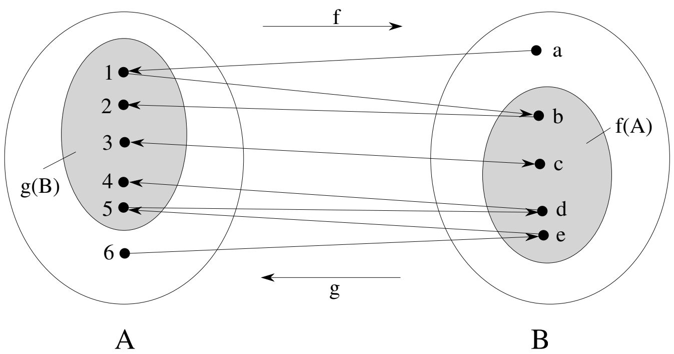


Abbildung 5.1: Zum Beweis des Satzes von Cantor-Bernstein: Element 3 hat Typ 1, Elemente 4, 5 und 6 haben Typ 2, Elemente 1 und 2 haben Typ 3.

Wir zeigen zunächst, daß F injektiv ist. Es seien $a_1, a_2 \in A$ und es gelte $F(a_1) = F(a_2)$. Wir müssen zeigen, daß dann $a_1 = a_2$ gilt. Sind a_1 und a_2 beide vom Typ 1 oder 2, so folgt aus der Injektivität von f wie gewünscht $a_1 = a_2$. Sind a_1 und a_2 beide vom Typ 3, so ist zu beachten, daß auch $g^{-1}\colon f(B) \to B$ injektiv ist. Dies folgt sofort daraus, daß g eine Funktion ist. Damit folgt aber aus $F(a_1) = g^{-1}(a_1) = F(a_2) = g^{-1}(a_2)$ wieder $a_1 = a_2$. Im letzten verbleibenden Fall können wir annehmen, daß a_1 Typ 1 oder 2 hat, und daß a_2 vom Typ 3 ist. Aus $f(a_1) = g^{-1}(a_2)$ ergibt sich durch Anwenden der Abbildung f^{-1} die Gleichung $a_1 = f^{-1}(g^{-1}(a_2))$. Es ist also a_1 ein 2-Vorgänger von a_2. Damit müssten aber offenkundig a_1 und a_2 denselben Typ haben, was einen Widerspruch darstellt. Wir können daher diesen Fall ausschließen.

Es bleibt noch zu zeigen, daß F surjektiv ist. Sei $b \in B$ gegeben. Wir setzen $a := g(b)$. Hat a Typ 3, so folgt $F(a) = g^{-1}(a) = b$ und b tritt als Bild unter F auf. Hat a den Typ 1 oder 2, so muß a einen 2-Vorgänger $a' := f^{-1}(b)$ haben, da ja b bereits ein 1-Vorgänger von a ist. Es hat dann auch a' Typ 1 oder 2. Damit gilt $F(a') = f(a') = b$ und b tritt auch hier als Bild unter F auf.

Wir haben gezeigt, daß F eine Bijektion zwischen A und B ist. Demzufolge sind A und B gleichmächtig. ■

5.2 Abzählbar unendliche Mengen

In diesem Kapitel untersuchen wir, welche Mengen mit $\mathbb{N}$ gleichmächtig sind, und unter welchen Mengenoperationen diese Mengen abgeschlossen sind.

Definition 5.4 Eine Menge A heißt *abzählbar unendlich* genau dann, wenn eine Bijektion $f\colon \mathbb{N} \to A$ existiert.

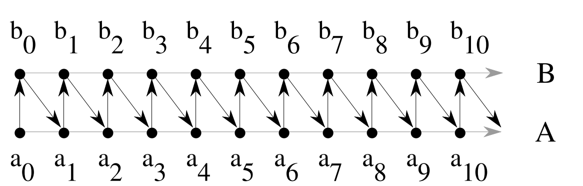
Bemerkung 5.5 Eine Menge, die endlich oder abzählbar unendlich ist, wird *abzählbar* genannt. Ist A abzählbar unendlich, und ist $f\colon \mathbb{N} \to A$ eine Bijektion, so können wir statt $f(n)$ auch einfacher a_n schreiben. Es ist dann $A = \{a_n \mid n \in \mathbb{N}\}$. Die Überlegung zeigt, daß eine Menge A genau dann abzählbar unendlich ist, wenn sie sich bei geeigneter Indizierung in der Form $A = \{a_n \mid n \in \mathbb{N}\}$ darstellen läßt, wobei für $n \neq m$ stets $a_n \neq a_m$ zu gelten hat.

Lemma 5.6 *Wenn A und B abzählbar unendliche Mengen sind, so ist auch die Vereinigung $A \cup B$ abzählbar unendlich.*

Beweis. Es seien A und B abzählbar unendliche Mengen. Nach Bemerkung 5.5 können wir die beiden Mengen in der Form $A = \{a_n \mid n \in \mathbb{N}\}$ beziehungsweise $B = \{b_n \mid n \in \mathbb{N}\}$ darstellen. Wir erhalten zunächst eine surjektive Funktion $g\colon \mathbb{N} \to A \cup B$, indem wir

$$g(n) = \begin{cases} a_{n/2} & \text{falls } n \text{ gerade} \\ b_{(n-1)/2} & \text{falls } n \text{ ungerade} \end{cases}$$

setzen. Die Funktion g läßt sich wie folgt bildlich darstellen—wir erhalten die Folge $g(0), g(1), \ldots$, indem wir den Pfeilen von $g(0) = a_0$ beginnend folgen:



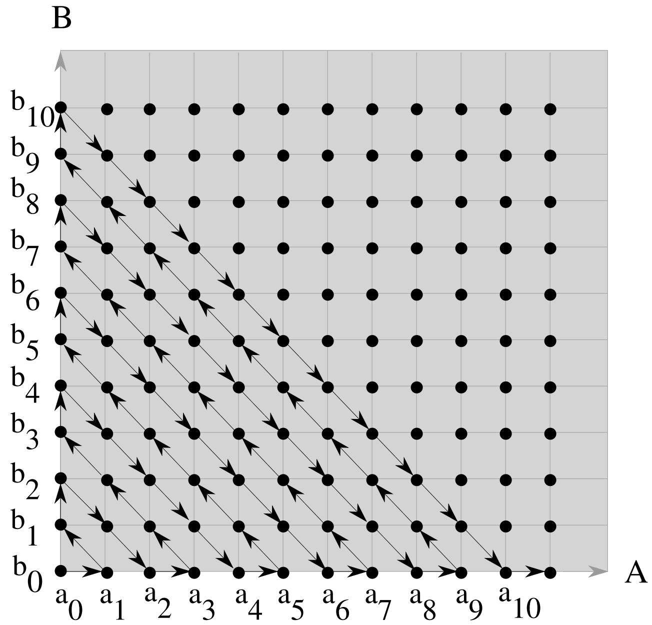


Abbildung 5.2: Das kartesische Produkt zweier abzählbar unendlicher Mengen ist abzählbar unendlich.

Offenkundig ist g injektiv (und damit bijektiv) genau dann, wenn A und B disjunkt sind. Da wir nicht voraussetzen, daß A und B disjunkte Mengen sind, müssen wir im Fall $A \cap B \neq \emptyset$ die Abbildung g noch etwas abändern, um eine bijektive Funktion zu erhalten. Es reicht, wenn wir induktiv definieren $f(0) := g(0) = a_0$ sowie $f(n) :=$ das erste Element der Folge $g(n), g(n+1), g(n+2), \ldots$, das nicht in $\{f_0, \ldots, f_{n-1}\}$ enthalten ist. ■

Lemma 5.7 *Ist $n \geq 1$ und sind $A_1, \ldots, A_n$ abzählbar unendliche Mengen, so ist auch die Vereinigung $\bigcup_{i=1,\ldots n} A_i$ abzählbar unendlich.*

Beweis. Dies folgt mit einer offenkundigen Induktion aus Lemma 5.6. ■

Lemma 5.8 *Wenn A und B abzählbar unendliche Mengen sind, so ist auch das kartesische Produkt $A \times B$ abzählbar unendlich.*

Beweis. Es seien A und B abzählbar unendliche Mengen. Nach Bemerkung 5.5 können wir die beiden Mengen in der Form $A = \{a_n \mid n \in \mathbb{N}\}$ beziehungsweise $B = \{b_n \mid n \in \mathbb{N}\}$ darstellen. Man kann nun eine Bijektion f von $\mathbb{N}$ auf $A \times B$ an Figur 5.2 ablesen, wo die Elemente des kartesischen Produkts $A \times B$ genau den Gitterpunkten entsprechen. Dazu beginnen wir mit $f(0) := \langle a_0, b_0 \rangle$. Die Aufzählung der weiteren Elemente $f(1), f(2)$ etc. erhalten wir, indem wir den Pfeilen folgen. Es ist offenkundig, daß die Abbildung f in der Tat eine Bijektion ist. Damit ist das Produkt $A \times B$ abzählbar unendlich. ■

Lemma 5.9 *Ist $n \geq 1$ und sind $A_1, \ldots, A_n$ abzählbar unendliche Mengen, so ist auch das kartesische Produkt $\prod_{i=1,\ldots n} A_i$ abzählbar unendlich.*

Beweis. Dies folgt leicht aus Lemma 5.8 mittels vollständiger Induktion. ■

Lemma 5.10 *Für jedes $n \in \mathbb{N}$ bezeichne A_n eine abzählbar unendliche Menge. Dann ist die Vereinigung $\bigcup_{n \in \mathbb{N}} A_n$ abzählbar unendlich.*

Beweis. Nach Bemerkung 5.5 können wir jede der Mengen A_n in der Form $A_n = \{a_{n,i} \mid i \in \mathbb{N}\}$ indexieren[1]. Wie in Figur 5.3 dargestellt, können wir eine Synthese der Beweisideen der Lemmata 5.6 und 5.8 verwenden, um zunächst eine surjektive Funktion $g\colon \mathbb{N} \to \bigcup_{n \in \mathbb{N}} A_n$ zu definieren. Wir erhalten die Folge der Elemente $g(0), g(1), \ldots$ indem wir von $a_{0,1}$ beginnend den Pfeilen folgen. Ausgehend von g können wir wie im Beweis von Lemma 5.6 leicht eine bijektive Funktion $f\colon \mathbb{N} \to \bigcup_{n \in \mathbb{N}} A_n$ gewinnen, indem wir mehrfach aufgezählte Elemente in $g(0), g(1), \ldots$ jeweils überspringen. ■

Lemma 5.11 *Es sei $A \neq \emptyset$ ein endliches oder abzählbar unendliches Alphabet. Dann ist die Menge A^* aller Wörter über A abzählbar unendlich.*

Beweis. 1. Ist A endlich und hat genau $k > 0$ Elemente, so gibt es für jede mögliche Länge n genau k^n viele Wörter der Länge n über A. In der Tat, da man jedes Wort der Länge n als eine Abbildung von $\{1, \ldots, n\}$ in A auffassen kann (vgl. hierzu die Diskussion in Abschnitt 3.6.3), folgt dies unmittelbar aus Lemma 3.33. Da jede der Mengen A^n ($n \in \mathbb{N}$) damit endlich

[1] Für Leser mit Hintergrundwissen in Mengenlehre sei bemerkt, daß wir hier eine schwache Form des sogenannten *Auswahlaxioms* benutzen.

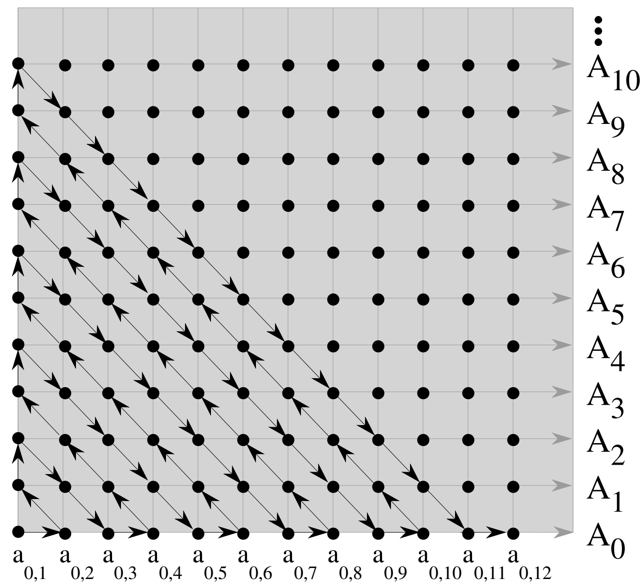


Abbildung 5.3: Die Vereinigung abzählbar unendlich vieler abzählbar unendlicher Mengen ist abzählbar unendlich.

ist, können wir als Folge $f(0), f(1), \ldots$ zuerst das leere Wort, danach alle Wörter der Länge 1 über A, danach alle Wörter der Länge 2 über A etc. aufzählen. Hierbei können wir etwa annehmen, daß die Menge der Wörter fester Länge jeweils in der lexikographischen Ordnung aufgeführt werden. Dadurch ergibt sich offenkundig eine Bijektion von $\mathbb{N}$ auf $A^* = \bigcup_{n\in\mathbb{N}} A^n$.

2. Ist A abzählbar unendlich, so folgt aus Lemma 5.9 zunächst, daß jede der Mengen A^n ($n \in \mathbb{N}$) abzählbar unendlich ist. Nun zeigt Lemma 5.10, daß auch $A^* = \bigcup_{n\in\mathbb{N}} A^n$ abzählbar unendlich ist. ■

5.3 Überabzählbare Mengen

Wir werden nun zeigen, daß es Mengen gibt, die echt mächtiger sind als $\mathbb{N}$. Mengen dieser Art heißen auch *überabzählbar*.

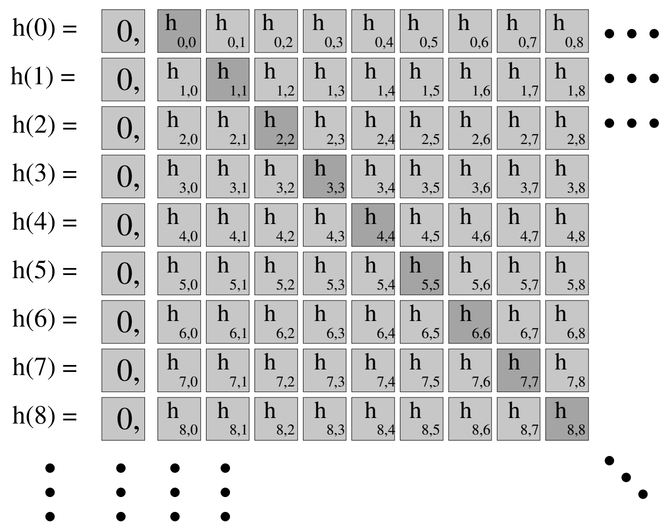
Satz 5.12 (Cantor) *Die Menge* $\mathbb{R}$ *der reellen Zahlen ist überabzählbar.*

Beweis. Wir zeigen, daß die Menge der reellen Zahlen im Intervall $[0, 1]$ überabzählbar ist. Damit ist offensichtlich $\mathbb{R}$ selbst überabzählbar.

Die Abbildung $f\colon \mathbb{N} \to [0, 1]$; $n \mapsto 1/(n + 1)$ ist offenkundig injektiv. Nach Definition 5.1 Nr. 2 bleibt zu zeigen, daß es keine Bijektion zwischen $\mathbb{N}$ und $[0, 1]$ gibt. Wir nehmen an, daß es eine solche Bijektion gibt und führen die Annahme zu einem Widerspruch. Es sei also h eine Bijektion von $\mathbb{N}$ auf $[0, 1]$. Wir wählen eine eindeutige Dezimaldarstellung und schreiben nun die Glieder der Folge

$$h(0), h(1), h(2), \ldots$$

in dieser Dezimaldarstellung untereinander. In der nachfolgenden Abbildung sind also die Einträge $h_{i,j}$ Ziffern aus $\{0, 1, \ldots, 9\}$ und jede Zahl aus $[0, 1]$ ist nach Voraussetzung in genau einer Zeile aufgeführt.



Wir konstruieren nun eine neue Zahl dadurch, daß wir die Ziffern der Diagonalen überall abändern. Wenn $h_{i,i}$ eine der Ziffern $\{0, 1, 2, 3, 4, 5\}$ ist, so sei $h'_{i,i}$ die Ziffer 8, falls aber $h_{i,i}$ eine der Ziffern $\{6, 7, 8, 9\}$ ist, so sei $h'_{i,i}$ die Ziffer 1. Nun liegt die Zahl

$$d = 0, h'_{0,0}\ h'_{1,1}\ h'_{2,2}\ h'_{3,3} \ldots$$

sicherlich im Intervall $[0, 1]$. Sie müßte also gemäß unserer Annahme mit einer der Zeilen $h(j)$ zusammenfallen. Dies ist jedoch nicht der Fall: die j-te Ziffer der Zahl $h(j)$ lautet $h_{j,j}$, die j-te Ziffer von d lautet $h'_{j,j}$ und diese Ziffern sind nach Konstruktion verschieden. Wenn aber die j-ten Ziffern zweier Zahlen verschieden sind, dann müssen natürlich die Zahlen selbst verschieden sein, somit gilt $d \neq h_j$. Da j beliebig war, kann d nicht in der Liste $h(0), h(1), \ldots$ sein. Wir erhalten also den erwünschten Widerspruch. Es kann somit keine Bijektion zwischen $\mathbb{N}$ und $[0, 1]$ geben. Damit sind $[0, 1]$ und $\mathbb{R}$ selbst überabzählbar. ■

Das hier angewandte Verfahren geht auf Georg Cantor zurück und ist unter der Bezeichung *Cantorsches Diagonalverfahren* bekannt. Man beachte, daß sich die kritische Zahl d durch Abänderung der Ziffern der Diagonalen ergibt. Durch ein ähnliches Diagonalargument kann das folgende Lemma bewiesen werden.

Lemma 5.13 *Die Menge* $\mathbb{N}^{\mathbb{N}}$ *aller Abbildungen von* $\mathbb{N}$ *in* $\mathbb{N}$ *ist überabzählbar.*

Beweis. Übung (vgl. Aufgabe 5.4). ■

Natürlich kann man im Bereich des Endlichen zu jeder Menge eine echt größere angeben. Die vorigen Resultate zeigen, daß es auch im Unendlichen zumindest gewisse Größenunterschiede gibt: die Mengen $\mathbb{R}$ und $\mathbb{N}^{\mathbb{N}}$ sind echt mächtiger als die Menge $\mathbb{N}$, im Sinn von Definition 5.1. Der nun folgende Satz von Cantor zeigt, daß es auch im Bereich des Unendlichen sogar unendlich viele Größenunterschiede gibt.

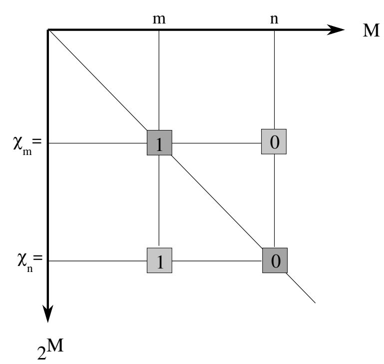
Satz 5.14 (Cantor) *Sei M eine beliebige Menge. Dann ist die Potenzmenge $\mathcal{P}(M)$ echt mächtiger als M.*

Beweis. Gemäß Definition 5.1 Nr. 2 ist zu zeigen, daß M zwar injektiv in $\mathcal{P}(M)$ abbildbar ist, aber nicht bijektiv. Die Abbildung $f\colon m \mapsto \{m\}$ liefert eine Injektion von M in $\mathcal{P}(M)$. Angenommen es gäbe eine Bijektion $b\colon M \to \mathcal{P}(M)$. Es sei

$$D = \{m \in M \mid m \notin b(m)\}.$$

Da D ein Element von $\mathcal{P}(M)$ ist, existiert ein $d \in M$ mit $D = b(d)$. Nun sind zwei Fälle denkbar, $d \in D$ oder $d \notin D$. Im ersten Fall folgt $d \in b(d)$, also $d \notin D$ nach Definition von D, daher ergibt sich ein Widerspruch. Im zweiten Fall gilt aber $d \notin b(d)$, woraus $d \in D$ folgt, wir erhalten ebenso einen Widerspruch. Damit ist die Annahme widerlegt, es könne eine Bijektion von M auf $\mathcal{P}(M)$ geben. ■

Man kann den vorhergehenden Beweis wieder als Diagonalargument auffassen. Um dies bildlich zu verdeutlichen, stellen wir uns die Elemente $m, n, \ldots$ von M in linearer Ordnung auf einer horizontalen Achse vor. Wenn es nun eine Bijektion b wie im obigen Beweis gäbe, so könnten wir alle Elemente der Potenzmenge 2^M in der entsprechenden Reihenfolge $b(m), b(n), \ldots$ auf einer vertikalen Achse untereinanderschreiben. Wir stellen uns dabei vor, daß diese Elemente in der Form charakteristischer Funktionen $\chi_m, \chi_n, \ldots$ beschrieben sind (vgl. Kap. 3.6.2). In unserer Abbildung ist also angedeutet, daß m ein Element von $b(n)$ ist, n hingegen nicht. Ähnlich ist m ein Element von $b(m)$, aber n nicht.



Die Menge d aus dem obigen Beweis, wenn wir sie als charakteristische Funktion darstellen, ist genau diejenige Funktion, die wir dadurch erhalten, daß wir auf der Diagonalen jede Null zu einer Eins abändern und umgekehrt. Deswegen kann die so erhaltene charakteristische Funktion bzw. Teilmenge aus M mit keiner der Mengen $b(m)$ für $m \in M$ übereinstimmen, und wir erhalten den üblichen Widerspruch.

5.4 Anwendungen

Am Schluß des Kapitels wollen wir zeigen, daß Cantors Satz 5.14 einige bemerkenswerte negative Konsequenzen für die Möglichkeit der Repräsentation von formalen Sprachen (vgl. Definition 3.11) mittels Grammatiken hat. Es sei A ein nichtleeres Alphabet. Es ist für die nachfolgenden Überlegungen natürlich, sich A als endliche Menge von Symbolen vorzustellen, auch wenn wir das nicht vorauszusetzen brauchen. In Lemma 5.11 hatten wir gezeigt, daß die Menge aller Wörter A^* abzählbar unendlich ist. Nun zeigt Satz 5.14 aber, daß es überabzählbar viele Sprachen über dem Alphabet A gibt.

Formale Sprachen können auf viele unterschiedliche Arten durch *Grammatiken* beschrieben werden. Wir werden nicht darauf eingehen, was genau unter einer Grammatik für eine Sprache zu verstehen ist. Zwei Prinzipien sind auch so einsichtig:

1. Jede Grammatik ist selbst eine *endliche* Symbolfolge über einem geeigneten endlichen oder abzählbar unendlichen Alphabet B,

2. Eine gegebene Grammatik repräsentiert maximal *eine* formale Sprache.

Das Alphabet B zum Schreiben von Grammatiken denken wir uns nachfolgend als fest vorgegeben. Aus Prinzip 1 folgt mit Lemma 5.11, daß es (höchstens) abzählbar unendlich viele verschiedene Grammatiken geben kann. Damit können aber aufgrund Prinzip 2 höchstens abzählbar unendlich viele formale Sprachen durch Grammatiken repräsentiert werden. Da es überabzählbar viele formale Sprachen gibt, sind also die meisten formalen Sprachen nicht durch Grammatiken beschreibbar.

Ein ähnliches negatives Resultat ergibt sich aus Lemma 5.13 bezüglich der Möglichkeit, alle Funktionen aus $\mathbb{N}^{\mathbb{N}}$ durch den Computer berechnen zu lassen. Wir lassen wiederum offen, was es bedeutet, eine Funktion $f\colon \mathbb{N} \to \mathbb{N}$ durch ein Computerprogramm berechnen zu können, und legen uns nur auf zwei Prinzipien fest.

1. Eine Computerprogramm ist selbst eine *endliche* Symbolfolge über einem geeigneten endlichen oder abzählbar unendlichen Alphabet B,

2. Eine gegebenes Computerprogramm berechnet maximal *eine* Funktion $f\colon \mathbb{N} \to \mathbb{N}$.

Wie im vorigen Fall ergibt sich nun aus Lemma 5.11, daß es (höchstens) abzählbar unendlich viele verschiedene Computerprogramme geben kann. Angesichts von Lemma 5.13 bedeutet dies, daß die meisten Funktionen $f\colon \mathbb{N} \to \mathbb{N}$ nicht durch ein Computerprogramm berechenbar sind.

5.5 Aufgaben zu Kapitel 5

Aufgabe 5.1 Es seien A, B und C Mengen. Die Schreibweise $|A| \leq |B|$ sei wie in Abschnitt 5.1 erklärt. Zeigen Sie: aus $|A| \leq |B|$ und $|B| \leq |C|$ folgt $|A| \leq |C|$.

Aufgabe 5.2 Zeigen Sie, daß die Menge $\mathcal{Q}$ der rationalen Zahlen abzählbar unendlich ist.

Aufgabe 5.3 Für $i \in \mathbb{N}$ sei A_i eine endliche nichtleere Menge. Für $i \neq j$ seien A_i und A_j stets disjunkt ($i, j \in \mathbb{N}$). Zeigen Sie, daß $\bigcup_{i \in \mathbb{N}} A_i$ abzählbar unendlich ist.

Aufgabe 5.4 Beweisen Sie Lemma 5.13: die Menge $\mathbb{N}^{\mathbb{N}}$ ist überabzählbar.

5.6 Bibliographische Angaben

Der Satz von Cantor-Bernstein wurde von Felix Bernstein bewiesen. Die obige Darstellung des Beweises folgt [Fel79] Band III (Kap. 8) und geht auf J. König [Kö06] zurück. In [Fel79] findet sich auch eine Verallgemeinerung dieses Satzes, die von Dedekind und Peano bewiesen wurde, sowie weiterer Hintergrund zu dem Themenbereich dieses Kapitels. Zwei interessante populärwissenschaftliche Bücher, in denen die Bedeutung von Diagonalmethoden in Hinblick auf die Möglichkeiten und Grenzen der sogenannten „künstlichen Intelligenz“ konträr diskutiert werden, sind [Hof79] und [Pen90].

6

Strukturen und Algebren

In unseren bisherigen Betrachtungen hatten wir—zumindest außerhalb der illustrierenden Beispiele—nie eine gegebene Menge mit ganz bestimmten Relationen oder Funktionen in Verbindung gebracht. Auch unter den Elementen einer Menge hatten wir keinerlei Unterschiede gemacht, alle hatten denselben Stellenwert. Diese einfache Perspektive ist sinnvoll, wenn es darum geht, die abstrakte Natur der Grundkonzepte zu erfassen. Sie ist aber zu farblos, wenn man wirklich interessante mathematische Gegenstände betrachtet, deren Analyse einem tatsächlichen Zweck dient.

Betrachten wir als Beispiel die natürlichen Zahlen. Wenn wir uns für zweistellige Funktionen auf dieser Menge interessieren, so gibt es unendlich viele[1]. Fast alle sind jedoch herzlich uninteressant. Zweifelsfrei wichtig hingegen ist die Addition, oder auch die Multiplikation. Wir verbinden die Grundmenge $\mathbb{N}$ fast „automatisch" mit diesen Funktionen. Nach einer wichtigen zweistelligen Relation gefragt, denken wir sofort an die kleiner-gleich-Relation. Es ist also diese Grundmenge in natürlicher Weise *assoziiert* mit *ganz bestimmten* Funktionen und Relationen. Eine entsprechende Beobachtung läßt sich auch über die Situationen und Sprechkontexte machen, die wir mittels natürlicher Sprache beschreiben. Nur für bestimmte, interessierende Relationen haben wir sprachliche Ausdrucksmittel (zum Beispiel Verben wie „kennt", „liebt", „sieht", oder Nomen wie „Frau", „Mann" oder „Kind") in der Umgangssprache, die Mehrzahl der mathematisch denkbaren Relationen hingegen ist irrelevant und daher auch sprachlich nur über Umwege zu

[1] In Kapitel 5 hatten wir gezeigt, daß es sogar überabzählbar viele sind.

beschreiben.

Es liegt also nahe, eine gegebene Grundmenge und die „*dazugehörigen*“ relevanten Relationen und Funktionen zu einem gemeinsamen Objekt zusammenzufassen. Eine solche Zusammenfassung leistet der Begriff der Struktur, den wir in diesem Kapitel beschreiben.

Nach einer formalen Definition des Begriffs der Struktur geben wir im ersten Teilkapitel einen kurzen Überblick über verschiedene Klassen von Strukturen, die in der mathematischen Literatur, insbesondere im Teilgebiet der Algebra, untersucht werden. Viele dieser Klassen sind auch für andere Fächer und Disziplinen von großer Bedeutung, einige Anmerkungen hierzu sind beigefügt. In Abschnitt 6.3 gehen wir auf Teilstrukturen und Erzeugung von Teilstrukturen ein. Im Anschluß werden in Kapitel 6.4 spezielle Funktionen zwischen Strukturen betrachtet, die mit den ausgezeichneten Funktionen, Relationen und Elementen in besonderer Weise verträglich sind. Funktionen dieser Art sind unter der Bezeichnung „Homomorphismen“ bekannt. Schließlich verfeinern wir in Abschnitt 6.5 das Konzept der Äquivalenzrelation und gehen auf sogenannte Kongruenzrelationen ein. Wir stellen dar, wie man mittels einer Kongruenzrelationen auf einer Struktur eine vereinfachte Variante dieser Struktur definieren kann.

6.1 Beispiele und Klassen von Strukturen

Um den Begriff der Struktur zu formalisieren, gibt es eine Reihe notationeller Alternativen. Wir beginnen mit einer Beschreibung, die sich aufgrund ihrer Einfachheit gut zur Darstellung einzelner Strukturen eignet. Später geben wir noch eine modifizierte Definition, die vom Begriff der Signatur Gebrauch macht. Diese zweite Darstellungsart ist etwa beim Vergleich „ähnlicher“ Strukturen vorteilhaft.

Definition 6.1 Eine *Struktur* ist ein Tupel $\mathcal{A}$, dessen erste Komponente eine nichtleere[2] Menge A ist, die *Grundmenge* oder *Grundbereich* von $\mathcal{A}$ genannt wird. Die weiteren Komponenten repräsentieren eine Folge von Funktionen, Relationen und Elementen von A.

[2]Den nichtleeren Grundbereich bei Strukturen fordert man, um pathologische Situationen auszuschließen. Sätze wie $(\forall x\colon R(x)) \Rightarrow (\exists x\colon R(x))$ sind damit in jeder Struktur stets erfüllt.

Die in der Beschreibung enthaltenen Funktionen und Relationen haben beliebige positive Stelligkeit. Sie werden manchmal als die *ausgezeichneten* Funktionen beziehungsweise Relationen bezeichnet. Entsprechend spricht man von den *ausgezeichneten Elementen* der Struktur. Die ausgezeichneten Funktionen werden oft auch als die *Operationen* der Struktur bezeichnet.

Die obige Definition beinhaltet eine notationelle Konvention, die wir stets verwenden, wenn wir über nicht näher festgelegte oder bekannte Strukturen reden: ist $\mathcal{A}$ ($\mathcal{B}$,...) eine Struktur, so verwenden wir A (resp. B,...) als Symbol für die zugehörige Grundmenge.

Hinter den bekannten Zahlbereichen wie den natürlichen Zahlen und den reellen Zahlen stehen Strukturen. Während der Grundbereich hier klar ist, kann man jeweils unterschiedliche Funktionen (etwa Addition, Multiplikation, Exponentiation,...), Relationen (kleiner-gleich und andere) und Elemente (0, 1, e, π, ...) in die formale Beschreibung mitaufnehmen. Die konkrete Auswahl hängt vom intendierten Anwendungsgebiet ab.

Beispiele 6.2 Die Menge der natürlichen Zahlen mit der Addition und der üblichen Ordnung $\langle \mathbb{N}, +, \leq \rangle$ ist eine Struktur. Eine vergleichbare Struktur ist $\langle \mathbb{R}, +, \leq \rangle$.[3]

Um das Konzept der Struktur besser zu verdeutlichen, stellen wir zunächst eine längere Liste bestimmter Klassen von Strukturen auf. Die nachfolgend beschriebenen Klassen sogenannter „Algebren" sind in der Mathematik wohluntersucht, wir werden auch nicht andeutungsweise auf die dort erhaltenen Resultate eingehen können und verweisen auf die sehr umfangreiche Spezialliteratur, zu der sich einige Anmerkungen am Kapitelende befinden.

6.1.1 Algebren

Algebren[4] zeichnen sich dadurch aus, daß sie keine ausgezeichneten Relationen enthalten.

[3] Üblicherweise wird für die Addition auf den bekannten Zahlbereichen stets das Symbol „+" verwendet, obwohl es sich im strikten Sinn um unterschiedliche Funktionen (i.e., Mengen von geordneten Paaren) handelt. Ähnliches gilt für die Verwendung des Symbols „$\leq$".

[4] Der Begriff „Algebra" wird in zwei unterschiedlichen Bedeutungen verwendet. Einerseits bezeichnet er ein Teilgebiet der Mathematik, andererseits die hier dargestellte Klasse von Strukturen.

Definition 6.3 Eine Struktur heißt *Algebra* genau dann, wenn sie nur ausgezeichnete Funktionen und ausgezeichnete Elemente enthält.

Wir wollen nachfolgend einige Typen von Algebren angeben, auf denen Funktionen ausgezeichnet sind, die sich ähnlich wie die Rechenoperationen — Multiplikation oder Addition — auf den oben erwähnten Zahlbereichen verhalten. Teilweise werden diese Funktionen daher als „Multiplikation" oder „Addition" bezeichnet, obwohl es nicht notwendigerweise eine direkte Beziehung zur üblichen Multiplikation oder Addition gibt. Zunächst betrachten wir Algebren, wo wir lediglich eine verallgemeinerte Art der Multiplikation haben. Diese werden als Halbgruppen bezeichnet.

Definition 6.4 Eine *Halbgruppe* ist eine Algebra $\mathcal{H} = \langle H, \cdot \rangle$, wo „$\cdot$" eine zweistellige und assoziative Funktion auf H ist. Es muß also gelten:

$$\forall a, b, c \in H\colon a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

Eine Halbgruppe $\mathcal{H} = \langle H, \cdot \rangle$ heißt *kommutativ* genau dann, wenn „$\cdot$" eine kommutative Funktion ist, wenn also gilt: $\forall h_1, h_2 \in H\colon h_1 \cdot h_2 = h_2 \cdot h_1$.

Offenkundig sind die oben genannten Zahlbereiche kommutative Halbgruppen, wenn wir uns auf die Addition oder die Multiplikation beschränken. Bezeichnet 2^A die Potenzmenge der Menge A, so sind $\langle 2^A, \cup \rangle$ und $\langle 2^A, \cap \rangle$ kommutative Halbgruppen.

Eine Halbgruppe, die in der Informatik besondere Bedeutung hat, stellt das nachfolgende Beispiel dar.

Beispiel 6.5 Es sei A ein nichtleeres Alphabet. Wie gewohnt schreiben wir A^* für die Menge aller Wörter über A (vgl. Def. 3.9). Bezeichnet „$\circ$" die in Beispiele 3.34 Nr. 5 definierte Konkatenation von Wörtern, so ist $\langle A^*, \circ \rangle$ eine Halbgruppe. In der Tat ist „$\circ$" offenkundig assoziativ. $\langle A^*, \circ \rangle$ wird auch die *von A frei erzeugte* Halbgruppe genannt.

Eine andere (in vieler Hinsicht verwandte) Halbgruppe erhalten wir, wenn wir die Menge aller Relationen auf einer Menge mit der Komposition betrachten.

Beispiel 6.6 Es sei A eine nichtleere Menge und $\mathcal{R}(A)$ die Menge aller zweistelligen Relationen auf A. Bezeichnet „$\circ$“ die in Definition 3.23 erklärte Komposition von Relationen, so ist $\langle \mathcal{R}(A), \circ \rangle$ eine Halbgruppe, wie sich sofort aus Lemma 3.27 ergibt.

Eine weitere verwandte Halbgruppe bildet die Menge aller Sprachen über einem Alphabet zusammen mit der Komposition von Sprachen.

Beispiel 6.7 Es bezeichne $\mathcal{L}(A)$ die in Definition 3.11 eingeführte Menge aller Sprachen über dem Alphabet A und „$\circ$“ die in Beispiele 3.34 Nr. 7 erklärte Komposition von Sprachen. Dann ist $\langle \mathcal{L}(A), \circ \rangle$ eine Halbgruppe. Aus der Definition der Komposition folgt leicht, daß es sich um eine assoziative Funktion handelt (vgl. Aufgabe 6.1).

Die genannten Beispiele von Halbgruppen $\mathcal{H} = \langle H, \cdot \rangle$ machen deutlich, daß man in sehr vielen Fällen ein sogenanntes *Neutralelement* bezüglich „$\cdot$“ hat, das heißt ein Element $1_{\mathcal{H}} \in H$, für das gilt:

$$\forall h \in H\colon 1_{\mathcal{H}} \cdot h = h \cdot 1_{\mathcal{H}} = h.$$

Oft wird man einem solchen Element einen besonderen Stellenwert zukommen lassen. Dies motiviert den nachfolgenden Begriff.

Definition 6.8 Ein *Monoid* ist eine Algebra $\mathcal{M} = \langle M, \cdot, 1_{\mathcal{M}} \rangle$, wo $\langle M, \cdot \rangle$ eine Halbgruppe ist, und wo $1_{\mathcal{M}}$ ein Neutralelement bezüglich „$\cdot$“ ist.

Alle oben genannten Halbgruppen besitzen ein Neutralelement und können auch als Monoid beschrieben werden. Es bleibt dem Leser überlassen, jeweils das Neutralelement anzugeben (vgl. Aufgabe 6.2).

Wenn wir von den Monoiden dadurch einen Schritt weitergehen, daß wir eine verallgemeinerte Division einführen, kommen wir zu den sogenannten Gruppen. Diese Algebren wurden erstmals von E. Galois[5] eingeführt im Zusammenhang mit dem Problem der Lösung von Gleichungen durch sogenannte Radikale. Heute spielen Gruppen in vielen Bereichen eine große Rolle, unter anderem in der Kristallographie, in der Quantenmechanik, und in der Relativitätstheorie.

[5] Évariste Galois, französischer Mathematiker (1811-1832).

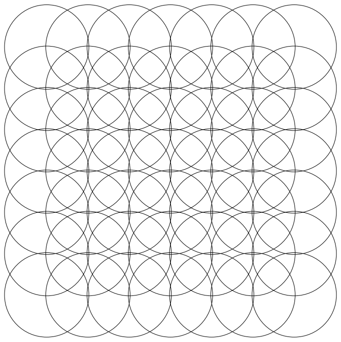
Definition 6.9 Eine *Gruppe* ist eine Algebra der Form $\mathcal{G} = \langle G, \cdot, 1_{\mathcal{G}}, ^{-1} \rangle$ wo gilt:

1. $\langle G, \cdot, 1_{\mathcal{G}} \rangle$ ist ein Monoid,

2. „$^{-1}$" ist eine einstellige Funktion $g \mapsto g^{-1}$ auf G, genannt *Inversenbildung*, die die folgende Bedingung erfüllt: $\forall g \in G$: $g \cdot g^{-1} = g^{-1} \cdot g = 1_{\mathcal{G}}$.

Wie im Falle von Halbgruppen spricht man von einer kommutativen (oder abelschen[6]) Gruppe, wenn die sogenannte Gruppenmultiplikation „$\cdot$" kommutativ ist.

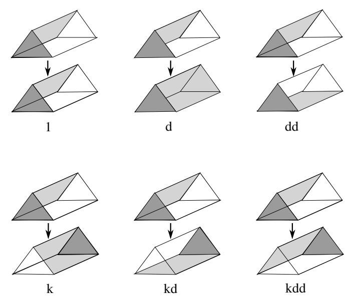
Beispiel 6.10 Zu periodischen Ornamenten gibt es eine Gruppe, die die möglichen Verschiebungen beschreibt, welche zur Deckungsgleiche führen. Betrachten wir als ein Beispiel das folgende Relief, das wir uns in alle vier Richtungen unendlich fortgesetzt denken.



Offenkundig gibt es zu jeder Verschiebung eine Umkehrungverschiebung, wodurch sich die für Gruppen typische Inversenbildung ergibt.

Beispiel 6.11 Verwandte Beispiele von Gruppen erhalten wir etwa, wenn wir bestimmte Drehungen eines regelmäßigen Körpers betrachten. Die nachfolgende Abbildung repräsentiert die sechs Elemente der sogenannten *Diedergruppe* des Grads 3, die mit D_3 bezeichnet wird.

[6]benannt nach dem norwegischen Mathematiker Niels Henrik Abel (1802-1829).



Wenn wir die mehrfache Multiplikation eines Gruppenelements mit sich selbst in der üblichen Exponentenschreibweise abkürzen, so gelten die Beziehungen $d^3 = 1$ und $k^2 = 1$. Jedes Element der Gruppe kann durch Multiplikation von Elementen k und d in geeigneter Vielfachheit dargestellt werden, wie in der Figur angedeutet ist. Allgemeiner bilden für jede Zahl $n \geq 1$ die Drehungen und Klappungen eines regelmäßigen n-Ecks in ähnlicher Weise eine Gruppe, die mit D_n bezeichnet wird.

Beispiel 6.12 Eine andere wichtige Klasse von Gruppen stellen die sogenannten *Permutationsgruppen* dar. Wie bereits erwähnt ist eine Permutation eine bijektive Abbildung einer Menge M in sich selbst. Wenn wir für M die Menge $\{1, 2, 3\}$ wählen, so erhalten wir die folgenden Permutationen $M \to M$:

$$\begin{bmatrix} 1 & \mapsto & 1 \\ 2 & \mapsto & 2 \\ 3 & \mapsto & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & \mapsto & 3 \\ 2 & \mapsto & 1 \\ 3 & \mapsto & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & \mapsto & 2 \\ 2 & \mapsto & 3 \\ 3 & \mapsto & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & \mapsto & 1 \\ 2 & \mapsto & 3 \\ 3 & \mapsto & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & \mapsto & 2 \\ 2 & \mapsto & 1 \\ 3 & \mapsto & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & \mapsto & 3 \\ 2 & \mapsto & 2 \\ 3 & \mapsto & 1 \end{bmatrix}$$

Die Menge dieser Abbildungen, mit der Komposition als Verknüpfung, bildet eine Gruppe, unabhängig von der Zahl der Elemente. Die Inversenbildung wird durch die jeweilige Umkehrabbildung einer Permutation geliefert. Dem Leser bleibt überlassen, im Spezialfall von drei Elementen die Diedergruppe D_3 und die obige Permutationsgruppe auf etwaige Ähnlichkeiten zu untersuchen (vgl. Aufgabe 6.23).

Ist $R \subseteq A \times A$ eine zweistellige Relation, so könnte man vielleicht vermuten, daß es sich beim Übergang zur inversen Relation R^{-1} um eine Inversenbildung im gruppentheoretischen Sinn handelt. Dies ist jedoch im allgemeinen nicht der Fall, da in der Regel die Gleichung $R \circ R^{-1} = Id_A$ *nicht* gilt.

Die nächste zu besprechende Klasse stellen die sogenannten Ringe dar, in denen man–grob gesprochen–in der gewohnten Weise addieren, subtrahieren und multiplizieren darf, aber nicht notwendigerweise dividieren.

Definition 6.13 Ein *Ring* ist eine Algebra der Form $\mathcal{R} = \langle R, +, 0_R, -, \cdot \rangle$ wo gilt:

1. $\langle R, +, 0_R, - \rangle$ ist eine kommutative Gruppe,
2. $\langle R, \cdot \rangle$ ist eine Halbgruppe,
3. es gelten die beidseitigen *Distributivgesetze*, das heißt es gilt

$$\forall x, y, z \in R\colon x \cdot (y + z) = x \cdot y + x \cdot z \ \wedge \ (x + y) \cdot z = x \cdot z + y \cdot z.$$

Ein Ring heißt kommutativ genau dann, wenn die Ringmultiplikation „$\cdot$“ kommutativ ist. Beispielsweise bilden die ganzen Zahlen mit der Addition und der Multiplikation einen kommutativen Ring. Ist R ein Ring, und bezeichnet $F := R^M$ die Menge aller Abbildungen einer nichtleeren Menge M in R, so können wir auf F durch die Definitionen $(f +_F g)(x) := f(x) + g(x)$ sowie $(f \cdot_F g)(x) = f(x) \cdot g(x)$ $(x \in M)$ eine Ringstruktur erhalten. Die Einzelheiten bleiben dem Leser überlassen.

Beispiel 6.14 Aus dem Ring der ganzen Zahlen können wir durch eine einfache Konstruktion zu jeder natürlichen Zahl $n > 0$ einen kommutativen Ring mit genau n Elementen erhalten. Dazu stellen wir uns die ganzen Zahlen auf einer Uhr mit den „Stunden“ $0, \ldots, n-1$ aufgewickelt vor. Dies ist in Abbildung 6.1 für den Fall $n = 6$, wo wir also die Stunden $0, \ldots, 5$ haben, dargestellt. Es wird deutlich, daß durch das Aufwickeln genau diejenigen ganzen Zahlen auf derselben Stunde zu liegen kommen, die beim Teilen durch n denselben Rest lassen. Wie in Beispiele 4.7 Nr. 2 und in Beispiel 4.12 angedeutet, wird hierdurch eine Äquivalenzrelation beschrieben. Die Äquivalenzklassen $[0], \ldots, [5]$ sind jeweils grau unterlegt und entsprechen den 5 Stunden. Auf der Menge $\{[0], \ldots, [5]\}$ erklären wir nun „Uhrenaddition $\oplus$“, „Uhrenmultiplikation $\odot$“ und Inversenbildung „$\ominus$“ in der Form

$$\begin{aligned} [i] \oplus [j] &:= [i + j], \\ [i] \odot [j] &:= [i \cdot j], \\ \ominus[i] &:= [-i]. \end{aligned}$$

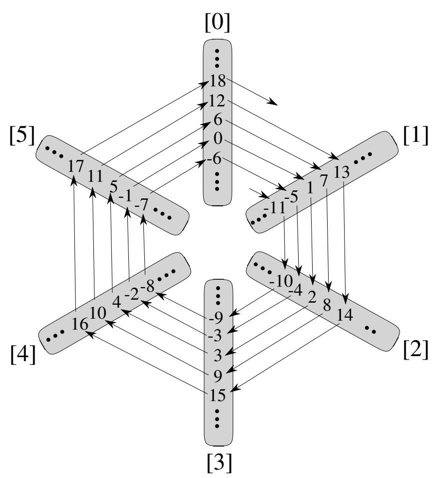


Abbildung 6.1: Der Ring Z_6.

Hierbei stehen „+“, „·“ und „−“ für die üblichen Operationen auf den ganzen Zahlen. Für die Uhrenaddition addiert man also die beiden Zahlen i und j zunächst wie gewohnt, und nimmt dann die zugeordnete Stunde oder Äquivalenzklasse $[i + j]$. Im dargestellten Fall $n = 6$ erhält man also etwa $[4] \oplus [4] = [8] = [2]$ und $[4] \odot [3] = [12] = [0]$.

Bei den Definitionen ist nachzuweisen, daß sie wohldefiniert, das heißt unabhängig von der Wahl der Repräsentanten i und j sind. Im Fall der Uhrenaddition „$\oplus$“ folgt dies daraus, daß Drehungen um Vielfache von 360 Grad irrelevant für die Zugehörigkeit zu einer Äquivalenzklasse sind. Ähnliches gilt für die Inversenbildung „$\ominus$“. Betrachten wir nun die Uhrenmultiplikation. Falls $[i] = [i']$ und $[j] = [j']$, so gibt es ganze Zahlen k und l mit $i = i' + k \cdot 5$ und $j = j' + l \cdot 5$. Es folgt

$$\begin{aligned} i \cdot j &= (i' + k \cdot 5) \cdot (j' + l \cdot 5) \\ &= i' \cdot j' + 5 \cdot (i' \cdot l + k \cdot j' + k \cdot l \cdot 5). \end{aligned}$$

Folglich ist $i \cdot j - i' \cdot j'$ ein ganzzahliges Vielfaches von 5 und $[i \cdot j] = [i‘ \cdot j']$.

Aufgrund der obigen Definitionen übertragen sich nun Rechengesetze wie das Kommutativgesetz und das Distributivgesetz sofort. Im Fall des Distributivgesetzes etwa rechnet man

$$[i] \odot ([j] \oplus [k]) \quad = \quad [i] \odot ([j + k]) = [i \cdot (j + k)]$$

$$= \quad [i \cdot j + i \cdot k] = [i \cdot j] \oplus [i \cdot k] = [i] \odot [j] \oplus [i] \odot [k].$$

Man kann somit leicht nachprüfen, daß die resultierende Algebra ein Ring mit additiver Null [0] ist. Den für eine gegebene Zahl $n > 0$ auf diese Weise erhaltenen Ring bezeichnet man als Z_n.

Die letzte Klasse von Algebren, die wir kurz darstellen, sind die Körper. Körper sind ähnlich wie Ringe, erlauben aber die Division durch von Null verschiedene Elemente. In Körpern können wir damit in ähnlicher Weise wie in den rationalen, reellen oder komplexen Zahlen rechnen. Diese Bereiche bilden auch Standardbeispiele für Körper.

Definition 6.15 Ein *Körper* ist eine Algebra der Form $\mathcal{K} = \langle K, +, 0_K, -, \cdot, 1_K, ^{-1} \rangle$ wo gilt

1. $1_K \neq 0_K$,
2. $\langle K, +, 0_K, -, \cdot \rangle$ ist ein kommutativer Ring,
3. $\langle K \setminus \{0_K\}, \cdot, 1_K, ^{-1} \rangle$ ist eine kommutative Gruppe.

Neben den oben erwähnten unendlichen Körpern, zu denen noch viele weitere Beispiele existieren, gibt es auch endliche Körper. Der in Beispiel 6.14 eingeführte Ring Z_n ist genau dann ein Körper, wenn n eine Primzahl ist. Für $n = 5$ ist dieser in Abbildung 6.2 angedeutet. Die Elemente von Z_5 sind einfachheitshalber in der Form $0, \ldots, 4$ (anstelle $[0], \ldots, [4]$) notiert. Um zu sehen, daß Z_n für jede Primzahl n stets ein Körper ist, besteht der einzig wesentliche Schritt darin, zu zeigen, daß jedes der Elemente $1, \ldots, n-1$ bezüglich der neuen Multiplikation ein Inverses hat. Hierzu vergleiche man Aufgabe 6.7.

Bemerkung 6.16 Ein interessanter Spezialfall ergibt sich für den Fall $n = 2$, wo sich die zwei „Stunden“ 0 und 1 auch als Wahrheitswerte interpretieren lassen. Die nachfolgende Tabelle gibt die Werte der Addition „$\oplus$“ und der Multiplikation „$\odot$“ in Z_2 wieder.

α	β	$\alpha \oplus \beta$	$\alpha \odot \beta$
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

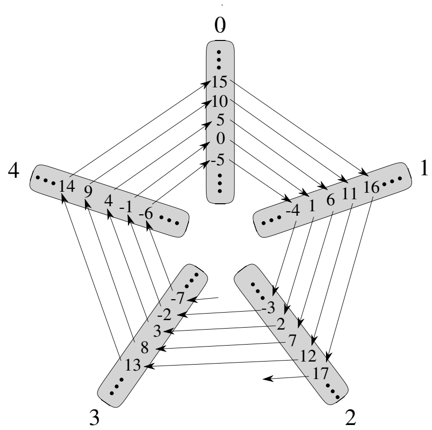


Abbildung 6.2: Endlicher Körper.

Man sieht sofort, daß die Multiplikation genau der logischen Konjunktion „$\wedge$“ entspricht, die Addition dem „ausschließenden oder“, das wir in Kapitel 1 (vgl. Aufg. 1.6) mit dem Symbol „$+$“ bezeichnet hatten. Dies erklärt, warum viele Gesetzmäßigkeiten beim Rechnen mit Wahrheitswerten direkt an Gesetzmäßigkeiten beim Rechnen mit Zahlen erinnern.

6.1.2 Relationalstrukturen

In verschiedenen Kontexten betrachtet man Strukturen, bei denen überhaupt keine ausgezeichneten Funktionen auftreten.

Definition 6.17 Eine Struktur heißt *Relationalstruktur*, falls sie nur ausgezeichnete Elemente und ausgezeichnete Relationen enthält.

Die in Kapitel 7 zu besprechenden Graphen bilden eine wichtige Klasse von Relationalstrukturen. Eine andere Klasse von großer Bedeutung stellen die Ordnungsstrukturen dar.

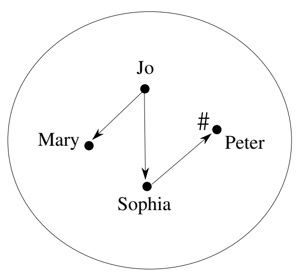
Definition 6.18 Eine *Ordnungsstruktur* ist eine Relationalstruktur der Form $\mathcal{O} = \langle O, \leq \rangle$ wo „$\leq$“ eine Quasi-Ordnung auf O ist. Ist „$\leq$“ eine partielle (lineare) Ordnung, so wird $\mathcal{O} = \langle O, \leq \rangle$ auch eine partiell (linear) geordnete Menge (englisch oft „poset“) genannt.

Beispiele von Ordnungsstrukturen werden durch die Zahlbereiche der natürlichen, ganzen, rationalen und reellen Zahlen geliefert, wenn wir die übliche (lineare) Ordnung auf diesen Grundmengen auszeichnen. Allgemeinere Ordnungsstrukturen, die sogenannten Verbände, werden wir in Kapitel 8 ausführlich diskutieren.

Relationalstrukturen treten auch in natürlicher Weise bei der Formalisierung des Inhalts einfacher natürlich-sprachlicher Sätze auf. Ein „Modell" des Satzes „*Maria liebt einen Geiger*" könnte man etwa dadurch erhalten, daß man auf einer Grundmenge wie $A = \{Jo, Mary, Sophia, Peter\}$ passende (zwei- beziehungsweise einstellige) Relationen $liebt_{\mathcal{A}}$, $Geiger_{\mathcal{A}}$ auszeichnet, wie zum Beispiel

$$liebt_{\mathcal{A}} := \{\langle Jo, Mary\rangle, \langle Jo, Sophia\rangle, \langle Sophia, Peter\rangle\}$$

sowie $Geiger_{\mathcal{A}} := \{Peter\}$. Die Grundmenge mit den zugehörigen Relationen bildet dann eine Relationalstruktur $\mathcal{A}$. Wenn wir die Relation $liebt_{\mathcal{A}}$ durch Pfeile markieren und den Geiger mit einem Symbol # kenntlich machen, so erhalten wir folgendes Bild.



Auch andere, bereits vertraute graphische Darstellungen können wir als Relationalstrukturen auffassen. Schachspieler verwenden bei der Analyse von Schachstellungen eine Reihe von ausgezeichneten Individuen (schwarze Dame, weißer König, etc.) und Relationen (bedroht, deckt, etc.). Wenn wir uns auf eine feste Auswahl solcher Relationen und Individuen einigen, so repräsentieren die in Abbildung 6.3 dargestellten Schachstellungen zwei Relationalstrukturen $\mathcal{A}$ und $\mathcal{B}$ auf der Grundmenge der jeweiligen Figuren auf dem Brett. In der links abgebildeten Struktur deckt die weiße Dame den weißen König, während dies dort nicht auf schwarze Dame und schwarzen König zutrifft.

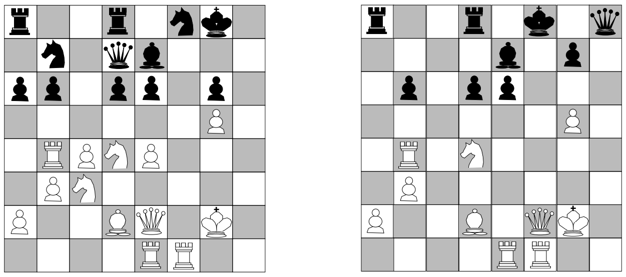


Abbildung 6.3: Schachkonfigurationen als Relationalstrukturen.

6.2 Signaturen

Die im Abschnitt 6.1 gegebene Definition einer Struktur erfüllt ihren Dienst, solange wir nur eine Struktur untersuchen. Eines unserer Ziele wird aber der systematische Vergleich von Strukturen sein. Ein solcher Vergleich zwischen zwei Strukturen macht erst Sinn, wenn klar ist, welche der Relationen aus beiden Strukturen in Beziehung gesetzt werden sollen. Insbesondere muß zu *jeder* Relation der einen Struktur eine „korrespondierende“ Relation der anderen Struktur festgelegt sein und umgekehrt. Entsprechendes muß natürlich auch für Funktionen und ausgezeichnete Elemente gelten. Um einen systematischen Vergleich zu erleichtern, führen wir nun den Begriff der Signatur (auch „Ähnlichkeitstyp“ genannt) ein, und geben dann darauf aufbauend eine modifizierte Definition des Begriffs einer Struktur.

Definition 6.19 Eine *Signatur* ist eine Menge Σ von Relationssymbolen, Funktionssymbolen und Individuennamen (auch Individuenkonstante genannt). Hierbei ist jedes Relations- und Funktionssymbol mit einer festen *Stelligkeit* in $\mathbb{N}$ versehen. Funktionssymbole haben positive Stelligkeit.

Ist Σ eine Signatur, so bezeichen wir mit $\Sigma_{\mathcal{R}}$ die Menge der Relationssymbole aus Σ, mit $\Sigma_{\mathcal{F}}$ die Menge der Funktionssymbole aus Σ, und mit $\Sigma_{\mathcal{E}}$ die Menge der Individuenkonstanten aus Σ. Die Mengen $\Sigma_{\mathcal{R}}$, $\Sigma_{\mathcal{F}}$ und $\Sigma_{\mathcal{E}}$ sollen natürlich stets paarweise disjunkt sein.

Beispiel 6.20 Als Beispiel einer Signatur betrachten wir die Menge $\Sigma = \{kgl, Add, Mult, Null, Eins\}$. Hierbei sollen das Relationssymbol „kgl“ und die Funktionssymbole „*Add*“ und „*Mult*“ zweistellig sein, *Null* und *Eins* sind Individuenkonstanten.

Es ist nachfolgend wichtig zu verstehen, daß die Elemente einer Signatur wirklich nur Symbole sind!! Diese Symbole stehen für Relationen, Funktionen und ausgezeichnete Elemente von „passenden“ Strukturen, wie die nachfolgende modifizierte Definition einer Struktur deutlich macht.

Definition 6.21 Es sei Σ eine Signatur. Eine Σ-*Struktur* ist ein Paar $\mathcal{A} = \langle A, I \rangle$, wobei A eine nichtleere Menge ist, die der Grundbereich von $\mathcal{A}$ genannt wird. Weiterhin ist I eine „Interpretationsfunktion“, das heißt eine Funktion mit Definitionsbereich Σ, die folgende Bedingungen erfüllt:

1. für jedes Relationssymbol $R \in \Sigma_{\mathcal{R}}$ der Stelligkeit m ist $I(R)$ eine m-stellige Relation auf A,

2. für jedes Funktionssymbol $f \in \Sigma_{\mathcal{F}}$ der Stelligkeit n ist $I(f)$ eine n-stellige Funktion auf A,

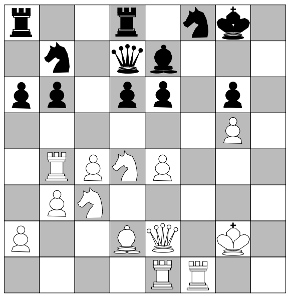
3. für jede Individuenkonstante $e \in \Sigma_{\mathcal{E}}$ ist $I(e)$ ein Element von A.

Beispiel 6.22 Es sei Σ die Signatur $\{kgl, Add, Mult, Null, Eins\}$ aus Beispiel 6.20. Setzen wir

$$\begin{array}{rcl} I(kgl) & = & \text{die übliche Ordnung „}\leq\text{“ auf } \mathbb{R}, \\ I(Add) & = & \text{die übliche Addition „}+\text{“ auf } \mathbb{R}, \\ I(Mult) & = & \text{die übliche Multiplikation „}\cdot\text{“ auf } \mathbb{R}, \\ I(Eins) & = & \text{die reelle Zahl } 1, \end{array}$$

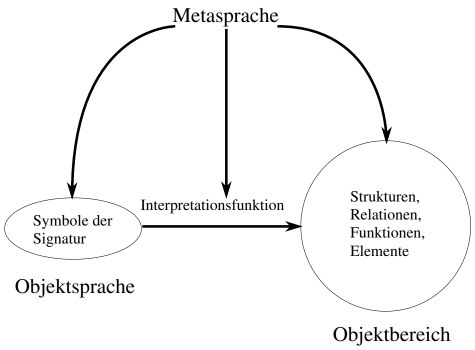
so ist $\langle \mathbb{R}, I \rangle$ eine Σ-Struktur.

Beispiel 6.23 Es sei Σ die Signatur mit den Individuenkonstanten *schwarze-Dame*, *weiße-Dame*, *schwarzer-König*, *weißer-König* und den zweistelligen Relationssymbolen *deckt* und *bedroht*. Die Schachkonfiguration



stellt eine Σ-Relationalstruktur dar. Die Interpretationsfunktion I ordnet hierbei der Individuenkonstante „*schwarze-Dame*“ die schwarze Dame zu und ist in analoger Weise für die anderen Individuenkonstanten definiert. Die Interpretation des Relationssymbols „*deckt*“ (resp. „*bedroht*“) ist diejenige Relation, die alle Paare gleichfarbiger Figuren umfaßt, wo die erste die zweite deckt (resp. bedroht), im üblichen Sinn des Schachspiels.

Wo wir nun zwischen den Symbolen der Signatur und ihrer Interpretation sauber trennen, sollte man sich klar machen, daß wir in unserer Darstellung drei Ebenen zusammenbringen. Die Elemente der Signatur lassen sich als die primitiven Symbole einer „Objektsprache“ auffassen, denen erst durch eine Interpretationsfunktion I in einer Struktur eine feste Bedeutung zugewiesen wird. Die mathematische Argumentation und das Räsonieren über Signaturen und Strukturen führen wir auf der Ebene der „normalen“ Sprache, die nun zur „Metasprache“ geworden ist.



Eine gewisse Schwierigkeit ergibt sich dadurch, daß wir auch auf der Ebene der Metasprache gezwungen sind, sprachliche Symbole zu verwenden, um auf die tatsächlichen Objekte (Strukturen etc.) zu referieren. Dies kann zu Verwechslungen der beiden Sprachebenen führen.

Bemerkung 6.24 In vielen Zusammenhängen wird man darauf verzichten wollen, die Interpretationsfunktion I einer Struktur explizit anzugeben. Eine einfache notationelle Konvention, von der wir nachfolgend Gebrauch machen, besteht darin, daß man für die Interpretation eines Funktionssymbols $f \in \Sigma_{\mathcal{F}}$ in der Σ-Struktur $\mathcal{A} = \langle A, I \rangle$ einfach $f_{\mathcal{A}}$ schreibt. Analog kann man die Relationen von $\mathcal{A}$ in der Form $R_{\mathcal{A}}$ ($R \in \Sigma_{\mathcal{R}}$) und die ausgezeichneten Elemente in der Form $e_{\mathcal{A}}$ ($e \in \Sigma_{\mathcal{E}}$) notieren.

6.3 Teilstrukturen und Teilalgebren

Wenn man sich ein besseres Bild von einer Struktur machen möchte, wird man sich oft zunächst überlegen, wie die Teilmengen der Struktur aussehen, die selbst wieder eine Struktur bilden.

Definition 6.25 Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Strukturen derselben Signatur Σ. $\mathcal{B}$ heißt *Teilstruktur* (oder *Substruktur*) von $\mathcal{A}$ genau dann, wenn gilt:

(i) $B \subseteq A$,

(ii) für jedes $R \in \Sigma_{\mathcal{R}}$ mit Stelligkeit m gilt:

$$\forall b_1, \ldots, b_m \in B\colon R_{\mathcal{A}}(b_1, \ldots, b_m) \Leftrightarrow R_{\mathcal{B}}(b_1, \ldots, b_m),$$

(iii) für jedes $f \in \Sigma_{\mathcal{F}}$ mit Stelligkeit n gilt:

$$\forall b_1, \ldots, b_n \in B\colon f_{\mathcal{A}}(b_1, \ldots, b_n) = f_{\mathcal{B}}(b_1, \ldots, b_n),$$

(iv) für jedes $e \in \Sigma_{\mathcal{E}}$ gilt $e_{\mathcal{A}} = e_{\mathcal{B}}$.

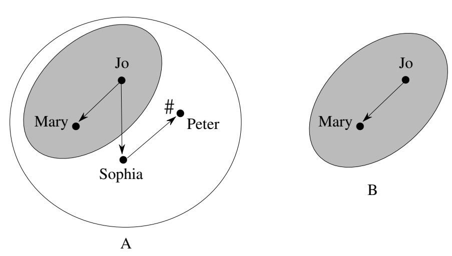
Sind die Eigenschaften (i)–(iii) erfüllt, so sagt man auch, daß alle Relationen und Funktionen der beiden Strukturen auf der Menge B übereinstimmen, und daß sich die Funktionen und Relationen von $\mathcal{B}$ durch *Einschränkung der Funktionen und Relationen von $\mathcal{A}$ auf die kleinere Grundmenge B* ergeben.

Es ist zu beachten, daß bei Teilstrukturen die Interpretation der Symbole aus Σ allein durch die Angabe des Grundbereichs bereits feststeht.

Teilstrukturen von Algebren werden auch *Teilalgebren* genannt. Für Algebren einer speziellen Art sind eigene Bezeichnungen für Teilalgebren gebräuchlich, man spricht etwa von *Unterhalbgruppen, Teilmonoiden, Untergruppen, Teilringen und Teilkörpern.*

Beispiele 6.26 Wir lassen bei den nachfolgenden Beispielen die Signatur weg, wenn klar ist, welche Funktionen, Relationen und ausgezeichneten Elemente zueinander korrespondieren.

1. Jede Struktur $\mathcal{A}$ ist stets eine Teilstruktur von sich selbst.
2. $\langle \mathbb{N}, \leq, +, 0\rangle$ ist eine Substruktur von $\langle \mathbb{R}, \leq, +, 0\rangle$.
3. Ist $A_1 \subseteq A_2$ ein Teilalphabet des Alphabets A_2, so ist $\langle A_1^*, \circ\rangle$ eine Unterhalbgruppe von $\langle A_2^*, \circ\rangle$. Hierbei bezeichnet „$\circ$" die Konkatenation von Wörtern.
4. Für $n > 1$ bezeichne D_n die in Beispiel 6.11 eingeführte Diedergruppe vom Grad n. Wenn n die Zahl m teilt, so ist D_n eine Untergruppe von D_m.
5. Ist M eine Teilmenge von N, so bildet die Gruppe all derjenigen Permutationen von M, die jedes Element aus M auf sich selbst abbilden, eine Untergruppe der Gruppe aller Permutationen von N.
6. Der Körper der rationalen Zahlen ist ein Teilkörper von $\langle \mathbb{R}, +, 0, -, 1, ^{-1}\rangle$.
7. Es sei Σ die Signatur $\{liebt, Geiger\}$ und $A = \{$Jo,Mary,Sophia,Peter$\}$. Die nachfolgende Figur repräsentiert eine Σ-Struktur $\mathcal{A}$, wobei die Paare der Relation $liebt_{\mathcal{A}}$ durch Pfeile dargestellt sind, und wo die Elemente der einstelligen Relation $Geiger_{\mathcal{A}}$ durch ein Symbol „#" markiert sind. In der rechten Hälfte ist eine von mehreren möglichen Substrukturen von $\mathcal{A}$ dargestellt (vgl. Aufgabe 6.9).



8. Ist $N \subseteq M$, so ist $\langle \mathcal{P}(N), \cup, \cap \rangle$ eine Teilalgebra von $\langle \mathcal{P}(M), \cup, \cap \rangle$.

Ist $\mathcal{A}$ eine Σ-Struktur, so können wir fragen, unter welchen Voraussetzungen eine gegebene nichtleere Teilmenge $B \subseteq A$ die Grundmenge einer Teilstruktur $\mathcal{B}$ von $\mathcal{A}$ ist. Die Verifikation des nachfolgenden Lemmas können wir dem Leser überlassen.

Lemma 6.27 *Eine nichtleere Teilmenge B einer Σ-Struktur $\mathcal{A}$ ist die Grundmenge einer Teilstruktur $\mathcal{B}$ von $\mathcal{A}$ genau dann, wenn gilt:*

1. *für jedes $f \in \Sigma_{\mathcal{F}}$ mit Stelligkeit n und für alle $b_1, \ldots, b_n \in B$ ist auch $f_{\mathcal{A}}(b_1, \ldots, b_n) \in B$,*
2. *für jedes $e \in \Sigma_{\mathcal{E}}$ gilt $e_{\mathcal{A}} \in B$.*

Die betreffende Teilstruktur ergibt sich durch die Einschränkung aller Funktionen und Relationen von $\mathcal{A}$ auf die Menge B.

Man sieht, daß eine nichtleere Teilmenge B einer *Relationalstruktur* $\mathcal{A}$ die Grundmenge einer Teilstruktur $\mathcal{B}$ bildet, sobald nur B alle Interpretationen von Individuenkonstanten der gegebenen Signatur enthält. Für Algebren hingegen gilt diese Aussage nicht, da im allgemeinen eine gegebene Teilmenge nicht unter Anwendung von Funktionen aus $\Sigma_{\mathcal{F}}$ abgeschlossen ist.

Die Menge der Teilstrukturen einer gegebenen Struktur ist im nachfolgend zu präzisierenden Sinn unter Durchschnittsbildung abgeschlossen.

Lemma 6.28 *Für jedes k aus der nichtleeren Indexmenge K sei $\mathcal{B}_k = \langle B_k, I_k \rangle$ eine Teilstruktur der Σ-Struktur $\mathcal{A} = \langle A, I \rangle$. Es sei $B := \bigcap_{k \in K} B_k$ nichtleer. Wir definieren eine Interpretationsfunktion J wie folgt:*

- *für $R \in \Sigma_{\mathcal{R}}$ sei $J(R) := \bigcap_{k \in K} I_k(R)$,*
- *für $f \in \Sigma_{\mathcal{F}}$ sei $J(f) := \bigcap_{k \in K} I_k(f)$,*
- *für $e \in \Sigma_{\mathcal{E}}$ sei $J(e) := I(e)$.*

Hierbei werden die Relationen $I_k(R)$ und die Funktionen $I_k(f)$ als Mengen geordneter Tupel aufgefaßt, für jedes $k \in K$. Dann ist die Struktur $\mathcal{B} := \langle B, J \rangle$ eine Teilstruktur von $\mathcal{A}$.

Beweis. Übung (vgl. Aufgabe 6.11). ■

Das vorausgegangene Lemma zeigt insbesondere, daß der Durchschnitt zweier Teilstrukturen $\mathcal{B}_1$ und $\mathcal{B}_2$ einer gegebenen Σ-Struktur $\mathcal{A}$ mit nichtdisjunkten Grundmengen B_1 und B_2 stets wieder eine Teilstruktur ist. Wenn die Signatur Σ zumindest eine Individuenkonstante enthält, ist diese Nichtdisjunktheitsforderung automatisch erfüllt.

Es bleibt dem Leser überlassen, sich ein Beispiel einer Struktur auszudenken, wo hingegen die *Vereinigung* zweier Teilstrukturen nicht selbst wieder eine Teilstruktur ergibt (vgl. Aufgabe 6.12).

Erzeugung von Teilstrukturen

Wir wollen nun der sich hier anschließenden Frage nachgehen, wie wir zu einer gegebenen nichtleeren Teilmenge M einer Σ-Struktur $\mathcal{A} = \langle A, I \rangle$ die kleinste Teilstruktur $\mathcal{B}$ von $\mathcal{A}$ finden können, die M umfaßt.

Eine erste, allerdings unbefriedigende Antwort liefert Lemma 6.28: wir können einfach den Durchschnitt aller M umfassenden Teilstrukturen von $\mathcal{A}$ nehmen. Dieser Durchschnitt ist vernünftig erklärt, da ja zumindest $\mathcal{A}$ selbst eine M umfassende Teilstruktur ist. Es ist daher trivial zu sehen, daß wir auf diese Art in der Tat die kleinste M umfassende Teilstruktur von $\mathcal{A}$ erhalten.

Aufschlußreicher ist hingegen die folgende induktive Konstruktion (vgl. Abschnitt 3.6.1), die man als *Erzeugung von Teilstrukturen* bezeichnet. Wir führen das Symbol $\langle M \rangle^\Sigma$ für die Grundmenge der gesuchten Teilstruktur ein.

- Die *Basisregel* besagt, daß die Elemente von M selbst und alle Individuen aus A der Form $e_\mathcal{A}$ für $e \in \Sigma_\mathcal{E}$ zu $\langle M \rangle^\Sigma$ gehören.

- Die *induktive Regel* besagt, daß für jedes Funktionssymbol $f \in \Sigma_F$ und für alle Elemente $a_1, \ldots, a_n$ aus $\langle M \rangle^\Sigma$ (wo n die Stelligkeit von f ist) stets auch $f_\mathcal{A}(a_1, \ldots, a_n)$ zu $\langle M \rangle^\Sigma$ gehört.

Bevor wir zeigen, daß die so definierte Menge $\langle M \rangle^\Sigma$ tatsächlich die Grundmenge der kleinsten M umfassenden Teilstruktur von $\mathcal{A}$ ist, wollen wir den in Abschnitt 3.6.1 beschriebenen schrittweisen Aufbau der induktiv definierten Menge im hier vorliegenden Fall nochmals genauer beschreiben. Hierzu führen wir die folgende Abbildung $F\colon \mathcal{P}(A) \to \mathcal{P}(A)$ ein: für $N \subseteq A$ sei

$$F(N) := N \cup \{f_{\mathcal{A}}(b_1, \ldots, b_n) \mid f \in \Sigma_{\mathcal{F}}, b_1, \ldots, b_n \in N\}.$$

Es ist also $F(N)$ genau diejenige Menge, die wir ausgehend von N erhalten, wenn wir zu N alle Elemente hinzufügen, die sich durch die Anwendung der induktiven Regel auf Elemente aus N ergeben. Wir iterieren diesen Schritt und definieren $F^0(N) := N$ und $F^{n+1}(N) := F(F^n(N))$. Es ist nun

$$F^0(N) \subseteq F^1(N) \ldots F^k(N) \subseteq F^{k+1}(N) \ldots$$

ein aufsteigender Turm von Teilmengen von A und es gilt $\langle M\rangle^\Sigma = \bigcup_{n\geq 0} F^n(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\})$.

Lemma 6.29 *Es seien Σ, $\mathcal{A}$ und $\emptyset \neq M \subseteq A$ wie oben. Dann ist $\langle M\rangle^\Sigma$ die Grundmenge der kleinsten M umfassenden Teilstruktur von $\mathcal{A}$.*

Beweis. Es sei B die Grundmenge der der kleinsten M umfassenden Teilstruktur von $\mathcal{A}$. Aus Lemma 6.27 folgt sofort, daß $\{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\} \subseteq B$ und $F(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\}) \subseteq B$ gilt. Mit einer einfachen Induktion folgt allgemeiner $F^n(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\}) \subseteq B$ für alle $n \geq 0$, damit $\langle M\rangle^\Sigma \subseteq B$. Um nun $\langle M\rangle^\Sigma = B$ zu verifizieren, reicht es offenkundig, zu zeigen, daß $\langle M\rangle^\Sigma$ bereits die Grundmenge einer M umfassenden Teilstruktur ist. Hierzu verwenden wir Lemma 6.27. Wegen $M \subseteq \langle M\rangle^\Sigma$ ist mit M auch $\langle M\rangle^\Sigma$ nichtleer. Aus $\{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\} \subseteq F(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\}) \subseteq \langle M\rangle^\Sigma$ folgt, daß $\langle M\rangle^\Sigma$ Bedingung 2 aus Lemma 2 erfüllt.

Sind nun $b_1, \ldots, b_n \in \langle M\rangle^\Sigma$ und ist $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol, so muß jedes der Elemente b_i in einer Schicht $F^{k_i}(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\})$ des Turms auftreten $(1 \leq i \leq n)$. Bezeichnet k das Maximum der Zahlen k_i $(1 \leq i \leq n)$, so gilt $b_1, \ldots, b_n \in F^k(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\})$ und damit gemäß der Definition von F auch $f_{\mathcal{A}}(b_1, \ldots, b_n) \in F^{k+1}(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\}) \subseteq \langle M\rangle^\Sigma$. Damit ist auch Bedingung 1 aus Lemma 6.27 für $\langle M\rangle^\Sigma$ erfüllt, was zeigt, daß $\langle M\rangle^\Sigma$ in der Tat die Grundmenge einer M umfassenden Teilstruktur von $\mathcal{A}$ ist. ■

Definition 6.30 Es sei $M \subseteq A$. Dann heißt die Menge

$$\langle M\rangle^\Sigma = \bigcup_{n\geq 0} F^n(M \cup \{e_{\mathcal{A}} \mid e \in \Sigma_{\mathcal{F}}\})$$

das Σ-*Erzeugnis* von M.

Wenn wir die Signatur nicht explizit machen, so reden wir einfacher vom *Erzeugnis* der Menge M und schreiben $\langle M\rangle$. Eine notationelle Verwechslung mit Einertupeln wird nachfolgend stets ausgeschlossen sein.

Beispiel 6.31 Wir betrachten die Struktur der reellen Zahlen mit der Addition „+“ und der Null 0 als ausgezeichnetem Element. Als Ausgangsmenge nehmen wir die Menge $M := \{2\}$. Damit wird

$$\begin{array}{rcl} F(M\cup\{0\}) & = & \{0,2,4\}, \\ F^2(M\cup\{0\}) & = & \{0,2,4,6,8\}, \\ F^3(M\cup\{0\}) & = & \{0,2,4,6,8,10,12,14,16\}, \\ \dots & & \dots \end{array}$$

Das Erzeugnis $\langle\{2\}\rangle$ besteht also aus allen geraden natürlichen Zahlen.

Beispiel 6.32 Es sei A ein Alphabet. Es bezeichne A^* die Menge aller Wörter über A, „$\circ$“ die Konkatenation und „ϵ“ das leere Wort. Auf A^* betrachten wir für jedes $a\in A$ die „a-Nachfolgerfunktion“ $succ_a\colon w\mapsto w\circ a$. Bezüglich dieser Menge von ausgezeichneten Funktionen ergibt sich

$$\begin{array}{rcl} F(\{\epsilon\}) & = & \{\epsilon\}\cup A, \\ F^2(\{\epsilon\}) & = & \{\epsilon\}\cup A\cup A^2, \\ F^3(\{\epsilon\}) & = & \{\epsilon\}\cup A\cup A^2\cup A^3, \\ \dots & & \dots \end{array}$$

Damit gilt $\langle\{\epsilon\}\rangle = A^*$.

Bemerkung 6.33 Die Abbildung, die jeder nichtleeren Teilmenge B der Grundmenge einer Struktur $\mathcal{A}$ ihr Erzeugnis zuordnet, ist eine Hüllenabbildung im Sinn von Definition 4.35.

6.4 Homomorphismen

Als wir Funktionen zwischen Mengen eingeführt hatten, hatten wir darauf hingewiesen, daß Abbildungen oft verwendet werden, um eine Korrespondenz zwischen den Elementen zweier Mengen zu formalisieren. Wenn man nun weitergehend eine Korrespondenz zwischen zwei *Strukturen* festhalten möchte, so sollte eine entsprechende Abbildung auch die Beziehungen der Elemente untereinander erhalten. Dies führt uns auf den Begriff des Homomorphismus. Kurz gesagt ist ein Homomorphismus eine „strukturerhaltende“ Abbildung. Bevor wir eine formale Definition geben, zunächst ein motivierendes halbformales Beispiel, mit dem wir versuchen, die grundlegende Bedeutung des Begriffs „Strukturerhaltung“ lebensnah wiederzugeben.

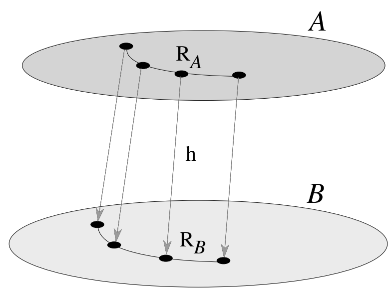
Beispiel 6.34 Sie, Frau, nett, sind in Las Vegas, jung verliebt, und erwarten von Ihrem Partner, schwerreich, Drillinge. Die Hochzeit ist beschlossene Sache, ebenso die baldige Auswanderung in ein anderes Land. Ihr Partner schlägt vor, auf der Stelle zu heiraten. Es gibt nun Auswanderungsländer X, wo Ihr Ortswechsel strukturerhaltend ist in dem Sinn, daß die Ehe aus Las Vegas in X juristisch anerkannt ist. Ein solcher Ortswechsel wäre unter dem Gesichtpunkt des Ehebündnisses „homomorph“. Wenn jedoch im Land X Ehen aus Las Vegas juristisch bedeutungslos sind, so wäre der Ortswechsel genau nicht strukturerhaltend oder homomorph.

Definition 6.35 Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Strukturen derselben Signatur Σ. Eine Abbildung $h\colon A \to B$ heißt *Homomorphismus* von $\mathcal{A}$ nach $\mathcal{B}$ genau dann, wenn gilt:

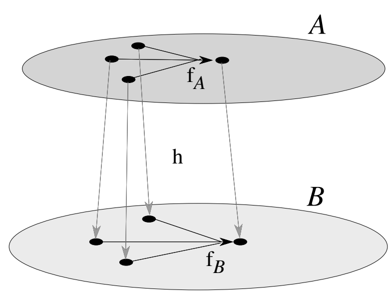
1. für alle $R \in \Sigma_{\mathcal{R}}$, m-stellig und für alle $a_1, \ldots, a_m \in A$:
$$R_{\mathcal{A}}(a_1, \ldots, a_m) \Rightarrow R_{\mathcal{B}}(h(a_1), \ldots, h(a_m)),$$
2. für alle $f \in \Sigma_{\mathcal{F}}$, n-stellig und für alle $a_1, \ldots, a_n \in A$:
$$h(f_{\mathcal{A}}(a_1, \ldots, a_n)) = f_{\mathcal{B}}(h(a_1), \ldots, h(a_n)),$$
3. für alle $e \in \Sigma_{\mathcal{E}}$ gilt $h(e_{\mathcal{A}}) = e_{\mathcal{B}}$.

Bemerkung 6.36 Die so eingeführten Homomorphismen werden manchmal auch als *schwache* Homomorphismen bezeichnet. Für *starke Homomorphismen* fordert man als Eigenschaft 1 weitergehend, daß $R_{\mathcal{A}}(a_1, \ldots, a_m)$ stets zu $R_{\mathcal{B}}(h(a_1), \ldots, h(a_m))$ *äquivalent* ist. Man beachte, daß für Algebren diese Unterscheidung irrelevant ist!

Bedingung 1 aus Definition 6.35 kann wie folgt ausgedrückt werden: wenn wir ein n-Tupel von Elementen aus A (resp. B), das zur Relation $R_{\mathcal{A}}$ (resp. $R_{\mathcal{B}}$) gehört, als ein R-*Tupel* aus $\mathcal{A}$ (resp. $\mathcal{B}$) bezeichnen, so müssen R-Tupel aus $\mathcal{A}$ durch h in R-Tupel aus $\mathcal{B}$ abgebildet werden. Bei starken Homomorphismen ist jedes Urbild eines R-Tupels aus $\mathcal{B}$ auch ein R-Tupel aus $\mathcal{A}$.



Eigenschaft 2 aus Definition 6.35 kann in verschiedener Weise gelesen werden: analog wie für Relationen können wir von f-Tupeln aus $\mathcal{A}$ und $\mathcal{B}$ reden, die Bedingung besagt dann, daß f-Tupel aus $\mathcal{A}$ durch h in f-Tupel aus $\mathcal{B}$ abgebildet werden. Aus anderer Sicht besagt die Bedingung, daß es dasselbe Resultat liefert, wenn man die Elemente zuerst mit $f_{\mathcal{A}}$ in der Struktur $\mathcal{A}$ verknüpft und dann mit h abbildet, oder wenn man zunächst die Argumente einzeln mit h abbildet und dann die Bilder mittels der korrespondierenden Funktion $f_{\mathcal{B}}$ in $\mathcal{B}$ verknüpft.



Beispiel 6.37 Sei Σ die Signatur, die nur das einstellige Funktionssymbol N enthält. Es sei A eine Menge von Personen, die in einem Raum an Zweiertischen sitzen. An jedem Tisch sollen zwei Personen sitzen. Wenn wir N durch die Funktion auf A interpretieren, die jeder Person ihren Nebensitzer zuordnet, so wird hierdurch eine Σ-Algebra $\mathcal{A} = \langle A, N_{\mathcal{A}} \rangle$ erklärt. Nun sei

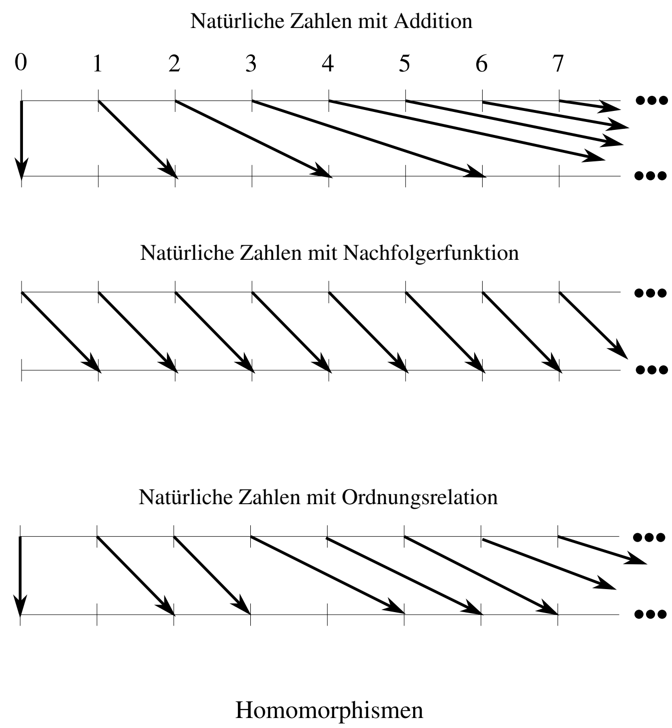


Abbildung 6.4: Abhängigkeit des Homomorphiebegriffs von der Signatur.

B eine zweite Menge von Personen, die in einem anderen Raum an Zweiertischen sitzen. Wir führen analog die Σ-Algebra $\mathcal{B} = \langle B, N_{\mathcal{B}} \rangle$ ein. Eine (nicht notwendigerweise injektive oder surjektive) Abbildung h: $A \to B$ erfüllt Bedingung 2 aus Definition 6.35 und ist ein Homomorphismus genau dann, wenn sie Paare von Nebensitzern in Paare von Nebensitzern überführt. Äquivalent: ausgehend von Person $a \in A$ erhalten wir stets dieselbe Person, wenn wir zunächst den Nebensitzer von a nehmen und dessen Bild betrachten, oder andererseits den Nebensitzer des Bilds von a betrachten.

Da man mit einer festen Grundmenge verschiedene Strukturen verbinden kann, ist zu beachten, daß der Begriff des Homomorphismus von den durch die gewählte Signatur ausgezeichneten Funktionen, Relationen und Elementen abhängig ist. Dies ist in Abbildung 6.4 illustriert und wird auch vom nachfolgenden Beispiel unterstrichen.

Beispiel 6.38 Es sei $A \neq \emptyset$ und $f\colon A \to B$ eine Funktion. Wir betrachten die Strukturen mit Grundmengen $\mathcal{P}(A)$ und $\mathcal{P}(B)$ (vgl. Definition 2.26) mit der Vereinigung „$\cup$" als jeweils einziger Funktion. Gemäß Definition 3.37 (a) können wir f auch als Funktion von $\mathcal{P}(A)$ nach $\mathcal{P}(B)$ auffassen, indem wir jedem $X \in \mathcal{P}(A)$ die Menge $f(X) \in \mathcal{P}(B)$ zuordnen. Lemma 3.40 (i) besagt gerade, daß die so betrachtete Funktion ein Homomorphismus ist. Wenn wir den Durchschnitt „$\cap$" als zusätzliche Funktion in den Strukturbegriff mitaufnehmen, so definiert f im obigen Sinn genau dann einen Homomorphismus, wenn f injektiv ist, wie Lemma 3.53 zeigt.

Lemma 6.39 *Es seien $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ Strukturen derselben Signatur. Es sei f ein Homomorphismus von $\mathcal{A}$ nach $\mathcal{B}$ und g ein Homomorphismus von $\mathcal{B}$ nach $\mathcal{C}$. Dann ist $f \circ g$ ein Homomorphismus von $\mathcal{A}$ nach $\mathcal{C}$.*

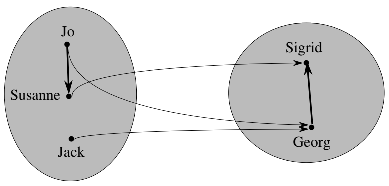
Beweis. Übung (vgl. Aufgabe 6.18). ■

Definition 6.40 Ein injektiver Homomorphismus wird *Monomorphismus* genannt, ein surjektiver Homomorphismus heißt auch *Epimorphismus*. Ein *Isomorphismus* ist ein bijektiver und starker Homomorphismus.

Beispiele 6.41 Einige Illustrationen zu den Begriffen Monomorphismus, Epimorphismus, Isomorphismus.

1. Ist $\mathcal{A}$ eine Teilstruktur von $\mathcal{B}$, so ist die identische Abbildung Id_A ein Monomorphismus von $\mathcal{A}$ nach $\mathcal{B}$.

2. Es sei $A = \{\text{Jo,Jack,Susanne}\}$ und $liebt_{\mathcal{A}} = \{\langle\text{Jo,Susanne}\rangle\}$. Weiter sei $B = \{\text{Georg,Sigrid}\}$ und $liebt_{\mathcal{B}} = \{\langle\text{Georg,Sigrid}\rangle\}$. Dann ist die Abbildung f:
$$\left[\begin{array}{ccc} \text{Jo} & \mapsto & \text{Georg} \\ \text{Jack} & \mapsto & \text{Georg} \\ \text{Susanne} & \mapsto & \text{Sigrid} \end{array}\right]$$
ein schwacher Epimorphismus von $\langle A, liebt_{\mathcal{A}}\rangle$ auf $\langle B, liebt_{\mathcal{B}}\rangle$. Offenkundig ist f surjektiv, und mit $liebt_{\mathcal{A}}(\text{Jo,Susanne})$ gilt auch $liebt_{\mathcal{B}}(\text{Georg,Sigrid})$.



Da $liebt_{\mathcal{A}}$(Jack,Susanne) nicht gilt, ist f kein starker Homomorphismus.

3. Es sei A das Alphabet mit dem einzigen Zeichen a. Es sei nun $\mathcal{M} := \langle A^*, \circ \rangle$ wo „$\circ$“ die Konkatenation bezeichnet. Die Abbildung h: $\mathbb{N} \to A^*$, die jede Zahl $n \in \mathbb{N}$ auf das eindeutig bestimmte Wort $a \cdots a$ der Länge n abbildet, ist ein Isomorphismus zwischen $\langle \mathbb{N}, + \rangle$ und $\mathcal{M}$. Für ein einelementiges Alphabet kann man die Konkatenation daher als eine Form der Addition ansehen.

4. Die Exponentiation exp: $\mathbb{R} \to \mathbb{R}^{>0}, x \mapsto e^x$ ist ein Isomorphismus von $\langle \mathbb{R}, +, 0 \rangle$ in $\langle \mathbb{R}^{>0}, \cdot, 1 \rangle$ (hierbei ist wieder $\mathbb{R}^{>0}$ die Menge der positiven reellen Zahlen). Es gilt nämlich

$$\forall x, y \in \mathbb{R}\colon exp(x+y) = e^{x+y} = e^x \cdot e^y = exp(x) \cdot exp(y)$$

sowie $exp(0) = e^0 = 1$. Außerdem ist exp eine Bijektion (ohne Beweis).

5. Die Abbildung f: $\mathbb{N} \to \mathcal{Z}$

$$f(n) = \begin{cases} n/2 & \text{falls } n \text{ gerade ist} \\ -(n+1)/2 & \text{falls } n \text{ ungerade ist} \end{cases}$$

stellt eine Bijektion zwischen $\mathbb{N}$ und $\mathcal{Z}$ dar. Es ist f jedoch *kein* Isomorphismus von $\langle \mathbb{N}, \leq \rangle$ nach $\langle \mathcal{Z}, \leq \rangle$: es gilt nämlich $0 \leq 1$ aber $f(0) = 0 \not\leq -1 = f(1)$. Man kann leicht sehen (vgl. Aufgabe 6.20), daß es zwar einen Monomorphismus von $\langle \mathbb{N}, \leq \rangle$ nach $\langle \mathcal{Z}, \leq \rangle$ gibt, aber keinen Epimorphismus.

Die nachfolgenden zwei Lemmas stellen zwei wichtige Beobachtungen über Homomorphismen zusammen.

Lemma 6.42 *Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Σ-Strukturen und h ein Homomorphismus von $\mathcal{A}$ nach $\mathcal{B}$. Dann bildet das Bild $h(A)$ die Grundmenge einer Σ-Teilstruktur von $\mathcal{B}$.*

Beweis. Wir verifizieren die Voraussetzungen von Lemma 6.27. Offenkundig ist mit A auch $h(A)$ nichtleer. Aus Bedingung 3 von Definition 6.35 folgt sofort, daß $h(A)$ alle ausgezeichneten Elemente $e_{\mathcal{B}}$ enthält ($e \in \Sigma_{\mathcal{E}}$). Seien nun $f \in \Sigma_{\mathcal{F}}$ mit Stelligkeit n und Elemente $b_1, \ldots, b_n \in h(A)$ gegeben. Dann existieren Elemente $a_1, \ldots, a_n \in A$ mit $b_i = h(a_i)$, $1 \leq i \leq n$. Aufgrund Bedingung 3 aus Definition 6.35 ist damit $f_{\mathcal{B}}(b_1, \ldots, b_n) = f_{\mathcal{B}}(h(a_1), \ldots, h(a_n)) = h(f_{\mathcal{A}}(a_1, \ldots, a_n))$ in $h(A)$. Die Behauptung folgt nun mit Lemma 6.27. ■

Lemma 6.43 *Es sei h: $A \to B$ ein Isomorphismus zwischen den Σ-Strukturen $\mathcal{A}$ und $\mathcal{B}$. Dann ist auch h^{-1}: $B \to A$ ein Isomorphismus und es gilt $h \circ h^{-1} = Id_A$ und $h^{-1} \circ h = Id_B$.*

Beweis. Es sei h: $A \to B$ ein Isomorphismus zwischen $\mathcal{A}$ und $\mathcal{B}$. Gemäß Lemma 3.54 ist h^{-1}: $B \to A$ eine Bijektion und es gilt $h \circ h^{-1} = Id_A$ und $h^{-1} \circ h = Id_B$. Es bleibt noch nachzuweisen, daß h^{-1} ein starker Homomorphismus ist.

Sei $R \in \Sigma_{\mathcal{R}}$ ein m-stelliges Relationssymbol und $b_1, \ldots, b_m \in B$. Weiter sei $a_1 := h^{-1}(b_1), \ldots, a_m := h^{-1}(b_m)$, wodurch $b_1 = h(a_1), \ldots, b_m = h(a_m)$ folgt. Da h als Isomorphismus nach Voraussetzung ein starker Homomorphismus ist, gilt

$$\begin{array}{rcl} R_{\mathcal{B}}(b_1, \ldots, b_m) & \Leftrightarrow & R_{\mathcal{B}}(h(a_1), \ldots, h(a_m)) \\ & \Leftrightarrow & R_{\mathcal{A}}(a_1, \ldots, a_m) \\ & \Leftrightarrow & R_{\mathcal{A}}(h^{-1}(b_1), \ldots, h^{-1}(b_m)). \end{array}$$

Dies zeigt, daß h^{-1} Eigenschaft 1 aus Definition 6.35 erfüllt.

Es sei $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol und $b_1, \ldots, b_m \in B$. Weiter sei $a_1 := h^{-1}(b_1), \ldots, a_n := h^{-1}(b_n)$. Da h ein Homomorphismus ist, gilt

$$\begin{array}{rcl} h^{-1}(f_{\mathcal{B}}(b_1, \ldots, b_n)) & = & h^{-1}(f_{\mathcal{B}}(h(a_1), \ldots, h(a_n))) \\ & = & h^{-1}(h(f_{\mathcal{A}}(a_1, \ldots, a_n))) \\ & = & f_{\mathcal{A}}(a_1, \ldots, a_n) \\ & = & f_{\mathcal{A}}(h^{-1}(b_1), \ldots, h^{-1}(b_n)). \end{array}$$

Dies zeigt, daß h^{-1} Eigenschaft 2 aus Definition 6.35 erfüllt. Die Überprüfung, daß h^{-1} auch Eigenschaft 3 erfüllt, ist einfach und wird dem Leser überlassen. ■

In der mathematischen Literatur gibt es noch weitere Namen für spezielle Homomorphismen. Ein Homomorphismus von einer Struktur in sich selbst heißt *Endomorphismus*. Ein Isomorphismus von einer Struktur in sich selbst heißt *Automorphismus*. Wir wollen nun den Begriffen Isomorphismus, Monomorphismus und Epimorphismus eine anschauliche Interpretation geben. Diese Aufgabe wird erleichtert, wenn wir uns hierbei auf starke Homomorphismen beschränken.

Gibt es einen Isomorphismus h zwischen den Strukturen $\mathcal{A} = \langle A, I\rangle$ und $\mathcal{B} = \langle B, J\rangle$, so bedeutet dies, daß $\mathcal{A}$ und $\mathcal{B}$ genaue Kopien voneinander sind. Sie unterscheiden sich nur durch die Namensgebung der Elemente der Grundmenge sowie gegebenenfalls durch die Namen der ausgezeichneten Relationen und Funktionen und der ausgezeichneten Elemente. In der Mathematik möchte man oft zwischen isomorphen Strukturen keinen Unterschied machen. Ein häufiges Ziel algebraischer Untersuchungen ist es, alle Strukturen einer bestimmten Art bis auf Isomorphie zu klassifizieren.

Gibt es einen starken Monomorphismus h von $\mathcal{A} = \langle A, I\rangle$ nach $\mathcal{B} = \langle B, F\rangle$, so bedeutet dies, daß $\mathcal{A}$ in $\mathcal{B}$ einbettbar ist. Das Bild $h(A)$ von A, versehen mit den Einschränkungen der ausgezeichneten Relationen und Funktionen von $\mathcal{B}$ auf $h(A)$ und mit den ausgezeichneten Elementen von $\mathcal{B}$, ist eine Teilstruktur von $\mathcal{B}$, die zu $\mathcal{A}$ isomorph ist (Aufgabe 6.19). Man kann also $\mathcal{A}$ bis auf Namensgebung in $\mathcal{B}$ als Teilstruktur wiederfinden.

Gibt es einen starken Epimorphismus h von $\mathcal{A} = \langle A, I\rangle$ nach $\mathcal{B} = \langle B, J\rangle$, so bedeutet dies, daß $\mathcal{B}$ als eine Vergröberung der Struktur $\mathcal{A}$ betrachtet werden kann, in der im Prinzip jedoch dieselben Beziehungen gelten, wenn man davon absieht, daß h verschiedene Elemente von A identifizieren kann. Die Bedeutung der Epimorphismen beruht darauf, daß man in der gröberen Struktur oft viel einfacher rechnen kann, ohne daß dabei etwas Entscheidendes an Information verlorenginge. Dies illustriert das folgende Beispiel.

Beispiel 6.44 Angenommen, wir möchten natürliche Zahlen addieren und multiplizieren, interessieren uns aber nur für die Frage, wie für das Rechenergebnis der beim Teilen durch 5 übrigbleibende Rest des jeweiligen Rechenergebnisses aussieht. Wir können ähnlich wie in Abbildung 6.2 durch Aufwicklung von $\mathbb{N}$ eine Äquivalenzrelation „$\sim$“ auf $\mathbb{N}$ einführen. Vermöge „$\sim$“ werden genau solche Zahlen identifiziert, zwischen denen wir keinen Unterschied machen wollen. Die Äquivalenzklassen sind in Abbildung 6.5 dargestellt. Die Abbildung, die jeder ganzen Zahl k die zugehörige Stun-

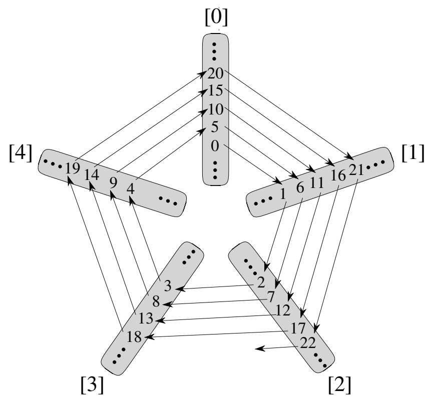


Abbildung 6.5: Vereinfachung von Rechenoperationen mittels Epimomorphismus.

de oder Äquivalenzklasse $[k]$ zuordnet, ist ein Epimorphismus bezüglich der Addition und Multiplikation. Die Homomorphie-Eigenschaft folgt sofort aus der Definition von Uhrenaddition und Uhrenmultiplikation

$$\begin{aligned}[i] \oplus [j] &:= [i+j], \\ [i] \odot [j] &:= [i \cdot j],\end{aligned}$$

deren Wohldefiniertheit wir in Beispiel 6.14 nachgewiesen hatten. Die Surjektivität ist trivial. Aus der Homomorphie-Eigenschaft ergibt sich sofort, daß sich sich gültige Rechengesetze der Ausgangsstruktur — wie die Kommutativität der Addition oder der Multiplikation — automatisch auf die epimorphe Bildstruktur Z_5 übertragen. Unter verschiedenen Aspekten ist das Rechnen in Z_5 aber zudem gleichzeitig einfacher und reichhaltiger. Wir können die Operationen „$\oplus$“ und „$\odot$“ in *endlichen* Tabellen darstellen, wie wir das in Bemerkung 6.16 bereits für Z_2 durchgeführt hatten. Es ist dann leichter, $[1] \odot [2]$ zu berechnen, als das Produkt $338459306 \cdot 834087547$. Während $\mathbb{N}$ zusammen mit der Addition oder der Multiplikation keine Gruppe bildet, da wir dort keine negativen Zahlen oder Brüche haben, ist in Z_5 eine Inversenbildung bezüglich der Uhrenaddition und bezüglich der Uhrenmultiplikation möglich. Während also die ursprüngliche Struktur nicht einmal einen Ring darstellte, ist das epimorphe Bild sogar ein Körper.

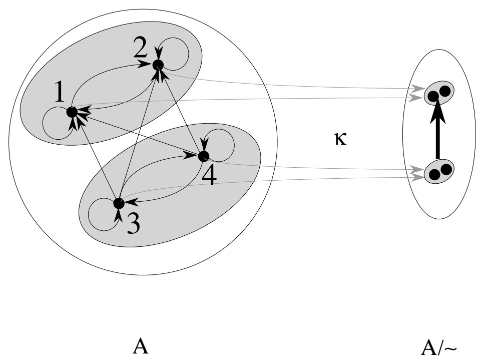


Abbildung 6.6: Kongruenzrelation auf einer Quasi-Ordnung und Quotientenstruktur.

6.5 Kongruenzrelationen und Quotientenstrukturen

Beispiel 6.44 zeigt, daß es unter bestimmten Voraussetzungen möglich ist, ausgehend von einer Struktur $\mathcal{A}$ und einer Äquivalenzrelation „$\sim$“ auf der Grundmenge A eine neue Struktur $\mathcal{B}$ mit Grundmenge $A/\sim$ anzugeben, die man als „Vereinfachung“ von $\mathcal{A}$ ansehen kann, und wo die Rechenoperationen sich durch Übertragung der Rechenoperationen von $\mathcal{A}$ ergeben. Die Einführung der Äquivalenzrelation „$\sim$“ folgte der Vorstellung, Elemente von A zu identifizieren, zwischen denen wir im aktuellen Kontext keinen Unterschied machen wollen.

Ein verwandtes Phänomen war uns bereits bei der Diskussion von Quasi-Ordnungen und partiellen Ordnungen in Beispiel 4.29 begegnet. Bei einer Quasi-Ordnung „$\preceq$“ möchte man manchmal Elemente a und b identifizieren, für die zugleich $a \preceq b$ und $b \preceq a$ gilt. Wie wir in Lemma 4.30 gesehen hatten, beschreibt diese symmetrische Art der Beziehung eine Äquivalenzrelation „$\sim$“ auf A. Wir konnten dann durch Übertragen der unsymmetrischen Ordnungsbeziehungen von Elementen auf Äquivalenzklassen eine korrespondierende (partielle) Ordnung „$\leq$“ auf $A/\sim$ eingeführen. Abbildung 6.6 illustriert dies nochmals.

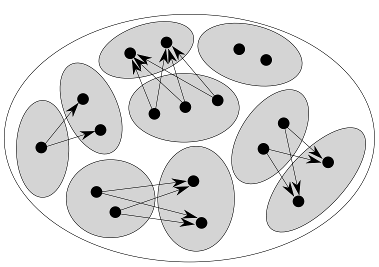
Wir werden nun zeigen, wie man diese Beispiele in systematischer Weise verallgemeinern kann, um damit eine Technik zur Vereinfachung von Strukturen zu erhalten. Dazu fragen wir zunächst nach den speziellen Eigenschaften, die eine Äquivalenzrelationen „$\sim$“ auf der Grundmenge A einer Struktur $\mathcal{A}$ aufweisen muß, um mit ihrer Hilfe eine vereinfachte Variante von $\mathcal{A}$ zu definieren.

Definition 6.45 Es sei Σ eine Signatur und $\mathcal{A}$ eine Σ-Struktur mit Grundmenge A. Eine Äquivalenzrelation „$\sim$“ auf A heißt *Kongruenzrelation* auf $\mathcal{A}$ genau dann, wenn gilt:

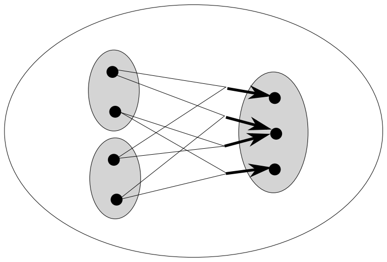
1. Für jedes Relationssymbol $R \in \Sigma_{\mathcal{R}}$ der Stelligkeit m und alle Elemente $a_1, a'_1, \ldots, a_m, a'_m$ aus A mit $a_1 \sim a'_1, \ldots, a_m \sim a'_m$ gilt $R_{\mathcal{A}}(a_1, \ldots, a_m)$ genau dann, wenn $R_{\mathcal{A}}(a'_1, \ldots, a'_m)$.

2. Für jedes Funktionssymbol $f \in \Sigma_{\mathcal{F}}$ der Stelligkeit n und alle Elemente $a_1, a'_1, \ldots, a_n, a'_n$ aus A mit $a_1 \sim a'_1, \ldots, a_n \sim a'_n$ gilt $f_{\mathcal{A}}(a_1, \ldots, a_n) = f_{\mathcal{A}}(a'_1, \ldots, a'_n)$.

Bedingung 1 besagt folgendes: wenn wir in einem Tupel irgendwelche Komponenten durch äquivalente Elemente ersetzen, so kann dies nichts an der Zugehörigkeit oder Nichtzugehörigkeit zur Relation $R_{\mathcal{A}}$ ändern. Die nachfolgende Abbildung illustriert dies am Beispiel einer zweistelligen Relation $R_{\mathcal{A}}$, die durch Pfeile angedeutet ist.



Bedingung 2 besagt, daß wir äquivalente Ergebnisse erhalten, wenn wir dieselbe ausgezeichnete Funktion auf äquivalente Argumente anwenden. In der nachfolgenden Abbildung sind vier Bilder einer zweistelligen Funktion angedeutet. Die das erste (zweite) Argument jeweils aus derselben Äquivalenzklasse kommt, müssen die Bilder in derselben Äquivalenzklasse liegen.



In Beispiele 4.7 Nr. 1 hatten wir gesehen, daß jede Abbildung $f\colon A \to B$ eine Äquivalenzrelation „$\sim_f$“ auf A definiert vermöge $a \sim_f a' :\Leftrightarrow f(a) = f(b)$. Diese Beobachtung kann wie folgt verallgemeinert werden.

Lemma 6.46 *Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Σ-Strukturen und $h\colon A \to B$ ein starker Homomorphismus. Die Relation „$\sim_h$“ auf A sei durch $a \sim_h a' :\Leftrightarrow h(a) = h(a')$ definiert. Dann ist „$\sim_h$“ eine Kongruenzrelation auf $\mathcal{A}$.*

Beweis. Gemäß Beispiele 4.7 Nr. 1 definiert „$\sim_h$“ eine Äquivalenzrelation auf A. Es sei $R \in \Sigma_{\mathcal{R}}$ ein m-stelliges Relationssymbol und $a_1, a_1', \ldots, a_m, a_m'$ Elemente aus A mit $a_i \sim_h a_i'$ für $1 \leq i \leq m$. Da h ein starker Homomorphismus ist, gilt

$$\begin{array}{rcl} R_{\mathcal{A}}(a_1, \ldots, a_m) & \Leftrightarrow & R_{\mathcal{B}}(h(a_1), \ldots, h(a_m)) \\ & \Leftrightarrow & R_{\mathcal{B}}(h(a_1'), \ldots, h(a_m')) \\ & \Leftrightarrow & R_{\mathcal{A}}(a_1', \ldots, a_m') \end{array}$$

Entsprechend weist man die Kongruenzbedingungen für Funktionen nach. ■

Wir nennen „$\sim_h$“ die durch den Homomorphismus h *induzierte* Kongruenzrelation auf $\mathcal{A}$. Weitere Beispiele für Kongruenzrelationen werden wir unten angeben. Zunächst beschreiben wir die Vereinfachung von Strukturen mittels Kongruenzrelationen.

Satz 6.47 *Es sei Σ eine Signatur, $\mathcal{A}$ eine Σ-Struktur und „$\sim$“ eine Kongruenzrelation auf $\mathcal{A}$. Definieren wir*

1. *für $R \in \Sigma_{\mathcal{R}}$ der Stelligkeit m und $[a_1], \ldots, [a_m] \in A/{\sim}$:*

$$R_{A/\sim}([a_1], \ldots, [a_m]) :\Leftrightarrow R_{\mathcal{A}}(a_1, \ldots, a_m),$$

2. für $f \in \Sigma_{\mathcal{F}}$ der Stelligkeit n und $[a_1], \ldots, [a_n] \in A/{\sim}$:

$$f_{A/\sim}([a_1], \ldots, [a_n]) := [f_{\mathcal{A}}(a_1, \ldots, a_n)],$$

3. für $e \in \Sigma_{\mathcal{E}}$: $e_{A/\sim} := [e_{\mathcal{A}}]$,

so hängen die Definitionen in 1 und 2 nicht von der Wahl der Repräsentanten ab. Die durch 1– 3 erklärte Interpretation der Symbole aus Σ definiert eine Σ-Struktur $\mathcal{A}/{\sim}$ mit Grundmenge $A/{\sim}$. Die kanonische Abbildung κ: $A \to A/{\sim}$; $a \mapsto [a]$ ist ein starker Homomorphismus von $\mathcal{A}$ nach $\mathcal{A}/{\sim}$.

Beweis. Wie man sich sofort vergewissern kann, drücken die Bedingungen 1 und 2 aus Definition 6.45 nichts anderes aus, als daß die obigen Definitionen nicht von der Wahl der Vertreter abhängen. Es ist damit klar, daß die obigen Definitionen eine Σ-Struktur $\mathcal{A}/{\sim}$ mit Grundmenge $A/{\sim}$ festlegen. Es bleibt zu zeigen, daß κ ein Homomorphismus ist. Sei $f \in \Sigma$ ein n-stelliges Funktionssymbol und $a_1, \ldots, a_n \in A$. Dann gilt nach Teil 2 der obigen Definition

$$\begin{aligned} \kappa(f_{\mathcal{A}}(a_1, \ldots, a_n)) &= [f_{\mathcal{A}}(a_1, \ldots, a_n)] \\ &= f_{A/\sim}([a_1], \ldots, [a_n]) \\ &= f_{A/\sim}(\kappa(a_1), \ldots, \kappa(a_n)), \end{aligned}$$

damit ist die Homomorphiebedingung bezüglich Funktionen erfüllt. Sei $R \in \Sigma$ ein m-stelliges Relationssymbol und $a_1, \ldots, a_m \in A$. Aus $R_{\mathcal{A}}(a_1, \ldots, a_m)$ folgt nach Teil 1 der obigen Definition $R_{A/\sim}([a_1], \ldots, [a_m])$ und damit $R_{A/\sim}(\kappa(a_1), \ldots, \kappa(a_m))$. Umgekehrt bedeutet $R_{A/\sim}(\kappa(a_1), \ldots, \kappa(a_m))$ gerade $R_{A/\sim}([a_1], \ldots, [a_m])$ und impliziert $R_{\mathcal{A}}(a_1, \ldots, a_m)$. Damit ist die Homomorphiebedingung für Relationen erfüllt. Weiter gilt für jede Individuenkonstante $e \in \Sigma_{\mathcal{E}}$ stets

$$\kappa(e_{\mathcal{A}}) = [e_{\mathcal{A}}] = e_{A/\sim}.$$

Damit ist auch die Homomorphiebedingung für Individienkonstante verifiziert. ■

Definition 6.48 In der Situation von Satz 6.47 wird $\mathcal{A}/{\sim}$ die *Quotientenstruktur* von $\mathcal{A}$ bezüglich der Kongruenzrelation „$\sim$“ genannt. Ist $\mathcal{A}$ eine Algebra, so spricht man von der *Quotientenalgebra*. Die Abbildung κ: $A \to A/{\sim}$; $a \mapsto [a]$ wird der *kanonische Homomorphismus* von $\mathcal{A}$ nach $\mathcal{A}/{\sim}$ genannt.

Im speziellen Fall, wo die Kongruenzrelation „$\sim$“ durch einen Epimorphismus induziert ist, ergibt sich ein Zusammenhang zwischen Bildstruktur und Quotientenstruktur.

Lemma 6.49 *Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Σ-Strukturen und h: $A \to B$ ein starker Epimorphismus. Ist „$\sim_h$“ die in Lemma 6.46 definierte Kongruenzrelation auf $\mathcal{A}$, so ist die Abbildung $A/{\sim_h} \to B$: $[a] \mapsto h(a)$ wohldefiniert und ein Isomorphismus von $\mathcal{A}/{\sim_h}$ auf $\mathcal{B}$.*

Beweis. Übung. ■

Um die vorausgegangenen Begriffe und Resultate zu illustrieren, kehren wir zu unseren motivierenden Beispielen zurück. Bei den nachfolgenden Beispielen verzichten wir auf die explizite Angabe der Signatur. Die Kongruenzbedingungen sind für alle ausgezeichneten Relationen und Funktionen nachzuweisen.

Beispiel 6.50 Für $i, j \in \mathbb{N}$ gelte $i \sim i'$ genau dann, wenn i und i' beim Teilen durch 5 denselben Rest lassen. Dann ist „$\sim$“ eine Kongruenzrelation auf $\langle \mathbb{N}, +, \cdot \rangle$. Wir könnten dies dadurch nachweisen, daß wir Bedingung 2 aus Definition 6.45 für die Funktionen „$+$“, und „$\cdot$“ verifizieren (Übung). Da jedoch $i \sim j$ genau dann gilt, wenn die Äquivalenzklassen $[i]$ und $[i']$ bezüglich „$\sim$“ identisch sind, und da wir in Beispiel 6.44 bereits gesehen hatten, daß die kanonische Abbildung $i \mapsto [i]$ ein (dann automatisch starker) Epimorphismus ist, zeigt eine Anwendung von Lemma 6.46 auch direkt, daß „$\sim$“ eine Kongruenzrelation auf $\mathcal{A}$ ist. Die resultierende Quotientenalgebra ist im wesentlichen mit dem Körper Z_5 identisch, wobei wir allerdings nur die Addition und die Multiplikation als Operationen augezeichnet haben. Wir hätten die Zahl 5 natürlich auch durch jede andere positive natürliche Zahl n ersetzen können.

Beispiel 6.51 Es sei $\mathcal{A}$ eine Ordnungsstruktur der Form $\langle A, \preceq \rangle$ wo „$\preceq$“ eine Quasi-Ordnung auf A ist. Die Äquivalenzrelation „$\sim$“ sei wie in Lemma 4.30 erklärt: es gilt $a \sim a'$ genau dann, wenn $a \preceq a'$ und $a' \preceq a$. Dann ist „$\sim$“ eine Kongruenzrelation auf $\mathcal{A}$. Um dies nachzuweisen, müssen wir lediglich Bedingung 1 aus Definition 6.45 für „$\preceq$“ verifizieren. Es gelte $a \sim a'$, $b \sim b'$ und $a \preceq b$. Aus

$$a' \preceq a \preceq b \preceq b'$$

folgt wegen der Transitivität von „$\preceq$“ sofort $a' \preceq b'$, was zu zeigen war.

Bei Halbgruppen ist eine vereinfachte Charakterisierung von Kongruenzrelationen möglich.

Beispiel 6.52 Es sei $\langle H, \cdot \rangle$ eine Halbgruppe. Eine Äquivalenzrelation „$\sim$" auf H heißt *linksverträglich* (bzw. *rechtsverträglich*) mit „$\cdot$" genau dann, wenn mit $b \sim b'$ und $a \in H$ (bzw. $c \in H$) stets auch $a \cdot b \sim a \cdot b'$ (bzw. $b \cdot c \sim b' \cdot c$) gilt. Die Äquivalenzrelation „$\sim$" heißt verträglich mit „$\cdot$" genau dann, wenn „$\sim$" links- und rechtsverträglich mit „$\cdot$" ist. Es ist „$\sim$" verträglich mit „$\cdot$" genau dann, wenn „$\sim$" eine Kongruenzrelation auf $\langle H, \cdot \rangle$ ist (Aufgabe 6.27).

Wenn es gelingt, durch Quotientenbildung von einer unendlichen Struktur zu einer *endlichen* Variante überzugehen, so kann dies aus algorithmischer Sicht sehr vorteilhaft sein, da sich Relationen und Funktionen auf endlichen Strukturen in Form von endlichen Tabellen darstellen lassen. Oftmals ist es mit Hilfe der Tabellen, die in geeigneter Weise implementiert werden, dann möglich, relevante Fragen algorithmisch zu klären. Dies soll ein Beispiel aus dem Bereich der formalen Sprachen illustriert werden.

Beispiel 6.53 Wir betrachten die Menge A^* aller Wörter über dem zweielementigen Alphabet $A = \{a, b\}$. Auf dieser Menge betrachten wir die einstellige „a-Nachfolger"-Funktion $succ_a : w \mapsto w \circ a$ und die einstellige „b-Nachfolger"-Funktion $succ_b : w \mapsto w \circ b$. Wir interessieren uns für Verfahren, mit denen wir feststellen können, ob ein Eingabewort $w \in A^*$ zur Sprache $L := \{a^n b^m \mid n, m \in \mathbb{N}\}$ gehört oder nicht. In Aufgabe 4.14 hatten wir die folgende Äquivalenzrelation „$\sim_L$" eingeführt: für $v, v' \in A^*$ gilt

$$v \sim_L v' :\Leftrightarrow (\forall u, w \in A^*\colon u \circ v \circ w \in L \Leftrightarrow u \circ v' \circ w \in L).$$

Es ist nicht schwer festzustellen, daß die Relation „$\sim_L$" für die hier gewählte Sprache L genau die drei Äquivalenzklassen

$$\begin{aligned}
[\epsilon] &= \{a^n \mid n \in \mathbb{N}\} \\
[b] &= \{a^m b^n \mid m, n \in \mathbb{N}, n > 0\} \\
[ba] &= \{a, b\}^* \setminus ([\epsilon] \cup [b])
\end{aligned}$$

hat (vgl. Aufgabe 4.15). Diese sind in Abbildung 6.7 dargestellt und verschieden unterlegt. In der Abbildung sind auch a-Nachfolger und b-Nachfolger mittels gelabelter Pfeile markiert. Man kann an der Abbildung ablesen, daß

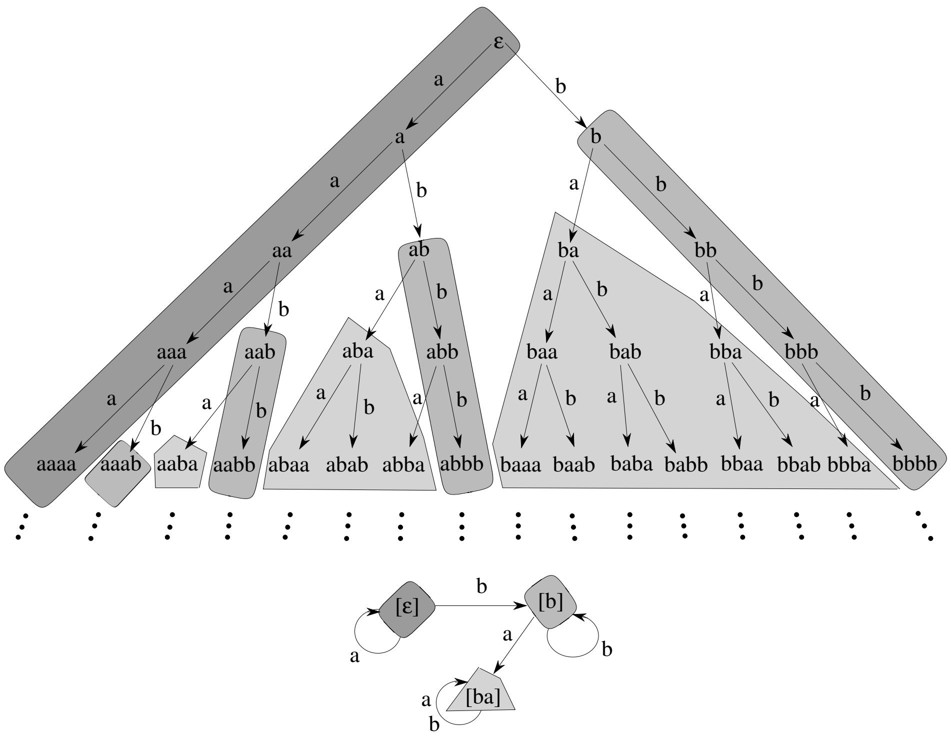


Abbildung 6.7: Kongruenz bezüglich a-Nachfolger, b-Nachfolger und Konkatenation, sowie Quotientenalgebra.

„$\sim_L$" bezüglich der beiden ausgezeichneten Funktionen auf A^* eine Kongruenzrelation ist. In der Tat, ist $[u]$ eine der drei Äquivalenzklassen, so führt jeder a-Nachfolgerschritt von einem beliebigen Element $u' \in [u]$ stets zur selben Äquivalenzklasse $[ua]$, unabhängig von der Wahl von u'. Analoges gilt für die b-Nachfolgerschritte. Wir können daher wie in Satz 6.47 beschrieben die zugehörige Quotientenstruktur berechnen. Im vorliegenden Fall definieren wir auf der Menge der Äquivalenzklassen $\{[\epsilon], [b], [ba]\}$ eine vererbte a-Nachfolgerfunktion $Succ_a$ und b-Nachfolgerfunktion $Succ_b$ vermöge

$$\begin{aligned} Succ_a([w]) &:= [wa] \\ Succ_b([w]) &:= [wb] \end{aligned}$$

Im unteren Teil der Figur 6.7 ist die resultierende vereinfachte Algebra mit den drei Elementen $[\epsilon]$, $[b]$, $[ba]$ und den zwei zugehörigen Nachfolgerfunktionen abgebildet. In Tabellenform stellen sich $Succ_a$ und $Succ_b$ wie folgt dar.

	$Succ_a$	$Succ_b$
$[\epsilon]$	$[\epsilon]$	$[b]$
$[b]$	$[ba]$	$[b]$
$[ba]$	$[ba]$	$[ba]$

Nun gilt aber $L = [\epsilon] \cup [b]$.[7] Daraus erhalten wir für beliebige Wörter $w = a_1 \cdots a_n \in A^*$ folgenden Zusammenhang:

$$\begin{aligned} w \in L \quad &\Leftrightarrow \quad w \in [\epsilon] \quad \vee \quad w \in [b] \\ &\overset{(i)}{\Leftrightarrow} \quad [w] = [\epsilon] \quad \vee \quad [w] = [b] \\ &\Leftrightarrow \quad [(succ_{a_1} \circ \cdots \circ succ_{a_n})(\epsilon)] = [\epsilon] \vee \\ &\qquad [(succ_{a_1} \circ \cdots \circ succ_{a_n})(\epsilon)] = [b] \\ &\overset{(ii)}{\Leftrightarrow} \quad (Succ_{a_1} \circ \cdots \circ Succ_{a_n})([\epsilon]) = [\epsilon] \vee \\ &\qquad (Succ_{a_1} \circ \cdots \circ Succ_{a_n})([\epsilon]) = [b]. \end{aligned}$$

Hierbei folgt (i) aus Lemma 4.10, (ii) ergibt sich daraus, daß die kanonische Abbildung κ: $v \mapsto [v]$ ein Homomorphismus bezüglich der beiden Nachfolgerfunktionen ist.

[7]Allgemeiner ist für beliebiges L stets L darstellbar als Vereinigung von Äquivalenzklassen von „$\sim_L$" siehe Lemma 6.54.

Wie die letzten zwei Zeilen zeigen, können wir also durch n-maliges Nachschlagen in den Tabellen für $Succ_a$ beziehungsweise $Succ_b$ feststellen, ob $w \in L$ gilt oder nicht. Optisch einfacher haben wir einfach ausgehend von $[\epsilon]$ entsprechend der Symbole des Eingabeworts w den Nachfolgerpfeilen in der bildlichen Repräsentation der Quotientenalgebra zu folgen. Es gehört w zu L genau dann, wenn wir nach dem letzten Eingabesymbol in einem der Knoten $[\epsilon]$ oder $[b]$ landen.

Es sei zu diesem Beispiel angemerkt, daß man nicht für jede Sprache L eine endliche Quotientenstruktur erhält. Genauer ist die Quotientenalgebra endlich genau dann, wenn L eine sogenannte *reguläre Sprache* ist. Als Verallgemeinerung des vorangegangenen Beispiels halten wir jedoch das folgende Lemma fest.

Lemma 6.54 *Es sei L eine formale Sprache über dem Alphabet A und $\mathcal{B}$ die Algebra mit der Grundmenge A^* und allen a-Nachfolgerfunktionen $succ_a$: $w \mapsto w \circ a$ für $a \in A$ als ausgezeichneten Funktionen. Die Relation „$\sim$“ auf A^* sei durch*

$$v \sim_L v' :\Leftrightarrow (\forall u, w \in A^*\colon u \circ v \circ w \in L \Leftrightarrow u \circ v' \circ w \in L).$$

für $v, v' \in A^*$ erklärt. Dann ist „$\sim$“ eine Kongruenzrelation auf $\mathcal{B}$. L läßt sich darstellen als Vereinigung von Äquivalenzklassen von „$\sim_L$“.

Beweis. Um die Kongruenzeigenschaft für die Nachfolgerfunktionen nachzuweisen, sei $a \in A$ gegeben. Gilt $v \sim_L v'$ für $v, v' \in A^*$, so gilt für alle $u, w \in A^*$ offenkundig

$$\begin{array}{rcl} u \circ succ_a(v) \circ w \in L & \Leftrightarrow & u \circ (v \circ a) \circ w \in L \\ & \Leftrightarrow & u \circ v \circ (a \circ w) \in L \\ & \Leftrightarrow & u \circ v' \circ (a \circ w) \in L \\ & \Leftrightarrow & u \circ (v' \circ a) \circ w \in L \\ u \circ succ_a(v') \circ w \in L & & \end{array}$$

woraus sich $succ_a(v) \sim_L succ_a(v')$ ergibt. Ist $v \in L$ und $v' \in [v]$ so folgt $v' \sim_L v$ und daher aus der Definition von „$\sim_L$“ durch die Wahl $u := w := \epsilon$ sofort $v' \in L$. Daher folgt aus $v \in L$ stets $[v] \subseteq L$ und $L = \bigcup\{[v] \mid v \in L\}$. ■

6.6 Ergänzung: Die Cantorsche Zick-Zack Methode

Wir schließen nochmals an Abschnitt 6.4 an und stellen eine berühmte Technik vor, die zum Nachweis der Isomorphie zweier abzählbar unendlicher Strukturen verwendet werden kann. Dieselbe Methode kann auch dazu dienen, um in systematischer Art und Weise isomorphie Strukturen aufzubauen. Der Kontext, in dem wir die Methode darstellen, ist noch unter einem weiteren Gesichtpunkt interessant. Es ist manchmal wünschenswert, eine gegebene Struktur durch die Angabe einiger einfacher Eigenschaften in möglichst eindeutiger Weise zu charakterisieren. Im Idealfall soll dann jede Struktur, die dieselben Eigenschaften hat, bereits zur Ausgangsstruktur isomorph sein. Der unten nachfolgende Satz gibt ein Beispiel dieser Vorgehensweise. Wir benötigen zuvor einige Begriffe.

Definition 6.55 Es sei $\mathcal{A} = \langle A, \leq \rangle$ eine linear geordnete Menge.

1. $\mathcal{A}$ ist *dicht* geordnet genau dann, wenn gilt:
$$\forall a, b \in A\colon (a < b \Rightarrow \exists c \in A : a < c < b),$$
2. $\mathcal{A}$ ist *diskret* geordnet genau dann, wenn gilt:
$$\forall a, b \in A\colon (a < b \Rightarrow \exists c, d \in A\colon a \dot{<} c \wedge d \dot{<} b).$$
Hierbei steht $a \dot{<} c$ für $a < c$ und $\neg \exists e \in A\colon a < e < c$.
3. Ein Element $a \in A$ heißt *Endpunkt* genau dann, wenn a das größte oder kleinste Element von A bezüglich $\leq$ ist.

Wir fragen, welche abzählbaren, dicht und linear geordneten Mengen ohne Endpunkte existieren. Ein Standardbeispiel liefern die rationalen Zahlen in der natürlichen Ordnung. Dies ist auch schon—bis auf Isomorphie—das einzige Beispiel.

Satz 6.56 (G. Cantor) *Je zwei abzählbare, dicht geordnete, linear geordnete Mengen $\mathcal{A} = \langle A, \leq_{\mathcal{A}} \rangle$ und $\mathcal{B} = \langle B, \leq_{\mathcal{B}} \rangle$ ohne Endpunkte sind isomorph.*

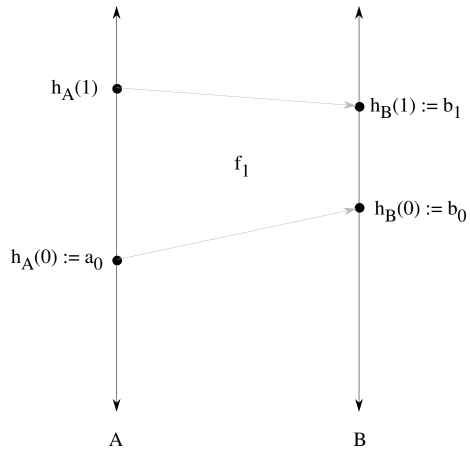
Beweis. Wir verwenden die sogenannte „Zick-Zack-Methode“ (englisch: „back-and-forth method“). Offenkundig müssen die Grundmengen A und B abzählbar unendlich sein. Es seien Aufzählungen aller Elemente von A und B in der Form

$$\begin{gathered} a_0, a_1, a_2, \ldots \\ b_0, b_1, b_2, \ldots \end{gathered}$$

vorgegeben. Wir konstruieren nun rekursiv „partielle“ Isomorphismen f_n ($n = 0, 1, 2, \ldots$), die jeweils genau n Elemente von A bijektiv und ordnungsisomorph auf n Elemente von B abbilden. Beim Schritt von f_n zu f_{n+1} wird jeweils ein neues Paar, bestehend aus einem Elements a aus A und seinem Bild $f_{n+1}(a)$ in B zu f_n hinzugefügt. Es gilt also stets $f_n \subseteq f_{n+1}$. Der Name „Zick-Zack-Methode“ beruht auf der Strategie, mit der sichergestellt wird, daß schließlich *jedes* Element von A und *jedes* Element aus B als Argument oder Bild unter einer Funktion f_n auftritt. Hierzu halten wir fortlaufend Buch über die Elemente aus A und B, die noch keinen Partner haben. Es wechseln nun sogenannte „Zickschritte“, wo wir das erste noch partnerlose Element der Folge $a_0, a_1, a_2, \ldots$ ein Bild in B konstruieren, ab mit sogenannten „Zackschritten“, wo wir für das erste noch partnerlose Element der Folge $b_0, b_1, b_2, \ldots$ ein Urbild in A angeben. Offenkundig garantiert dieses Vorgehen, daß tatsächlich alle Elemente aus A und aus B einen Partner erhalten. Die nachfolgend verwendeten Funktionen $h_A\colon \mathbb{N} \to A$ und $h_B\colon \mathbb{N} \to B$ geben an, in welcher Reihenfolge die Elemente aus A und B in diesem Sinn „verhäkelt“ werden. Es bezeichnet $h_A(n)$ (resp. $h_B(n)$) das n-te Element von A (resp. B), für das wir ein Bild (resp. Urbild) festsetzen.

Schritt 0 („Zick“). Wir setzen $h_A(0) := a_0$, $A_0 := \{h_A(0)\}$, $h_B(0) := b_0$, $B_0 := \{h_B(0)\}$ und $f_0 = \{\langle h_A(0), h_B(0)\rangle\}$; Offensichtlich ist f_0 ein Isomorphismus von $\langle A_0, \leq_{\mathcal{A}}\rangle$ auf $\langle B_0, \leq_{\mathcal{B}}\rangle$.

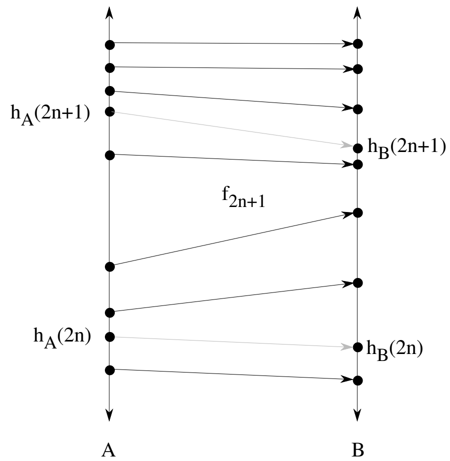
Schritt 1 („Zack“). Zum ersten noch nicht verhäkelten Element der Folge $b_0, b_1, b_2, \ldots$, also zu $h_B(1) := b_1$, ist ein Urbild gesucht. Sei $h_A(1)$ das erste Element in der Aufzählung $a_0, a_1, a_2, \ldots$ von A mit der Eigenschaft, daß $\langle \{h_A(0), h_A(1)\}, \leq_{\mathcal{A}}\rangle$ durch $f_1 := f_0 \cup \{\langle h_A(1), h_B(1)\rangle\}$ isomorph auf $\langle h_B(0), h_B(1), \leq_{\mathcal{B}}\rangle$ abgebildet wird. Natürlich ist dann $h_A(1)$ „neu“, das heißt es gilt $h_A(0) \neq h_A(1)$. Wir setzen $A_1 := \{h_A(0), h_A(1)\}$ sowie $B_1 := \{h_B(0), h_B(1)\}$. Die nachfolgende Abbildung gibt die Situation nach den ersten beiden Schritten wieder.



Annahme: Es seien die Elemente $h_A(0), \ldots, h_A(2n-1)$ von A und $h_B(0), \ldots, h_B(2n-1)$ von B definiert. Es sei $A_{2n-1} := \{h_A(i) \mid i \leq 2n-1\}$ und $B_{2n-1} := \{h_B(i) \mid i \leq 2n-1\}$. Weiterhin sei f_{2n-1}: $A_{2n-1} \to B_{2n-1}$; $h_A(i) \mapsto h_B(i)$ $(1 \leq i \leq 2n-1)$ ein Isomorphismus von $\langle A_{2n-1}, \leq_{\mathcal{A}} \rangle$ auf $\langle B_{2n-1}, \leq_{\mathcal{B}} \rangle$.

Schritt 2n, („Zick“): Sei $h_A(2n)$ das erste Element von $A \setminus A_{2n-1}$ in der oben gegebenen Aufzählung. Wir wählen als Bild $h_B(2n)$ von $h_A(2n)$ das erste Element von $B \setminus B_{2n-1}$ in der oben gegebenen Aufzählung, so daß $f_{2n} := f_{2n-1} \cup \{\langle h_A(2n), h_B(2n) \rangle\}$ ein Isomorphismus von $\langle A_{2n-1} \cup \{h_A(2n)\}, \leq_{\mathcal{A}} \rangle$ auf $\langle B_{2n-1} \cup \{h_B(2n)\}, \leq_{\mathcal{B}} \rangle$ wird. Ein Element $h_B(2n)$ mit dieser Eigenschaft existiert, da $\mathcal{B}$ dicht geordnet und ohne Randpunkte ist. Natürlich gilt $h_B(2n) \notin B_{2n-1}$. Wir setzen $A_{2n} := \{h_A(i) \mid i \leq 2n\}$ und $B_{2n} := \{h_B(i) \mid i \leq 2n\}$.

Schritt 2n+1 („Zack): Sei $h_B(2n+1)$ das erste Element von $B \setminus B_{2n}$ in der oben gegebenen Aufzählung. Wir wählen als Urbild $h_A(2n+1)$ von $h_B(2n+1)$ das kleinste Element von $A \setminus A_{2n}$ in der oben gegebenen Aufzählung, so daß $f_{2n+1} := f_{2n} \cup \{\langle h_A(2n+1), h_B(2n+1) \rangle\}$ ein Isomorphismus von $\langle A_{2n} \cup \{h_A(2n+1)\}, \leq_{\mathcal{A}} \rangle$ auf $\langle B_{2n} \cup \{h_B(2n+1)\}, \leq_{\mathcal{B}} \rangle$ wird. Ein Element $h_A(2n+1)$ mit dieser Eigenschaft existiert, da $\mathcal{A}$ dicht geordnet und ohne Randpunkte ist. Natürlich gilt $h_A(2n+1) \notin A_{2n}$. Wir setzen $A_{2n+1} := \{h_A(i) \mid i \leq 2n+1\}$ und $B_{2n+1} := \{h_B(i) \mid i \leq 2n+1\}$. Die nachfolgende Abbildung illustriert die beiden Schritte. Die Auswahl eines Partnerelements auf der jeweils anderen Seite erfolgt immer so, daß es keine Überschneidungen der Linien gibt, die die Abbildung f_{2n+1} darstellen.



Per Rekursion sind somit für jede natürliche Zahl k Teilmengen $A_k = \{h_A(i) \mid i \leq k\}$ und $B_k = \{h_B(i) \mid i \leq k\}$ sowie Isomorphismen f_k: $A_k \to B_k$; $h_A(i) \mapsto h_B(i)$ $(1 \leq i \leq k)$ von $\langle A_k, \leq_{\mathcal{A}}\rangle$ auf $\langle B_k, \leq_{\mathcal{B}}\rangle$ definiert. Für $k \leq l$ gilt $f_k \subseteq f_l$. Wir setzen nun $f := \bigcup_{k \in \mathbb{N}} f_k$. Damit ist f—als Vereinigung über einen aufsteigenden Turm von Funktionen—offenkundig wieder eine Funktion. Wegen $\bigcup_{k \in \mathbb{N}} A_k = A$ ist offenkundig $A = \mathit{Def}(f)$. Nach Konstruktion ist f injektiv und surjektiv (um Surjektivität zu erreichen, wurden die Zack-Schritte gemacht). Es seien nun $a_n, a_m \in A$. Nach endlich vielen Konstruktionsschritten sind beide verarbeitet: $\exists k \in \mathbb{N}$: $a_n, a_m \in \mathit{Def}(f_k)$. Da f_k ein Isomorphismus ist, folgt daß $a_n \leq_{\mathcal{A}} a_m$ gilt genau dann, wenn $f_k(a_n) \leq_{\mathcal{B}} f_k(a_m)$. Somit ist auch f ein Isomorphismus. ■

6.7 Aufgaben zu Kapitel 6

Aufgaben zu Teilkapitel 6.1

Aufgabe 6.1 Zeigen Sie, daß die in Beispiele 3.34 Nr. 7 erklärte Komposition von Sprachen eine assoziative Funktion ist.

Aufgabe 6.2 Geben Sie die Neutralelemente aller Halbgruppen an, die wir in Abschnitt 6.1.1 als Beispiele eingeführt hatten.

Aufgabe 6.3 Geben Sie eine Halbgruppe an, die kein Neutralelement besitzt.

Aufgabe 6.4 Es bezeichne D_3 die in Beispiel 6.11 eingeführte Diedergruppe des Grads 3. Geben Sie eine vollständige „Multiplikationstafel“ an, indem Sie die folgende Tabelle ergänzen:

$\cdot$	1	d	dd	k	kd	kdd
1						
d						
dd						
k						
kd						
kdd						

Geben Sie zu jedem Element sein Inverses an.

Aufgabe 6.5 Stellen Sie in Analogie zu Aufgabe 6.4 eine vollständige Tafel der Gruppenmultiplikation für die in Beispiel 6.12 eingeführte Gruppe aller Permutationen einer dreielementigen Menge auf.

Aufgabe 6.6 Zeigen Sie: Ist $\mathcal{R}$ ein Ring, und bezeichnet $F := R^M$ die Menge aller Abbildungen einer nichtleeren Menge M in R, so erhalten wir auf F durch die Definitionen $(f +_F g)(x) := f(x) + g(x)$ sowie $(f \cdot_F g)(x) = f(x) \cdot g(x)$ $(x \in M)$ eine Ringstruktur. Geben Sie hierzu das Neutralelement der Addition „$+_F$“ an und charakterisieren die die Inversenbildung bezüglich dieser Addition. Ist $\mathcal{R}$ kommutativ, so auch der resultierende Ring.

Aufgabe 6.7 Es sei $p > 1$ eine Primzahl. Wenn dann p ein Produkt $k \cdot l$ natürlicher Zahlen teilt, so teilt p zumindest eine der Zahlen k und l. Verwenden sie diese Eigenschaft, um zu zeigen:

- für $1 \leq i, j \leq p-1$ läßt das Produkt $i \cdot j$ beim Teilen durch p nie den Rest 0,
- gilt $1 \leq i \leq p-1$, so sind die Reste, die beim Teilen der Zahlen $1 \cdot i, 2 \cdot i, \ldots, (p-1) \cdot i$ entstehenden, paarweise verschieden.

Weisen Sie nun nach, daß im Ring Z_p jedes Element $[i] \neq [0]$ bezüglich der Ringmultiplikation ein Inverses besitzt.

Aufgaben zu Teilkapitel 6.3

Aufgabe 6.8 Geben Sie alle Untergruppen der Diedergruppe des Grads 3 (vgl. Beispiel 6.11) an.

Aufgabe 6.9 Geben Sie alle Teilstrukturen der in Beispiele 6.26 Nr. 7 abgebildeten Relationalstruktur $\mathcal{A}$ an.

Aufgabe 6.10 Beweisen Sie Lemma 6.27.

Aufgabe 6.11 Beweisen Sie Lemma 6.28.

Aufgabe 6.12 Geben Sie ein Beispiel einer Struktur $\mathcal{A}$ mit Teilstrukturen $\mathcal{B}_1$ und $\mathcal{B}_2$ mit jeweiligen Grundmengen B_1 und B_2, wo $B_1 \cup B_2$ nicht selbst die Grundmenge einer Teilstruktur von $\mathcal{A}$ bildet.

Aufgabe 6.13 Es sei $\mathcal{A} = \langle A, I \rangle$ eine Struktur. Zeigen Sie: die Abbildung, die jeder nichtleeren Teilmenge M von A ihr Erzeugnis $\langle M \rangle$ zuordnet, ist eine Hüllenabbildung im Sinn von Kapitel 4.4.

Aufgabe 6.14 Zeigen Sie: ist Σ eine endliche Signatur, und ist B eine abzählbar unendliche Teilmenge der Σ-Struktur $\mathcal{A}$, so ist das Σ-Erzeugnis $\langle B \rangle^{\Sigma}$ auch wieder abzählbar unendlich.

Aufgabe 6.15 Es bezeichne $\mathcal{Z}$ die Menge der ganzen Zahlen. Berechnen Sie das Erzeugnis von $\{4, 6\}$ in der Gruppe $\langle \mathcal{Z}, +, 0, - \rangle$.

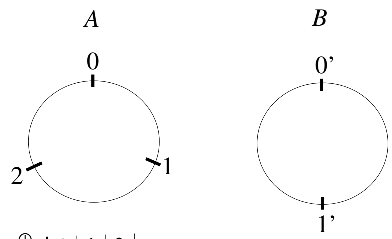
Aufgaben zu Teilkapitel 6.4

Aufgabe 6.16 Es sei $WW := \{0, 1\}$ und $f_{\wedge}\colon WW \to WW$ durch

$f_\wedge$	0	1
0	0	0
1	0	1

erklärt. Geben Sie alle Homomorphismen von $\langle WW, f_\wedge \rangle$ in sich selbst an.

Aufgabe 6.17 Es sei $\mathcal{A}$ die Algebra mit der Grundmenge $\{0, 1, 2\}$ und mit der im Bild links erklärten Uhrenaddition „$\oplus_\mathcal{A}$". Weiter sei $\mathcal{B}$ die Algebra mit der Grundmenge $\{0', 1'\}$ und der im Bild rechts angedeuteten Uhrenaddition „$\oplus_\mathcal{B}$".



$\oplus_A$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\oplus_B$	0'	1'
0'	0'	1'
1'	1'	0'

Zeigen Sie: es gibt einen und nur einen Homomorphismus von $\mathcal{A}$ nach $\mathcal{B}$.

Aufgabe 6.18 Es seien $\mathcal{A}$, $\mathcal{B}$ und $\mathcal{C}$ drei Σ-Strukturen und h_1: $A \to B$ und h_2: $B \to C$ Homomorphismen. Zeigen Sie, daß die Komposition $h_1 \circ h_2$ wieder ein Homomorphismus ist.

Aufgabe 6.19 Es seien $\mathcal{A}$ und $\mathcal{B}$ zwei Strukturen der Signatur Σ und h ein starker Monomorphismus von $\mathcal{A}$ nach $\mathcal{B}$. Zeigen Sie: das Bild $h(A)$ von A, versehen mit den Einschränkungen der Relationen und Funktionen von $\mathcal{B}$ auf $h(A)$ und den ausgezeichneten Elementen aus $\mathcal{B}$, ist eine Teilstruktur von $\mathcal{B}$, die zu $\mathcal{A}$ isomorph ist. Welche Rolle spielt hierbei unsere Voraussetzung, daß der Homomorphismus h stark ist?

Aufgabe 6.20 Beweisen Sie, daß es keinen Epimorphismus von $\langle \mathbb{N}, \leq \rangle$ auf $\langle \mathcal{Z}, \leq \rangle$ geben kann.

Aufgabe 6.21 Es sei $\mathcal{B} = \langle B, + \rangle$ eine beliebige Teilalgebra von $\mathcal{A} := \langle \mathbb{N}, + \rangle$ und $h\colon \mathbb{N} \to B$ ein Homomorphismus. Zeigen Sie: wenn ein von 0 verschiedenes Element $b \in B$ als Bild unter h auftritt, so ist h stets injektiv, also ein Monomorphismus.

Aufgabe 6.22 Bestimmen Sie bis auf Isomorphie alle Strukturen mit einer zweistelligen symmetrischen Relation R auf einer zweielementigen Grundmenge M.

Aufgabe 6.23 Zeigen Sie, daß die in Kapitel 6.1.1 angegebene Permutationsgruppe einer dreielementigen Menge zur Diedergruppe D_3 isomorph ist. Verwenden Sie gegebenenfalls die Aufgaben 6.4 und 6.5.

Aufgabe 6.24 Es sei M eine nichtleere Menge und h ein Automorphismus der Algebra $\langle \mathcal{P}(M), \cup, \cap \rangle$. Zeigen Sie:

1. Stets gilt $h(\emptyset) = \emptyset$.
2. Für jedes $m \in M$ existiert ein $n \in M$ mit $h(\{m\}) = \{n\}$.

Aufgabe 6.25 Die Σ-Algebra $\mathcal{A}$ sei von der Teilmenge $B \subseteq A$ erzeugt. Es sei $\mathcal{C}$ eine weitere Σ-Algebra. Zeigen Sie: je zwei Homomorphismen von $\mathcal{A}$ nach $\mathcal{C}$, die auf B übereinstimmen, sind identisch.

Aufgaben zu Teilkapitel 6.5

Aufgabe 6.26 Für $i, j \in \mathbb{N}$ gelte $i \sim i'$ genau dann, wenn i und i' beim Teilen durch 5 denselben Rest lassen. Zeigen Sie, daß „$\sim$" eine Kongruenzrelation auf $\langle \mathbb{N}, +, \cdot \rangle$ darstellt, indem Sie Bedingung 2 aus Definition 6.45 für die Funktionen „$+$", und „$\cdot$" verifizieren.

Aufgabe 6.27 Es sei $\langle H, \circ \rangle$ eine Halbgruppe und „$\sim$" eine Äquivalenzrelation auf H. Zeigen Sie: „$\sim$" ist verträglich (vgl. Beispiel 6.52) mit „$\circ$" genau dann, wenn „$\sim$" eine Kongruenzrelation auf $\langle H, \circ \rangle$ ist.

Aufgabe 6.28 In Lemma 4.17 hatten wir gezeigt, daß der Durchschnitt einer nichtleeren Menge von Äquivalenzrelationen wieder eine Äquivalenzrelation ist. Zeigen Sie, daß auch Durchschnitt einer nichtleeren Menge von Kongruenzrelationen wieder eine Kongruenzrelation ist.

Aufgabe 6.29 Beweisen Sie Lemma 6.49.

Aufgabe 6.30 Es sei $A := \{a, b\}$ und $L := \{ab^n a \mid n \in \mathbb{N}\}$. Berechnen Sie die Menge der Äquivalenzklassen der in Lemma 6.54 definierten Kongruenzrelation „$\sim_L$“. Die Quotientenalgebra $\mathcal{B}$ sei wie in Lemma 6.54 erklärt. Geben Sie wie in Beispiel 6.53 die Tabelle für die übertragenen Nachfolgerfunktionen „$Succ_a$“ und „$Succ_b$“ auf $\mathcal{B}$ an.

Aufgaben zu Teilkapitel 6.6

Aufgabe 6.31 Wäre die Aussage von Satz 6.56 auch für diskret geordnete Mengen richtig? Geben Sie einen Beweis oder Gegenbeispiel.

6.8 Bibliographische Angaben

Die in Abschnitt 6.1.1 dargestellten Klassen von Algebren sind in zahlreichen Lehrbüchern sehr viel detaillierter dargestellt. Exemplarisch seien [Lan65, Jac85, Mey75, RSV69, dW67] genannt. Eine ausführliche Diskussion von endlichen Automaten und ihren Algebren bietet [Büc98]. Mehr Hintergrund zum Thema Algebren, Homomorphismen, Kongruenzrelationen und Quotientenalgebren findet sich in allen Büchern und Abhandlungen der „universellen Algebra“. Als Beispiel seien [Grä68, MT92] genannt.

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.

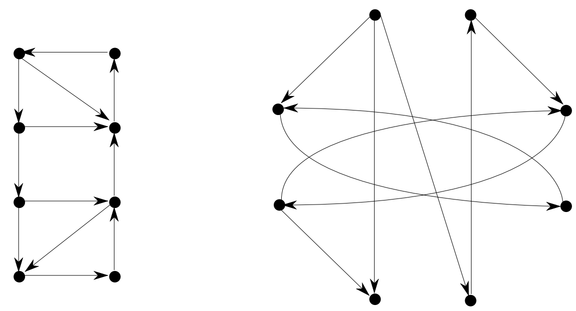
7.1 Beispiele und Typen von Graphen

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

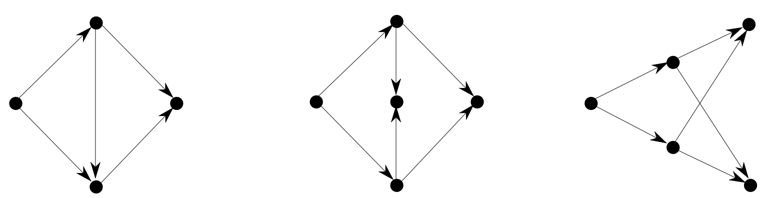
Gerichtete Graphen

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.



1 2 3

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.

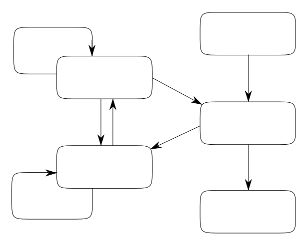


Diagramme ihrerseits werden häufig verwendet, um die Struktur von Programmen zu beschreiben. Graphen können damit die Struktur von Programmen auf einer abstrakten Ebene widerspiegeln.

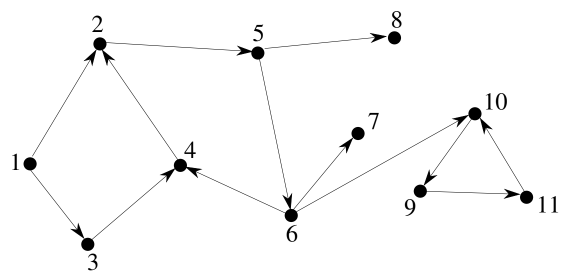
Definition 7.5 Sei $\mathcal{G} = \langle N, E\rangle$ ein gerichteter Graph. Der Knoten m heißt *direkter Nachfolger* des Knotens n genau dann, wenn $E(n,m)$ gilt. Umgekehrt heißt in dieser Situation n ein *direkter Vorgänger* von m. Eine Folge von Kanten $e_1, \ldots, e_k$ $(k \geq 1)$ heißt *Pfad* 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.

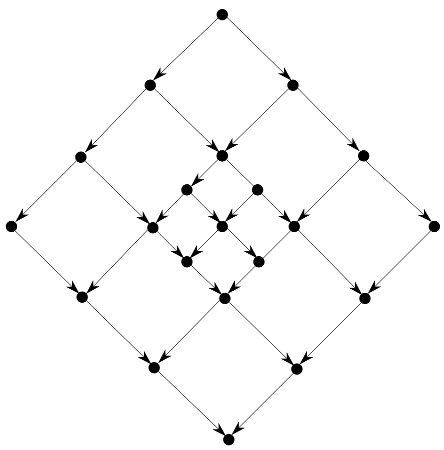
Lemma 7.10 *Ein gerichteter Graph $\mathcal{G}$ ist ein DAG genau dann, wenn für jeden Pfad $e_1, \ldots, e_k$ von $\mathcal{G}$ die Knoten $start(e_1)$, $start(e_2)$, …, $start(e_k)$, $ende(e_k)$ stets paarweise verschieden sind.*

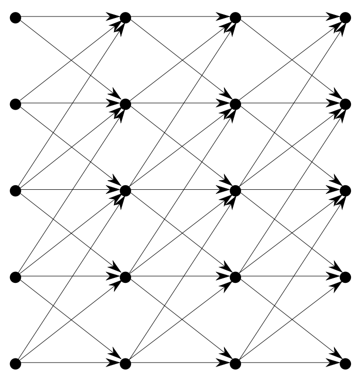
Beweis. Wären zwei Knoten der Folge $start(e_1)$, $start(e_2)$, …, $start(e_k)$, $ende(e_k)$ identisch, so hätte $\mathcal{G}$ offenkundig einen geschlossenen Pfad und wäre kein DAG. Sind für jeden Pfad $e_1, \ldots, e_k$ von $\mathcal{G}$ die Knoten $start(e_1), \ldots, start(e_k), ende(e_k)$ stets 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.*

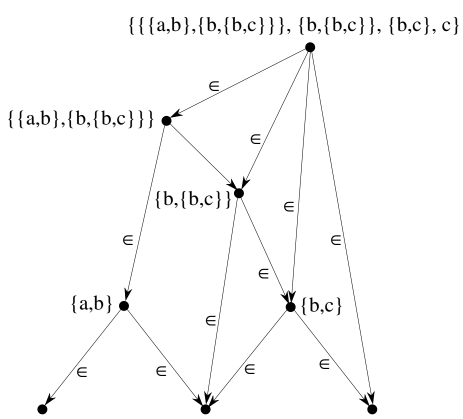
Beispiele 7.12 Die nachfolgende Abbildung zeigt zwei DAGs.





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. ■

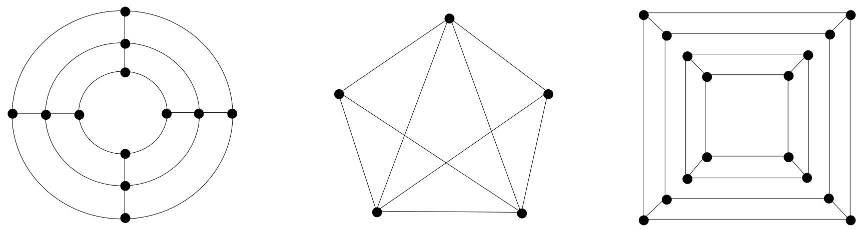
Lemma 7.15 *Es sei $\mathcal{G} = \langle N, E\rangle$ ein DAG mit endlicher Grundmenge N. Dann besitzt $\mathcal{G}$ zumindest ein Element ohne direkten Vorgänger (direkten Nachfolger).*

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.

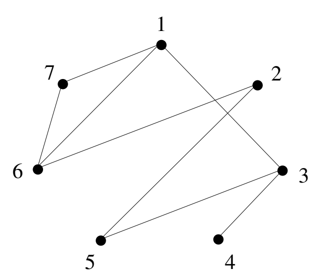
Symmetrische Graphen werden manchmal auch *ungerichtete Graphen* genannt. Mit $\langle m, n\rangle$ liegt auch stets $\langle n, m\rangle$ in der Kantenmenge. Ein Paar $\{\langle m, n\rangle, \langle n, m\rangle\}$ kann alternativ als Menge $\{m, n\}$ repräsentiert werden und wird oft als *ungerichtete 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.

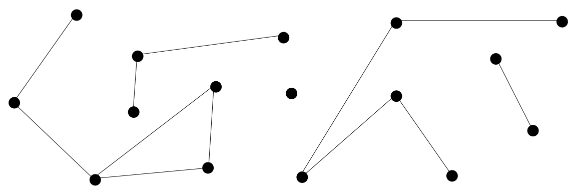
Definition 7.19 Ein symmetrischer Graph $\mathcal{G} = \langle N, E \rangle$ heißt *zusammenhängend* (engl. „connected") genau dann, wenn es zu je zwei verschiedenen Knoten $m, n \in N$ stets einen Pfad von m nach n gibt.

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. ∎

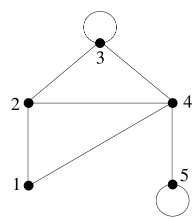
Als Übungsaufgabe (vgl. Aufgabe 7.7). werden wir zeigen, daß ein symmetrischer Graph $\mathcal{G}$ zusammenhängend ist genau dann, wenn die Quotientenmenge N/E^* genau ein Element hat. Die Äquivalenzklassen von E^* werden auch die *Zusammenhangskomponenten* des Graphen $\mathcal{G}$ genannt. Die Zusammenhangskomponenten 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.

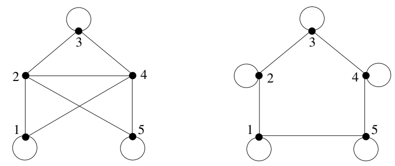
Bei der üblichen bildlichen Darstellung eines symmetrischen Graphen entspricht dem Grad eines Knotens damit die Zahl der Linien, die von dem Knoten ausgehen. Im nachfolgenden Beispiel haben die Knoten 2 und 5 den Grad 3, die Knoten 3 und 4 haben den Grad 4, der Knoten 1 hat den Grad 2.



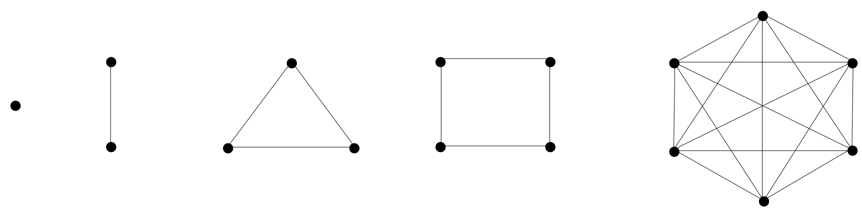
Da jede ungerichtete Kante eines symmetrischen Graphen entweder bei einem Knoten zwei Beiträge zum Grad oder bei zwei Knoten genau einen Beitrag zum Grad liefert, gilt das folgende Lemma.

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.*

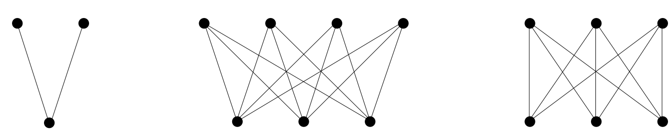
Beispiel 7.23 Ein symmetrischer Graph $\mathcal{G}$ heißt *regulär* genau dann, wenn jeder Knoten von $\mathcal{G}$ denselben Grad hat. Die nachfolgende Abbildung stellt zwei reguläre Graphen dar, wo jeder Knoten den Grad 4 besitzt.



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

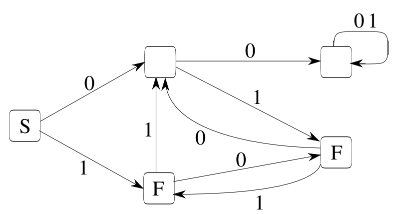
Am Ende dieses Teilkapitels wollen wir noch kurz auf einige Erweiterungen des Grundkonzepts eines Graphen eingehen. Häufig werden etwa *gelabel-*

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 $e_1, \dots, e_n$ durch den Graphen gibt, der beim Startzustand beginnt und bei einem Finalzustand endet, so daß w genau die zum Passieren von $e_1, \dots, e_n$ benötigte Ticketfolge ist. Beim obigen Beispielautomaten gehören etwa die Wörter 1, 111, 10, 101 zur charakterisierten Sprache L, alle mit dem Präfix 00 beginnenden

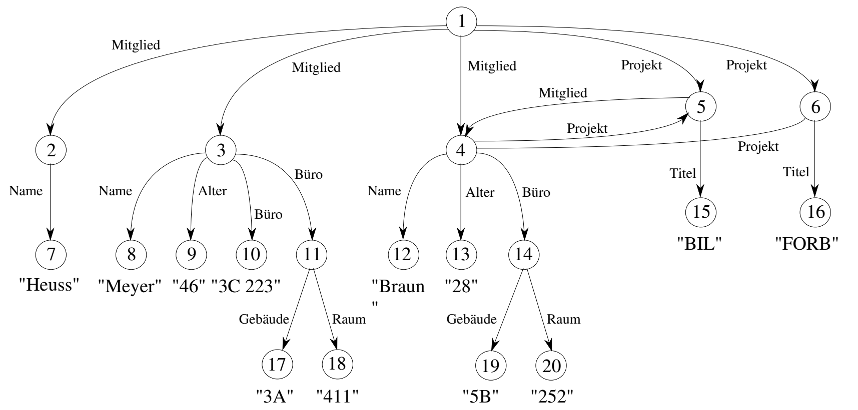


Abbildung 7.1: Datenbank als gelabelter Graph.

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

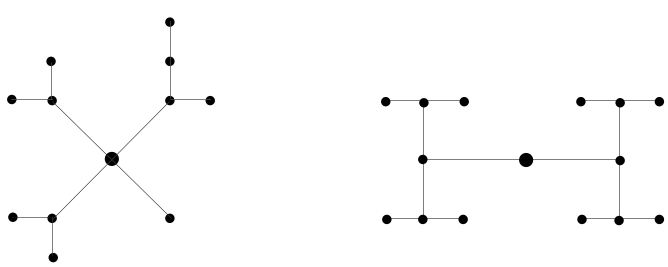
Ungeordnete Bäume mit Wurzeln kann man sowohl als symmetrische Graphen als auch als gerichtete Graphen formalisieren. Wir stellen zunächst die erste Alternative dar.

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.

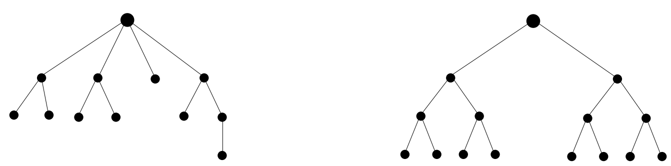
Die nachfolgende Abbildung zeigt zwei ungeordnete Bäume. Die Wurzel ist jeweils hervorgehoben.



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:

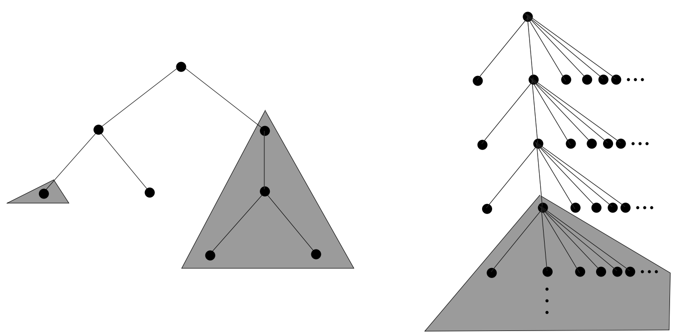


Abbildung 7.2: Drei Teilbäume.

1. $n_1 \in N$,
2. $N_1 = \{n_1\} \cup \{n \in N \mid n \text{ ist ein Nachfolger von } n_1 \text{ in } \mathcal{T}\}$,
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.

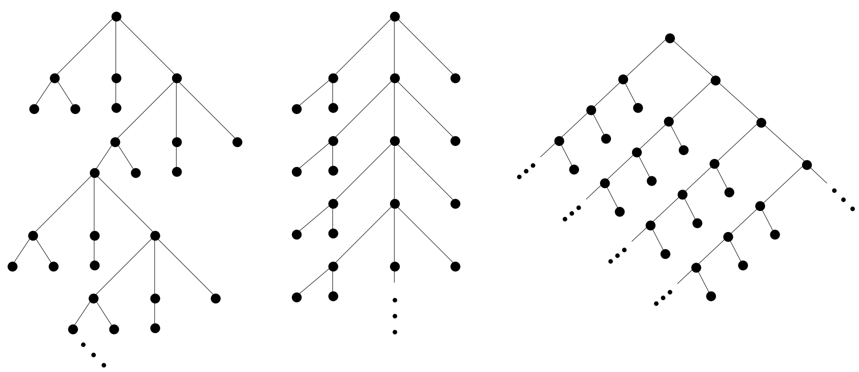


Abbildung 7.3: Unendliche rationale Bäume.

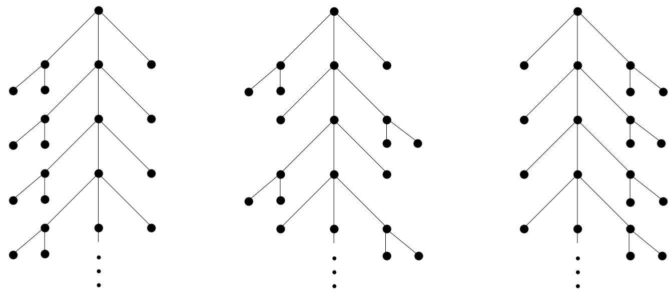
Abbildung 7.3 zeigt drei unendliche rationale Bäume. Ein sehr häufig verwendetes kombinatorisches Prinzip ist das nachfolgende.

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 $n_k := ende(e_k)$ 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.

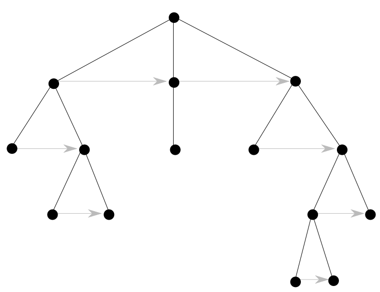


In vielen Zusammenhängen ist es indes durchaus erwünscht, tatsächlich eine Reihenfolge zwischen den verschiedenen Kindern eines Knotens explizit festzulegen. Dies führt auf den Begriff des geordneten Baums.

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.

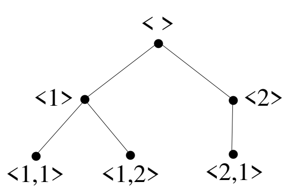
Definition 7.36 Ein *Baumskelett* ist eine endliche Menge $S \neq \emptyset$ endlicher Folgen von positiven natürlichen Zahlen mit den folgenden Eigenschaften:

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:

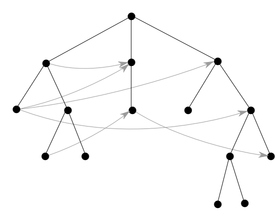


Umgekehrt kann man zu jedem endlichen geordneten Baum $\mathcal{T}$ sofort ein Baumskelett S angeben, so daß $\mathcal{T}$ zu $\mathcal{T}(S)$ isomorph ist.

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.

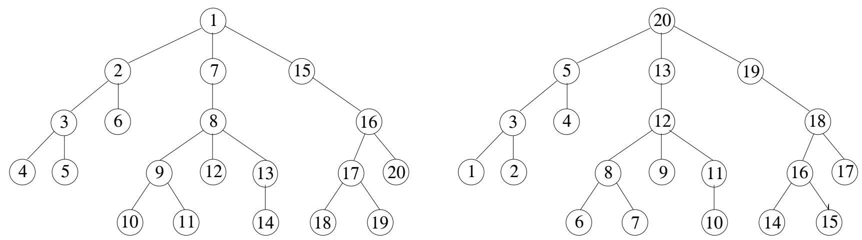


Abbildung 7.4: Präordnung und Postordnung.

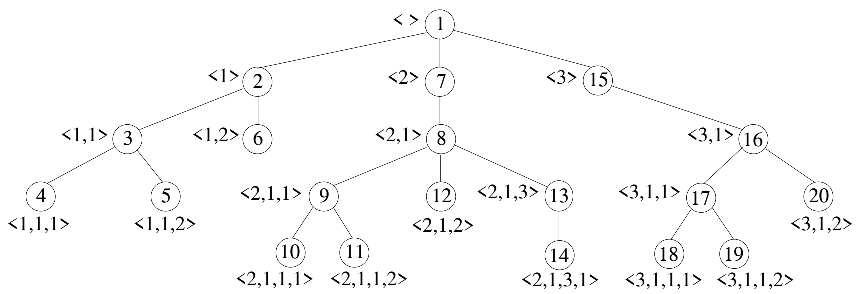


Abbildung 7.5: Präordnung als lexikographische Ordnung des Baumskeletts.

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.*

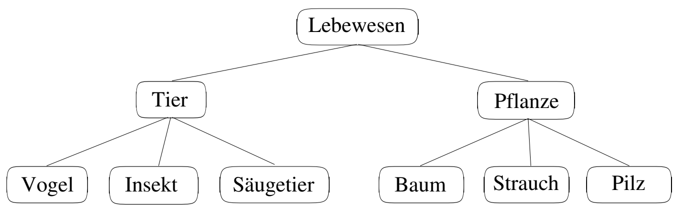
1. *Es gilt $m <_{lr} n$ genau dann, wenn $m <_p n$ und wenn n kein Nachfolger von m ist.*

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

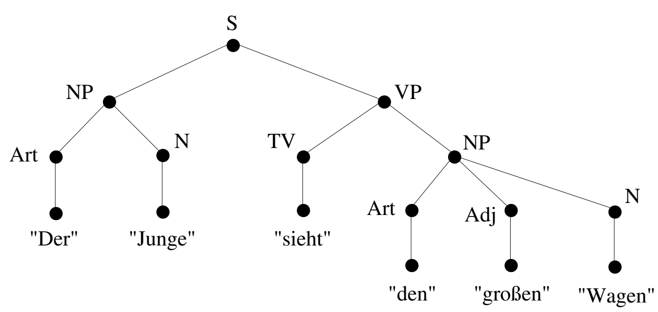
Wie bei allgemeinen Graphen werden auch bei ungeordneten oder geordneten Bäumen häufig Kanten oder Knoten gelabelt.

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 $D(\eta) \in \Sigma_{\mathcal{E}}$.

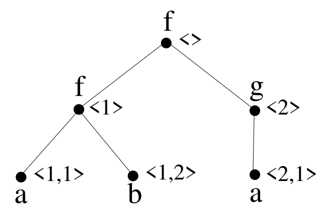
Mit $B(\Sigma)$ bezeichnen wir die Menge aller Bäume über der Signatur Σ.

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

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

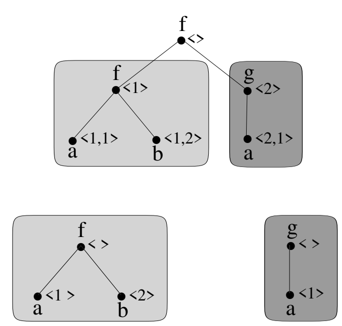
Dann ist $T := \langle S, D \rangle$ ein Baum in $B(\Sigma)$. Hier eine graphische Repräsentation.



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 $T_1, \dots, T_n$ seien Σ-Bäume. Dann gibt es einen eindeutig bestimmten Baum $T \in B(\Sigma)$, dessen Wurzelknoten mit f dekoriert ist, und wo $T_1, \dots, T_n$ die unmittelbaren Teilbäume der Positionen $\langle 1 \rangle, \dots, \langle n \rangle$ sind.*

Beweis. Wir beweisen zunächst die Existenz. Es habe T_i die Form $\langle S_i, D_i \rangle$, für $i = 1, \dots, n$. 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 $i = 1, \dots n$, so erhalten wir eine Dekoration von S über Σ. Es ist einfach nachzurechnen, daß $T := \langle S, D \rangle$ die oben genannten Bedingungen erfüllt.

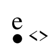
Der Nachweis der Eindeutigkeit bleibt dem Leser überlassen. Man zeigt zunächst, daß das Baumskelett eindeutig bestimmt ist, danach die Eindeutigkeit der Dekoration. ■

Offenkundig gilt auch der folgende einfache Sachverhalt.

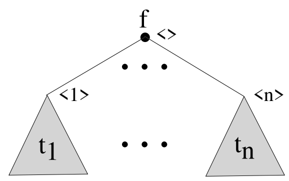
Lemma 7.48 *Jeder Baum $T \in B(\Sigma)$ hat eine eindeutig bestimmte Wurzeldekoration. Diese ist entweder eine Individuenkonstante, wobei T dann Höhe 0 hat, oder sie ist ein Funktionssymbol $f \in \Sigma$ der Stelligkeit $n \geq 1$. In letzterem Fall hat T eine Höhe $k > 0$, die unmittelbaren Teilbäume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

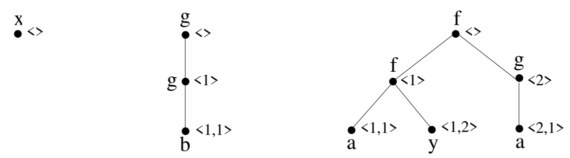
Beispiel 7.55 Die Signatur Σ enthalte ein zweistelliges Funktionssymbol f, ein einstelliges Funktionssymbol g und die Individuenkonstante a und b. Weiter seien x und y Variablen aus X. Dann sind x, a, $g(g(b))$ und $f(f(a,y), g(a))$ Terme über Σ und X. Es gilt $w(x) = w(a) = 0$ und $w(f(f(a,y), g(a))) = 3$. 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 $T_1, \ldots, T_n$, so ist $\mathit{fl}(T) := f(\mathit{fl}(T_1), \ldots, \mathit{fl}(T_n))$.

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 $f(t_1, \ldots, t_n)$, 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 $\mathit{tree}(t_1), \ldots, \mathit{tree}(t_n)$ 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

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

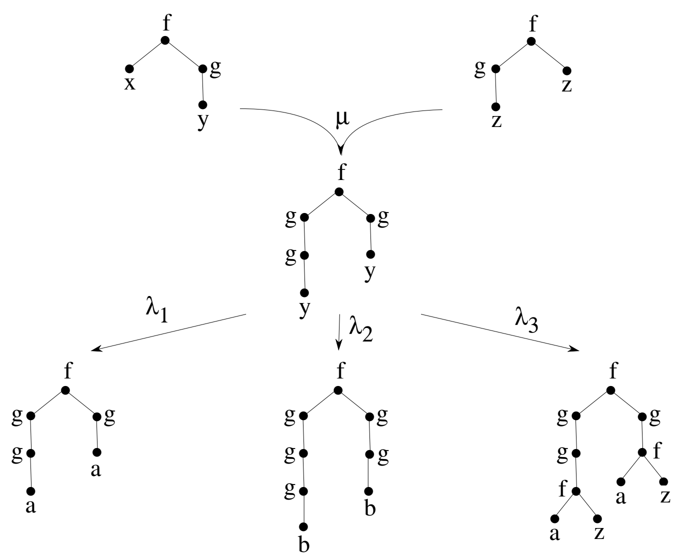
Alle drei genannten Substitutionen sind jedoch keine allgemeinsten Unifikatoren von t_1 und t_2. Ein allgemeinster Unifikator ist zum Beispiel

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

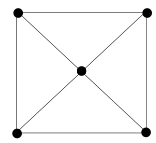
Bei endlichen zusammenhängenden Graphen stellt sich das bekannte Problem, ob es möglich ist, alle Kanten in einem Pfad zusammenzufassen, ohne dabei eine Kante zweimal zu durchlaufen.

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.

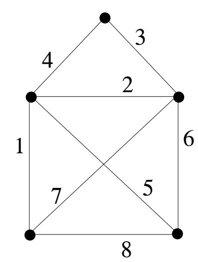
Die Definition beinhaltet insbesondere, daß in einem Eulerschen Pfad nie zwei Kanten $\langle n, m\rangle$ und $\langle m, n\rangle$ auftreten. Es wird damit jede ungerichtete Kante genau einmal und jeweils nur in einer Richtung durchlaufen. Bei

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

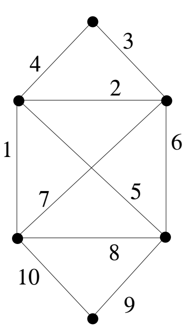
einem Eulerschen Kreis muß man am Ende wieder am Anfangspunkt ankommen. Man kann nicht immer einen Eulerschen Pfad finden, wie das folgende Beispiel zeigt.



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.

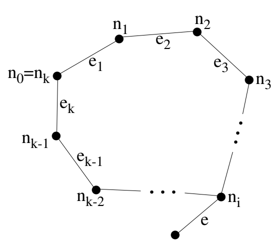
Lemma 7.72 (Satz von Euler) *Es sein $\mathcal{G}$ ein endlicher zusammenhängender Graph. Wenn in $\mathcal{G}$ jeder Knoten geraden Grad hat, so besitzt $\mathcal{G}$ einen Eulerschen Kreis.*

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 $n_{i-1} := start(e_i)$ und $n_i := ende(e_i)$. 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 $n_i := start(e)$ 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 $n_i := start(e)$ 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).

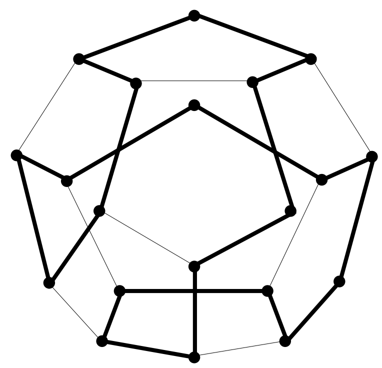
Neben den Eulerschen Kreisen, wo das Traversieren aller Kanten die Hauptbedingung darstellt, gibt es das verwandte Problem, in einem Graphen alle Punkte zu durchlaufen, ohne einen Knoten zweimal zu besuchen.

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.

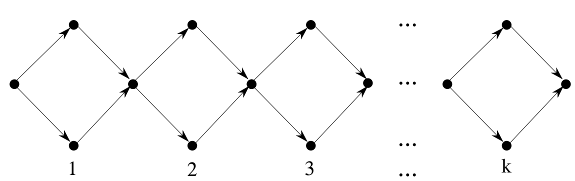


7.6 Aufgaben zu Kapitel 7

Aufgaben zu Teilkapitel 7.1

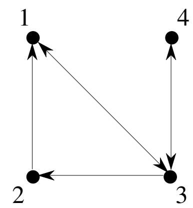
Aufgabe 7.1 Zeigen Sie, daß die Abbildungen 2 und 3 in Bemerkung 7.2 isomorphe gerichtete Graphen repräsentieren. Benennen Sie hierzu die Knoten und geben Sie dann einen Isomorphismus an.

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?

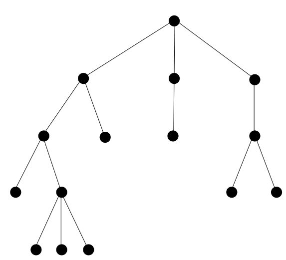
Aufgabe 7.10 Es sei $\mathcal{T}$ ein endlicher ungeordneter Baum mit m Knoten und n Kanten. Beweisen Sie, daß $n = m - 1$ gilt.

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.

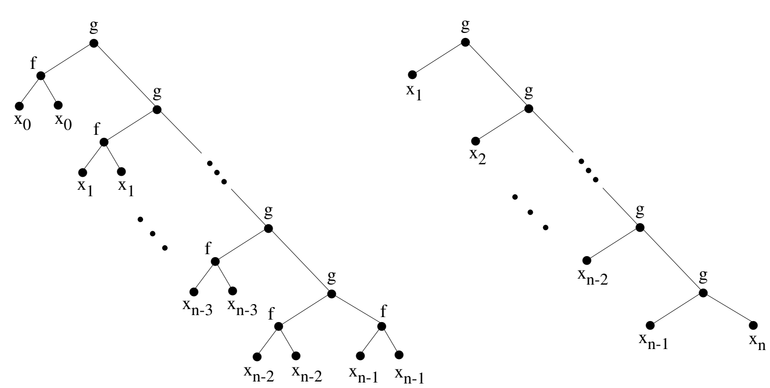
Aufgaben zu Teilkapitel 7.5.1

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.

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

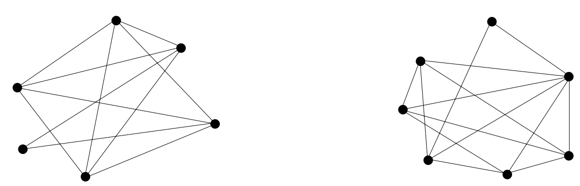
Aufgabe 7.28 Unifizieren Sie nachfolgend in Baumform dargestellten Terme. Es sei μ ein allgemeinster Unifikator.



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.

8

Verbände

In diesem Kapitel wollen wir nun eine Klasse von Ordnungsstrukturen betrachten, die uns wieder näher an ein zu Beginn angesprochenes Thema—die Beziehung zwischen Aussagen, Junktoren und Mengenoperationen—heranführt. Viele der Gesetze, die wir beim Rechnen mit aussagenlogischen Junktoren und mit Mengenoperationen beobachtet haben, gelten in ähnlicher Form auch beim Rechnen in den nun zu besprechenden Verbänden. Im nächsten Kapitel werden uns dann die Booleschen Algebren—eine spezielle Klasse von Verbänden—endgültig zu den Aussagen und Mengen zurückführen. Verbände treten in der Mathematik an vielen Stellen auf. Aber auch in der Informatik, in Bereichen der Wissensverarbeitung in der formalen Linguistik spielen Verbände (oder Halbverbände) eine wichtige Rolle.

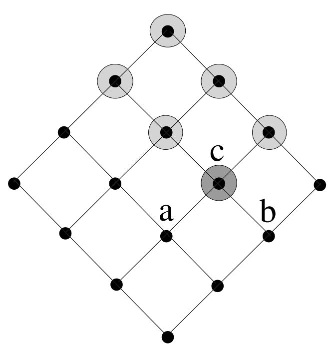
In Abschnitt 8.1 definieren wir Verbände als Ordnungsstrukturen und illustrieren das Konzept anhand einer Reihe von Beispielen. In Kapitel 8.3 zeigen wir, daß sich Verbände auch als Algebren darstellen lassen, und geben Übersetzungen zwischen beiden Betrachtungsweisen an. In Abschnitt 8.4 zeigen wir, daß die Begriffe des Verbandshomomorphismus und des Teilverbands unterschiedlich sind, wenn wir einerseits ordnungstheoretische Verbände, andererseits algebraische Verbände zur Grundlage nehmen. Die in Abschnitt 8.5 dargestellten distributiven Verbände dienen der Vorbereitung auf das nachfolgende Kapitel über Boolesche Algebren. Diese stellen eine Teilklasse der distributiven Verbände dar. Die in Abschnitt 8.6 eingeführten Ideale und Filter sind spezielle Teilverbände, die im Zusammenhang mit Quotientenbildung bei Booleschen Algebren relevant werden. Im ergänzenden Kapitel 8.7 stellen wir kurz das Konzept des Halbverbands dar und gehen auf den Fixpunktsatz von Tarski und Knaster ein.

8.1 Beispiele von Verbänden

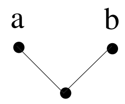
Bevor wir eine Definition von Verbänden als Ordnungsstrukturen geben können, benötigen wir einige einfache Begriffe.

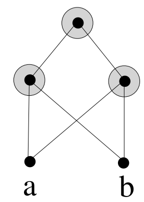
Definition 8.1 Es sei $\langle M, \leq\rangle$ eine partiell geordnete Menge und $X \subseteq M$. Das Element $a \in M$ heißt *obere* (*untere*) *Schranke von* X genau dann, wenn gilt: $\forall x \in X$: $x \leq a$ $(a \leq x)$. Das Element a heißt *Supremum* oder oder *kleinste obere Schranke* von X genau dann, wenn a obere Schranke von X ist und wenn $a \leq b$ für jede obere Schranke b von X gilt. Wir schreiben $a = sup\ X$. Dual heißt a *Infimum* oder *größte untere Schranke* von X genau dann, wenn a untere Schranke von X ist und wenn $b \leq a$ für jede untere Schranke von X gilt. Wir schreiben $a = inf\ X$.

Beispiele 8.2 In der nachfolgend in der Form eines Hasse-Diagramms abgebildeten Ordnungsstruktur sind die oberen Schranken der Menge $\{a, b\}$ hervorgehoben, wobei c das Supremum von $\{a, b\}$ ist.



Die Elemente a und b der nachfolgend links abgebildeten Ordnungsstruktur zeigen, daß es im allgemeinen keine obere Schranken für eine vorgegebene Teilmenge X einer Ordnungsstruktur gibt.





Wie das Beispiel auf der rechten Seite klar macht, braucht es keine *kleinste* obere Schranke geben, selbst wenn obere Schranken existieren.

Wenn es in einer partiell geordneten Menge $\langle M, \leq\rangle$ ein Supremum über die Teilmenge X gibt, so ist dieses eindeutig bestimmt. Falls nämlich $a_1 = \mathit{sup}\, X$ und $a_2 = \mathit{sup}\, X$, so folgt $a_1 \leq a_2$ und $a_2 \leq a_1$, da ja a_1 und a_2 insbesondere obere Schranken sind. Wegen der Antisymmetrie folgt nun $a_1 = a_2$. Symmetrisch folgt die Eindeutigkeit von $\mathit{inf}\, X$, falls $\mathit{inf}\, X$ existiert. Man beachte, daß obere (untere) Schranken und Suprema (Infima) einer Teilmenge $X \subseteq M$ nicht selbst in X liegen zu brauchen.

Bemerkung 8.3 Falls $a = \mathit{sup}\, \emptyset$ existiert, so ist a das kleinste Element von M, da jedes $b \in M$ obere Schranke von $\emptyset$ ist. In diesem Fall wird $\mathit{sup}\, \emptyset$ auch oft als "0" notiert. Falls $\mathit{inf}\, \emptyset$ existiert, so ist $\mathit{inf}\, \emptyset$ das größte Element von M und wird als"1" notiert.

Wir kommen nun zur zentralen Definition dieses Kapitels.

Definition 8.4 Eine partiell geordnete Menge $\mathcal{V} = \langle V, \leq\rangle$ heißt *Verband* genau dann, wenn $\mathit{sup}\, \{a, b\}$ und $\mathit{inf}\, \{a, b\}$ für alle $a, b \in V$ existieren. Wenn weitergehend $\mathit{sup}\, M$ und $\mathit{inf}\, M$ für jede Teilmenge M von V existieren, so wird $\mathcal{V}$ ein *vollständiger* Verband genannt.

Auch wenn gemäß Definition 8.4 in einem Verband nur die Existenz von Suprema und Infima über *zweielementige* Teilmengen gefordert wird, so gilt eine entsprechende Eigenschaft dann doch automatisch für alle *endlichen* Teilmengen.

Lemma 8.5 *Ist $\mathcal{V} = \langle V, \leq\rangle$ ist ein Verband, so existieren für je endlich viele Elemente $a_1, \ldots, a_n \in V$ $(n > 0)$ stets $\mathit{sup}\{a_1, \ldots, a_n\}$ und $\mathit{inf}\, \{a_1, \ldots, a_n\}$.*

Ist $n > 1$, so gilt stets

$$\begin{aligned} sup\{a_1, \ldots, a_n\} &= sup\ \{sup\{a_1, \ldots, a_{n-1}\}, a_n\}, \\ inf\,\{a_1, \ldots, a_n\} &= inf\,\{inf\,\{a_1, \ldots, a_{n-1}\}, a_n\}. \end{aligned}$$

Beweis. Durch vollständige Induktion (Übung). ■

Insbesondere folgt aus Lemma 8.5, daß jeder endliche Verband vollständig ist und eine 1 und eine 0 besitzt. Unendliche Verbände hingegen sind nicht notwendigerweise vollständig. Als ein Gegenbeispiel betrachte man etwa $\langle \mathbb{N}, \leq \rangle$ (wieso?). Das nachfolgende Kriterium ist manchmal hilfreich, um nachzuweisen, daß eine gegebene partiell geordnete Menge ein vollständiger Verband ist.

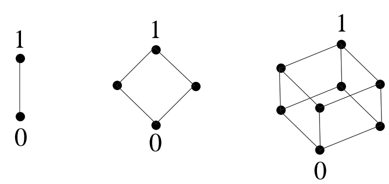
Lemma 8.6 *Es sei $\mathcal{V} = \langle V, \leq \rangle$ eine partiell geordnete Menge mit einem größten Element 1. Falls für jede nichtleere Teilmenge X von V stets $inf\,X$ existiert, so ist $\mathcal{V}$ ein vollständiger Verband.*

Beweis. Da $1 = inf\,\emptyset$, existiert gemäß der Voraussetzung für jede Teilmenge $X \subseteq V$ das Infimum. Es reicht daher nachzuweisen, daß auch für jede Teilmenge Y von V stets $sup\ Y$ existiert.

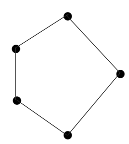
Sei also $Y \subseteq V$. Die Menge S aller oberen Schranken von Y ist nichtleer, da 1 eine obere Schranke von Y ist. Nach Vorausetzung existiert damit $s := inf\,S$. Wir zeigen, daß s selbst eine obere Schranke für Y ist. Damit gilt offenkundig $s = sup\ Y$.

Ist $y \in Y$, so ist offenkundig y eine untere Schranke für S. Da s die größte untere Schranke für s ist, folgt $y \leq s$. Da dies für jedes $y \in Y$ gilt, ist s eine obere Schranke für Y. ■

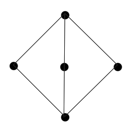
Beispiele 8.7 Die nachfolgende Abbildung zeigt einige endliche Verbände, die wir in Form von Hasse-Diagrammen beschreiben.



In diesen drei Beispielen können die Elemente der Verbände auf natürliche Weise mit einem Rang versehen werden können: wir können jedem Element x eine natürliche Zahl $Rang(x)$ zuordnen derart, daß stets $Rang(x) = Rang(y) + 1$ gilt, falls y unmittelbarer Vorgänger von x ist. Hier noch zwei weitere endliche Verbände, „Pentagon" und „Diamant". Das Pentagon zeigt, daß es auch bei endlichen Verbänden nicht immer möglich ist, eine Rangfunktion einzuführen.



Pentagon



Diamant

Bevor wir auf Beispiele für Verbände innerhalb der Mathematik eingehen, beschreibt das nächste Beispiel eine Anwendung in der Wissensverarbeitung.

Beispiel 8.8 Verbände werden in der sogenannten formalen Konzeptanalyse verwendet, um Wissensbereiche begrifflich zu strukturieren und den gegenseitigen Bezug von Konzepten graphisch zu verdeutlichen. Um diese Rolle zu erläutern, stellen wir uns als Ausgangspunkt eine Tabelle der Form vor,

	a	b	c	d	e	f
1	×			×	×	×
2	×	×	×	×		×
3	×		×			×
4	×	×	×			
5		×			×	

wo $O = \{1, 2, 3, 4, 5\}$ eine Menge von Objekten repräsentiert und $A = \{a, b, c, d, e, f\}$ eine Menge von Attributen. Beispielsweise könnte O eine Menge von Tiergattungen sein und A eine Menge möglicher genetischer Eigenschaften. Ein Kreuz „$\times$" in Zeile i und Spalte j bedeutet, daß Objekt i das Attribut (Merkmal) j besitzt, wir schreiben i *hat* j. für eine Menge $U \subseteq O$ sei

$$Attr(U) := \{m \in A \mid i \; hat \; m \text{ für alle Objekte } i \in U\}$$

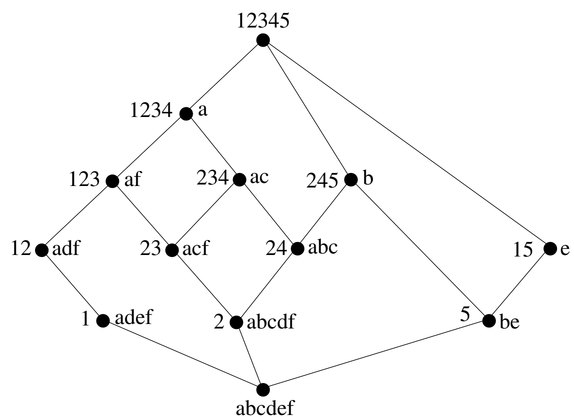


Abbildung 8.1: Konzeptverband für Beispiel 8.8.

die Menge aller Attribute, die alle Objekte in U besitzen. Umgekehrt sei für eine Menge $B \subseteq A$

$$Obj(B) := \{i \in O \mid i \ hat \ m \text{ für alle Attribute } m \in B\}$$

die Menge aller Objekte, die alle Eigenschaften aus B besitzen. Ein *formales Konzept* ist ein Paar (U, B) wo $U = Obj(B)$ und $B = Attr(U)$. Beispielsweise sind $(\{1,2\}, \{a,d,f\})$ und $(\{2,4,5\}, \{b\})$ formale Konzepte. Intuitiv faßt ein formales Konzept eine Menge von Objekten zusammen, die sich durch eine Menge gemeinsamer Attribute charakterisieren läßt. Alle Objekte, auf die diese Attribute zutreffen, müssen in der entsprechenden Objektmenge mitaufgenommen sein. Zum Beispiel tritt die Menge $\{2,5\}$ nicht als Konzept auf, weil das einzige gemeinsame Attribut b auch auf Objekt 5 zutrifft.

Definieren wir nun eine Ordnung zwischen Konzepten vermöge

$$(U_1, B_1) \leq (U_2, B_2) :\Leftrightarrow U_1 \subseteq U_2,$$

so erhalten wir dadurch eine Verbandsstruktur auf der Menge aller Konzepte. Wir verzichten auf den Beweis. Der für das Beispiel resultierende Konzeptverband ist in Abbildung 8.1 dargestellt. Er ist automatisch berechenbar und kann zur Strukturierung und Analyse des durch die Ausgangstabelle erfaßten thematischen Gebiets verwendet werden.

Beispiele 8.9 Die nachfolgende Liste enthält nur einige der unzähligen Beispiele, wo Verbände in der Mathematik auftreten.

1. Jede linear geordnete Menge $\langle M, \leq \rangle$ ist ein Verband.

2. Die Menge der natürlichen Zahlen $\mathbb{N}$ mit der Teilbarkeitsrelation „|“ bildet einen Verband. Bezeichnet nämlich $ggT(x,y)$ den größten gemeinsamen Teiler von x und y, und $kgV(x,y)$ deren kleinstes gemeinsames Vielfaches, so gilt für beliebige natürliche Zahlen a, b, c: aus $c|a$ und $c|b$ folgt $c|ggT(a,b)$, aus $a|c$ und $b|c$ folgt $kgV(a,b)|c$. Damit ist $\mathit{inf}\,\{a,b\} = ggT(a,b)$ und $\mathit{sup}\ \{a,b\} = kgV(a,b)$.

3. Ist M eine Menge, so bildet $\langle \mathcal{P}(M), \subseteq \rangle$ einen vollständigen Verband.

4. für je zwei Äquivalenzrelationen $\sim_1$ und $\sim_2$ auf einer Menge M definieren wir $\sim_1 \leq \sim_2$ genau dann, wenn „$\sim_1$“ feiner ist als „$\sim_2$“ im Sinn von Definition 4.15. Es sei $\ddot{A}q(M)$ die Menge aller Äquivalenzrelationen auf M. Dann ist $\langle \ddot{A}q(M), \leq \rangle$ ein vollständiger Verband.

5. Es sei A eine beliebige Menge und $B \subseteq \mathbb{N}$. für $f, g \in B^A$ sei $f \preceq g$ gdw. $f(a) \leq g(a)$ für alle $a \in A$. Dann ist $\langle B^A, \preceq \rangle$ ein Verband.

Der Nachweis, daß die angegebenen Strukturen Verbände sind, ist in der Regel einfach und kann als Übung überlassen werden. Das unter 4 genannte Beispiel besitzt eine interessante Verallgemeinerung.

Lemma 8.10 *Es sei M eine Menge und $H : 2^M \to 2^M$ eine Hüllenabbildung im Sinn von Definition 4.35. Dann bildet die Menge Q aller Fixpunkte von H bezüglich der Inklusion „$\subseteq$“ einen vollständigen Verband.*

Beweis. Da aufgrund der Erweiterungseigenschaft (vgl. Def. 4.35 (i)) stets M selbst Fixpunkt von H ist, gilt $Q \neq \emptyset$. Offenkundig stellt „$\subseteq$“ eine partielle Ordnung auf Q dar und $1 := M$ ist das größte Element. Es sei nun $X = \{A_i \mid i \in I\}$ eine nichtleere Teilmenge von Q. Gemäß Lemma 4.37 ist auch $\bigcap X$ wieder ein Fixpunkt von H und damit in Q. Offenkundig ist dann $\bigcap X$ das Infimum von X bezüglich der Inklusion. Nach Lemma 8.6 ist $\langle Q, \subseteq \rangle$ ein vollständiger Verband. ■

Aus Lemma 8.10 ergibt sich beispielsweise, daß die Menge aller reflexiven (symmetrischen, transitiven) Relationen auf einer nichtleeren Menge M stets einen vollständigen Verband bezüglich der Inklusion „$\subseteq$“ bilden. Hierzu vergleiche man Definition 4.38 und Aufgabe 4.25.

8.2 Das Dualitätsprinzip

Im Kontext der Verbandstheorie ist es interessant, auf eine Beobachtung aus Beispiel 4.28 Nr. 6 zurückzukommen: ist „$\leq$“ eine partielle Ordnung auf einer Menge M, so auch ist auch die duale Relation „$\geq$“ eine partielle Ordnung ist. Wir nennen $\langle M, \geq \rangle$ die zu $\langle M, \leq \rangle$ *duale* Ordnungsstruktur. Beim Übergang von einer Ordnungsstruktur zur dualen Struktur vertauschen sich offenkundig genau die Rollen von oberen und unteren Schranken und die Rollen von Suprema und Infima, wie folgende Illustration deutlich macht.

Dualität bei Ordnungsstrukturen

Ist $\mathcal{V} = \langle V, \leq \rangle$ ein Verband, so ist daher auch $\mathcal{V}' := \langle V, \geq \rangle$ ein Verband. $\mathcal{V}'$ heißt der zu $\mathcal{V}$ *duale Verband.*

Zu vielen Aussagen, die wir nachfolgend über partiell geordnete Mengen oder Verbände machen, existiert eine duale Version, die wir dadurch erhalten, daß wir „$\leq$“ und „$\geq$“, Supremum und Infimum, sowie gegebenenfalls „0“ und „1“ miteinander vertauschen. Als Beispiel einer solchen Dualisierung erhalten wir aus Lemma 8.6 die folgende duale Version.

Lemma 8.11 *Es sei $\mathcal{V} = \langle V, \leq \rangle$ eine partiell geordnete Menge mit einem kleinsten Element* 0*. Falls für jede nichtleere Teilmenge X von V stets* sup X *existiert, so ist $\mathcal{V}$ ein vollständiger Verband.*

Ein Beweis ergibt sich durch Dualisierung des Beweises von Lemma 8.6, ist aber mit der folgenden allgemeinen Beobachtung sogar überflüssig.

Bemerkung 8.12 [Dualitätsprinzip für Verbände] Wenn eine gegebene Aussage in allen partiell geordneten Mengen (resp. Verbänden) gilt, so gilt offenkundig die duale Aussage in allen dualen partiell geordneten Mengen (resp. Verbänden). Da aber jede partiell geordnete Menge (jeder Verband) durch Dualisierung aus einer entsprechenden Struktur zu erhalten ist, kann man sich stets darauf beschränken, die Gültigkeit einer der beiden Versionen von zwei dualen Aussagen nachzuweisen, die duale Version gilt dann automatisch.

8.3 Verbände als Ordnungsstrukturen und als Algebren

Wie in der Einleitung angedeutet wurde, können Verbände auch rein algebraisch definiert werden. Um den Zusammenhang zwischen beiden Betrachtungsweisen darzustellen, leiten wir nachfolgend ausgehend von der Beschreibung eines Verbands als Ordnungsstruktur zwei binäre Operationen „join" und „meet" ab, für die wir vier charakteristische Eigenschaften festhalten. Aus diesen Eigenschaften ergibt sich die Charakterisierung von Verbänden als Algebren. Wir zeigen dann, daß eine Übersetzung auch in umgekehrter Weise möglich ist: man kann auf einem algebraischen Verband ausgehend von den Operationen „join" und „meet" sofort eine partielle Ordnung definieren, die die Struktur zu einem ordnungstheoretischen Verband macht.

Wir führen nun die Abkürzungen

$$\begin{aligned} a \sqcup b &= \mathit{sup}\,\{a, b\} \\ a \sqcap b &= \mathit{inf}\,\{a, b\} \end{aligned}$$

ein und nennen „$\sqcup$" das *join* und „$\sqcap$" das *meet*. In einem Verband sind also „join" und „meet" binäre Operationen. Das folgenden Lemma haben wir in einem Spezialfall (vgl. Lemma 2.19 und Beispiel 8.9 Nr. 3) bereits kennengelernt.

Lemma 8.13 *Sei $\mathcal{V} = \langle V, \leq \rangle$ ein Verband. Dann gilt:*

$$\forall a, b \in V : a \leq b \quad \Leftrightarrow \quad a \sqcap b = a \quad \Leftrightarrow \quad a \sqcup b = b.$$

Beweis. trivial. ■

für die Operationen „join" und „meet" ergeben sich die folgenden charakteristischen Eigenschaften.

Lemma 8.14 *In einem Verband $\mathcal{V} = \langle V, \leq \rangle$ sind die Operationen „meet" und „join" stets (i) idempotent, (ii) kommutativ und (iii) assoziativ, und sie erfüllen die Absorptionseigenschaft (iv), das heißt es gilt*

(i) $\forall a \in V\colon a \sqcup a = a, a \sqcap a = a,$

(ii) $\forall a, b \in V\colon a \sqcup b = b \sqcup a, a \sqcap b = b \sqcap a,$

(iii) $\forall a, b, c \in V\colon (a \sqcup b) \sqcup c = a \sqcup (b \sqcup c), (a \sqcap b) \sqcap c = a \sqcap (b \sqcap c).$

(iv) $\forall a, b \in V\colon a \sqcup (a \sqcap b) = a, a \sqcap (a \sqcup b) = a.$

Beweis. Natürlich gilt für alle $a, b \in V$ stets $\mathit{sup}\ \{a, a\} = \mathit{sup}\ \{a\} = a$ und $\mathit{sup}\ \{a, b\} = \mathit{sup}\ \{b, a\}$, wodurch sich (i) und (ii) ergibt. Nach Lemma 8.5 gilt für Elemente $a, b, c \in V$ stets

$$\begin{aligned} \mathit{sup}\ \{a, b, c\} &= \mathit{sup}\ \{\mathit{sup}\ \{a, b\}, c\} = \mathit{sup}\ \{a, \mathit{sup}\ \{b, c\}\}, \\ \mathit{inf}\ \{a, b, c\} &= \mathit{inf}\ \{\mathit{inf}\ \{a, b\}, c\} = \mathit{inf}\ \{a, \mathit{inf}\ \{b, c\}\}. \end{aligned}$$

Hieraus folgt (iii). Da für $a, b \in V$ stets $a \leq \mathit{sup}\ \{a, b\} = a \sqcup b$ und $a \sqcap b = \mathit{inf}\ \{a, b\} \leq a$ ergibt sich (iv) aus Lemma 8.13. ∎

Es ist instruktiv, sich ins Gedächtnis zu rufen, daß wir aus (i)–(iv) gültige Aussagen erhalten, wenn wir „$\sqcup$" und „$\sqcap$" durch die aussagenlogischen Junktoren „$\wedge$" und „$\vee$" ersetzen, a, b, c als Aussagen α, β, γ interpretieren und „$=$" als Äquivalenz „$\Leftrightarrow$".

Algebren $\langle V, \sqcup, \sqcap \rangle$ mit zwei binären Operationen gibt es unzählige. Welche besonderen Eigenschaften der Operationen „$\sqcup$" und „$\sqcap$" lassen uns mit Sicherheit schließen, daß es sich um das „meet" und „join" eines Verbands handelt? Wie wir gesehen haben, müssen hierzu notwendig beide Operationen idempotent, assoziativ und kommutativ sein, und sie müssen die Absorptionseigenschaft haben. Wir werden nun sehen, daß dies auch schon ausreicht. Wir kommen zur folgenden *algebraischen Definition* eines Verbands. Um beide Definitionen auseinanderhalten zu können, sprechen wir im Bedarfsfall von einem ordnungstheoretischen Verband beziehungsweise von einem algebraischen Verband.

Definition 8.15 Ein *algebraischer Verband* ist eine Algebra $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$, wo $\sqcup, \sqcap\colon V \times V \to M$ binäre Operationen auf V sind, die die Eigenschaften (ii)–(iv) aus Lemma 8.14 erfüllen.

Das nachfolgende Lemma zeigt, daß in einem Verband das „join" und das „meet" auch Eigenschaft (i) aus Lemma 8.14 erfüllen.

Lemma 8.16 *In einem algebraischen Verband $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ sind die Operationen join „$\sqcup$" und meet „$\sqcap$" idempotent.*

Beweis. Sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein Verband und $a \in V$. Dann gilt wegen der Absorptionseigenschaften $a \sqcap a = (a \sqcap (a \sqcup a)) \sqcap (a \sqcup (a \sqcap a))$. Wegen der Kommutativität und Assoziativität von „meet" können wir den rechten Ausdruck zu $(a \sqcap (a \sqcup (a \sqcap a))) \sqcap (a \sqcup a)$ umformen. Durch Anwenden der Absorptionseigenschaft, mit $b = (a \sqcap a)$, erhalten wir zunächst $a \sqcap (a \sqcup a)$ und dann a. Also gilt $a \sqcap a = a$. Der Nachweis der Idempotenz des „join" geht analog. ■

Theorem 8.17 *Algebraische und ordnungstheoretische Verbände entsprechen sich im folgenden Sinn:*

1. *Es sei $\mathcal{V} = \langle V, \leq \rangle$ ein ordnungstheoretischer Verband, die Operationen join „$\sqcup$" und meet „$\sqcap$" seien durch $a \sqcup b := \sup\{a, b\}$ und $a \sqcap b := \inf\{a, b\}$ erklärt. Dann ist $\mathcal{V}^* := \langle V, \sqcup, \sqcap \rangle$ ein algebraischer Verband.*
2. *Es sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein algebraischer Verband. Definiert man $a \leq b$ gdw. $a \sqcap b = a$, so ist $\mathcal{V}^{\leq} := \langle V, \leq \rangle$ ein ordnungstheoretischer Verband.*
3. *Sei $\mathcal{V} = \langle V, \leq \rangle$ ein ordnungstheoretischer Verband. Dann gilt $(\mathcal{V}^*)^{\leq} = \mathcal{V}$.*
4. *Sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein algebraischer Verband. Dann gilt $(\mathcal{V}^{\leq})^* = \mathcal{V}$.*

Beweis. Teil 1 des Theorems wurde bereits in Lemma 8.14 gezeigt.

2. Wir zeigen zunächst, daß „$\leq$" eine partielle Ordnung ist. Nachfolgend seien a, b, c beliebige Elemente aus V.

Reflexivität. Da „$\sqcap$" idempotent ist (Lemma 8.16), ist „$\leq$" reflexiv.

Antisymmetrie. Falls $a \leq b$ und $b \leq a$, so gilt $a \sqcap b = a$ und $b \sqcap a = b$. Da „$\sqcap$" kommutativ ist, folgt

$$a = a \sqcap b = b \sqcap a = b,$$

somit ist „$\leq$" antisymmetrisch.

Transitivität. Falls $a \leq b$ und $b \leq c$ gilt, so folgt $a \sqcap b = a$ und $b \sqcap c = b$. Unter Verwendung der Assoziativität erhalten wir daraus

$$a = a \sqcap b = a \sqcap (b \sqcap c) = (a \sqcap b) \sqcap c = a \sqcap c$$

woraus sich $a \leq c$ ergibt. Somit ist „$\leq$" auch transitiv, also eine partielle Ordnung.

Um weiter zu zeigen, daß $\langle M, \leq \rangle$ ein Verband ist, beweisen wir, daß $a \sqcap b = \mathit{inf}\,\{a,b\}$ und $a \sqcup b = \mathit{sup}\,\{a,b\}$. Sicher ist $a \sqcap b$ eine untere Schranke für $\{a,b\}$, denn wegen der Eigenschaften (i)–(iii) aus Definition 8.15 gilt

$$(a \sqcap b) \sqcap a = a \sqcap (b \sqcap a) = a \sqcap (a \sqcap b) = (a \sqcap a) \sqcap b = a \sqcap b,$$

also $a \sqcap b \leq a$, genauso folgt $a \sqcap b \leq b$. Ist nun c eine weitere untere Schranke für $\{a,b\}$, so gilt $c \leq a, c \leq b$, in anderen Worten $c \sqcap a = c$ und $c \sqcap b = c$, daher

$$c \sqcap (a \sqcap b)) = (c \sqcap a) \sqcap b = c \sqcap b = c,$$

daher gilt $a \sqcap b = \mathit{inf}\,\{a,b\}$. Der Beweis für $a \sqcup b = \mathit{sup}\,\{a,b\}$, der eine zusätzliche kleine Hilfsüberlegung benötigt, wird als Übung dem Leser überlassen (vgl. Aufgabe 8.4).

3. Dies folgt sofort daraus, daß gemäß Lemma 8.13 die partiellen Ordnungen von $\mathcal{V}$ und von $(\mathcal{V}^*)^{\leq}$ übereinstimmen.

4. In Teil 2 hatten wir gesehen, daß stets $a \sqcup b$ (resp. $a \sqcap b$) das Supremum (resp. Infimum) von $\{a,b\}$ bezüglich der definierten Ordnung „$\leq$" von $\mathcal{V}^{\leq}$ ist $(a, b \in V)$. Daraus ergibt sich die Behauptung. ■

8.4 Verbandshomomorphismen und Teilverbände

Bei der Diskussion von Homomorphismen in Kapitel 6.4 hatten wir darauf hingewiesen, daß der Begriff stark abhängig von der gewählten Signatur ist. Da wir nun für Verbände zwei stark unterschiedliche Beschreibungsformen haben, stellt sich die Frage, ob ein Homomorphismus zwischen zwei algebraischen Verbänden stets auch ein Homomorphismus zwischen den dazu assoziierten ordnungstheoretischen Verbänden ist und umgekehrt. Es gilt dies allerdings nur in einer Richtung.

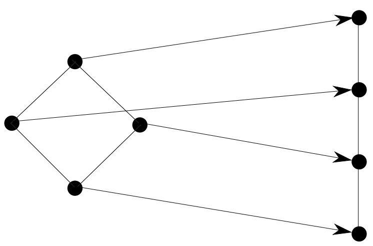


Abbildung 8.2: Isotone Abbildungen müssen nicht Suprema und Infima erhalten.

Lemma 8.18 *Es seien $\mathcal{V}_1 = \langle V_1, \sqcup_1, \sqcap_1 \rangle$ und $\mathcal{V}_2 = \langle V_2, \sqcup_2, \sqcap_2 \rangle$ algebraische Verbände. Dann ist jeder Homomorphismus von $\mathcal{V}_1$ nach $\mathcal{V}_2$ auch ein Homomorphismus zwischen den assoziierten ordnungstheoretischen Verbänden $\langle V_1, \leq_1 \rangle$ und $\langle V_1, \leq_2 \rangle$.*

Beweis. Es sei h ein Homomorphismus von $\mathcal{V}_1$ nach $\mathcal{V}_2$ und $a, b \in V_1$. Gilt $a \leq_1 b$, so folgt $a \sqcap_1 b = a$. Da h Homomorphismus ist, folgt

$$h(a) = h(a \sqcap_1 b) = h(a) \sqcap_2 h(b)$$

und somit $h(a) \leq_2 h(b)$. ■

Ein Homomorphismus h zwischen zwei partiell geordneten Mengen $\langle V_1, \leq_1 \rangle$ und $\langle V_1, \leq_2 \rangle$ wird oft einfach eine *ordnungserhaltende* oder *isotone* Abbildung genannt. Das Beispiel in Abbildung 8.2 zeigt, daß selbst eine bijektive isotone Abbildung zwischen ordnungstheoretischen Verbänden nicht notwendig Suprema und Infima erhält. Damit sind isotone Abbildungen nicht notwendig Homomorphismen zwischen den assoziierten algebraischen Verbänden. Allerdings gilt die folgende eingeschränkte Aussage.

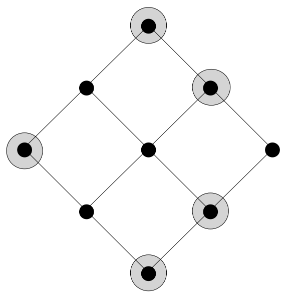
Lemma 8.19 *Es seien $\mathcal{V}_1 = \langle V_1, \leq_1 \rangle$ und $\mathcal{V}_2 = \langle V_1, \leq_2 \rangle$ ordnungstheoretische Verbände. Ist h ein Isomorphismus von $\mathcal{V}_1$ auf $\mathcal{V}_2$, so ist h auch ein Isomorphismus zwischen den assoziierten algebraischen Verbänden $\langle V_1, \sqcup_1, \sqcap_1 \rangle$ und $\langle V_2, \sqcup_2, \sqcap_2 \rangle$.*

Beweis. Es seien $a, b \in V_1$. Aus $a, b \leq_1 a \sqcup_1 b$ folgt $h(a), h(b) \leq_2 h(a \sqcup_1 b)$, da h isoton ist. Damit ist $h(a \sqcup_1 b)$ eine obere Schranke von $\{h(a), h(b)\}$ bezüglich

„$\leq_2$“. Es sei $c \in V_2$ irgendeine obere Schranke von $\{h(a), h(b)\}$ bezüglich „$\leq_2$“. Da h ein Ordnungsisomorphismus ist, ist auch h^{-1} isoton. Somit ist $h^{-1}(c)$ eine obere Schranke der Elemente $h^{-1}(h(a)) = a$ und $h^{-1}(h(b)) = b$ bezüglich „$\leq_1$“. Es folgt $a \sqcup_1 b \leq_1 h^{-1}(c)$. Da h isoton ist erhalten wir $h(a \sqcup_1 b) \leq_2 h(h^{-1}(c)) = c$. Also ist $h(a \sqcup_1 b)$ die kleinste obere Schranke von $\{h(a), h(b)\}$ bezüglich „$\leq_2$“ und es gilt $h(a \sqcup_1 b) = h(a) \sqcup_2 h(b)$.

Der Nachweis, daß $h(a \sqcap_1 b) = h(a) \sqcap_2 h(b)$ gilt, erfolgt entsprechend durch Dualisierung. ■

Auch der Begriff des Teilverbands erhält eine unterschiedliche Auslegung, je nachdem ob man algebraische Verbände oder Verbände als Ordnungsstrukturen betrachtet. In der nachfolgenden Abbildung bilden die hervorgehobenen Elemente einen zum Pentagon isomorphen Teilverbands des zugrundegelegten ordnungstheoretischen Verbands. Da jedoch „join“ und „meet“ nicht erhalten bleiben, liegt kein Teilverband im algebraischen Sinn vor.



Wir sprechen nachfolgend daher von algebraischen Teilverbänden, wenn wir verlangen, daß „meet“ und „join“ des Teilverbands sich durch Einschränkung der entsprechenden Operationen des Gesamtverbands ergeben.

Gemäß Lemma 6.27 stellt jede nichtleere Teilmenge eines Verbands $\mathcal{V}$ *immer* zumindest eine Teilstruktur (nicht notwendig einen Teilverband!) im ordungstheoretischen Sinn dar. Daher sind algebraische Teilverbände, die ja Suprema und Infima über zweielementige Teilmengen stets enthalten, dann immer Teilverbände im ordnungstheoretischen Sinn.

8.5 Distributive und modulare Verbände

Angesichts der Parallelität zwischen Verbandsregeln und gewissen aussagenlogischen Tautologien kann man sich fragen, ob „meet" und „join" nicht auch distributiv sind. Das folgende Lemma zeigt, daß Teilaussagen gelten.

Lemma 8.20 *Es sei $\langle V, \sqcup, \sqcap \rangle$ ein Verband. Dann gelten für alle $a, b, c \in V$ stets die folgenden Beziehungen:*

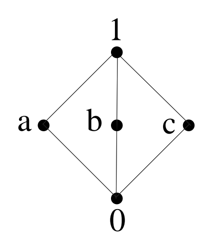
$$\begin{aligned} a \sqcup (b \sqcap c) &\leq (a \sqcup b) \sqcap (a \sqcup c) \\ (a \sqcap b) \sqcup (a \sqcap c) &\leq a \sqcap (b \sqcup c). \end{aligned}$$

Beweis. Wir weisen die erste Ungleichung nach, die zweite folgt ähnlich. Aus $b \sqcap c \leq c$ ergibt sich beispielsweise leicht $a \sqcup (b \sqcap c) \leq a \sqcup c$ (vgl. Aufgabe 8.8), und ähnlich $a \sqcup (b \sqcap c) \leq a \sqcup b$. Damit erhält man durch zweimaliges Anwenden von Lemma 8.13

$$\begin{aligned} (a \sqcup (b \sqcap c)) \sqcap ((a \sqcup b) \sqcap (a \sqcup c)) &= ((a \sqcup (b \sqcap c)) \sqcap (a \sqcup b)) \sqcap (a \sqcup c) \\ &= ((a \sqcup (b \sqcap c)) \sqcap (a \sqcup c) \\ &= a \sqcup (b \sqcap c). \end{aligned}$$

Wiederum mit Lemma 8.13 folgt nun $a \sqcup (b \sqcap c) \leq (a \sqcup b) \sqcap (a \sqcup c)$. ∎

Im allgemeinen gilt aber bei den oben genannten zwei Beziehungen nicht die Gleichheit. Beispielsweise gilt im Diamant



die Beziehung $a \sqcup (b \sqcap c) = a < 1 = (a \sqcup b) \sqcap (a \sqcup c)$.

Definition 8.21 Ein Verband $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ heißt *distributiv* genau dann, wenn für beliebige Elemente $a, b, c \in V$ die Gleichungen

$$\begin{aligned} a \sqcup (b \sqcap c) &= (a \sqcup b) \sqcap (a \sqcup c) \\ (a \sqcap b) \sqcup (a \sqcap c) &= a \sqcap (b \sqcup c) \end{aligned}$$

gelten.

Es reicht allerdings, in dieser Definition *eine* der beiden Identitäten zu fordern. Die jeweils andere läßt sich dann beweisen.

Bemerkung 8.22 [Dualitätsprinzip] Ist $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein distributiver Verband, so ist offenkundig auch der zu $\mathcal{V}$ duale Verband distributiv. Man kann daher das Dualitätsprinzip auf distributive Verbände spezialisieren: Aus jedem Gesetz, das in jedem distributiven Verband gilt, erhält man durch Dualisierung (Austausch von $\sqcup$ und $\sqcap$, $\leq$ und $\geq$ sowie von 0 und 1) wieder ein Gesetz über distributive Verbände.

Man kann ein sehr schönes geometrisches Kriterium angeben, das zeigt, wann ein Verband distributiv ist. Dazu führen wir zunächst einen etwas schwächeren Begriff ein.

Definition 8.23 Ein Verband $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ heißt *modular* genau dann, wenn für beliebige Elemente $a, b, c \in V$ stets $(a \sqcap b) \sqcup (a \sqcap c) = a \sqcap (b \sqcup (a \sqcap c))$ gilt.

In der Literatur werden modulare Verbände oft auch dadurch definiert, daß man fordert:

$$(\dagger) \quad \forall a, b, c\colon a \leq b \Rightarrow (a \sqcup (b \sqcap c) = b \sqcap (a \sqcup c)).$$

Wir überlassen es dem Leser nachzuweisen, daß beide Charakterisierungen der Modularität äquivalent sind (Aufgabe 8.11).

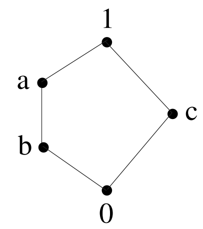
Lemma 8.24 *Jeder distributive Verband ist modular.*

Beweis. In einem distributiven Verband gilt unter Verwendung der Absorptionseigenschaft für beliebige Elemente a, b, c stets

$$a \sqcap (b \sqcup (a \sqcap c)) = a \sqcap (b \sqcup a) \sqcap (b \sqcup c) = a \sqcap (b \sqcup c) = (a \sqcap b) \sqcup (a \sqcap c).$$

Daraus ergibt sich die Behauptung. ■

Die Umkehrung von Lemma 8.24 gilt jedoch nicht: der oben abgebildete Diamant ist modular, wie man leicht nachweisen kann, jedoch nicht distributiv, wie wir gesehen hatten. Man kann zeigen, daß man auf modularen Verbänden stets eine Rangfunktion einführen kann. Damit ist das Pentagon



nicht modular. Wir können nun eine einfache Charakterisierung modularer und distributiver Verbände angegeben.

Theorem 8.25 (Birkhoff, Dedekind) *Ein Verband ist modular genau dann, wenn er keinen algebraischen Teilverband enthält, der zum Pentagon isomorph ist. Ein* modularer *Verband ist distributiv genau dann, wenn er keinen algebraischen Teilverband enthält, der zum Diamanten isomorph ist.*

Beweis. Enthält ein Verband einen nichtmodularen (nicht distributiven) algebraischen Teilverband, so kann er offenkundig nicht selbst modular (resp. distributiv) sein. Demzufolge ist eine Richtung der zwei Teilaussagen jeweils trivial. Es bleibt zu zeigen, (1) daß ein nichtmodularer Verband stets einen algebraischen Teilverband enthält, der zum Pentagon isomorph ist, und (2) daß ein modularer und nicht distributiver Verband stets einen algebraischen Teilverband enthält, der zum Diamanten isomorph ist.

(1) Es sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ nichtmodular. Wir verwenden die zweite Charakterisierung (†). Demnach existieren $a, b, c \in V$ mit $a \leq b$ aber $a \sqcup (b \sqcap c) \neq b \sqcap (a \sqcup c)$. Allerdings gilt

$$\begin{aligned} a \sqcup (b \sqcap c) &\overset{(i)}{\leq} (a \sqcup b) \sqcap (a \sqcup c) \\ &\overset{(ii)}{\leq} b \sqcap (a \sqcup c) \end{aligned}$$

Ungleichung (i) gilt nach Lemma 8.20, Gleichung (ii) folgt, da wegen $a \leq b$ auch $a \sqcup b = b$ gilt. Wir erhalten $a \sqcup (b \sqcap c) < b \sqcap (a \sqcup c)$. Wir setzen nun $a' := b \sqcap (a \sqcup c)$, $b' := a \sqcup (b \sqcap c)$, $1' := c \sqcup a$ und $0' := c \sqcap b$. Es ist offenkundig, daß die Ordnungsbeziehungen

$$0' \leq b' < a' \leq 1' \quad \wedge \quad 0' \leq c \leq 1'$$

vorliegen.

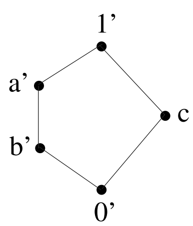
Wäre $b \sqcap c = 0' = b' = a \sqcup (b \sqcap c)$, so folgte $a \leq b \sqcap c$, also $a \leq b, c$ und $(a \sqcup c) \sqcap b = c \sqcap b = a \sqcup (b \sqcap c)$, was einen Widerspruch darstellt. Wäre $a \sqcup c = 1' = a' = b \sqcap (a \sqcup c)$, so folgte $b \geq a \sqcup c$, also $b \geq a, c$ und $a \sqcup (b \sqcap c) = a \sqcup c = b \sqcap (a \sqcup c)$, was abermals einen Widerspruch darstellt. Es folgt

$$0' < b' < a' < 1'.$$

Weiter gilt

$$\begin{aligned} c \sqcap a' &= c \sqcap (b \sqcap (a \sqcup c)) \overset{(i)}{=} (c \sqcap (c \sqcup a)) \sqcap b \overset{(ii)}{=} c \sqcap b = 0' \\ c \sqcup b' &= c \sqcup (a \sqcup (b \sqcap c)) \overset{(i)}{=} (c \sqcup (b \sqcap c)) \sqcup a \overset{(ii)}{=} c \sqcup a = 1'. \end{aligned}$$

Gleichungen (i) ergeben sich aus der Kommutativität und Assoziativität von „join“ und „meet“. Gleichungen (ii) ergeben sich aus der Absorptionseigenschaft. Der Rest des Beweises bleibt dem Leser überlassen (Aufgabe 8.12): Es ist nun leicht, zu zeigen, daß $0' < c < 1'$, $c \sqcap b' = 0'$ und $c \sqcup a' = 1'$. Daraus ergibt sich sofort, daß die Elemente $1', a', b', c, 0'$ einen algebraischen Teilverband der Form



bilden, der zum Pentagon isomorph ist.

(2) Es sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein modularer und nicht distributiver Verband. Demnach existieren $a, b, c \in V$ mit $(a \sqcap b) \sqcup (a \sqcap c) \neq a \sqcap (b \sqcup c)$. Mit Lemma 8.20 folgt

$$(a \sqcap b) \sqcup (a \sqcap c) < a \sqcap (b \sqcup c) \quad (*).$$

Wir definieren

$$\begin{aligned} 1' &:= (a \sqcup b) \sqcap (a \sqcup c) \sqcap (b \sqcup c), \\ 0' &:= (a \sqcap b) \sqcup (a \sqcap c) \sqcup (b \sqcap c), \\ a' &:= (a \sqcap 1') \sqcup 0', \\ b' &:= (b \sqcap 1') \sqcup 0', \\ c' &:= (c \sqcap 1') \sqcup 0'. \end{aligned}$$

Offenkundig gilt $0' \leq a', b', c'$. Im nächsten Schritt leiten wir vereinfachte Darstellungen der Elemente a', b', c' ab. Es gilt

$$\begin{aligned} a' &= (a \sqcap 1') \sqcup 0' \\ &= (a \sqcap (a \sqcup b) \sqcap (a \sqcup c) \sqcap (b \sqcup c)) \sqcup 0' \\ &\overset{(i)}{=} (a \sqcap (b \sqcup c)) \sqcup 0' \\ &= (a \sqcap (b \sqcup c)) \sqcup (a \sqcap b) \sqcup (a \sqcap c) \sqcup (b \sqcap c) \\ &\overset{(ii)}{=} (a \sqcap (b \sqcup c)) \sqcup (b \sqcap c). \end{aligned}$$

Bei (i) verwenden wir $a \leq a \sqcup b, a \sqcup c$, um in zwei Schritten $a \sqcap (a \sqcup b) \sqcap (a \sqcup c)$ zu a zu vereinfachen. Bei (ii) verwenden wir $a \sqcap b, a \sqcap c \leq a \sqcap (b \sqcup c)$, um die Elemente $(a \sqcap b)$ und $(a \sqcap c)$ aus der Supremumsbildung zu eliminieren. Wenn wir von der letzten Darstellung von a' ausgehen, ergibt sich mit $b \sqcap c \leq b \sqcup c$ aufgrund der Modularität (†) die duale Darstellung $a' = (a \sqcup (b \sqcap c)) \sqcap (b \sqcup c)$. Wenn wir dieselben Umformungen auch für b' und c' verwenden, erhalten wir zusammenfassend

$$\begin{aligned} a' &= (a \sqcap (b \sqcup c)) \sqcup (b \sqcap c) = (a \sqcup (b \sqcap c)) \sqcap (b \sqcup c), \\ b' &= (b \sqcap (c \sqcup a)) \sqcup (c \sqcap a) = (b \sqcup (c \sqcap a)) \sqcap (c \sqcup a), \\ c' &= (c \sqcap (a \sqcup b)) \sqcup (a \sqcap b) = (c \sqcup (a \sqcap b)) \sqcap (a \sqcup b). \end{aligned}$$

Weiter gilt

$$\begin{aligned} a' &= (a \sqcap (b \sqcup c)) \sqcup (b \sqcap c) \\ &\overset{(i)}{\leq} (a \sqcup (b \sqcap c)) \sqcap ((b \sqcup c) \sqcup (b \sqcap c)) \\ &\overset{(ii)}{=} (a \sqcup (b \sqcap c)) \sqcap (b \sqcup c) \\ &\overset{(iii)}{\leq} (a \sqcup b) \sqcap (a \sqcup c) \sqcap (b \sqcup c) \\ &= 1'. \end{aligned}$$

Bei (i) wenden wir Lemma 8.20 auf den Gesamtterm an, bei (ii) verwenden wir die Absorptionseigenschaft und $b \sqcap c \leq b \sqcup c$, (iii) folgt durch Anwendung von Lemma 8.20 auf $a \sqcup (b \sqcap c)$ und Isotonie von „$\sqcap$“ (vgl. Aufgabe 8.8). Die entsprechenden Rechnungen für b', c' ergeben nun

$$0' \leq a', b', c' \leq 1'.$$

Das nächste Ziel ist es, $0' < 1'$ nachzuweisen. Hierzu reicht es offenkundig, $a \sqcap 0' < a \sqcap 1'$ zu zeigen. Es gilt

$$a \sqcap 1' = a \sqcap (a \sqcup b) \sqcap (a \sqcup c) \sqcap (b \sqcup c)$$

$$
\begin{aligned}
&\overset{(i)}{=} && a \sqcap (a \sqcup c) \sqcap (b \sqcup c)\\
&\overset{(ii)}{=} && a \sqcap (b \sqcup c),\\
a \sqcap 0' &= && a \sqcap [(a \sqcap b) \sqcup (a \sqcap c) \sqcup (b \sqcap c)]\\
&= && a \sqcap [((a \sqcap b) \sqcup (b \sqcap c)) \sqcup (a \sqcap c)]\\
&\overset{(j)}{=} && [a \sqcap ((a \sqcap b) \sqcup (b \sqcap c))] \sqcup (a \sqcap c)\\
&= && [a \sqcap ((b \sqcap c) \sqcup (a \sqcap b))] \sqcup (a \sqcap c)\\
&\overset{(jj)}{=} && [(a \sqcap b \sqcap c) \sqcup (a \sqcap b)] \sqcup (a \sqcap c)\\
&= && (a \sqcap b) \sqcup (a \sqcap c).
\end{aligned}
$$

Hierbei sind (i), (ii) Anwendungen der Absorptionseigenschaft, bei (j) und (jj) wird die Modularitätsregel (Def. 8.23) auf den Gesamtterm beziehungsweise auf $a \sqcap ((b \sqcap c) \sqcup (a \sqcap b))$ angewandt. Nach $(*)$ folgt nun $a \sqcap 0' < a \sqcap 1'$ und damit $0' < 1'$.

Um nun nachzuweisen, daß die Elemente $1', a', b', c', 0'$ einen algebraischen Teilverband bilden, der zum Diamanten isomorph ist, zeigen wir, daß $a' \sqcap b' = a' \sqcap c' = b' \sqcap a' = 0'$ und $a' \sqcup b' = a' \sqcup c' = b' \sqcup a' = 1'$ gilt. Man beachte, daß dies auch wegen $0' < 1'$ die paarweise Verschiedenheit der Elemente $1', a', b', c', 0'$ impliziert. Es gilt

$$
\begin{aligned}
a' \sqcup b' &= && (b \sqcap c) \sqcup [(a \sqcap (b \sqcup c)) \sqcup (b \sqcap (c \sqcup a))] \sqcup (c \sqcap a)\\
&\overset{(l)}{=} && (b \sqcap c) \sqcup [[(a \sqcap (b \sqcup c)) \sqcup b] \sqcap (c \sqcup a)] \sqcup (c \sqcap a)\\
&\overset{(ll)}{=} && (b \sqcap c) \sqcup [[(b \sqcup a) \sqcap (b \sqcup c)] \sqcap (c \sqcup a)] \sqcup (c \sqcap a)\\
&\overset{(lll)}{=} && (b \sqcap c) \sqcup [(b \sqcup a) \sqcap (b \sqcup c) \sqcap (c \sqcup a)] \sqcup (c \sqcap a)\\
&\overset{(lv)}{=} && (b \sqcup a) \sqcap (b \sqcup c) \sqcap (c \sqcup a)\\
&= && 1'.
\end{aligned}
$$

Hierbei ist (l) eine Anwendung der Modularität (†) mit $a \sqcap (b \sqcup c) \leq a \leq c \sqcup a$, (ll) eine Anwendung der Modularität mit $b \leq b \sqcup c$. Bei (lv) wird ausgenutzt, daß $(b \sqcap c), (c \sqcap a) \leq (b \sqcup a), (b \sqcup c), (c \sqcup a)$. Damit spielen diese Terme bei der Supremumsbildung keine Rolle.

Durch Ausnutzen der dualen Repräsentation von a', b' erhalten wir durch Dualisierung $a' \sqcap b' = 0'$. Analog ergeben sich die anderen oben erwähnten Gleichungen. ■

8.6 Ideale und Filter

Ideale und Filter sind spezielle Teilverbände, sowohl im ordnungstheoretischen wie auch im algebraischen Sinn. Ihre volle Bedeutung wird erst im nächsten Kapitel klar werden, wenn wir Boolesche Algebren besprechen.

Definition 8.26 Es sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ ein Verband, die partielle Ordnung $\leq$ auf V sei wie in Teil 2 von Satz 8.17 definiert. Eine nichtleere Teilmenge $I \subseteq V$ heißt *Ideal* von $\mathcal{V}$ genau dann, wenn gilt:

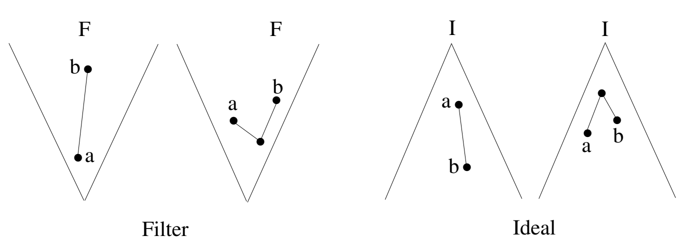
1. $\forall a, b \in V$: $a \in I$ und $b \leq a$ impliziert $b \in I$,
2. $\forall a, b \in V$: $a, b \in I$ impliziert $a \sqcup b \in I$.

I heißt *eigentliches* Ideal genau dann, wenn $I \neq V$. Dual heißt eine nichtleere Teilmenge $F \subseteq V$ *Filter* von $\mathcal{V}$ genau dann, wenn gilt:

1. $\forall a, b \in V$: $a \in F$ und $a \leq b$ impliziert $b \in F$,
2. $\forall a, b \in V$: $a, b \in F$ impliziert $a \sqcap b \in F$.

F heißt *eigentlicher* Filter genau dann, wenn $F \neq V$.

Die nachfolgende Abbildung gibt eine Veranschaulichung dieser Begriffe.



Lemma 8.27 *Ideale und Filter sind Teilverbände, das heißt unter „$\sqcup$“ und „$\sqcap$“ abgeschlossen.*

Beweis. Sei I ein Ideal des Verbands $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ und „$\leq$“ wie in Teil 2 von Theorem 8.17 definiert. Mit $a, b \in I$ folgt $a \sqcup b \in I$ nach Idealeigenschaft 2.

Aus $a \sqcap b \leq a$ folgt mit Eigenschaft 1 auch $a \sqcap b \in I$. für Filter können wir das Dualitätsprinzip verwenden. ■

Lemma 8.28 *Sei $\mathcal{V} = \langle V, \sqcup, \sqcap \rangle$ wie oben und $\emptyset \neq X \subseteq V$.*

1. *Die Menge $I(X) := \{a \in V \mid \exists x_1, \ldots, x_n \in X, a \leq x_1 \sqcup \ldots \sqcup x_n\}$ ist das kleinste X umfassende Ideal,*

2. *Die Menge $F(X) := \{a \in V \mid \exists x_1, \ldots, x_n \in X, a \geq x_1 \sqcap \ldots \sqcap x_n\}$ ist der kleinste X umfassende Filter,*

3. *Wenn 1 (bzw. 0) existiert, und wenn I ein Ideal (bzw. F ein Filter) ist, dann ist I (bzw. F) eigentlich genau dann, wenn $1 \notin I$ (bzw. $0 \notin F$).*

Beweis. 1. Es ist klar, daß jedes Ideal, das X umfaßt, auch $I(X)$ enthalten muß. Offenkundig ist mit X auch $I(X)$ nichtleer. Um zu zeigen, daß $I(X)$ bereits ein Ideal ist, seien a, b zwei Elemente der Menge. Dann existieren $x_1, \ldots, x_n$ und $y_1, \ldots, y_m$ in X mit

$$\begin{aligned} a &\leq x_1 \sqcup \ldots \sqcup x_n, \\ b &\leq y_1 \sqcup \ldots \sqcup y_m. \end{aligned}$$

Mit Hilfe von Aufgabe 8.4 folgt nun leicht

$$a \sqcup b \leq x_1 \sqcup \ldots \sqcup x_n \sqcup y_1 \sqcup \ldots \sqcup y_m,$$

also gilt $a \sqcup b \in I(X)$ und $I(X)$ erfüllt Idealeigenschaft 2. Aus der Transitivität von $\leq$ folgt sofort, daß $I(X)$ auch Idealeigenschaft 1 erfüllt. Damit ist Teil 1 gezeigt. Teil 2 folgt dual, Teil 3 ist trivial. ■

8.7 Ergänzungen

Im ersten Teil dieses Abschnitts stellen wir kurz die sogenannten Halbverbände vor. Sie sind ähnlich zu den Verbänden und erlauben Supremumsbildung *oder* Infimumsbildung über beliebige endliche Teilmengen, aber nicht notwendig beides. Im zweiten Abschnitt stellen wir den Fixpunktsatz von Tarski und Knaster dar.

8.7.1 Halbverbände

In manchen Bereichen — etwa in der Computerlinguistik — spielen neben den Verbänden auch sogenannte *Halbverbände* eine Rolle.

Definition 8.29 Eine partiell geordnete Menge $\mathcal{H} = \langle H, \leq \rangle$ heißt *join-Halbverband* (bzw. *meet-Halbverband*) genau dann, wenn $sup\,\{a, b\}$ (bzw. $inf\,\{a, b\}$) für je zwei Elemente $a, b \in H$ stets existiert.

Ist $\mathcal{H}$ ein join-Halbverband, so ist offensichtlich die zu $\mathcal{H}$ duale partiell geordnete Menge ein meet-Halbverband und umgekehrt.

Beispiel 8.30 In Klassifikationen existiert oft zu je zwei Begriffen der kleinste gemeinsame Oberbegriff, aber nicht notwendig immer ein (größter) gemeinsamer Unterbegriff. Es liegt dann ein echter Halbverband vor.

Auch die Halbverbände lassen sich auf algebraische Weise charakterisieren. Die Definition des Halbverbands als Algebra ist wie folgt:

Definition 8.31 Ein Halbverband ist eine Algebra $\mathcal{H} = \langle H, \sqcup \rangle$, wo $\sqcup$: $H \times H \to H$ eine idempotente, kommutative und assoziative Funktion ist.

Lemma 8.32 *Algebraische und ordnungtheoretische Definitionen von Halbverbänden sind äquivalent im folgenden Sinn:*

1. *Sei $\mathcal{H} = \langle H, \leq \rangle$ ein join-Halbverband. für $a, b \in H$ definieren wir $a \sqcup b := sup\,\{a, b\}$. Dann ist $H^* := \langle H, \sqcup \rangle$ ein algebraischer Halbverband.*

2. *Es sei $\mathcal{H} = \langle H, \sqcup \rangle$ ein algebraischer Halbverband. Es sei $\leq_1$ definiert durch $a \leq_1 b$ gdw. $a = a \sqcup b$ und $\leq_2$ durch $a \leq_2 b$ gdw. $b = a \sqcup b$. Dann ist $\mathcal{H}^{\leq_1} := \langle H, \leq_1 \rangle$ ein meet-Halbverband und $\mathcal{H}^{\leq_2} := \langle H, \leq_2 \rangle$ ein join-Halbverband.*

3. *Sei $\mathcal{H} = \langle H, \leq \rangle$ ein join-Halbverband. Dann gilt $(\mathcal{H}^*)^{\leq_i} = \mathcal{H}$ $(i = 1, 2)$.*

4. *Es sei $\mathcal{H} = \langle H, \sqcup \rangle$ ein algebraischer Halbverband. Dann gilt $(\mathcal{H}^{\leq_i})^* = \mathcal{H}$ $(i = 1, 2)$.*

Beweis. 1. Natürlich ist „$\sqcup$“ eine idempotente und kommutative binäre Operation auf A. Die Assoziativität ergibt sich aus einer leicht zu beweisenden Variante von Lemma 8.5.

2. Es seien $a, b \in A$. Wir zeigen, daß $a \sqcup b = \mathit{inf}\,\{a, b\}$ bezüglich „$\leq_1$“ gilt. Zunächst gilt, da „$\sqcup$“ kommutativ, assoziativ und idempotent ist,

$$(a \sqcup b) \sqcup a = (b \sqcup a) \sqcup a = b \sqcup (a \sqcup a) = b \sqcup a = a \sqcup b,$$

in anderen Worten $a \sqcup b \leq_1 a$. Symmetrisch folgt $a \sqcup b \leq_1 b$, also ist $a \sqcup b$ eine untere Schranke von $\{a, b\}$. Sei nun c irgendeine untere Schranke von $\{a, b\}$, also $c \leq_1 a$ und $c \leq_1 b$. Dann folgt $c = c \sqcup a$ und $c = c \sqcup b$. Wir erhalten

$$c = c \sqcup c = (c \sqcup a) \sqcup (c \sqcup b) = \ldots = c \sqcup (a \sqcup b),$$

also $c \leq a \sqcup b$. Somit ist $a \sqcup b$ größte untere Schranke von $\{a, b\}$.

Wir haben gesehen, daß $\mathcal{H}^{\leq_1} = \langle H, \leq_1 \rangle$ ein meet-Halbverband ist. Analog zeigt man, daß $a \sqcup b = \mathit{sup}\,\{a, b\}$ bezüglich $\leq_2$ gilt, damit ist $\mathcal{H}^{\leq_2} = \langle H, \leq_2 \rangle$ ein join-Halbverband. Der Beweis der Teile 3 und 4 ist einfach. ∎

Natürlich läßt sich ein duales Lemma für meet-Halbverbände zeigen.

8.7.2 Der Fixpunktsatz von Tarski und Knaster

für isotone Abbildungen eines vollständigen Verbands in sich selbst läßt sich eine interessante Eigenschaften verifizieren, auf die wir in diesem Abschnitt kurz eingehen möchten. Wir benötigen zunächst einen Hilfsbegriff.

Definition 8.33 Es sei M eine nichtleere Menge und $f \colon M \to M$ eine Abbildung. Das Element $a \in M$ heißt *Fixpunkt* der Abbildung f, falls $f(a) = a$ gilt.

Theorem 8.34 (Fixpunktsatz von Tarski und Knaster) *Jede isotone Abbildung f eines* vollständigen *Verbands $\mathcal{V}$ in sich selbst besitzt einen Fixpunkt. Unter den Fixpunkten von f gibt es einen größten und einen kleinsten.*

Beweis. Es sei f eine isotone Abbildung des vollständigen Verbands $\mathcal{V} = \langle V, \leq \rangle$ in sich. Wir sagen, ein Element $b \in V$ wird unter f gerichtet aufwärts (abwärts) abgebildet, falls $b \leq f(b)$ (respektive $f(b) \leq b$) gilt. Wir zeigen die folgende

Hilfsbehauptung: Ist a das Supremum (Infimum) über eine Menge $B \subseteq V$ von Elementen, die gerichtet aufwärts (abwärts) abgebildet werden, so werden a und $f(a)$ gerichtet aufwärts (abwärts) abgebildet.

Gilt nämlich $a = \mathit{sup}\, B$, wo die Elemente von B gerichtet aufwärts abgebildet werden, so gilt $b \leq a$ für alle $b \in B$. Wegen der Isotonie folgt $b \leq f(b) \leq f(a)$. Damit ist $f(a)$ ist eine obere Schranke von B, woraus sich $a \leq f(a)$ ergibt. Mit nochmaliger Anwendung der Isotonie folgt $f(a) \leq f(f(a))$. Der Beweis für die duale Behauptung folgt analog.

Um einen größten Fixpunkt zu erhalten, betrachten wir die Teilmenge B aller Elemente, die gerichtet aufwärts abgebildet werden. Es sei $a := \mathit{sup}\, B$. Die Hilfsbehauptung zeigt, daß $a \leq f(a)$. Andererseits ist nach der Hilfsbehauptung auch $f(a) \in B$ und somit $f(a) \leq a$. Daher gilt $a = f(a)$. Ist c ein Fixpunkt von f, so ist $c \in B$ und es gilt $c \leq a$.

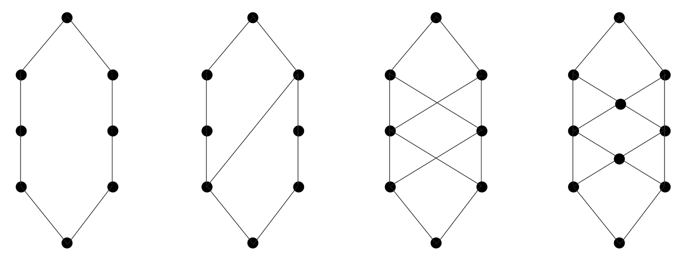
Um einen kleinsten Fixpunkt zu erhalten, betrachten wir die Teilmenge B' aller Elemente, die gerichtet abwärts abgebildet werden, und setzen $a' := \mathit{inf}\, B'$. Die Hilfsbehauptung zeigt, daß $f(a)' \leq a$. Andererseits ist $f(a') \in B'$ und somit $a' \leq f(a')$. Somit gilt $f(a') = a'$. Ist c ein Fixpunkt von f, so ist $c \in B'$ und es gilt $a' \leq c$. ■

Um eine Anwendungsform des Fixpunktsatzes und den Zusammenhang zu anderen Begriffsbildungen anzudeuten, denken wir uns eine Situation, wo wir aus den Elementen einer gegebenen Teilmenge A einer Obermenge M vermöge eines geeigneten Erzeugungsprozesses (z.B. durch Anwendung induktiver Regeln) neue Elemente aus M erhalten, die wir zu A hinzufügen. Typischerweise ist in dieser Situation die Abbildung f, die jeder Menge A die vergrößerte Menge A' zuordnet, isoton bezüglich der Mengeninklusion. Wie in Beispiel 8.9 Nr. 3 erwähnt, ist $\langle \mathcal{P}(M), \subseteq \rangle$ ein vollständiger Verband. Daher besitzt f nach dem Fixpunktsatz einen kleinsten Fixpunkt. Dieser repräsentiert dann die kleinste Teilmenge von M, die unter dem gegebenen Erzeugungsprozeß abgeschlossen ist.

8.8 Aufgaben zu Kapitel 8

Aufgaben zu Teilkapitel 8.1

Aufgabe 8.1 Welche der nachfolgend in Form von Hasse-Diagrammen dargestellten Ordnungsstrukturen sind Verbände, welche nicht?



Aufgabe 8.2 Geben Sie alle Verbände mit genau vier Elementen an. Geben Sie alle Verbände mit genau fünf Elementen an.

Aufgabe 8.3 Beweisen Sie Lemma 8.5.

Aufgabe 8.4 Verifizieren Sie in den Beispielen 8.9, Teil 1, 3 und 5, daß die angegebenen Strukturen Verbände sind.

Aufgabe 8.5 Eine Teilmenge $M \subseteq \mathbb{N}$ heiße *co-endlich* genau dann, wenn $\mathbb{N} \setminus M$ endlich ist. Zeigen Sie: Die Menge $\mathcal{K}$ aller endlichen oder co-endlichen Teilmengen von $\mathbb{N}$ bildet mit der Inklusionsbeziehung „$\subseteq$" einen Verband. Ist dieser Verband vollständig?

Aufgabe 8.6 Ist $\mathcal{V} = \langle V, \leq \rangle$ ist ein Verband mit Null-Element 0, so wird ein Element $a \in V$ *Atom* genannt genau dann, wenn a ein direkter Nachfolger von 0 ist, falls also 0 das einzige Element von V ist, das echt kleiner ist als a. Ein Verband $\mathcal{V}$ heißt *atomar* genau dann, wenn für jedes $b \in V$, $b \neq 0$, stets ein Atom a existiert mit $a \leq b$. Offenkundig ist jeder endliche Verband atomar. Geben Sie einige Beispiele für nicht-atomare Verbände.

Aufgaben zu Teilkapitel 8.3

Aufgabe 8.7 Vervollständigen Sie den Beweis von Satz 8.17, Teil 2. Verifizieren Sie hierzu zunächst die folgende Hilfsbehauptung: ist die partielle Ordnung „$\leq$“ auf V wie angegeben definiert, so gilt stets $a \leq b$ genau dann, wenn $a \sqcup b = b$, für alle $a, b \in V$.

Aufgabe 8.8 Zeigen sie, daß für beliebige Elemente a, b, c eines Verbands stets gilt: $a \leq b$ impliziert $a \sqcup c \leq b \sqcup c$ und $a \sqcap c \leq b \sqcap c$. Die Eigenschaft wird oft als *Isotonie* von „$\sqcup$“ und „$\sqcap$“ bezeichnet.

Aufgaben zu Teilkapitel 8.4

Aufgabe 8.9 Geben Sie alle — starken und schwachen — Homomorphismen zwischen den Ordnungsstrukturen $\mathbf{P}_2 := \langle \mathcal{P}(\{0,1\}), \subseteq \rangle$ und $\mathbf{2} = \langle \{0,1\}, \leq \rangle$ an (wobei $0 \leq 1$ gilt). Wie sehen in diesem Fall die möglichen Homomorphismen zwischen den zugehörigen algebraischen Verbänden aus?

Aufgabe 8.10 Geben Sie in Form von Hasse-Diagrammen einige Beispiele von Verbänden mit ordnungstheoretischen Teilverbänden, die keine algebraischen Teilverbände darstellen.

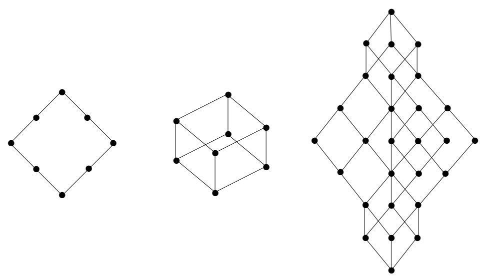
Aufgaben zu Teilkapitel 8.5

Aufgabe 8.11 Weisen Sie nach, daß die beiden Charakterisierungen der Modularität von Verbänden aus Abschnitt 8.5 (Definition 8.23 bzw. (†)) äquivalent sind.

Aufgabe 8.12 Beenden Sie wie angegeben Teil (1) des Beweises von Theorem 8.25.

Aufgabe 8.13 Beweisen Sie, daß das Pentagon nicht modular ist.

Aufgabe 8.14 Verwenden Sie Theorem 8.25, um zu entscheiden, welche der folgenden Verbände modular oder distributiv sind.



(a) (b) (c)

Aufgabe 8.15 Zeigen Sie: ist $n > 0$ eine natürliche Zahl, so bildet die Menge der Teiler von n einen vollständigen und distributiven Verband bezüglich der Ordnung "$|$" (vgl. Beispiel 8.9, Teil 2). Geben Sie eine Rangfunktion an. Zeichnen Sie ein Hasse-Diagramm des Verbands aller Teiler von 60.

Aufgabe 8.16 Es sei M eine Menge von 4 Elementen. Geben Sie ein Hasse-Diagramm für den Verband aller Äquivalenzrelationen auf M mit der Verfeinerungsordnung (vgl. Beispiel 8.9 Nr. 4) an. Ist dieser Verband modular bzw. distributiv?

Aufgabe 8.17 Es sei $A := \{a, b\}$ und $B \subseteq \mathbb{N}$ eine endliche Menge. Geben Sie ein Hasse-Diagramm für den in Beispiel 8.9, Teil 5, eingeführten Verband $\langle B^A, \preceq\rangle$. Geben Sie eine passende Rangfunktion. Ist dieser Verband modular bzw. distributiv?

Aufgaben zu Teilkapitel 8.6

Aufgabe 8.18 Es sei $\mathcal{V}$ der Diamant (das Pentagon). Geben Sie alle Ideale und Filter von $\mathcal{V}$ an.

Aufgabe 8.19 Es sei M eine nichtleere Menge und $\mathcal{V} := \langle \mathcal{P}(M), \subseteq\rangle$. Wie sehen die maximalen eigentlichen Ideale und Filter von $\mathcal{V}$ aus?

Aufgabe 8.20 Es sei X eine nichtleere Teilmenge des Verbands $\mathcal{V}$ und M eine Menge von Idealen (Filter) von V, die alle X umfassen. Zeigen Sie, daß auch $\bigcap M$ ein X umfassendes Ideal (X umfassender Filter) ist.

Aufgaben zu Teilkapitel 8.7

Aufgabe 8.21 Zeigen Sie, daß die Menge aller Wörter über einem nichtleeren Alphabet A zusammen mit der Präfixrelation (vgl. Beispiel 3.16 Nr. 1) einen Meet-Halbverband bildet. Warum entsteht i.a. kein Verband? Gibt es einen Sonderfall, wo sich ein Verband ergibt?

8.9 Bibliographische Angaben

Die Standardreferenz für Verbandstheorie ist [Bir84]. Die in Beispiel 8.8 erwähnten Konzeptverbände und darauf basierende Anwendungen sind in [GW99] und [SE00] beschrieben.

9

Boolesche Algebren

Im vorangegangenen Kapitel hatten wir gesehen, daß beim Rechnen in Verbänden viele Gesetzmäßigkeiten auftreten, die wir auch vom Umgang mit Mengen und Aussagen kennen. Selbst wenn wir uns auf distributive Verbände einschränken, bleibt jedoch noch ein wesentlicher Unterschied bestehen. In Verbänden gibt es im allgemeinen keine Komplementbildung, die eine Parallele zur logischen Negation — oder zur Komplementbildung bei Mengen — erlaubt. Die Booleschen Algebren stellen eine spezielle Klasse distributiver Verbände dar, bei denen eine solche Komplementbildung möglich ist. In der Tat werden wir sehen, daß das Rechnen in Booleschen Algebren als eine algebraische Beschreibung des Rechnens mit Aussagen aufgefaßt werden kann. In einem zu klärenden Sinn können darüberhinaus Boolesche Algebren bis auf Isomorphie stets mit Hilfe von Mengen und den üblichen Mengenoperationen repräsentiert werden.

Im ersten Teilkapitel führen wir Boolesche Algebren ein und stellen einige Regeln für das Rechnen mit Komplementen zusammen. Im zweiten Teilkapitel werden wir den Zusammenhang zwischen Idealen, Homomorphismen und Kongruenzrelationen bei Booleschen Algebren genauer studieren und die Bildung von Quotientenalgebren betrachten. Das Teilkapitel kann damit als eine konkrete Fallstudie zu den in Kapitel 6.5 dargestellten allgemeinen Techniken zur Vereinfachung von Strukturen gesehen werden. Im dritten Teilkapitel betrachten wir Primideale und kommen dann auf die Repräsentierbarkeit von Booleschen Algebren durch Mengen und Mengenoperationen zu sprechen.

9.1 Komplemente

Wir beginnen mit dem zentralen Begriff dieses Kapitels.

Definition 9.1 Eine *Boolesche Algebra* ist eine Algebra $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ wo gilt

1. $\langle B, \sqcup, \sqcap \rangle$ ist ein distributiver Verband,

2. die Abbildung „$-$" ist eine einstellige Operation auf B, und es gilt

$$\forall a, b \in B\colon (a \sqcap -a) \sqcup b = b, (a \sqcup -a) \sqcap b = b.$$

Das Element $-a$ wird *Komplement* von a in $\mathcal{B}$ genannt.

Bemerkung 9.2 Ist $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra, so können wir—wie von den Verbänden bereits gewohnt—eine partielle Ordnung „$\leq$" auf B definieren vermöge $a \leq b :\Leftrightarrow a \sqcap b = a$. Die Bedingung ist äquivalent zu $a \sqcup b = b$. Im nachfolgenden werden wir von dieser Ordnung ohne besondere Erwähnung Gebrauch machen.

Bemerkung 9.3 [Dualitätsprinzip für Boolesche Algebren] Gemäß Bemerkung 8.22 ist der zu einer Booleschen Algebra $\mathcal{B}$ duale Verband wieder distributiv. Wenn wir auf diesem die Komplementbildung wie auf $\mathcal{B}$ selbst definieren, so erfüllt diese die Regeln aus Bedingung 2 von Definition 9.1. Damit ist der duale Verband selbst wieder eine Boolesche Algebra, mit derselben Komplementbildung. Man kann daher das Dualitätsprinzip auf Boolesche Algebren spezialisieren: Aus jedem Gesetz, das in jeder Booleschen Algebra gilt, erhält man durch Dualisierung (Austausch von $\sqcup$ und $\sqcap$, $\leq$ und $\geq$ sowie von 0 und 1, Erhaltung von $-$) wieder ein Gesetz über Boolesche Algebren.

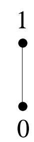
Nachfolgend nun einige Beispiele für Boolesche Algebren.

Beispiel 9.4 Es sei $B = 2 = \{0, 1\}$. Die Operationen $\sqcap, \sqcup$ und „$-$" seien wie in folgender Tabelle festgelegt:

a	$-a$
0	1
1	0

a	b	$a \sqcap b$	$a \sqcup b$	$-a \sqcup b$	$(-a \sqcup b) \sqcap (-b \sqcup a)$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

Dann ist $\mathbf{2} := \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra.



$\mathbf{2} = \mathbf{P}_1$

Wenn man diese Tabellen mit den in Kapitel 1 gegebenen Tabellen für die aussagenlogischen Junktoren vergleicht, sieht man, daß die Boolesche Algebra $\mathbf{2}$ genau das Rechnen mit Wahrheitswerten beschreibt. Die den Junktoren „$\Rightarrow$“ und „$\Leftrightarrow$“ entsprechenden Verbandsoperationen sind in den letzten beiden Spalten wiedergegeben.

Beispiel 9.5 Es sei $n \geq 1$ und $B = \{0,1\}^n = \{0,1\} \times \cdots \times \{0,1\}$ (n Faktoren). Wir definieren die Operation „join“ vermöge

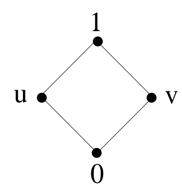
$$\langle a_1, \ldots, a_n \rangle \sqcup \langle b_1, \ldots, b_n \rangle := \langle a_1 \sqcup b_1, \ldots, a_n \sqcup b_n \rangle.$$

Analog seien alle andereren Booleschen Operationen koordinatenweise definiert. Dann ist die resultierende Algebra eine Boolesche Algebra. Wir notieren sie in der Form $\mathbf{2}^n$. Die Elemente von $\{0,1\}^n$ können als Bitvektoren aufgefaßt werden. Es entspricht „$\sqcap$“ dann der (koordinatenweisen) Multiplikation der Bitvektoren, „$\sqcup$“ der Maximalwertberechnung. Beispielsweise gilt

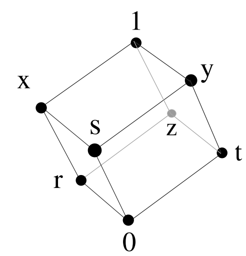
$$\begin{array}{r} \langle 0,0,1,0,1,0,0 \rangle \\ \sqcup \langle 1,0,1,1,0,0,0 \rangle \\ = \langle 1,0,1,1,1,0,0 \rangle \end{array} \qquad \begin{array}{r} \langle 0,0,1,0,1,0,0 \rangle \\ \sqcap \langle 1,0,1,1,0,0,0 \rangle \\ = \langle 0,0,1,0,0,0,0 \rangle \end{array}$$

Beispiel 9.6 Sei M eine Menge und $\mathcal{P}(M)$ die Potenzmenge von M. Für $X \subseteq M$ sei $-X$ das Komplement $M \setminus X$ von X. Dann ist

$\langle \mathcal{P}(M), \cup, \cap, - \rangle$ eine Boolesche Algebra. Offenkundig sind zwei Boolesche Algebren $\langle \mathcal{P}(M), \cup, \cap, - \rangle$ und $\langle \mathcal{P}(N), \cup, \cap, - \rangle$ isomorph, falls M und N dieselbe Zahl von Elementen haben. Mit $\mathbf{P_n}$ bezeichnen wir die Boolesche Algebra $\langle \mathcal{P}(M), \cup, \cap, - \rangle$ für eine n-elementige Menge M. Die Komplementbildung in diesen Algebren läßt sich jeweils durch Diagonalbildung geometrisch interpretieren.



$\mathbf{P}_2$



$\mathbf{P}_3$

Für den Rest dieses Kapitels kehren wir zur Konvention zurück, das eindeutig bestimmte größte (resp. kleinste) Element mit 1 (bzw. 0) zu bezeichnen. Während ein solches Element bei Verbänden nicht notwendig existieren muß, haben Boolesche Algebren stets eine 0 und eine 1.

Lemma 9.7 *Es sei $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra. Dann existieren* 0 *und* 1 *in* $\mathcal{B}$ *und es gilt:*

$$\forall b \in B\colon \quad 0 = b \sqcap -b \quad \wedge \quad 1 = b \sqcup -b.$$

Beweis. Die zweite Bedingung in Definition 9.1 impliziert sofort, daß für jedes $a \in B$ die Beziehung $b \sqcap -b \leq a \leq b \sqcup -b$ gilt. ∎

Das folgende Lemma besagt, daß Komplemente durch die in Lemma 9.7 beschriebenen Eigenschaften eindeutig bestimmt sind.

Lemma 9.8 *Sei $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra und $a, b \in B$. Falls $0 = b \sqcap a$ und $1 = b \sqcup a$, so gilt $a = -b$.*

Beweis. Wir haben

$$a = a \sqcap 1 = a \sqcap (b \sqcup -b) = (a \sqcap b) \sqcup (a \sqcap -b) = 0 \sqcup (a \sqcap -b) = a \sqcap -b,$$

das heißt $a \leq -b$. Analog folgt $-b \leq a$. ■

Wir stellen noch einige einfache Regeln für das Rechnen mit Komplementen zusammen.

Lemma 9.9 *In einer Booleschen Algebra* $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ *gilt für beliebige Elemente* $a, b \in B$ *stets:*

1. $-(a \sqcap b) = -a \sqcup -b$ *und* $-(a \sqcup b) = -a \sqcap -b$,
2. $-(-a) = a$,
3. $a \leq b$ *gdw.* $-b \leq -a$,
4. $-1 = 0$ *und* $-0 = 1$,
5. $a \leq b$ *gdw.* $-a \sqcup b = 1$.

Beweis. Wir beweisen Teil 1, der Rest wird als Übung empfohlen (vgl. Aufgabe 9.2). Nach Lemma 9.8 reicht es, für die erste Gleichung die folgenden Identitäten zu zeigen:

$$(a \sqcap b) \sqcup (-a \sqcup -b) = 1 \text{ und } (a \sqcap b) \sqcap (-a \sqcup -b) = 0.$$

Es gilt

$$(a \sqcap b) \sqcup (-a \sqcup -b) = (a \sqcup -a \sqcup -b) \sqcap (b \sqcup -a \sqcup -b) = 1 \sqcup 1 = 1$$

und

$$(a \sqcap b) \sqcap (-a \sqcup -b) = (a \sqcap b \sqcap -a) \sqcup (a \sqcap b \sqcap -b) = 0 \sqcup 0 = 0.$$

Die zweite Gleichheit folgt dual nach Bemerkung 9.3. ■

9.2 Ideale, Kongruenzrelationen und Homomorphismen

Im vorigen Kapitel hatten wir den Begriff des Ideals und des Filters eines Verbands definiert. Da Boolesche Algebren spezielle Verbände sind, ist damit auch der Begriff eines Ideals bzw. Filters einer Booleschen Algebra erklärt. Wir zeigen, daß bei Booleschen Algebren die Ideale und Filter in einer direkten Korrespondenz zu den Kongruenzrelationen stehen. Da Kongruenzrelationen eng mit Homomorphismen verbunden sind, ergibt sich zwischen allen drei Begriffen eine Reihe interessanter Querbeziehungen. Diese werden nachfolgend dargestellt, wobei wir auch auf die in Kapitel 6.5 vorgestellten Techniken zur Vereinfachung von Strukturen mittels Quotientenbildung eingehen.

Zunächst halten wir fest, wie sich die allgemeine Definition von Homomorphismen im speziellen Fall der Booleschen Algebren liest.

Bemerkung 9.10 Es seien $\mathcal{B} = \langle B, \sqcup_B, \sqcap_B, -_B \rangle$ und $\mathcal{C} = \langle C, \sqcup_C, \sqcap_C, -_C \rangle$ Boolesche Algebren mit Nullelementen 0_B bzw. 0_C. Eine Abbildung $h\colon B \to C$ ist ein Homomorphismus von $\mathcal{B}$ nach $\mathcal{C}$ genau dann, wenn gilt:

$$\begin{aligned} \forall a, b \in B\colon h(a \sqcup_B b) &= h(a) \sqcup_C h(b), \\ \forall a, b \in B\colon h(a \sqcap_B b) &= h(a) \sqcap_C h(b), \\ \forall b \in B\colon h(-_B b) &= -_C h(b). \end{aligned}$$

Definition 9.11 In der Situation von Bemerkung 9.10 heißt die Menge $ker(h) := \{b \in B \mid h(b) = 0_C\}$ der *Kern* der Abbildung h.

In Bemerkung 9.10 waren wir so genau, für die in der Tat im allgemeinen verschiedenen Operationen auf $\mathcal{B}$ und $\mathcal{C}$ auch verschiedene Symbole zu verwenden. Im nachfolgenden werden wir darauf verzichten, wenn klar ist, auf welche Algebra sich die Operationen beziehen.

Beispiel 9.12 Die in Abbildung 9.1 dargestellte Abbildung h ist ein Homomorphismus von $\mathbf{P_3}$ in $\mathbf{P_1}$. Der Kern ist das Ideal $I := \{0, r, z, t\}$ von $\mathbf{P_3}$.

Gemäß Definition 6.45 ist eine Äquivalenzrelation „$\sim$“ auf der Grundmenge B einer Booleschen Algebra $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Kongruenzrelation auf $\mathcal{B}$ genau dann, wenn gilt:

1. $\forall a, a', b, b' \in B\colon (a \sim a' \wedge b \sim b') \Rightarrow a \sqcup b \sim a' \sqcup b'$,

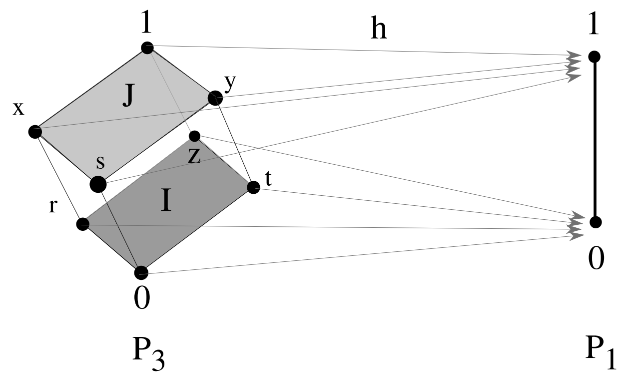


Abbildung 9.1: Homomorphismus von $\mathbf{P_3}$ in $\mathbf{P_1}$.

2. $\forall a, a', b, b' \in B$: $(a \sim a' \land b \sim b') \Rightarrow a \sqcap b \sim a' \sqcap b'$,

3. $\forall b, b' \in B$: $(b \sim b') \Rightarrow -b \sim -b'$.

Nachfolgend bezeichnet $[0]_\sim$ stets die Äquivalenzklasse der Null 0 von $\mathcal{B}$ bezüglich „$\sim$“.

Beispiel 9.13 In Abbildung 9.1 repräsentieren die unterlegten Teilmengen eine Kongruenzrelation „$\sim_h$“ auf $\mathbf{P_3}$ mit den Äquivalenzklassen $I = \{0, r, z, t\}$ und $J = \{1, x, s, y\}$. Es ist „$\sim_h$“ die durch den Homomorphismus h induzierte Kongruenzrelation auf $\mathbf{P_3}$ im Sinn von Lemma 6.46. Der Kern von h ist das Ideal I, das mit der Äquivalenzklasse $[0]_{\sim_h}$ übereinstimmt.

Lemma 9.14 *Ist „$\sim$“ eine Kongruenzrelation der Booleschen Algebra $\mathcal{B}$ mit Nullelement* 0, *so ist* $[0]_\sim$ *ein Ideal.*

Beweis. Sei „$\sim$“ eine Kongruenzrelation auf $\mathcal{B}$. Wir zeigen, daß $[0]_\sim$ die drei Bedingungen aus Definition 8.26 erfüllt.

1. Offenkundig ist $0 \in [0]_\sim \neq \emptyset$.

2. Es gelte $a, b \in [0]_\sim$. Dann folgt $a \sim 0$ und $b \sim 0$ und aufgrund der Kongruenzregeln auch $a \sqcup b \sim 0 \sqcup 0$, also $a \sqcup b \sim 0$ und $a \sqcup b \in [0]_\sim$.

3. Es gelte $a \in [0]_\sim$, beziehungsweise $a \sim 0$, und $b \leq a$. Dann folgt aus $a \sqcap b \sim 0 \sqcap b$ und $a \sqcap b = b$ sowie $b \sqcap 0 = 0$ nun $b \sim 0$ und damit $b \in [0]_\sim$. ■

Lemma 9.15 *Es sei I ein Ideal der Booleschen Algebra $\mathcal{B}$. Dann ist die Relation „$\sim_I$“, die durch*

$$a \sim_I b \quad :\Leftrightarrow \quad \exists x, y \in I\colon a \sqcup x = b \sqcup y$$

definiert ist, eine Kongruenzrelation auf $\mathcal{B}$. Es ist $[0]_{\sim_I}$ gerade das Ideal I.

Beweis. Wir zeigen zunächst, daß „$\sim_I$“ eine Äquivalenzrelation ist. Die *Reflexivität* folgt aus $I \neq \emptyset$. Die *Symmetrie* ergibt sich sofort aus der Definition von „$\sim_I$“. Zur *Transitivität*: es gelte $a \sim_I b$ und $b \sim_I c$. Damit existieren $x, y, r, s \in I$ so daß $a \sqcup x = b \sqcup y$ sowie $b \sqcup r = c \sqcup s$. Da I Ideal ist, sind auch $x \sqcup r$ und $s \sqcup y$ in I. Es gilt

$$\begin{aligned} a \sqcup (x \sqcup r) &= (a \sqcup x) \sqcup r = (b \sqcup y) \sqcup r \\ &= (b \sqcup r) \sqcup y = (c \sqcup s) \sqcup y = c \sqcup (s \sqcup y), \end{aligned}$$

woraus $a \sim_I c$ folgt.

Nun verifizieren wir die Gültigkeit der Kongruenzregeln. Es gelte $a \sim_I a'$ und $b \sim_I b'$. Damit existieren $x, x', y, y' \in I$ so daß $a \sqcup x = a' \sqcup x'$ sowie $b \sqcup y = b' \sqcup y'$. Zunächst zeigen wir, daß „$\sim_I$“ die *Kongruenzregel für „$\sqcup$“* erfüllt. Da I Ideal ist, sind $x \sqcup y$ und $x' \sqcup y' \in I$. Es gilt

$$\begin{aligned} (a \sqcup b) \sqcup (x \sqcup y) &= \ldots = (a \sqcup x) \sqcup (b \sqcup y) \\ &= (a' \sqcup x') \sqcup (b' \sqcup y') = \ldots = (a' \sqcup b') \sqcup (x' \sqcup y'), \end{aligned}$$

also $a \sqcup b \sim_I a' \sqcup b'$.

Wir zeigen nun, daß „$\sim_I$“ die *Kongruenzregel für „$\sqcap$“* erfüllt. Da I Ideal ist, sind auch $(a \sqcap y) \sqcup (x \sqcap b) \sqcup (x \sqcap y)$ sowie $(a' \sqcap y') \sqcup (x' \sqcap b') \sqcup (x' \sqcap y')$ in I. Es gilt

$$\begin{aligned} & (a \sqcap b) \sqcup (a \sqcap y) \sqcup (x \sqcap b) \sqcup (x \sqcap y) \\ = \; & (a \sqcap (b \sqcup y)) \sqcup (x \sqcap (b \sqcup y)) \\ = \; & (a \sqcup x) \sqcap (b \sqcup y) = (a' \sqcup x') \sqcap (b' \sqcup y') \\ = \; & (a' \sqcap b') \sqcup (a' \sqcap y') \sqcup (x' \sqcap b') \sqcup (x' \sqcap y'), \end{aligned}$$

somit gilt auch $a \sqcap b \sim_I a' \sqcap b'$.

Schließlich zeigen wir, daß „$\sim_I$“ die *Kongruenzregel für die Komplementbildung „$-$“* erfüllt. Aus $a \sqcup x = a' \sqcup x'$ folgt gemäß Lemma 9.9, Teil 1, daß $-a \sqcap -x = -a' \sqcap -x'$ gilt. Hieraus ergibt sich

$$(-a \sqcap -x) \sqcup (x \sqcup x') = (-a' \sqcap -x') \sqcup (x \sqcup x'),$$

damit folgt wegen der Distributivität

$$(-a \sqcup x \sqcup x') \sqcap (-x \sqcup x \sqcup x') = (-a' \sqcup x \sqcup x') \sqcap (-x' \sqcup x \sqcup x').$$

Aus Lemma 9.7 ergibt sich nun $(-a \sqcup x \sqcup x') = (-a' \sqcup x \sqcup x')$. Somit gilt $-a \sim_I -a'$. Wir haben somit gezeigt, daß „$\sim_I$“ eine Kongruenzrelation ist.

Für Idealelemente $i \in I$ gilt $i \sqcup 0 = 0 \sqcup i$, woraus $i \sim_I 0$ folgt. Gilt andererseits $i \sim_I 0$, so existieren $x, x' \in I$ wo $i \sqcup x = 0 \sqcup x' \in I$. Es folgt $i = i \sqcap (i \sqcup x) = i \sqcap (0 \sqcup x') \in I$. ■

Beispiel 9.16 In Abbildung 9.1 stimmt die Kongruenzrelation „$\sim_h$“ auf $\mathbf{P_3}$ mit der durch das Ideal I definierten Kongruenzrelation „$\sim_I$“ überein.

Das nachfolgende Lemma zeigt, daß es zu einem Ideal I immer genau eine Kongruenzrelation „$\sim$“ gibt, so daß $I = [0]_\sim$ gilt. In diesem Sinn ist jede Kongruenzrelation „$\sim$“ vollständig durch die Angabe des Ideals $[0]_\sim$ festgelegt.

Lemma 9.17 *Es sei „$\sim$“ eine Kongruenzrelation auf der Booleschen Algebra $\mathcal{B}$. Bezeichnet I das Ideal $[0]_\sim$, und ist „$\sim_I$“ die gemäß Lemma 9.15 durch I bestimmte Kongruenzrelation, so sind „$\sim$“ und „$\sim_I$“ identisch.*

Beweis. Um die Gleichheit der Relationen $\sim$ und $\sim_I$ zu beweisen, zeigen wir zuerst $\sim \subseteq \sim_I$. Falls $a \sim b$, so folgt aufgrund der Kongruenzregeln $a \sqcap -b \sim b \sqcap -b = 0$, also gilt $a \sqcap -b \in I$. Analog folgt $b \sqcap -a \in I$. Nun erhalten wir

$$\begin{aligned} b \sqcup (a \sqcap -b) &= (b \sqcup a) \sqcap (b \sqcup -b) = b \sqcup a \\ &= a \sqcup b = (a \sqcup b) \sqcap (a \sqcup -a) = a \sqcup (b \sqcap -a), \end{aligned}$$

wodurch sich $a \sim_I b$ ergibt.

Gilt umgekehrt $a \sim_I b$, so existieren $x, y \in [0]_\sim = I$ mit $a \sqcup x = b \sqcup y$. Es folgt wegen der Kongruenzregeln

$$a = a \sqcup 0 \sim_S a \sqcup x = b \sqcup y \sim_S b \sqcup 0 = b,$$

damit gilt aber $a \sim b$. ■

Gemäß Lemma 6.46 induziert jeder Homomorphismus h zwischen Booleschen Algebren $\mathcal{B}$ und $\mathcal{C}$ eine Kongruenzrelation „$\sim_h$“ auf B vermöge $b \sim_h b' \;:\Leftrightarrow h(b) = h(b')$. Aus den Lemmas 9.14, 9.15 und 9.17 erhalten wir die folgende Beobachtung.

Korollar 9.18 *Ist h ein Homomorphismus zwischen Booleschen Algebren $\mathcal{B}$ und $\mathcal{C}$, so ist der Kern $\ker(h)$ ein Ideal von $\mathcal{B}$. Für $b, b' \in B$ gilt $b \sim_{\ker(h)} b'$ genau dann, wenn $h(b) = h(b')$.* ■

Ist umgekehrt I ein Ideal der Booleschen Algebra $\mathcal{B}$, und ist „$\sim_I$" die gemäß Lemma 9.15 abgeleitete Kongruenzrelation, so können wir wie in Lemma 6.47 beschrieben die Quotientenalgebra $\mathcal{B}/\sim_I$ bilden. Offenkundig ist dann I gerade der Kern des kanonischen Homomorphismus $\kappa : B \to B/\sim_I, b \mapsto [b]_{\sim_I}$. Mit Korollar 9.18 folgt nun

Lemma 9.19 *Jedes Ideal einer Booleschen Algebra ist der Kern eines Homomorphismus und umgekehrt.*

Als Spezialfall des in Lemma 6.49 dargestellten allgemeinen Sachverhalts halten wir noch folgende Beobachtung fest.

Lemma 9.20 *In der Situation von Korollar 9.18 ist $\mathcal{C}' = \langle h(B), \sqcup, \sqcap, - \rangle$ wieder eine Boolesche Algebra. Die Abbildung $g : [b]_{\sim_{\ker(h)}} \mapsto h(b)$ ist ein Isomorphismus von $\mathcal{B}/\sim_{\ker(h)}$ auf $\mathcal{C}'$.*

Ein abschließendes Beispiel soll diesen Zusammenhang illustrieren.

Beispiel 9.21 Wir betrachten den in Abbildung 9.2 dargestellten Homomorphismus h von $\mathbf{P_3}$ in $\mathbf{P_1}$. Der Kern von h ist das Ideal $I := \{0, r, z, t\}$ von $\mathbf{P_3}$. Die entsprechende Kongruenzrelation „$\sim_I$" ist durch die Partition $\{I, J\}$ bestimmt, wo $J := \{x, s, y, 1\}$. Die Quotientenalgebra $\mathbf{P_3}/\sim_I$ ist isomorph zu $\mathbf{P_1}$.

9.3 Primideale

Eine besondere Rolle spielen diejenigen Ideale, die als Kerne von Homomorphismen in die Boolesche Algebra $\mathbf{2}$ auftreten. Der folgende Satz zeigt, daß sie auf sehr unterschiedliche Art und Weise charakterisiert werden können.

Lemma 9.22 *Sei $I \subset B$ ein eigentliches Ideal von $\mathcal{B}$. Dann sind äquivalent:*

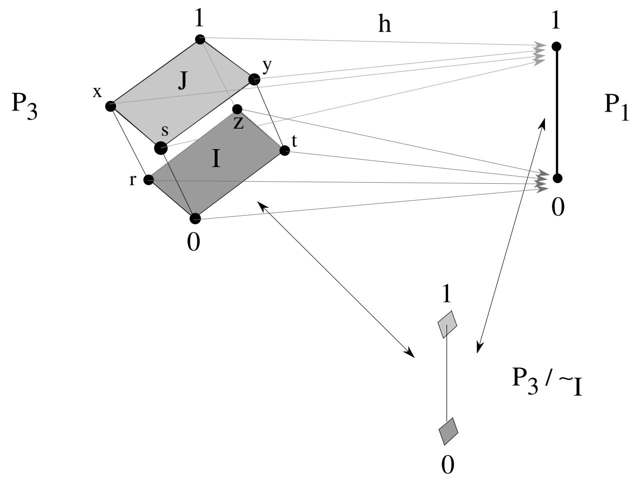


Abbildung 9.2: Isomorphie zwischen homomorphem Bild und Quotientenstruktur.

1. *Es existiert ein Homomorphismus g von $\mathcal{B}$ auf* $\mathbf{2}$ *mit* $I = \ker(g)$,

2. *Für alle* $b \in B$ *ist entweder* $b \in I$ *oder* $-b \in I$,

3. *für alle* $a, b \in B$ *gilt: wenn* $a \sqcap b \in I$, *dann ist* $a \in I$ *oder* $b \in I$,

4. I *ist maximal, das heißt es gibt kein eigentliches Ideal* J *von* $\mathcal{B}$ *mit* $I \subset J$.

Beweis. „$1 \to 2$“ : Wären $b \in I$ und $-b \in I$, so wäre $1_B = b \sqcup -b$ in I, und I wäre damit nach Lemma 8.28, Teil 3, kein eigentliches Ideal. Also liegen nicht b und $-b$ beide in I. Würde nun weder b noch $-b$ in $I = ker(g)$ liegen, so wäre $g(b) = g(-b) = 1_2 = g(b) \sqcap g(-b)$. Weil g homomorph ist, ergibt sich $1_2 = g(b) \sqcap -g(b) = 0_2$, also ein Widerspruch.

„$2 \to 3$“ : Wenn $a \notin I$ und $b \notin I$, so gilt $-a \in I$ und $-b \in I$, also $-a \sqcup -b = -(a \sqcap b) \in I$ und $a \sqcap b \notin I$.

„$3 \to 4$“ : Es sei J ein Ideal von $\mathcal{B}$ mit $I \subset J$. Wir zeigen, daß $J = B$ nicht eigentlich ist: sei $a \in J \setminus I$. Da $a \sqcap -a = 0 \in I$ folgt nach Voraussetzung $-a \in I$, somit $-a \in J$ und $1 = a \sqcup -a \in J$. Somit $J = B$ nach Lemma 8.28, Teil 3.

„4 → 1" : Sei I maximal. Für $b \notin I$ ist das kleinste $\{b\} \cup I$ umfassende Ideal gerade

$$\{a \in B \mid \exists i \in I : a \leq b \sqcup i\}.$$

Wegen der Maximalität von I gilt $1 = b \sqcup i$ für ein $i \in I$. Wir zeigen nun:

$$(*)\text{: mit } b \notin I \text{ und } b' \notin I \text{ folgt auch } b \sqcap b' \notin I.$$

Es sei nämlich $1 = b \sqcup i = b' \sqcup i'$ mit passenden $i, i' \in I$. Dann folgt

$$1 = (b \sqcup i) \sqcap (b' \sqcup i') = (b \sqcap b') \sqcup (b \sqcap i') \sqcup (i \sqcap b') \sqcup (i \sqcap i').$$

Wäre $b \sqcap b' \in I$, so auch 1, was nicht sein kann. Mit Hilfe von (*) läßt sich nun leicht verifizieren, daß die Abbildung g, die I auf 0_2 und $B \setminus I$ auf 1_2 wirft, ein Homomorphismus von $\mathcal{B}$ auf $\mathbf{2}$ ist (vgl. Aufgabe 5). ■

Definition 9.23 Ein *Primideal* ist ein eigentliches Ideal I, das eine der Anforderungen 1–4 aus Lemma 9.22 erfüllt.

Beispiel 9.24 Es sei M eine nichtleere Menge, für $X \subseteq M$ sei $-X$ das Komplement $M \setminus X$ von X. Sei $b \in M$. Dann die Menge aller Teilmengen von $M \setminus \{m\}$ ein Primideal der Booleschen Algebra $\langle \mathcal{P}(M), \cup, \cap, - \rangle$.

Korollar 9.25 *Ein Ideal I einer Booleschen Algebra $\mathcal{B}$ ist genau dann ein Primideal, wenn $\mathcal{B}/R(I)$ isomorph zur Booleschen Algebra $\mathbf{2}$ ist.*

Beweis. Ist I Primideal, so gibt es einen Homomorphismus g von $\mathcal{B}$ auf $\mathbf{2}$ mit $I = ker(g)$. Offenkundig ist g surjektiv. Aus Lemma 9.20 folgt, daß $\mathcal{B}/R(I)$ isomorph zu $\mathbf{2}$ ist. Es sei umgekehrt $\mathcal{B}/R(I)$ isomorph zu $\mathbf{2}$. Dann existiert offenkundig ein Homomorphismus h von $\mathcal{B}$ über $\mathcal{B}/R(I)$ nach $\mathbf{2}$, so daß $ker(h)$ der Kern des kanonischen Homomorphismus $\pi_{R(I)}$, also nach Korollar **??** gerade I ist. Also ist I ein Primideal. ■

Lemma 9.26 *Ist $I \subset B$ ein eigentliches Ideal einer Booleschen Algebra $\mathcal{B}$, so gibt es ein Primideal $\hat{I}$ von $\mathcal{B}$ mit $I \subseteq \hat{I}$.*

Beweis. Der einfachste Beweis benutzt das sogenannte Zornsche Lemma.[1] Es besagt: wenn in einer partiell geordneten Menge $\langle M, \leq \rangle$ jede linear geordnete Teilmenge eine obere Schranke besitzt, dann hat $\langle M, \leq \rangle$ ein maximales Element.
Sei nun I ein eigentliches Ideal von $\mathcal{B}$, weiterhin sei M die Menge der eigentlichen Ideale J von $\mathcal{B}$ mit $I \subseteq J$, geordnet durch die mengentheoretische Inklusion. Ist $K \subseteq M$ eine linear geordnete Teilmenge, so ist $\bigcup K$ wieder ein I enthaltendes Ideal. Da $1 \notin \bigcup K$ ist $\bigcup K$ eigentliches Ideal, also eine obere Schranke von K in M. Nach Zorns Lemma enthält nun M ein maximales eigentliches Ideal $\hat{I}$. ■

Lemma 9.27 *Es sei $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra, $1 \neq b \in B$ und $a \in B$, es gelte* nicht *$a \leq b$. Dann gibt es ein Primideal J von $\mathcal{B}$ mit $b \in J$, $a \notin J$.*

Beweis. Da $a \not\leq b$ gilt $-a \sqcup b \neq 1$ nach Lemma 9.9, Teil 5. Das von $\{-a \sqcup b\}$ erzeugte Ideal ist $I = \{c \in B \mid c \leq -a \sqcup b\}$. I ist ein eigentliches Ideal und enthält offensichtlich $-a$ und b. Andererseits gilt $a \notin I$, andernfalls läge nämlich auch $1 = a \sqcup -a \in I$. Wir betrachten die Menge $M(I, a)$ der eigentlichen Ideale von $\mathcal{B}$, die I enthalten, nicht aber a. Ist $K \subseteq M(I, a)$ eine totalgeordnete Teilmenge, so ist $\bigcup K$ wieder ein I enthaltendes eigentliches Ideal mit $a \notin \bigcup K$, also eine obere Schranke. Nach Zorns Lemma enthält nun $M(I, a)$ ein maximales eigentliches Ideal J. Offenkundig gilt $b \in J$, $a \notin J$. Wäre nun J kein Primideal, so könnte man nach dem vorausgegangenen Lemma J in ein Primideal $\hat{J}$ einbetten. Nach Definition von J muß $\hat{J}$ dann das Element a enthalten, also auch $1 = a \sqcup -a$. Somit wäre $\hat{J}$ kein eigentliches Ideal. Daher ist J ein Primideal. ■

9.4 Mengenalgebren und Stones Theorem

In diesem Abschnitt werden wir zeigen, daß man alle Boolesche Algebren mit Hilfe geeigneter Mengen in einem zu klärenden Sinn darstellen kann.

Definition 9.28 Eine *Mengenalgebra* ist eine Boolesche Teilalgebra einer Booleschen Algebra der Form $\langle P(M), \cup, \cap, - \rangle$, wo M Menge ist.

[1] Das Zornsche Lemma ist zum sogenannten Auswahlaxiom äquivalent und von den Standard-Axiomen der Zermelo-Fraenkel Mengenlehre unabhängig, mit diesen aber verträglich.

Beispiel 9.29 Es sei B die Menge aller endlichen oder co-endlichen Teilmengen von $\mathbb{N}$. Hierbei heißt eine Menge co-endlich, wenn ihr Komplement in $\mathbb{N}$ endlich ist. B ist unter Durchschnitten, Vereinigungen und Komplementbildung (in $\mathbb{N}$) abgeschlossen. Daher ist $\langle B, \cup, \cap, - \rangle$ eine Mengenalgebra.

Das nachfolgende Theorem besagt, daß sich jede Boolesche Algebra als eine Mengenalgebra auffassen läßt.

Theorem 9.30 (Stonesches Repräsentationstheorem) *Jede Boolesche Algebra ist isomorph zu einer Mengenalgebra.*

Beweis. Es sei $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine beliebige Boolesche Algebra. Wir ordnen jedem $b \in B$ die Menge

$$g(b) := \{P \mid P \text{ ist ein Primideal von } B \text{ mit } b \notin P\}$$

zu. Wir zeigen nun, daß g ein Monomorphismus von $\mathcal{B}$ in die Mengenalgebra der Menge aller Primideale von $\mathcal{B}$ ist.

Homomorphie bzgl. „$\sqcap$“: wir zeigen, daß stets gilt:

$$\begin{aligned}\{P \mid P \text{ Primideal}, a \notin P\} \cap \{P \mid P \text{ Primideal}, b \notin P\} \\ = \{P \mid P \text{ Primideal}, a \sqcap b \notin P\}.\end{aligned}$$

Daraus folgt sofort $g(a \sqcap b) = g(a) \cap g(b)$. Die Inklusion „$\subseteq$“ folgt aus Lemma 9.22, Teil 3, die umgekehrte Inklusion ist trivial, da mit $a \in P$ oder $b \in P$ stets $a \sqcap b \in P$ gilt.

Homomorphie bzgl. „$\sqcup$“: wir zeigen, daß stets gilt:

$$\begin{aligned}\{P \mid P \text{ Primideal}, a \notin P\} \cup \{P \mid P \text{ Primideal}, b \notin P\} \\ = \{P \mid P \text{ Primideal}, a \sqcup b \notin P\}.\end{aligned}$$

Daraus folgt sofort $g(a \sqcup b) = g(a) \cup g(b)$. Die Inklusion „$\subseteq$“ folgt aus $a, b \leq a \sqcup b$, die umgekehrte Inklusion ist trivial.

Homomorphie bzgl. „$-$“: Mit Lemma 9.22, Teil 2, folgt sofort $g(-b) = -g(b)$.

Injektivität: Es seien $a, b \in B$, $a \neq b$. Falls $a = 1$, so ist $g(a) =$ Menge aller Primideale von B. Nach Lemma 9.27 gibt es ein Primideal, das b enthält, das heißt $g(b) \neq g(a)$. Umgekehrt folgt aus $b = 1$ genauso $g(b) \neq g(a)$. Gilt $a \neq 1 \neq b$, so gilt ohne Beschränkung der Allgemeinheit, daß nicht $a \leq b$. Aus Lemma 9.27 folgt nun wieder, daß $g(a) \neq g(b)$. ■

9.5 Ergänzungen

Zum Abschluß des Kapitels geben wir nachfolgend eine vollständige Klassifikation aller *endlichen* Booleschen Algebren. Schließlich stellen wir noch kurz eine alternative Beschreibung Boolescher Algebren in Gestalt sogenannter Boolescher Ringe vor.

9.5.1 Endliche Boolesche Algebren

Die Standardbeispiele für Boolesche Algebren sind die Algebren der Form $\langle \mathcal{P}(M), \cup, \cap, - \rangle$ wo $\mathcal{P}(M)$ die Potenzmenge der Menge M bezeichnet. Da hier die Supremumsbildung mit der Vereinigung übereinstimmt, läßt sich jedes Element $N \neq \emptyset$ von $\mathcal{P}(M)$ in eindeutiger Weise als Supremum der nichtleeren Menge $\{\{n\} \mid n \in N\}$ darstellen. Die beteiligten Elemente $\{n\}$ sind unmittelbare Nachfolger der Null $0 = \emptyset$. Aus diesem Blickwinkel kann man die Elemente der Form $\{m\}$ für $m \in M$ als atomare Bausteine der Booleschen Algebra betrachten. Alle anderen Elemente sind in eindeutiger Weise aus solchen Bausteinen zusammengesetzt. In diesem Abschnitt zeigen wir, daß diese Perspektive auf alle anderen endlichen Booleschen Algebren übertragbar ist und zu einer vollständigen Klassifikation der endlichen Booleschen Algebren führt.

Definition 9.31 Ein Element $b \neq 0$ einer Booleschen Algebra $\mathcal{B}$ heißt *Atom* von $\mathcal{B}$ genau dann, wenn b nicht in der Form $b = a \sqcup c$ mit $a \neq b \neq c$ dargestellt werden kann.

Lemma 9.32 *Ein Element $b \neq 0$ einer Booleschen Algebra $\mathcal{B}$ ist ein Atom genau dann, wenn es kein Element $c \in B$ gibt mit $0 < c < b$.*

Beweis. „$\Rightarrow$“ : Sei b ein Atom. Angenommen es existiert ein $c \in B$ mit $0 < c < b$. Dann ist $c \sqcup b = b$ und damit

$$\begin{aligned} b &= b \sqcap 1 = (c \sqcup b) \sqcap (c \sqcup -c) \\ &= ((c \sqcup b) \sqcap c) \sqcup ((c \sqcup b) \sqcap -c) = c \sqcup 0 \sqcup (b \sqcap -c) \\ &= c \sqcup (b \sqcap -c). \end{aligned}$$

Da b ein Atom und von c verschieden ist, folgt $b = b \sqcap -c$. Wir erhalten $c = b \sqcap c = (b \sqcap -c) \sqcap c = b \sqcap 0 = 0$, was einen Widerspruch darstellt. Also kann es kein $c \in B$ geben mit $0 < c < b$.

„$\Leftarrow$“ : Es sei $0 \neq b \in B$ und es existiere kein $c \in B$ mit $0 < c < b$. Sei $b = a \sqcup c$ und $b \neq a$. Dann folgt $a, c \leq b$, $a = 0$ und $b = 0 \sqcup c = c$. Also ist b ein Atom. ■

Aus Lemma 9.32 folgt, daß jede endliche Boolesche Algebra, die zumindest zwei Elemente hat, dann auch zumindest ein Atom besitzt.

Lemma 9.33 *Es seien a und b Atome der Booleschen Algebra $\mathcal{B}$. Falls $a \neq b$ gilt, so ist $a \sqcap b = 0$.*

Beweis. Offenkundig gilt $0 \leq a \sqcap b \leq a, b$. Da a Atom ist, folgt im Fall $a \sqcap b \neq 0$ gemäß Lemma 9.32 $a \sqcap b = a$, damit $0 < a \leq b$. Da b Atom ist, zeigt Lemma 9.32 nun $a = b$. ■

Lemma 9.34 *Sei $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine endliche Boolesche Algebra. Dann kann jedes Element $b \neq 0$ von B in der Form $b = c_1 \sqcup \ldots \sqcup c_n$ $(n \geq 1)$ dargestellt werden, wo die Elemente $c_1, \ldots, c_n$ Atome sind. Diese Darstellung ist bis auf die Reihenfolge der Atome eindeutig und umfaßt genau die Atome c mit $c \leq b$.*

Beweis. Wir zeigen zunächst (*Existenz*), daß jedes Element von B als join von Atomen dargestellt werden kann. Angenommen, es gibt ein von Null verschiedenes Element von B, das sich nicht in der obigen Form darstellen läßt. Unter allen solchen Elementen wählen wir ein Element b, das bezüglich der partiellen Ordnung „$\leq$“ (siehe Bem. 9.2) auf B minimal ist. Offenkundig kann b kein Atom sein. Damit gibt es eine Darstellung $b = b_1 \sqcup b_2$, wo b_1 und b_2 beide von b verschieden sind, woraus sich $b_1, b_2 < b$ ergibt. Wäre etwa $b_1 = 0$, so folgte $b = 0 \sqcup b_2 = b_2$, was ausgeschlossen ist. Also gilt $b_1 \neq 0$ und symmetrisch $b_2 \neq 0$. Nach Wahl von b gibt es damit Darstellungen

$$\begin{aligned} b_1 &= c_1 \sqcup \ldots \sqcup c_m \\ b_2 &= c'_1 \sqcup \ldots \sqcup c'_n \\ b = b_1 \sqcup b_2 &= c_1 \sqcup \ldots \sqcup c_m \sqcup c'_1 \sqcup \ldots \sqcup c'_n \end{aligned}$$

was einen Widerspruch darstellt. Daher lassen sich alle von Null verschiedenen Elemente von B in der angegebenen Form als join von Atomen darstellen.

Wir zeigen nun (*Form*), daß für jedes Element b eine Darstellung $b = c_1 \sqcup \ldots \sqcup c_m$ existiert, wo die beteiligten Atome c_i genau die Atome c mit $c \leq b$ sind. Ist $\{a_1, \ldots, a_k\}$ die Menge aller Atome von B, so kann nach dem vorausgegangenen Argument die Eins 1 als join einer Menge von Atomen, daher auch als join über alle Atome in der Form $1 = a_1 \sqcup \ldots \sqcup a_k$ dargestellt werden. Für $b \in B$ folgt

$$b = b \sqcap 1 = b \sqcap (a_1 \sqcup \ldots \sqcup a_k) = (b \sqcap a_1) \sqcup \ldots \sqcup (b \sqcap a_k).$$

Da die Elemente a_i Atome sind, gilt entweder $b \sqcap a_i = 0$, in diesem Fall kann der Teilausdruck $b \sqcap a_1$ aus der Darstellung von b eliminiert werden, oder $b \sqcap a_i = a_i$, woraus sich $a_i \leq b$ ergibt ($1 \leq i \leq k$). Demzufolge gilt $b = c_1 \sqcup \ldots \sqcup c_m$ wo $\{c_1, \ldots, c_m\}$ die Menge aller Atome c mit $c \leq b$ ist.

Schließlich bleibt die *Eindeutigkeit* dieser Darstellung bis auf Reihenfolge der Atome zu verifizieren. Hierzu sei $b = c'_1 \sqcup \ldots \sqcup c'_n$ eine Darstellung von b als join über die Menge von Atomen $\{c'_1, \ldots, c'_n\}$. Dann gilt $c'_i \leq b$ für $1 \leq i \leq n$, also $\{c'_1, \ldots, c'_n\} \subseteq \{c_1, \ldots, c_m\}$. Für $1 \leq i \leq m$ gilt andererseits

$$0 \neq c_i = c_i \sqcap b = c_i \sqcap (c'_1 \sqcup \ldots \sqcup c'_n) = (c_i \sqcap c'_1) \sqcup \ldots \sqcup (c_i \sqcap c'_n)$$

Es muß daher zumindest ein Glied $c_i \sqcap c'_j$ geben, das von Null verschieden ist. Da c_i Atom ist, gilt $c_i = c'_j$. Daraus folgt $\{c_1, \ldots, c_m\} \subseteq \{c'_1, \ldots, c'_n\}$. ■

Im nachfolgenden sei $\mathcal{B}$ eine endliche Boolesche Algebra. Für jedes $b \in B$ bezeichne $At(b)$ die Menge der Atome a von $\mathcal{B}$ mit $a \leq b$. Ist $A = \{a_1, \ldots, a_n\}$ eine nichtleere Menge von Atomen, so steht $\bigsqcup_{a \in A} a$ für $a_1 \sqcup \ldots \sqcup a_n$. Das Element 0 fassen wir auf als join über die leere Menge von Atomen.

Lemma 9.35 *Es seien b_1, b_2 beliebige Elemente von B. Dann gilt*

$$\begin{aligned} At(b_1 \sqcup b_2) &= At(b_1) \cup At(b_2), \\ At(b_1 \sqcap b_2) &= At(b_1) \cap At(b_2), \\ At(-b_1) &= At(1) \setminus At(b_1). \end{aligned}$$

Beweis. Wir können $b_1 \sqcup b_2$ darstellen in der Form

$$b_1 \sqcup b_2 = (\bigsqcup_{a \in At(b_1)} a) \sqcup (\bigsqcup_{a \in At(b_2)} a).$$

Aufgrund der Assoziativität, Kommutativität und Idempotenz von „$\sqcup$“ folgt $b_1 \sqcup b_2 = \bigsqcup_{a \in At(b_1) \cup At(b_2)} a$. Nach Lemma 9.34 ist dann $At(b_1) \cup At(b_2)$ die Menge der Atome a mit $a \leq b_1 \sqcup b_2$.

Ähnlich stellen wir $b_1 \sqcap b_2$ dar als

$$b_1 \sqcap b_2 = (\bigsqcup_{a \in At(b_1)} a) \sqcap (\bigsqcup_{a \in At(b_2)} a).$$

Mittels der zusätzlichen Verwendung des Distributivgesetzes erhalten wir

$$b_1 \sqcap b_2 = \bigsqcup_{a_1 \in At(b_1), a_2 \in At(b_2)} (a_1 \sqcap a_2).$$

Nach Lemma 9.33 können wir alle Summanden $a_1 \sqcap a_2$ mit $a_1 \neq a_2$ durch 0 ersetzen. Es bleiben die Ausdrücke der Form $a \sqcap a$ übrig, wo $a \in At(b_1) \cap At(b_2)$. Nach Lemma 9.34 ist dann $At(b_1) \cap At(b_2)$ die Menge der Atome a mit $a \leq b_1 \sqcap b_2$.

Mit denselben Methoden ergibt sich

$$\begin{aligned} b_1 \sqcup (\bigsqcup_{a \in At(1) \setminus At(b_1)} a) &= 1 \\ b_1 \sqcap (\bigsqcup_{a \in At(1) \setminus At(b_1)} a) &= 0. \end{aligned}$$

Nach Lemma 9.7 folgt $-b_1 = \bigsqcup_{a \in At(1) \setminus At(b_1)} a$ und $At(-b_1) = At(1) \setminus At(b_1)$. ■

Wir erhalten nun die gewünschte Klassifikation aller endlichen Booleschen Algebren: die Zahl der Atome legt die Algebra bis auf Isomorphie bereits fest.

Satz 9.36 *Jede endliche Boolesche Algebra mit n Atomen ist isomorph zur Algebra $\mathbf{P_n}$ und besitzt genau 2^n Elemente.*

Beweis. Es sei $\mathcal{B}$ eine endliche Boolesche Algebra mit den Atomen $a_1, \ldots, a_n$. Sei $M := \{a_1, \ldots, a_n\}$. Die Abbildung $h : b \mapsto At(b)$ ist nach Lemma 9.35 ein Homomorphismus von $\mathcal{B}$ in $\langle \mathcal{P}(M), \cup, \cap, -_M \rangle$. Für $N \subseteq M$ gilt $N = h(\bigsqcup_{a \in N} a)$, somit ist h surjektiv. Aus $h(b_1) = h(b_2)$ folgt nach Lemma 9.34 auch $b_1 = b_2$, also ist h injektiv. ■

9.5.2 Boolesche Ringe

Man kann Boolesche Algebren in leicht modifizierter Form auch als Ringe im Sinn von Definition 6.13 charakterisieren. Da von dieser Beschreibungsweise häufig Gebrauch gemacht wird, geben wir kurz die Definition eines Booleschen Rings und die Übersetzungen zwischen beiden Begriffen an. Boolesche Ringe weisen im Vergleich zu anderen Ringen folgende Besonderheiten auf:

- Jedes Element b ist sein eigenes Inverses bezüglich der Ringaddition „$+$“.
- Die Ringmultiplikation „$*$“ ist kommutativ und idempotent, und
- es gibt ein Neutralelement 1 der Multiplikation.

Durch Hinzufügen der üblichen Regeln für Ringe (cf. Definition 6.13) ergibt sich folgende Definition.

Definition 9.37 Ein *Boolescher Ring* ist eine Algebra $\mathcal{B} = \langle B, +, 0, *, 1\rangle$ wo für beliebige Elemente $b, c, d \in B$ die folgenden Gleichungen erfüllt sind:

$$\begin{array}{rclcrcl} b+c &=& c+b & \qquad & b*c &=& c*b \\ (b+c)+d &=& b+(c+d) & & (b*c)*d &=& b*(c*d) \\ b+b &=& 0 & & b*b &=& b \\ 0+b &=& b & & 0*b &=& 0 \\ b*(c+d) &=& (b*c)+(b*d) & & 1*b &=& b. \end{array}$$

Um für Ausdrücke einer Booleschen Algebra die entsprechenden Ringausdrücke anzugeben, wendet man folgende Übersetzung an.

$$\begin{array}{rcl} b \sqcap c &\mapsto& b*c, \\ b \sqcup c &\mapsto& b+c+(b*c), \\ -b &\mapsto& 1+b. \end{array}$$

Die umgekehrte Übersetzung lautet wie folgt:

$$\begin{array}{rcl} b*c &\mapsto& b \sqcap c, \\ b+c &\mapsto& (b \sqcap -c) \sqcup (-b \sqcap c). \end{array}$$

9.6 Aufgaben zu Kapitel 9

Aufgaben zu Teilkapitel 9.1

Aufgabe 9.1 Stellen Sie $\mathbf{P_4}$ in Form eines Hasse-Diagramms dar.

Aufgabe 9.2 Ergänzen Sie die fehlenden Teile des Beweises von Lemma 9.9.

Aufgaben zu Teilkapitel 9.2

Aufgabe 9.3 Geben Sie alle Ideale der Booleschen Algebra $\mathbf{P_3}$ an. Wie sehen die zugehörigen Kongruenzrelationen aus, wie die daraus abzuleitenden Quotientenalgebren?

Aufgabe 9.4 Es seien I und J zwei Ideale der Booleschen Algebra $\mathcal{B}$. Zeigen Sie: falls I eine (echte) Teilmenge von J ist, so ist die zugehörige Kongruenzrelation „$\sim_I$“ eine (echte) Verfeinerung von „$\sim_J$“ im Sinn von Definition 4.15. Geben Sie einen Homomorphismus von $\mathcal{B}/{\sim_I}$ auf $\mathcal{B}/{\sim_J}$ an.

Aufgabe 9.5 Es sei $\mathcal{B}$ eine *endliche* Boolesche Algebra. Zeigen Sie:

1. Ist I ein Ideal von $\mathcal{B}$, so existiert $sup\ I \in I$ und ist das größte Element von I.

2. Definiert man die Relation $\equiv_I \subseteq B \times B$ durch
$$a \equiv_I b \;:\Leftrightarrow\; a \sqcup sup\ I = b \sqcup sup\ I,$$
so stimmt „$\equiv_I$“ mit der in Lemma 9.15 eingeführten Kongruenzelation „$\sim_I$“ überein.

3. Für jedes $b \in B$ ist die Kongruenzklasse $[b]_{\sim_I}$ stets ein Verband, das heißt unter Suprema und Infima abgeschlossen.

Aufgaben zu Teilkapitel 9.3

Aufgabe 9.6 Beenden Sie den Beweis von Lemma 9.22, indem Sie zeigen, daß die am Ende angegebene Abbildung g ein Homomorphismus von $\mathcal{B}$ auf $\mathbf{2}$ ist.

Aufgaben zu Teilkapitel 9.4

Aufgabe 9.7 Was passiert, wenn man im Beweis des Stoneschen Repräsentationstheorems die Definition von g abändert zu

$$g(b) := \{P \mid P \ \text{ ist ein Primideal von } B \ \text{ mit } \ b \in P\}?$$

Bei welchen Beweisschritten ergeben sich Probleme?

Aufgaben zu Teilkapitel 9.5

Aufgabe 9.8 Wie sehen die Atome der Booleschen Algebra $\mathbf{2}^n$ (vergl. Beispiel 9.5) aus? Wie sieht für ein Element b die Menge der Atome a mit $a \leq b$ aus? Geben Sie einen Isomorphismus auf die Boolesche Algebra $\mathbf{P_n}$ an.

9.7 Bibliographische Angaben

Eine ausführliche Darstellung Boolescher Algebren findet sich in [Sik70].

10

Klassische Aussagenlogik

In der *formalen Logik* versucht man, die Gesetzmäßigkeiten einer korrekten Argumentations- oder Schlußweise zu formalisieren. Zumindest drei Teilgebiete haben sich heute als eigene Forschungsfelder herauskristallisiert. Die *mathematische Logik* beschäftigt sich in erster Linie mit der Analyse mathematischer Schlußweisen und stellt damit neben der Mengenlehre ein zweites wichtiges Gebiet mathematischer Grundlagenforschung dar. In der *philosophischen Logik* werden allgemeiner Gesetzmäßigkeiten des Denkens und der Verwendung natürlicher Sprache untersucht. Die semantische Analyse natürlichsprachlicher Sätze und Diskurse bildet ein wichtiges Teilgebiet im Schnittpunkt von philosophischer Logik und (Computer-) Linguistik. In den letzten Jahrzehnten erlangte darüberhinaus die *Logik in der Informatik* eine immer stärkere Bedeutung. Hauptgegenstand hier ist die logische Analyse von Hard- und Software.

Die *klassische Logik* unterscheidet sich von anderen, historisch später entwickelten „konstruktiven" Logiken vor allem dadurch, daß die Existenz eines Objekts dadurch bewiesen werden kann, daß man die Annahme der Nichtexistenz zu einem Widerspruch führt. Es muß also kein geeignetes Beispielobjekt explizit angegeben oder konstruiert werden. Dies geht einher mit der Annahme der klassischen Logik, daß ein doppelt negierter Satz zum unnegierten Satz logisch äquivalent ist. Mathematisches Schließen beruht in der Regel auf klassischer Logik, wir werden uns im nachfolgenden auf klassische Logik beschränken.

Neben der Unterscheidung in konstruktive und nichtkonstruktive Logiksysteme gibt es ferner wichtige Unterschiede in Bezug auf die sprachlichen Ausdrucksmittel, die einer formalen Analyse unterzogen werden. Im Rahmen der *klassischen Aussagenlogik* beschäftigt man sich mit den Gesetzmäßigkeiten, die bei der Verbindung von Aussagen mit Bindewörtern (Junktoren) auftreten. Vom Inhalt der kleinsten, unzerlegbaren Aussagen wird dabei abstrahiert. Erst später bei der Prädikatenlogik werden wir auch die Verwendung von Quantoren formal studieren. Die explizite Verwendung von Quantoren ist nicht die einzige Form der Verallgemeinerung der klassischen Aussagenlogik. So werden etwa in Systemen der temporalen Aussagenlogik und modalen Aussagenlogiken spezielle Operatoren zur Erhöhung der Ausdrucksstärke eingeführt. Damit sollen allerdings nur einige Dimensionen angedeutet sein, in der man den nachfolgend betrachteten Gegenstand erweitern kann.

10.1 Sprache der Aussagenlogik

In Kapitel 1 hatten wir anhand verschiedener Beispiele gezeigt, daß aussagenlogische Junktoren in ihrer natürlichsprachlichen Verwendung manchmal ambig sind. Im Rahmen der klassischen Aussagenlogik wird zunächst eine formale Sprache eingeführt, die vollständig frei von vergleichbaren Ambiguitäten ist. Um die Syntax dieser Sprache geht es in diesem Abschnitt.

Der *Zeichenvorrat*, mit dem wir die Ausdrücke der Sprache aufbauen wollen, hat drei Teile:

(1) $At = \{A_n \mid n \in \mathbb{N}\}$, eine abzählbar unendliche Menge von Symbolen, die wir *atomare Aussagen, Atomformeln* oder *Aussagenvariablen* nennen,

(2) $J = \{\vee, \wedge, \Rightarrow, \neg\}$ ist die Menge der verwendeten aussagenlogischen Junktoren,

(3) Weiter verwenden wir die Klammern „(“ und „)“ als Hilfsmittel.

Wenn wir sagen, daß die Elemente von At „Symbole“ sind, so soll dies betonen, daß diese selbst keine eigene innere Struktur haben.

Es soll an dieser Stelle ausdrücklich festgehalten werden, daß die Junktoren $\vee$, $\wedge$, $\Rightarrow$ und $\neg$ von dieser Stelle an nicht mehr wie bisher als metasprachliche Ausdrücke verwendet werden, sondern nur noch als Zeichen der zu erklärenden formalen (Objekt-) Sprache.

Die *Ausdrücke* der Sprache werden wir aussagenlogische Formeln nennen.

Definition 10.1 Die Menge $\mathcal{L}_0$ der *aussagenlogischen Formeln* ist induktiv definiert als die kleinste Menge, die unter folgenden Bildungsregeln abgeschlossen ist:

1. jede Atomformel aus At ist eine aussagenlogische Formel,
2. sind α und β aussagenlogische Formeln, so auch $\neg\alpha$, $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$ und $(\alpha \Rightarrow \beta)$.

Aussagenlogische Formeln nennen wir auch $\mathcal{L}_0$-Formeln. Wir werden die Biimplikation „$\Leftrightarrow$" hier nur als Abkürzung zulassen: es steht $(\alpha \Leftrightarrow \beta)$ für die aussagenlogische Formel $((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha))$. Es sei hinzugefügt, daß die in Definition 10.1 verwendete Junktorenmenge zwar gebräuchlich ist, daß aber andere Auswahlen denkbar sind.

Beispiel 10.2 Es sind A_0, $(A_1 \wedge \neg A_2)$, $((A_1 \wedge \neg A_2) \Rightarrow (A_3 \vee \neg A_5))$ sowie

$$((A_0 \wedge (A_1 \vee A_2)) \Rightarrow ((A_1 \wedge \neg A_1) \vee A_0))$$

aussagenlogische Formeln.

Da Formeln endliche Zeichenfolgen sind, ist es zunächst wichtig, zu zeigen, daß sie in eindeutiger Weise zu lesen sind.

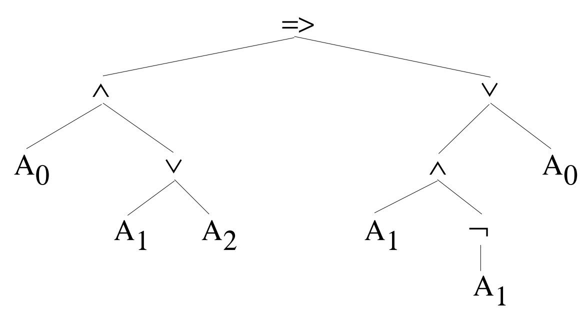
Satz 10.3 (Eindeutige Lesbarkeit aussagenlogischer Formeln) *Jede aussagenlogische Formel φ ist entweder atomar oder läßt sich auf genau eine Weise in einer der Formen $\neg\alpha, (\alpha \wedge \beta), (\alpha \vee \beta)$ oder $(\alpha \Rightarrow \beta)$ darstellen.*

Bemerkung 10.4 Wir werden Satz 10.3 nicht beweisen. Es sei aber bemerkt, daß wir die Junktoren als Funktionssymbole auffassen können, die in Infixschreibweise notiert sind. Damit ergibt sich eine direkte Entsprechung zwischen den hier definierten Formeln und den in 7.51 eingeführten Termen über einer geeigneten Signatur. Satz 10.3 ist lediglich eine Variante von Satz 7.52. Ein Beweis von Satz 10.3 würde sich ähnlich wie der Beweis von 7.52 auf die Analyse der Klammerstruktur stützen.

Nach Satz 10.3 kann man jeder aussagenlogischen Formel einen eindeutig bestimmten „Formelbaum" zuordnen. Der Formelbaum für

$$((A_0 \wedge (A_1 \vee A_2)) \Rightarrow ((A_1 \wedge \neg A_1) \vee A_0))$$

hat beispielsweise die folgende Form:



Im nachfolgenden werden wir erlauben, in aussagenlogischen Formeln Klammern wegzulassen, wenn klar bleibt, welche Formel gemeint ist.

Definition 10.5 Das *Gewicht* $w(\varphi)$ einer aussagenlogischen Formel φ ist induktiv wie folgt erklärt:

1. jede Atomformel A_n hat Gewicht 0.
2. Das Gewicht einer Formel der Form $\neg\alpha$ ist $w(\alpha)+1$. Das Gewicht einer Formel der Form $(\alpha\wedge\beta)$, $(\alpha\vee\beta)$ oder $(\alpha \Rightarrow \beta)$ ist $\max\{w(\alpha), w(\beta)\}+1$.

Das Gewicht einer Formel stimmt mit der Höhe des Formelbaums überein. Die Wohldefiniertheit des Gewichts ergibt sich aus Satz 10.3.

Beispiel 10.6 Die Formel A_0 hat Gewicht 0, $(A_1 \wedge \neg A_2)$ hat Gewicht 2. Die Formel $((A_0 \wedge (A_1 \vee A_2)) \Rightarrow ((A_1 \wedge \neg A_1) \vee A_0))$, deren Formelbaum oben abgebildet ist, hat Gewicht 4.

Prinzip der strukturellen Induktion. Bezeichnet $\mathcal{L}_0^{(n)}$ die Menge der aussagenlogischen Formeln des Gewichts n, so ist offensichtlich $\mathcal{L}_0$ die disjunkte Vereinigung der Mengen $\mathcal{L}_0^{(n)}$ für $n \in \mathbb{N}$. Wir können damit Behauptungen über die Gesamtheit aller Formeln induktiv über das Gewicht wie folgt beweisen.

1. Zunächst zeigen wir als *Induktionsanfang*, daß alle Formeln des Gewichts 0, also alle Atomformeln, die zu beweisende Eigenschaft haben.

2. Im *Induktionsschritt* nehmen an, daß für ein beliebiges $n \geq 0$ alle Formeln des Gewichts $\leq n$ die betreffende Eigenschaft haben. Wir geben uns nun eine beliebige Formel φ des Gewichts $n + 1 \geq 1$ vor und verifizieren mit Hilfe dieser Annahme, daß auch γ die zu zeigende Eigenschaft besitzt.

In der Tat folgt aus dem Gezeigten insgesamt, daß *alle* aussagenlogischen Formeln die betreffende Eigenschaft haben. Zur Begründung sei auf die Erläuterung der vollständigen Induktion in Abschnitt 2.5 verwiesen. In Schritt 2 nützen wir bei konkreten Beweisen in aller Regel aus, daß φ eine der Formen $\neg\alpha$, $(\alpha\wedge\beta)$, $(\alpha\vee\beta)$ oder $(\alpha \Rightarrow \beta)$ haben muß, wobei α und (gegebenenfalls) β Gewicht $\leq n$ haben und damit nach Induktionsannahme die zu beweisende Eigenschaft besitzen. Das so umschriebene Beweisprinzip wird auch als „strukturelle Induktion“ oder „Induktion über den Formelaufbau“ bezeichnet. Konkrete Anwendungen werden wir unten kennenlernen.

Definition 10.7 Die Menge der *Teilformeln* einer Formel φ, $TF(\varphi)$, ist induktiv wie folgt erklärt:

1. $TF(\varphi) = \{A_i\}$, falls φ ein Atom $A_i \in At$ ist,
2. $TF(\varphi) = TF(\alpha) \cup \{\neg\alpha\}$, falls φ die Form $\neg\alpha$ hat,
3. $TF(\varphi) = TF(\alpha) \cup TF(\beta) \cup \{(\alpha\mathcal{J}\beta)\}$, falls φ die Form $(\alpha\mathcal{J}\beta)$ mit $\mathcal{J} \in \{\vee, \wedge, \Rightarrow\}$ hat.

Satz 10.3 garantiert, daß die Funktion TF wohldefiniert ist. Die Teilformel α einer Formel der Form $\neg\alpha$ sowie die Teilformeln α und β einer Formel der Form $(\alpha\mathcal{J}\beta)$ mit $\mathcal{J} \in \{\vee, \wedge, \Rightarrow\}$ werden auch *unmittelbare Teilformeln* genannt.

Beispiel 10.8 Es gilt

$$\begin{array}{rcl} TF(A_0) & = & \{A_0\}, \\ TF(A_1 \wedge \neg A_2) & = & \{A_1, A_2, \neg A_2, (A_1 \wedge \neg A_2)\}. \end{array}$$

$TF((A_0 \wedge (A_1 \vee A_2)) \Rightarrow ((A_1 \wedge \neg A_1) \vee A_0))$ enthält genau die Formeln A_0, A_1, A_2, $\neg A_1$, $(A_1 \wedge \neg A_1)$, $((A_1 \wedge \neg A_1) \vee A_0)$, $(A_1 \vee A_2)$, $(A_0 \wedge (A_1 \vee A_2))$ und $((A_0 \wedge (A_1 \vee A_2)) \Rightarrow ((A_1 \wedge \neg A_1) \vee A_0))$.

10.2 Erfüllbarkeit und Tautologiebegriff

Um allgemeingültige Formeln formal zu definieren, wird das folgende Konzept eingeführt.

Definition 10.9 Eine 0-1 *Bewertungsfunktion* ist eine Funktion $g\colon \mathcal{L}_0 \to \{0,1\}$ mit den folgenden Eigenschaften: für beliebige Formeln $\alpha, \beta \in \mathcal{L}_0$ gilt:

(i) $g(\alpha \vee \beta) = \max\{g(\alpha), g(\beta)\}$,

(ii) $g(\alpha \wedge \beta) = \min\{g(\alpha), g(\beta)\}$,

(iii) $g(\alpha \Rightarrow \beta) = \max\{1 - g(\alpha), g(\beta)\}$,

(iv) $g(\neg\alpha) = 1 - g(\alpha)$,

0-1 Bewertungsfunktionen werden auch *Wahrheitswertzuordnungen* genannt. Natürlich läßt sich der Inhalt von Definition 10.9 wieder durch die aus dem ersten Kapitel bekannten Wahrheitswert-Tabellen für die Junktoren wiedergeben. Es folgt unmittelbar das folgende Kompositionalitätsprinzip:

Lemma 10.10 (Aussagenlogisches Kompositionalitätsprinzip) *Der Wahrheitswert einer aussagenlogischen Formel φ des Gewichts $n > 0$ bezüglich einer 0-1 Bewertungsfunktion g ergibt sich allein aus den Wahrheitswerten der unmittelbaren Teilformeln von φ und aus dem Junktor, mit dem diese zusammengesetzt sind.*

Lemma 10.11 *Jede Funktion $g_0\colon At \to \{0,1\}$ läßt sich auf genau eine Weise zu einer 0-1 Bewertungsfunktion g auf $\mathcal{L}_0$ fortsetzen.*

Der einfache Beweis sei hier nur angedeutet. Zunächst zeigt man durch strukturelle Induktion, daß eine Fortsetzung g von g_0 auf Formeln beliebigen Gewichts existiert. Ist g für Formeln des Gewichts $\leq n$ bereits erklärt, so verwendet man Definition 10.9, um die Definition von g auch auf Formeln des Gewichts $n+1$ zu erweitern. Hieraus ergibt sich eine auf der Formelmenge $\mathcal{L}_0$ definierte 0-1 Bewertungsfunktion. Um die Eindeutigkeit zu verifizieren, zeigt man induktiv für beliebiges Gewicht $n \geq 0$, daß je zwei Erweiterungen von g_0 auf allen Formeln des Gewichts n übereinstimmen müssen. Hieraus ergibt sich die Behauptung.

Für Leser, die mit dem Inhalt von Abschnitt 7.4 vertraut sind, sei erwähnt, daß Lemma 10.11 als eine spezielle Instanz von Theorem 7.60 aufgefasst werden kann. Man vergleiche hierzu Bemerkungen 10.4 und 10.27.

Lemma 10.12 (Aussagenlogisches Koinzidenzlemma) *Sind* g_1, g_2 *zwei* 0-1 *Bewertungsfunktionen, die auf allen atomaren Teilformeln einer aussagenlogischen Formel* φ *übereinstimmen, so gilt* $g_1(\varphi) = g_2(\varphi)$.

Beweis. Wir verwenden strukturelle Induktion.
Induktionsanfang. Falls $w(\varphi) = 0$ gilt, so ist φ atomar. Es gilt nach Voraussetzung $g_1(\varphi) = g_2(\varphi)$.
Als *Induktionsvoraussetzung* nehmen wir nun an, es sei nun die Behauptung für alle aussagenlogischen Formeln des Gewichts $\leq n$ gezeigt. Sei φ eine Formel des Gewichts $n+1$. Nach Lemma 10.3 kann man φ eindeutig in der Form $\neg\alpha$ oder $(\alpha J \beta)$ darstellen (für ein $J \in \{\wedge, \vee, \Rightarrow\}$), wobei die unmittelbaren Teilformeln α und (gegebenenfalls) β Gewicht $\leq n$ haben. Natürlich stimmen g_1 und g_2 auch auf allen atomaren Teilformeln von α und (gegebenenfalls) β überein. Damit können wir die Induktionsvoraussetzung auf α und (gegebenenfalls) β anwenden. Falls φ die Form $\neg\alpha$ hat, so gilt nach Induktionsvoraussetzung $g_1(\varphi) = 1 - g_1(\alpha) = 1 - g_2(\alpha) = g_2(\varphi)$. Falls φ die Form $(\alpha \wedge \beta)$ hat, so folgt analog $g_1(\varphi) = \min\{g_1(\alpha), g_1(\beta)\} = \min\{g_2(\alpha), g_2(\beta)\} = g_2(\varphi)$. Die anderen Fälle folgen analog. ■

Lemma 10.12 besagt in anderen Worten, daß der Wahrheitswert einer Formel φ unter einer 0-1-Bewertung g lediglich von den Wahrheitswerten der in φ auftretenden Atomformeln abhängen kann, nicht jedoch von den Werten von Atomformeln, die nicht in φ auftreten. Da jede Formel durch Anwendung endlich vieler Bildungsregeln aus Atomformeln aufgebaut ist, kann man dies auch als eine Folge des Kompositionalitätsprinzips betrachten.

Wir kommen zur zentralen Definition dieses Teilkapitels.

Definition 10.13 Es sollen nachfolgend φ, α, β beliebige Formeln aus $\mathcal{L}_0$ bezeichnen und $\Phi \subseteq \mathcal{L}_0$ eine Formelmenge.

1. Es heißt φ eine *aussagenlogische Tautologie* genau dann, wenn $g(\varphi) = 1$ für jede 0-1 Bewertungsfunktion g auf $\mathcal{L}_0$ gilt.

2. Es heißt φ *erfüllbar* genau dann, wenn es eine 0-1 Bewertungsfunktion g auf $\mathcal{L}_0$ gibt mit $g(\varphi) = 1$.

3. Φ heißt *erfüllbar* genau dann, wenn es eine 0-1 Bewertungsfunktion g auf $\mathcal{L}_0$ gibt mit $g(\varphi) = 1$ für alle $\varphi \in \Phi$.

4. Es heißen α und β *aussagenlogisch äquivalent* genau dann, wenn die Formel $(\alpha \Leftrightarrow \beta)$ eine aussagenlogische Tautologie ist.

Zwei Beobachtungen folgen sofort. Der Beweis bleibt dem Leser überlassen.

Lemma 10.14 *Zwei Formeln α und β sind aussagenlogisch äquivalent genau dann, wenn sie unter jeder* 0-1 *Bewertungsfunktion g denselben Wert erhalten.*

Von dieser Charakterisierung werden wir nachfolgend ohne besondere Erwähnung Gebrauch machen.

Lemma 10.15 *Eine Formel φ ist eine Tautologie genau dann, wenn $\neg\varphi$ unerfüllbar ist.*

Beispiel 10.16 Wenn wir in Satz 1.1 die Ausdrücke α, β und γ nun als aussagenlogische Formeln lesen, und die Junktoren als Zeichen der Sprache der Aussagenlogik, so stellen alle dort aufgeführten Formeln aussagenlogische Tautologien dar.

Durch strukturelle Induktion folgt leicht, daß eine Formel φ des Gewichts n höchstens 2^n viele atomare Teilformeln haben kann (vgl. Aufgabe 10.2). Nach Lemma 10.12 hängt der Wahrheitswert einer Formel φ in keiner Weise von solchen Atomformeln ab, die nicht φ auftreten. Um festzustellen, ob φ Tautologie oder erfüllbar ist, kommen wir daher mit einer endlichen Fallunterscheidung über die möglichen Wahrheitswerte der Atomformeln von φ aus. Man kann demnach mit der Methode der Wahrheitswert-Tabellen (vgl. Kapitel 1) entscheiden, ob φ eine aussagenlogische Tautologie (bzw. erfüllbar) ist. Entscheidbarkeit bedeutet hier, daß es ein Rechenverfahren gibt, das

- als Eingabe eine beliebige aussagenlogische Formel φ akzeptiert,
- stets nach endlicher Zeit terminiert, und

- die Antwort „ja“ liefert genau dann, wenn φ eine Tautologie (bzw. erfüllbar) ist.

Hieraus ergibt sich das folgende Ergebnis.

Satz 10.17 (Entscheidbarkeit der Aussagenlogik) *Es für eine gegebene Formel $\varphi \in \mathcal{L}_0$ entscheidbar, ob φ eine Tautologie (erfüllbar) ist.* ■

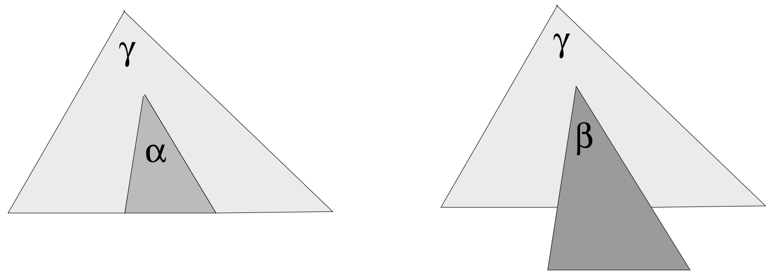
Beispiel 10.18 Wir fragen nach der Erfüllbarkeit der Formel $\beta := A_1 \wedge \neg(A_0 \Rightarrow (A_1 \Rightarrow A_2))$. Wir setzen $\alpha := (A_0 \Rightarrow (A_1 \Rightarrow A_2))$. Die Fallunterscheidung über die möglichen Wahrheitswerte der in β auftretenden Atomformeln führt auf folgende Wahrheitswert-Tabelle.

A_0	A_1	A_2	$A_1 \Rightarrow A_2$	α	$\neg\alpha$	β
0	0	0	1	1	0	0
0	0	1	1	1	0	0
0	1	0	0	1	0	0
1	0	0	1	1	0	0
0	1	1	1	1	0	0
1	0	1	1	1	0	0
1	1	0	0	0	1	1
1	1	1	1	1	0	0

Die vorletzte Zeile zeigt, daß β erfüllbar ist. Hingegen wäre die Formel $A_1 \wedge \neg(A_0 \Rightarrow (A_1 \Rightarrow A_0))$ nicht erfüllbar.

Interessantere Verfahren als die Methode der Wahrheitswerte-Tabellen werden wir später kennenlernen.

Zum Abschluß dieses Teilkapitels wollen wir nun noch zeigen, wie man mit den nun erreichten Techniken eine formale Berechtigung für die Bemerkungen in Abschnitt 1.4 zur Ersetzung äquivalenter Aussagen geben kann. Wir beginnen mit einer Schreibweise, die wir nicht vollständig formalisieren werden. Mit $\gamma(\alpha)$ bezeichnen wir eine aussagenlogische Formel γ, wo wir ein bestimmtes Vorkommen einer Teilformel α ausgezeichnet haben. Mit $\gamma(\alpha/\beta)$ notieren wir die Formel, die wir erhalten, wenn wir das Vorkommen der Teilformel α durch β ersetzen. Das Vorgehen wird schematisch durch die nachfolgenden Formelbäume angedeutet.



Lemma 10.19 (Ersetzungslemma) *Seien $\alpha, \beta, \gamma \in \mathcal{L}$. Es sei α eine Teilformel der aussagenlogischen Formel γ. Sind α und β aussagenlogisch äquivalent, so auch $\gamma(\alpha)$ und $\gamma(\alpha/\beta)$.*

Beweis. Wir verwenden wieder strukturelle Induktion.
Induktionsanfang. Es sei γ vom Gewicht 0, also atomar. Dann gilt offensichtlich $\gamma = \alpha$. Es folgt $\gamma(\alpha) = \alpha$ und $\gamma(\alpha/\beta) = \beta$. Somit sind $\gamma(\alpha)$ und $\gamma(\alpha/\beta)$ nach Voraussetzung aussagenlogisch äquivalent.
Als *Induktionshypothese* wollen wir annehmen, daß die Behauptung für alle Formeln des Gewichts $\leq n$ richtig sei. Es sei nun γ eine Formel des Gewichts $n+1$. Nach Lemma 10.3 kann man γ eindeutig in der Form $\neg\gamma_0$ oder $(\gamma_0 J \gamma_1)$ darstellen, für ein $J \in \{\wedge, \vee, \Rightarrow\}$. Hierbei haben die unmittelbaren Teilformeln γ_0 und (gegebenenfalls) γ_1 Gewicht $\leq n$. Wir beschränken uns auf den Fall, wo γ die Form $(\gamma_0 \wedge \gamma_1)$ hat (alle anderen Fälle gehen analog). Hier unterscheiden wir drei Teilfälle.
Fall 1. Das Vorkommen von α liegt in der Teilformel γ_0. Dann sind nach Induktionsvoraussetzung die Formeln $\gamma_0(\alpha)$ und $\gamma_0(\alpha/\beta)$ aussagenlogisch äquivalent. Ist nun g eine beliebige 0-1 Bewertungsfunktion, so folgt hieraus

$$g(\gamma(\alpha)) = g(\gamma_0(\alpha)) \sqcup g(\gamma_1) = g(\gamma_0(\alpha/\beta)) \sqcup g(\gamma_1) = g(\gamma(\alpha/\beta)).$$

Damit sind $\gamma(\alpha)$ und $\gamma(\alpha/\beta)$ aussagenlogisch äquivalent.
Fall 2. Das Vorkommen von α liegt in der Teilformel γ_1. Hier geht man analog wie in Fall 1 vor.
Fall 3. α ist γ selbst. In diesem Fall gilt wieder $\gamma(\alpha) = \alpha$ und $\gamma(\alpha/\beta) = \beta$ und es sind $\gamma(\alpha)$ und $\gamma(\alpha/\beta)$ nach Voraussetzung aussagenlogisch äquivalent. ■

Natürlich kann man Lemma 10.19 sofort dahingehend verallgemeinern, daß man mehrere (Vorkommen von) Teilformeln durch äquivalente Formeln ersetzt. Auch hierbei erhält man immer eine Formel, die zur Ausgangsformel äquivalent ist.

10.3 Der semantische Folgerungsbegriff

Der Begriff der Tautologie beschreibt diejenigen Formeln, die logisch allgemeingültig sind. In vielen Zusammenhängen macht es Sinn, allgemeiner nach denjenigen Formeln zu fragen, die gültig sind, wenn eine gegebene Menge von Prämissen wahr ist. Dies führt auf den folgenden Folgerungsbegriff.

Definition 10.20 Es sei $\Phi \subseteq \mathcal{L}_0$. Die Formel $\varphi \in \mathcal{L}_0$ *folgt aus* Φ, im Zeichen $\Phi \models \varphi$, genau dann, wenn für jede 0-1 Bewertung g auf $\mathcal{L}_0$ gilt:

$$(\forall \sigma \in \Phi\colon g(\sigma) = 1) \Rightarrow g(\varphi) = 1.$$

Die Menge $\mathit{Cons}(\Phi) = \{\varphi \in \mathcal{L}_0 \mid \Phi \models \varphi\}$ wird die Menge der *Konsequenzen* von Φ genannt.

Etwas anders ausgedrückt bedeutet $\Phi \models \varphi$, daß jede 0-1 Bewertung, die alle Formeln aus Φ wahr macht, stets auch φ auf 1 abbildet. Es sei darauf hingewiesen, daß „$\models$" kein Zeichen der Sprache der Aussagenlogik, sondern ein metasprachliches Zeichen ist. Im speziellen Fall $\Phi = \emptyset$ schreibt man meist kurz $\models \varphi$ für $\emptyset \models \varphi$.

Bemerkung 10.21 *Cons* ist eine Hüllenabbildung im Sinn von Definition 4.35 (Übung).

Beispiel 10.22 Es gilt beispielsweise

$$\begin{aligned} \{A_0, A_0 \Rightarrow A_1, A_1 \Rightarrow A_2\} &\models A_2, \\ \{A_0 \vee A_1, A_0 \Rightarrow A_1\} &\models A_1 \end{aligned}$$

Ein häufig verwendeter einfacher Zusammenhang ist der folgende (vgl. Lemma 10.15).

Lemma 10.23 *Es sei* $\Phi \subseteq \mathcal{L}_0$ *und* $\varphi \in \mathcal{L}_0$. *Dann gilt* $\Phi \models \varphi$ *genau dann, wenn* $\Phi \cup \{\neg\varphi\}$ *unerfüllbar ist.*

Beweis. Übung (cf. Aufgabe 10.13). ■

Das nachfolgende Lemma stellt eine Verbindung her zwischen metasprachlichem Folgerungsbegriff „$\models$" und objektsprachlicher Implikation „$\Rightarrow$".

Lemma 10.24 (Aussagenlogisches Deduktionstheorem) *Es sei* $\Phi \subset \mathcal{L}$ *und* $\delta, \gamma \in \mathcal{L}$. *Dann gilt* $\Phi \cup \{\delta\} \models \gamma$ *genau dann, wenn* $\Phi \models (\delta \Rightarrow \gamma)$.

Beweis. Es gelte $\Phi \cup \{\delta\} \models \gamma$. Es sei nun g eine 0-1 Bewertungsfunktion, die alle Formeln aus Φ auf 1 abbildet. Zu zeigen ist $g(\delta \Rightarrow \gamma) = 1$. Gilt nun $g(\delta) = 0$, so folgt dies unmittelbar aus Definition 10.9. Gilt aber $g(\delta) = 1$, so folgt aus $\Phi \cup \{\delta\} \models \gamma$ nun $g(\gamma) = 1$, damit auch $g(\delta \Rightarrow \gamma) = 1$. Gilt umgekehrt $\Phi \models (\delta \Rightarrow \gamma)$ und ist g eine 0-1 Bewertungsfunktion, die alle Formeln aus $\Phi \cup \{\delta\}$ auf 1 abbildet, so folgt aus $\Phi \models (\delta \Rightarrow \gamma)$ nun $g(\gamma) = 1$. Damit gilt $\Phi \cup \{\delta\} \models \gamma$. ■

Bezüglich der im nachfolgenden Korollar verwendeten Notation sei angemerkt, daß es Lemma 10.19 unter Verwendung der Assoziativität der Konjunktion erlaubt, in aussagenlogischen Teilformeln der Form $(\alpha_1 \wedge \ldots \wedge \alpha_n)$ die internen Klammern wegzulassen. Alle expliziten Klammerungen würden zu äquivalenten Formeln führen.

Korollar 10.25 *Es sei* $n \geq 1$, $\Phi = \{\sigma_1, \ldots, \sigma_n\} \subset \mathcal{L}_0$ *und* $\varphi \in \mathcal{L}_0$. *Dann gilt* $\Phi \models \varphi$ *genau dann, wenn* $((\sigma_1 \wedge \ldots \wedge \sigma_n) \Rightarrow \varphi)$ *eine aussagenlogische Tautologie ist. Es gilt* $\models \varphi$ *genau dann, wenn* φ *aussagenlogische Tautologie ist.*

Beweis. Wir verwenden vollständige Induktion über $n \geq 0$, wobei der Fall $n = 0$ die zweite Behauptung abdeckt.

Induktionsanfang. Es gelte $n = 0$. Es bedeutet $\models \varphi$, daß φ unter jeder 0-1 Bewertungsfunktion auf 1 abgebildet wird. Dies heißt gerade, daß φ eine aussagenlogische Tautologie ist.

Induktionsschritt. Es sei die Behauptung richtig für ein $n \geq 0$. Hat nun Φ die Form $\{\sigma_1, \ldots, \sigma_n, \sigma_{n+1}\}$, so sind die folgenden Aussagen äquivalent:

(1) $\{\sigma_1, \ldots, \sigma_n, \sigma_{n+1}\} \models \varphi$,

(2) $\{\sigma_1, \ldots, \sigma_n\} \models (\sigma_{n+1} \Rightarrow \varphi)$,

(3) $(\sigma_1 \wedge \ldots \wedge \sigma_n) \Rightarrow (\sigma_{n+1} \Rightarrow \varphi)$ ist Tautologie,

(4) $(\sigma_1 \wedge \ldots \wedge \sigma_n \wedge \sigma_{n+1}) \Rightarrow \varphi$ ist Tautologie.

Die Äquivalenz von (1) und (2) folgt aus Lemma 10.24. Die Äquivalenz von (2) und (3) folgt gemäß Induktionsvoraussetzung. Die Äquivalenz von (3) und (4) ist einfach und bleibt dem Leser überlassen. ■

10.4 Algebraischer Hintergrund

Die Aussagenlogik besitzt einen interessanten algebraischen Hintergrund. Wir haben ihn bislang bewußt außer acht gelassen, da wir für die vorausgegangenen Darstellungen den Inhalt von Kapitel 9 nicht voraussetzen wollten. Im nachfolgenden müssen wir nun Vertrautheit mit dem Begriff der Booleschen Algebra annehmen. Es ist für die folgende Diskussion zweckmäßig, Boolesche Algebren um eine weitere abgeleitete Operation zu erweitern: ist $\mathcal{B} = \langle B, \sqcup, \sqcap, - \rangle$ eine Boolesche Algebra im bisherigen Sinn, so wollen wir die durch $a \sqsubset b := -a \sqcup b$ definierte binäre Operation hinzunehmen. Boolesche Algebren haben also nun die Form $\mathcal{B} = \langle B, \sqcup, \sqcap, \sqsubset, - \rangle$, in Definition 9.1 wäre die Bedingung $\forall a, b \in B$: $a \sqsubset b = -a \sqcup b$ zu ergänzen.

Als ersten Schritt führen wir auf der Menge der Formeln eine einfache algebraische Struktur ein.

Definition 10.26 Wir definieren auf der Menge der Formeln $\mathcal{L}_0$ die folgenden Operationen:

1. $f_\vee$: $\mathcal{L}_0 \times \mathcal{L}_0 \to \mathcal{L}_0$: $\langle \alpha, \beta \rangle \mapsto (\alpha \vee \beta)$,
2. $f_\wedge$: $\mathcal{L}_0 \times \mathcal{L}_0 \to \mathcal{L}_0$: $\langle \alpha, \beta \rangle \mapsto (\alpha \wedge \beta)$,
3. $f_\Rightarrow$: $\mathcal{L}_0 \times \mathcal{L}_0 \to \mathcal{L}_0$: $\langle \alpha, \beta \rangle \mapsto (\alpha \Rightarrow \beta)$,
4. $f_\neg$: $\mathcal{L}_0 \to \mathcal{L}_0$: $\alpha \mapsto \neg\alpha$,

Die Algebra $\mathit{Form}(At) = \langle \mathcal{L}_0, f_\vee, f_\wedge, f_\Rightarrow, f_\neg \rangle$ nennen wir die *Formelalgebra* über At.

Bemerkung 10.27 Die in Bemerkung 10.4 angesprochene Parallele zwischen Formeln und Termen wird mit Definition 10.26 noch erweitert. Wir können Aussagenvariable schlicht als Variable im Sinn von Abschnitt 7.4 auffassen. Bei geeigneter Festlegung der Signatur Σ—deren Symbole wir durch $f_\vee$, $f_\wedge$, $f_\Rightarrow$ und $f_\neg$ interpretieren—ist die Formelalgebra isomorph zur Termalgebra $\mathcal{T}(\Sigma, At)$. Sie besitzt damit die in Theorem 7.60 angegebene absolute Freiheitseigenschaft.

Bemerkung 10.28 Nach der vorausgegangenen Lektüre von Kapitel 9 ist es wichtig, sich klarzumachen, daß *Form*(At) *keine* Boolesche Algebra ist. So gilt zum Beispiel nicht, daß $f_\wedge$ kommutativ ist. Die Formeln $(\alpha \wedge \beta)$ und $(\beta \wedge \alpha)$ sind nämlich nicht gleich, sondern nur logisch äquivalent.

Das nachfolgende Lemma könnte bei einer algebraischen Herangehensweise zur Definition der 0-1 Bewertungsfunktion verwendet werden.

Lemma 10.29 *Es sei* g: $\mathcal{L}_0 \to \{0,1\}$ *eine* 0-1 *Bewertungsfunktion. Mit* $\sqcup$, $\sqcap$, $\sqsubset$ *und* $-$ *seien die Operationen der Booleschen Algebra* **2** *bezeichnet. Dann gilt für beliebige Formeln* $\alpha, \beta \in \mathcal{L}_0$*:*

(i) $g(\alpha \vee \beta) = g(\alpha) \sqcup g(\beta)$,

(ii) $g(\alpha \wedge \beta) = g(\alpha) \sqcap g(\beta)$,

(iii) $g(\alpha \Rightarrow \beta) = g(\alpha) \sqsubset g(\beta)$,

(iv) $g(\neg\alpha) = -g(\alpha)$,

Der Beweis folgt umittelbar aus der Definition der Booleschen Operationen. Aus Lemma 10.29 ergibt sich folgende Charakterisierung von 0-1 Bewertungsfunktionen.

Lemma 10.30 *Die* 0-1 *Bewertungsfunktionen sind genau die Homomorphismen von Form*$(At) = \langle \mathcal{L}_0, f_\vee, f_\wedge, f_\Rightarrow, f_\neg \rangle$ *in* $\mathbf{2} = \langle \{0,1\}, \sqcup, \sqcap, \sqsubset, - \rangle$.

Beweis. trivial. ∎

Wenn wir statt 0-1 Bewertungen allgemeine Bewertungen zulassen, können wir einen entsprechenden Folgerungsbegriff aufstellen. Dieser erweist sich später als nützlich.

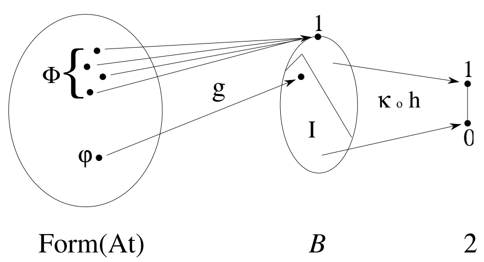
Definition 10.31 Eine *Bewertung* auf $\mathcal{L}_0$ ist ein Homomorphismus von *Form*(At) in eine Boolesche Algebra.

Der folgende Satz zeigt, daß der semantische Folgerungsbegriff nicht verändert wird, wenn wir 0-1 Bewertungen durch allgemeine Bewertungen ersetzen.

Satz 10.32 *Es sei $\Phi \subseteq \mathcal{L}_0$, $\varphi \in \mathcal{L}_0$. Dann gilt $\Phi \models \varphi$ genau dann, wenn für jede Bewertung g auf $\mathcal{L}_0$ gilt:*

$$(\forall \sigma \in \Phi\colon g(\sigma) = 1) \Rightarrow g(\varphi) = 1.$$

Beweis. Wenn für jede Bewertung g auf $\mathcal{L}_0$ mit der Eigenschaft $g(\sigma) = 1$ für alle $\sigma \in \Phi$ auch $g(\varphi) = 1$ gilt, so insbesondere für die 0-1 Bewertungen. Daher gilt in diesem Fall trivial $\Phi \models \varphi$. Es gelte nun umgekehrt $\Phi \models \varphi$. Es sei $\mathcal{B} = \langle B, \sqcup, \sqcap, \sqsubset, - \rangle$ eine Boolesche Algebra und g: $\mathit{Form}(At) \to B$ ein Homomorphismus mit $g(\sigma) = 1$ für alle $\sigma \in \Phi$. Wäre $b := g(\varphi) \neq 1$, so könnte nach Lemma 9.27 ein Primideal I mit $b \in I$ gefunden werden. Nach Korollar 9.25 ist $\mathcal{B}/R(I)$ isomorph zu $\mathbf{2}$, es sei etwa h ein Isomorphismus.



Bezeichnet κ: $B \to B/R(I)$ den kanonischen Homomorphismus, so liefert

$$g \circ \kappa \circ h\colon \mathit{Form}(At) \to B \to B/R(I) \to \{0, 1\}$$

nach Lemma 6.39 eine 0-1 Bewertung, die Φ auf 1 abbildet, nicht aber φ, das heißt $\Phi \models \varphi$ würde nicht gelten. Dies ist ein Widerspruch. Daher gilt $g(\varphi) = 1$. ■

Die Lindenbaumalgebra

Wir hatten in Bemerkung 10.28 bemerkt, daß die Formelalgebra selbst keine Boolesche Algebra ist. Man erhält jedoch mittels der in Abschnitt 6.5 beschriebenen Quotientenbildung eine Boolesche Algebra, indem man logisch äquivalente Formeln identifiziert. Dieser Zusammenhang soll nachfolgend in etwas allgemeinerer Form dargestellt werden.

Lemma 10.33 *Es sei $\Phi \subseteq \mathcal{L}_0$ eine Formelmenge. Dann wird durch*

$$\alpha \equiv_\Phi \beta \quad \textit{genau dann, wenn} \quad \Phi \models \alpha \Leftrightarrow \beta$$

eine Kongruenzrelation auf der Formelalgebra Form(At) *definiert. Die Quotientenalgebra* Form$(At)/\equiv_\Phi$ *mit den Operationen*

$$\begin{aligned}
[\alpha]_{\equiv_\Phi} \sqcup [\beta]_{\equiv_\Phi} &:= [f_\vee(\alpha,\beta)]_{\equiv_\Phi} = [(\alpha \vee \beta)]_{\equiv_\Phi}, \\
[\alpha]_{\equiv_\Phi} \sqcap [\beta]_{\equiv_\Phi} &:= [f_\wedge(\alpha,\beta)]_{\equiv_\Phi} = [(\alpha \wedge \beta)]_{\equiv_\Phi}, \\
[\alpha]_{\equiv_\Phi} \sqsubset [\beta]_{\equiv_\Phi} &:= [f_\Rightarrow(\alpha,\beta)]_{\equiv_\Phi} = [(\alpha \Rightarrow \beta)]_{\equiv_\Phi}, \\
-[\alpha]_{\equiv_\Phi} &:= [f_\neg(\alpha)]_{\equiv_\Phi} = [\neg\alpha]_{\equiv_\Phi}
\end{aligned}$$

ist eine Boolesche Algebra. Ist α irgendeine aus Φ folgende Formel im Sinn von Definition 10.20, so bildet $[\alpha]_{\equiv_\Phi}$ die Eins von Form$(At)/\equiv_\Phi$. *Ist $\neg\beta$ eine aus Φ folgende Formel, so bildet $[\beta]_{\equiv_\Phi}$ die Null von* Form$(At)/\equiv_\Phi$.

Da wir später (vgl. Lemma 11.6) einen ähnlichen Sachverhalt beweisen werden, wo der semantische Folgerungsbegriff „$\models$“ durch einen syntaktischen Ableitungsbegriff ersetzt wird, überlassen wir den Beweis an dieser Stelle dem Leser.

Definition 10.34 Die Quotientenalgebra *Form*$(At)/\equiv_\Phi$ heißt die *Lindenbaumalgebra* zur Formelmenge Φ.

Wir gehen nachfolgend etwas näher auf den speziellen Fall ein, wo $\Phi = \emptyset$ ist. Für ein $n \geq 1$ bezeichne $\mathcal{L}_0(A_1, \dots, A_n)$ die Menge aller Formeln, deren sämtliche atomare Teilformeln in $\{A_1, \dots, A_n\}$ liegen. Das nachfolgende Lemma ist trivial zu beweisen.

Lemma 10.35 *Für jedes $n \geq 1$ ist $\mathcal{L}_0(A_1, \dots, A_n)$ die Grundmenge einer Teilalgebra der Formelalgebra* Form(At).

Nachfolgend sei $n \geq 1$ fest gewählt. Wir bezeichnen die Teilalgebra mit Grundmenge $\mathcal{L}_0(A_1, \dots, A_n)$ in der Form *Form*$(A_1, \dots, A_n)$. Es stellt dann $\equiv_\emptyset$ eine Kongruenzrelation auf *Form*$(A_1, \dots, A_n)$ dar, für die wir einfachheitshalber „$\equiv$“ schreiben. Die Quotientenalgebra *Form*$(A_1, \dots, A_n)/\equiv$ ist wieder eine Boolesche Algebra. Wir verzichten auf einen Beweis und wollen stattdessen die Struktur von *Form*$(A_1, \dots, A_n)/\equiv$ ermitteln. Jede Formel der Gestalt A_i oder $\neg A_i$ $(1 \leq i \leq n)$ nennen wir nachfolgend ein *Literal*, jede Formel der Gestalt $L_1 \wedge \ldots \wedge L_n$ wo L_i ein Literal der Form A_i oder $\neg A_i$ ist $(1 \leq i \leq n)$, nennen wir eine *Bewertungsformel*. Man beachte, daß es genau 2^n verschiedene Bewertungsformeln gibt.

Lemma 10.36 *Die Äquivalenzklassen* $[L_1 \wedge \ldots \wedge L_n]_\equiv$ *der Bewertungsformeln sind Atome der Booleschen Algebra* $\mathrm{Form}(A_1, \ldots, A_n)/\equiv$.

Beweis. Sei $\alpha \in \mathcal{L}_0(A_1, \ldots, A_n)$ und $L_1 \wedge \ldots \wedge L_n$ eine Bewertungsformel. Es gelte $0 \leq [\alpha]_\equiv \leq [L_1 \wedge \ldots \wedge L_n]_\equiv$, wo die Ordnung „$\leq$“ wie in Bemerkung 9.2 erklärt ist. Wir müssen zeigen, daß $[\alpha]_\equiv = 0$ oder $[\alpha]_\equiv = [L_1 \wedge \ldots \wedge L_n]_\equiv$ gilt. Aus $[\alpha]_\equiv \leq [L_1 \wedge \ldots \wedge L_n]_\equiv$ ergibt sich $[\alpha]_\equiv \sqcap [L_1 \wedge \ldots \wedge L_n]_\equiv = [\alpha \wedge (L_1 \wedge \ldots \wedge L_n)]_\equiv = [\alpha]_\equiv$. Gemäß der Definition von „$\equiv$“ sind dann α und $\alpha \wedge (L_1 \wedge \ldots \wedge L_n)$ logisch äquivalent. Somit ist $\alpha \Rightarrow (L_1 \wedge \ldots \wedge L_n)$ eine Tautologie. Wir können annehmen, daß α nicht unerfüllbar ist, da wir ansonsten wegen $[\alpha]_\equiv = 0$ fertig sind. Es sei g irgendeine 0-1 Bewertung mit $g(\alpha) = 1$. Zumindest eine solche 0-1 Bewertung existiert nach Annahme. Für diese gilt auch $g(L_1 \wedge \ldots \wedge L_n) = 1$. Demnach gilt $g(A_i) = 1$ genau dann, wenn $L_i = A_i$ $(1 \leq i \leq n)$. Ist jedoch g' irgendeine 0-1 Bewertung mit $g(L_1 \wedge \ldots \wedge L_n) = 1$, so hat g' offenkundig dieselbe Eigenschaft, und es stimmen g und g' auf $A_1, \ldots, A_n$ überein. Das Koinzidenzlemma zeigt, daß $(L_1 \wedge \ldots \wedge L_n) \Rightarrow \alpha$ eine Tautologie ist. Es folgt, daß $(L_1 \wedge \ldots \wedge L_n)$ und α logisch äquivalent sind. Wir erhalten damit $[\alpha]_\equiv = [L_1 \wedge \ldots \wedge L_n]_\equiv$. ■

Der Beweis des nachfolgenden Lemmas bleibt als Übungsaufgabe offen.

Lemma 10.37 *Zwei verschiedene Bewertungsformeln sind nie logisch äquivalent. Jede erfüllbare Formel in* $\mathcal{L}_0(A_1, \ldots, A_n)$ *ist logisch äquivalent zu einer Disjunktion über* $k \geq 1$ *Bewertungsformeln.*

Wir erhalten damit eine präzise Charakterisierung der Struktur der Booleschen Algebra $\mathit{Form}(A_1, \ldots, A_n)/\equiv$.

Satz 10.38 *Die Algebra* $\mathrm{Form}(A_1, \ldots, A_n)/\equiv$ *ist die bis auf Isomorphie eindeutig bestimmte endliche Boolesche Algebra mit* 2^n *Atomen. Sie hat* 2^{2^n} *Elemente.*

Beweis. Aus Lemma 10.37 ergibt sich, daß $\mathit{Form}(A_1, \ldots, A_n)/\equiv$ eine endliche Boolesche Algebra mit genau 2^n Atomen ist. Die Behauptung folgt nach Satz 9.36. ■

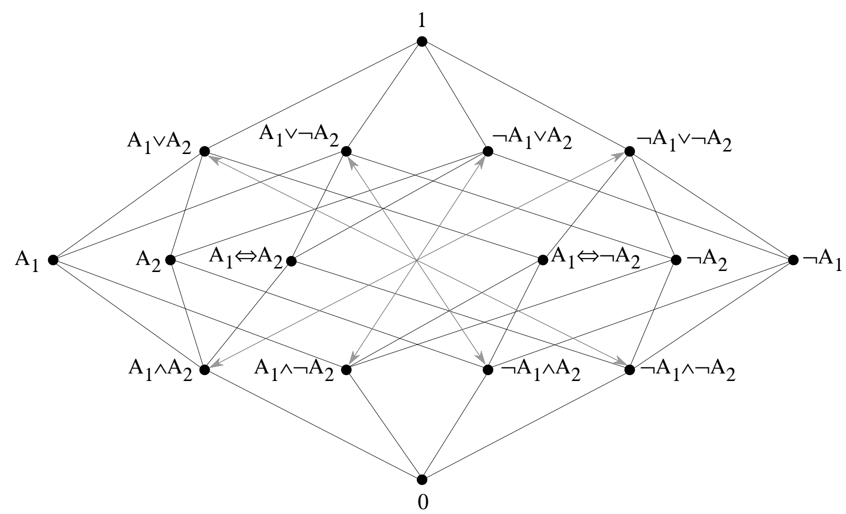


Abbildung 10.1: Die von zwei Aussagenvariablen erzeugte Boolesche Algebra.

In Abbildung 10.1 ist die Boolesche Algebra $\mathit{Form}(A_1, A_2)/\equiv$ dargestellt. Statt der $16 = 2^{2^2}$ Äquivalenzklassen sind jeweils Vertreter angegeben. Die Atome sind die über der Null angeordneten Bewertungsformeln. Alle Formeln der mittleren Zeile sind zu Disjunktionen mit zwei Bewertungsformeln logisch äquivalent. Alle Formeln direkt unterhalb der Eins sind zu Disjunktionen mit drei Bewertungsformeln logisch äquivalent. Mit grauen Doppelpfeilen sind die Komplemente von acht Elementen markiert.

10.5 Ergänzung: Mengentheoretische Identitäten und aussagenlogische Tautologien

Zum Abschluß dieses Kapitels wollen wir ein zusammenfassendes Resultat zur Ähnlichkeit zwischen dem Rechnen mit Mengen einerseits und dem Rechnen mit Aussagen andererseits beweisen. In den vorangegangenen Kapiteln hatten wir bereits verschiedentlich darauf hingewiesen, daß sich aussagenlogische Tautologien ergeben, wenn man mengentheoretische Identitäten oder die Axiome für Verbände und Boolesche Algebren passend „übersetzt“. Die optische Ähnlichkeit der Symbole $\vee, \cup, \sqcup$ bzw. der Symbole $\wedge, \cap, \sqcap$

erfährt dadurch eine Berechtigung. Wir wollen nun zeigen, daß sogenannte „mengentheoretische Identitäten" stets aussagenlogischen Tautologien entsprechen und umgekehrt. Eine Variable, die für eine Mengen steht, wollen wir nachfolgend Mengenvariable nennen. Zunächst müssen wir klären, was wir unter einer mengentheoretischen Identität verstehen.

Definition 10.39 *Mengenausdrücke* sind wie folgt erklärt. Jede Mengenvariable ist ein Mengenausdruck. Ist Π ein Mengenausdruck, so auch $-\Pi$ (Komplement). Sind Π und Ψ Mengenausdrücke, so auch $\Pi \cup \Psi$ und $\Pi \cap \Psi$. Eine Gleichung $\Pi = \Psi$ zwischen Mengenausdrücken Π und Ψ heißt *mengentheoretische Identität*, falls Π und Ψ für jede Belegung der Mengenvariablen in Π und Ψ mit Mengen identische Mengen repräsentieren. Hierbei wird „$-$" als Komplementbildung in einer Obermenge, die alle Interpretationen der Mengenvariablen enthält, aufgefaßt.

Beispiel 10.40 Wir fassen die Symbole M_1 und M_2 als Mengenvariablen auf. Die Gleichung $-(M_1 \cup M_2) = -M_1 \cap -M_2$ ist eine mengentheoretische Identität. Gemäß der de Morganschen Regel (Lemma 2.16) gilt die Gleichung für beliebige Mengen.

Aus einer Gleichung $\Pi = \Psi$ zwischen Mengenausdrücken mit Mengenvariablen $M_1, \ldots, M_k$ erhält man eine aussagenlogische Formel $\phi \Leftrightarrow \psi$, wenn man jede Mengenvariable M_i durch eine aussagenlogische Variable A_i ersetzt $(1 \leq i \leq n)$ und die mengentheoretischen Operatoren $-$, $\cup$ und $\cap$ durch die Booleschen Junktoren $\neg$, $\vee$, und $\wedge$. Durch die umgekehrte Übersetzung läßt sich offenkundig jede aussagenlogische Formel der Form $\phi \Leftrightarrow \psi$, wo die Teilausdrücke ϕ und ψ nur die Booleschen Junktoren $\neg$, $\vee$ und $\wedge$ enthalten, in eine Gleichung $\Pi = \Psi$ zwischen Mengenausdrücken umwandeln.

Unser Ziel ist zu zeigen, daß man bei dieser Übersetzung aus mengentheoretischen Identitäten stets aussagenlogische Tautologien erhält, bei der Rückübersetzung ergeben sich aus aussagenlogischen Tautologien stets mengentheoretische Identitäten. Wir beginnen mit einer Vorbereitung.

Lemma 10.41 *Es seien A, B zwei beliebige Teilmengen einer Menge C. Dann gilt $A = B$ genau dann, wenn $(-A \cup B) \cap (-B \cup A) = C$ gilt.*

Beweis. Da Mengenalgebren Boolesche Algebren sind, erhalten wir aus den Lemmas 9.7, 9.8 und 9.9 direkt die folgenden Äquivalenzen:

$$\begin{array}{rcl} A = B \Leftrightarrow & A = -(-B) & \\ & \Leftrightarrow & A \cup -B = C \quad \wedge \quad A \cap -B = \emptyset \\ & \Leftrightarrow & A \cup -B = C \quad \wedge \quad -(A \cap -B) = C \\ & \Leftrightarrow & A \cup -B = C \quad \wedge \quad -A \cup -(-B) = C \\ & \Leftrightarrow & A \cup -B = C \quad \wedge \quad -A \cup B = C \\ & \Leftrightarrow & (-A \cup B) \cap (-B \cup A) = C. \blacksquare \end{array}$$

Theorem 10.42 *Es sei* $\Pi = \Psi$ *eine mengentheoretische Identität. Dann ergibt sich durch die oben angegebene Übersetzung in eine aussagenlogische Formel stets eine aussagenlogische Tautologie. Ist umgekehrt* α *eine aussagenlogische Tautologie der Form* $\phi \Leftrightarrow \psi$, *wo* ϕ *und* ψ *nur die Junktoren* $\vee, \wedge$ *und* $\neg$ *enthalten, so ergibt sich durch obige Übersetzung stets eine mengentheoretische Identität.*

Beweis. Es sei $\phi \Leftrightarrow \psi$ die aussagenlogische Formel, die sich als Übersetzung der mengentheoretischen Identität $\Pi = \Psi$ ergibt. Angenommen $\phi \Leftrightarrow \psi$ ist keine aussagenlogische Tautologie. Dann gibt es gemäß Definition 10.13 eine 0-1 Bewertungsfunktion g: $\mathit{Form}(At) \to \mathbf{2}$ in die Boolesche Algebra $\mathbf{2}$ mit den Elementen 0 und 1 so daß $g(\phi \Leftrightarrow \psi) = 0$ ist. Somit gilt

$$g((\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)) = 0,$$
$$g((-\phi \vee \psi) \wedge (-\psi \vee \phi)) = 0$$

bzw.

$$(-g(\phi) \sqcup g(\psi)) \sqcap (-g(\psi) \sqcup g(\phi)) = 0.$$

Die Boolesche Algebra $\mathbf{2}$ ist jedoch offenkundig zur Potenzmengenalgebra mit den Elementen $\emptyset$ und $\{1\}$ isomorph. Ist h ein Isomorphismus, so ergibt sich nun

$$\begin{array}{rcl} \emptyset & = & h(0) = h((-g(\phi) \sqcup g(\psi)) \sqcap (-g(\psi) \sqcup g(\phi))) \\ & = & h((-g(\phi) \sqcup g(\psi))) \cap h((-g(\psi) \sqcup g(\phi))) \\ & = & (-h(g(\phi)) \cup h(g(\psi))) \cap (-h(g(\psi)) \cup h(g(\phi))). \end{array}$$

Setzen wir nun $A := h(g(\phi))$, $B := h(g(\psi))$ so ergibt sich

$$(-A \cup B) \cap (-B \cup A) = \emptyset \neq \{1\}.$$

Aus Lemma 10.41 folgt $h(g(\phi)) = A \neq B = h(g(\psi))$. Interpretieren wir in der mengentheoretischen Identität $\Pi = \Psi$ jede Mengenvariable M_i–die in $\phi \Leftrightarrow \psi$ in die aussagenlogische Variable A_i übersetzt ist—durch die Menge $h(g(A_i))$, so erhalten wir gerade die gültige Gleichung $h(g(\phi)) = h(g(\psi))$. Der so erhaltene Widerspruch zeigt, daß $\phi \Leftrightarrow \psi$ doch eine aussagenlogische Tautologie sein muß.

Der zweite Teil der Behauptung folgt unmittelbar aus der Definition der mengentheoretischen Operationen $\cup, \cap$ und $-$ mit Hilfe der Junktoren $\vee, \wedge$ und $\neg$. ∎

Eine andere Variante des Beweises ergibt sich daraus, daß nach dem Stoneschen Satz jede mengentheoretische Identität eine allgemeingültige Boolesche Formel darstellt und damit aus den Axiomen für Boolesche Algebren herleitbar ist, die ihrerseits aussagenlogischen Tautologien entsprechen.

10.6 Aufgaben zu Kapitel 10

Aufgaben zu Teilkapitel 10.1

Aufgabe 10.1 Geben Sie alle Teilformeln und den Formelbaum der Formel $((A_0 \vee A_1) \wedge A_2) \Rightarrow (A_0 \wedge \neg A_1)$ an.

Aufgabe 10.2 Zeigen Sie durch strukturelle Induktion: eine aussagenlogische Formel des Gewichts n besitzt höchstens 2^n viele atomare Teilformeln.

Aufgabe 10.3 Beweisen Sie: Die Menge $\mathcal{L}_0$ der aussagenlogischen Formeln ist abzählbar unendlich. Hinweis: Verwenden Sie zunächst Induktion über n um zu zeigen, daß für jedes $n \in \mathbb{N}$ die Menge der aussagenlogischen Formeln des Gewichts 0 abzählbar unendlich ist. Benutzen Sie dann Lemma 5.10.

Aufgaben zu Teilkapitel 10.2

Aufgabe 10.4 Es seien α, β, γ aussagenlogische Formeln. Zeigen Sie:
die Formeln $(\alpha \Rightarrow (\beta \Rightarrow \gamma))$ und $((\alpha \wedge \beta) \Rightarrow \gamma)$ sind aussagenlogisch äquivalent.

Aufgabe 10.5 Wieviel paarweise nicht äquivalente Formeln kann man durch Klammerungen des Ausdrucks $A_0 \Rightarrow A_1 \Rightarrow A_2 \Rightarrow A_3$ erhalten?

Aufgabe 10.6 Es seien α, β, γ Formeln. Sind dann stets $(\neg(\alpha \vee \beta) \Rightarrow \gamma)$ und $\neg((\alpha \vee \beta) \Rightarrow \gamma)$ aussagenlogisch äquivalent?

Aufgabe 10.7 Geben Sie einen ausführlichen Beweis für Lemma 10.11.

Aufgabe 10.8 Geben Sie eine möglichst große Menge aussagenlogischer Formeln an, die als einzige atomare Teilformel die Atomformel A_0 (unter Umständen mehrfach) enthalten, und wo alle Formeln untereinander paarweise nicht aussagenlogisch äquivalent sind.

Aufgabe 10.9 Geben Sie eine möglichst große Menge aussagenlogischer Formeln an, die als einzige atomare Teilformeln die Atomformeln A_0 und A_1 (unter Umständen mehrfach) enthalten, und wo alle Formeln untereinander paarweise nicht aussagenlogisch äquivalent sind.

Aufgabe 10.10 Zeigen Sie, es bis auf logische Äquivalenz höchstens 2^{2^n} viele verschiedene aussagenlogische Formeln geben kann, deren atomare Teilformeln in der Menge $\{A_i \mid i = 1, \ldots, n\}$ liegen. (Verwenden Sie zunächst das Koinzidenzlemma um die Zahl der relevanten 0-1 Bewertungsfunktionen zu erhalten.)

Aufgaben zu Teilkapitel 10.3

Aufgabe 10.11 In welchen der nachfolgenden Fälle gilt $\Phi \models \gamma$?

1. $\Phi = \{A_0, A_1 \Rightarrow A_0\}$, $\gamma = A_1$,
2. $\Phi = \{(\neg A_1 \Rightarrow A_2) \vee (A_0 \Rightarrow A_1), A_0, (A_2 \Rightarrow A_3)\}$, $\gamma = A_1 \vee A_3$,

Aufgabe 10.12 Welche der folgenden Folgerungsbeziehungen sind für beliebige Formeln α, β, γ richtig?

1. $\{(\alpha \Rightarrow (\beta \Rightarrow \gamma))\} \models ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma))$,

2. $\{((\alpha \Leftrightarrow \beta) \Leftrightarrow \gamma)\} \models (\alpha \Leftrightarrow (\beta \Leftrightarrow \gamma))$.

Aufgabe 10.13 Es sei $\Phi \subseteq \mathcal{L}_0$ und $\gamma \in \mathcal{L}_0$. Zeigen Sie: $\Phi \models \gamma$ gilt genau dann, wenn $\Phi \cup \{\neg\gamma\}$ unerfüllbar ist.

Aufgabe 10.14 Zeigen Sie, daß der in 10.20 definierte *Cons*-Operator ein Hüllenoperator ist.

Aufgabe 10.15 Gibt es aussagenlogische Formeln γ mit der Eigenschaft $\{\gamma\} \models \neg\gamma$? Wenn nein, warum nicht? Wenn ja, welche? Gibt es Formeln γ, wo γ zu $\gamma \vee \neg\gamma$ aussagenlogisch äquivalent ist? Wenn nein, warum nicht? Wenn ja, welche?

Aufgaben zu Teilkapitel 10.4

Aufgabe 10.16 Beweisen Sie Lemma 10.33.

Aufgabe 10.17 Beweisen Sie: Zwei verschiedene Bewertungsformeln sind nie logisch äquivalent. Jede erfüllbare Formel in $\mathcal{L}_0(A_1, \ldots, A_n)$ ist logisch äquivalent zu einer Disjunktion über $k \geq 1$ Bewertungsformeln.

Aufgabe 10.18 Geben Sie für jede der Formeln in Abbildung 10.1 eine andere, logisch äquivalente Formel aus $\mathcal{L}_0(A_1, \ldots, A_n)$ an.

10.7 Bibliographische Angaben

Unter vielen Büchern zur mathematischen Logik seien [**?**] erwähnt.

11

Beweissysteme für die klassische Aussagenlogik

Wir wenden uns nun der Frage zu, welche mechanischen Verfahren es gibt, um zu zeigen, daß eine gegebene Formel eine aussagenlogische Tautologie ist, oder daß eine Formel aus einer gegebenen Menge von Prämissen folgt. Im Prinzip kann man für diese Probleme die bereits in Kapitel 1 vorgestellte Methode der Wahrheitswert-Tabellen verwenden, wenn man Probleme mit unendlichen Prämissenmengen, die zusätzliche Überlegungen erfordern, einmal außer acht läßt. Ein entscheidender Nachteil dieser Methode besteht aber darin, daß sie sich nicht gut auf anspruchsvollere Logiksysteme erweitern läßt. Auch kann man über die Eleganz der Methode streiten, da sie letztlich doch auf ein primitives Austesten hinausläuft. Aus diesen Beweggründen sind eine ganze Reihe anderer Beweisverfahren entwickelt worden. Einige dieser Systeme sollen in diesem Kapitel vorgestellt werden.

Die nun zu besprechenden formalen Beweissysteme, auch *Kalküle* genannt, erlauben es, nach bestimmten Regeln Formeln (gegebenenfalls aus einer Prämissenmenge) herzuleiten. Die folgenden Begriffe sind hierbei von zentraler Bedeutung.

Definition 11.1 Ein Kalkül heißt *korrekt* genau dann, wenn jede mit den Regeln des Kalküls herleitbare Formel eine aussagenlogische Tautologie ist (oder gegebenenfalls eine Konsequenz der Prämissenmenge). Der Kalkül heißt *vollständig* genau dann, wenn jede aussagenlogische Tautologie (und gegebenenfalls jede Konsequenz der Prämissenmenge) mit den Regeln des Kalküls herleitbar ist.

Korrektheit und Vollständigkeit sind also zwei wesentliche Eigenschaften, die man von jedem vernünftigen Kalkül erwarten wird.

11.1 Hilbert-Kalkül für die klassische Aussagenlogik

Die Bedeutung des *Hilbert*[1]*-Kalküls*, der manchmal auch Frege[2]-Kalkül genannt wird, beruht vor allem darauf, daß man „Hilbert-Typ“ Kalküle für viele andere Logiken (und nicht nur für das aussagenlogische Fragment) angeben kann, wobei das Schema, mit dem man die Vollständigkeit des Kalküls beweist, jeweils im Prinzip ähnlich ist. Der Kalkül eignet sich hingegen nicht als Beweissystem, wenn man an einer effizienten Implementierung interessiert ist.

Axiome des Kalküls sind alle $\mathcal{L}_0$-Formeln der Gestalt:

$$
\begin{array}{ll}
(1) & \alpha \Rightarrow \alpha \\
(2) & \alpha \Rightarrow (\beta \Rightarrow \alpha) \\
(3) & (\alpha \Rightarrow \beta) \Rightarrow ((\beta \Rightarrow \gamma) \Rightarrow (\alpha \Rightarrow \gamma)) \\
(4) & (\alpha \Rightarrow (\beta \Rightarrow \gamma)) \Rightarrow ((\alpha \Rightarrow \beta) \Rightarrow (\alpha \Rightarrow \gamma)) \\
(5) & (\alpha \Rightarrow (\alpha \vee \beta)), (\beta \Rightarrow (\alpha \vee \beta)) \\
(6) & (\alpha \Rightarrow \gamma) \Rightarrow ((\beta \Rightarrow \gamma) \Rightarrow ((\alpha \vee \beta) \Rightarrow \gamma)) \\
(7) & ((\alpha \wedge \beta) \Rightarrow \alpha), ((\alpha \wedge \beta) \Rightarrow \beta) \\
(8) & (\gamma \Rightarrow \alpha) \Rightarrow ((\gamma \Rightarrow \beta) \Rightarrow (\gamma \Rightarrow (\alpha \wedge \beta))) \\
(9) & ((\alpha \wedge \beta) \vee \gamma) \Rightarrow ((\alpha \vee \gamma) \wedge (\beta \vee \gamma)) \\
 & ((\alpha \vee \gamma) \wedge (\beta \vee \gamma)) \Rightarrow ((\alpha \wedge \beta) \vee \gamma) \\
(10) & ((\alpha \vee \beta) \wedge \gamma) \Rightarrow ((\alpha \wedge \gamma) \vee (\beta \wedge \gamma)) \\
 & ((\alpha \wedge \gamma) \vee (\beta \wedge \gamma)) \Rightarrow ((\alpha \vee \beta) \wedge \gamma) \\
(11) & (\alpha \Rightarrow \beta) \Rightarrow (\neg\beta \Rightarrow \neg\alpha) \\
(12) & (\alpha \wedge \neg\alpha) \Rightarrow \beta \\
(13) & \beta \Rightarrow (\alpha \vee \neg\alpha) \\
(14) & (\alpha \Rightarrow \beta) \Rightarrow (\beta \vee \neg\alpha) \\
(15) & (\beta \vee \neg\alpha) \Rightarrow (\alpha \Rightarrow \beta)
\end{array}
$$

[1]David Hilbert, deutscher Mathematiker, 1862-1943.
[2]G. Frege, 1848-1925.

Regeln: Der Kalkül hat nur eine Regel, den *Modus Ponens*: Sind $\mathcal{L}_0$-Formeln der Form α und $\alpha \Rightarrow \beta$ abgeleitet, so darf β hingeschrieben werden.

Man beachte, daß man in (1)–(15) für α, β und γ beliebige Formeln einsetzen darf. Dadurch ergeben sich aus diesen 15 *Axiomenschemata* unendlich viele *Axiome*. In derselben Weise kann man den Modus Ponens auf beliebige Formeln α und β anwenden.

Definition 11.2 Eine *Herleitung der $\mathcal{L}_0$-Formel φ aus der Formelmenge Φ* ist eine endliche Folge $\langle \gamma_1, \ldots, \gamma_n \rangle$ $(n \geq 1)$ von Formeln aus $\mathcal{L}_0$ derart, daß gilt:

- $\gamma_n = \varphi$,
- jedes γ_i ist ein Axiom oder eine Formel aus Φ oder folgt durch Anwendung des Modus Ponens auf Formeln γ_j, γ_k mit $j, k < i$ $(i = 1, \ldots, n)$.

Wenn φ aus Φ herleitbar ist, so schreiben wir kurz $\Phi \vdash \varphi$.

Selbst bei offensichtlich gültigen Folgerungsbeziehungen ist es meist recht schwierig, im Hilbert-Kalkül eine Herleitung zu finden.

Beispiel 11.3 Als Beispiel geben wir in Abbildung 11.1 eine Herleitung von $\varphi := A_1$ aus $\Phi := \{A_0 \vee A_1, \neg A_0\}$ an. Als Abkürzungen verwenden wir $\alpha := (A_0 \vee A_1)$, $\beta := \neg A_0$ und $\gamma := (A_3 \Rightarrow A_3)$ sowie $\alpha_1 := (A_0 \wedge \neg A_0)$, und $\beta_1 := (A_1 \wedge \neg A_0)$.

Korrektheit und Vollständigkeit

Die *Korrektheit* des Hilbert-Kalküls ist einfach zu zeigen.

Lemma 11.4 (Korrektheit des Hilbert-Kalküls) *Für alle $\Phi \subseteq \mathcal{L}_0$ und $\varphi \in \mathcal{L}_0$ folgt aus $\Phi \vdash \varphi$ stets $\Phi \models \varphi$.*

Beweis. Wir verwenden Induktion über die Herleitungslänge.
Induktionsanfang. Es sei γ mit einer Herleitung der Länge 1 aus Φ herleitbar.

[1]	α	ist aus Φ,
[2]	β	ist aus Φ,
[3]	$(\gamma \Rightarrow \alpha) \Rightarrow ((\gamma \Rightarrow \beta) \Rightarrow (\gamma \Rightarrow (\alpha \wedge \beta)))$	Axiom (8),
[4]	$\alpha \Rightarrow (\gamma \Rightarrow \alpha)$	Axiom (2),
[5]	$\gamma \Rightarrow \alpha$	MP aus [4], [1],
[6]	$((\gamma \Rightarrow \beta) \Rightarrow (\gamma \Rightarrow (\alpha \wedge \beta)))$	MP aus [5], [3],
[7]	$\beta \Rightarrow (\gamma \Rightarrow \beta)$	Axiom (2),
[8]	$\gamma \Rightarrow \beta$	MP aus [2], [7],
[9]	$\gamma \Rightarrow (\alpha \wedge \beta)$	MP aus [6], [8],
[10]	γ	Axiom (1),
[11]	$\alpha \wedge \beta$	MP aus [10], [9],
[12]	$(\alpha \wedge \beta) \Rightarrow ((A_0 \wedge \neg A_0) \vee (A_1 \wedge \neg A_0))$	Axiom (10,1),
[13]	$(A_0 \wedge \neg A_0) \vee (A_1 \wedge \neg A_0)$	MP aus [11], [12],
[14]	$(A_0 \wedge \neg A_0) \Rightarrow A_1$	Axiom (12),
[15]	$(A_1 \wedge \neg A_0) \Rightarrow A_1$	Axiom (7,1),
[16]	$(\alpha_1 \Rightarrow A_1) \Rightarrow ((\beta_1 \Rightarrow A_1) \Rightarrow ((\alpha_1 \vee \beta_1) \Rightarrow A_1))$	Axiom (6),
[17]	$((\beta_1 \Rightarrow A_1) \Rightarrow ((\alpha_1 \vee \beta_1) \Rightarrow A_1))$	MP aus [16],[14],
[18]	$((\alpha_1 \vee \beta_1) \Rightarrow A_1))$	MP aus [17], [15],
[19]	A_1	MP aus [18],[13],

Abbildung 11.1: Ableitung von A_1 aus $\{A_0 \vee A_1, \neg A_0\}$ im Hilbert-Kalkül; Bezeichnungen vgl. Beispiel 11.3.

Dann ist γ ein Axiom oder eine Formel aus Φ. Man verifiziert sofort mit der üblichen Wahrheitswert-Tabellen Methode, daß alle Axiome Tautologien sind (vgl. Satz 10.17). In beiden Fällen folgt somit, daß $\Phi \models \gamma$ gilt.
Als *Induktionsvoraussetzung* nehmen wir an, daß $\Phi \models \delta$ gilt für alle Formeln δ, die aus Φ mit einer Herleitung der Länge $\leq n$ ableitbar sind. Es sei nun $\langle \gamma_1, \ldots, \gamma_{n+1} \rangle$ eine Herleitung der Länge $n+1$ von $\gamma = \gamma_{n+1}$ aus Φ. Dann ist γ_{n+1} ein Axiom, eine Formel aus Φ, oder durch Anwendung des Modus Ponens auf Formeln γ_i und $\gamma_j = \gamma_i \Rightarrow \gamma_{n+1}$ mit $i, j \leq n$ entstanden. In den ersten beiden Fällen folgt wie beim Induktionsanfang $\Phi \models \gamma_{n+1}$. Im dritten Fall gilt nach Induktionsvoraussetzung $\Phi \models \gamma_i$ und $\Phi \models \gamma_i \Rightarrow \gamma_{n+1}$. Daraus folgt aber sofort $\Phi \models \gamma_{n+1}$. ■

Schwieriger ist es, die *Vollständigkeit* des Hilbert-Kalküls nachzuweisen. Wir stellen zuerst die Idee kurz dar, bevor wir uns an die technische Durchführung machen. Wir wollen zeigen, daß für jede Formelmenge Φ und jede Formel φ stets aus $\Phi \models \varphi$ auch $\Phi \vdash \varphi$ folgt. Wir verwenden Kontraposition und zeigen, daß in jeder Situation, wo $\Phi \vdash \varphi$ nicht gilt, auch $\Phi \models \varphi$ nicht gelten kann. Die Idee ist hierbei, eine Bewertung in eine Boolesche Algebra $\mathcal{B}$ anzugeben, die alle Formeln aus Φ auf die 1 von $\mathcal{B}$ abbildet, φ hingegen auf ein Element $b \neq 1$. Die Boolesche Algebra $\mathcal{B}$ wird ein *Gegenmodell* zu $\Phi \models \varphi$ genannt. Der Hintergrund beruht auf Satz 10.32, aus dem sich ergibt, daß aus der Existenz eines Gegenmodells folgt, daß $\Phi \models \varphi$ nicht gelten kann.

Die entscheidende Frage ist nun, wie wir ein Gegenmodell finden können. Die Grundidee, die auch bei vielen Vollständigkeitsbeweisen anderer Logiksysteme verwendet wurde, besteht darin, das gesuchte Gegenmodell aus dem syntaktischen Material — hier also aus den Formeln — zu bilden, wobei der Ableitungsbegriff des Kalküls verwendet wird, um eine geeignete Struktur auf diesem Material zu erklären. Im hier vorliegenden Fall verwenden wir den Ableitungsbegriff, um auf der Formelalgebra *Form*(At) eine geeignete Kongruenzrelation zu definieren, wo die Quotientenstruktur eine Boolesche Algebra ist. Von dieser werden wir dann zeigen, daß sie in der Tat ein Gegenmodell darstellt.

Wir definieren nun also zunächst die Relation $\sim_\Phi$ als Teilmenge von $\mathcal{L}_0 \times \mathcal{L}_0$ wie folgt: es gelte $\sigma \sim_\Phi \gamma$ genau dann, wenn $\Phi \vdash \sigma \Rightarrow \gamma$ und $\Phi \vdash \gamma \Rightarrow \sigma$. Vereinfachend könnten wir beide Bedingungen in der Form $\Phi \vdash \sigma \Leftrightarrow \gamma$ zusammenfassen (vgl. Aufgabe 11.1). Die nachfolgenden Propositionen charakterisieren diese Relation näher.

Proposition 11.5 $\sim_\Phi$ *ist eine Kongruenzrelation auf* $\mathrm{Form}(At)$.

Beweis. 1. Wir zeigen zunächst, daß $\sim_\Phi$ eine Äquivalenzrelation auf $\mathcal{L}_0$ ist: die Symmetrie von $\sim_\Phi$ ist trivial, die Reflexivität folgt aus Axiom (1). Zur Transitivität: es gelte $\Phi \vdash \sigma \Rightarrow \gamma$ und $\Phi \vdash \gamma \Rightarrow \delta$. Ist eine Herleitung von $\sigma \Rightarrow \gamma$ und $\gamma \Rightarrow \delta$ aus Φ gegeben, so fügen wir zunächst

$$(\sigma \Rightarrow \gamma) \Rightarrow ((\gamma \Rightarrow \delta) \Rightarrow (\sigma \Rightarrow \delta))$$

als Instanz von Axiom (3) hinzu. Durch Anwendung des Modus Ponens erhalten wir nun $(\gamma \Rightarrow \delta) \Rightarrow (\sigma \Rightarrow \delta)$, durch eine nochmalige Anwendung dann $\sigma \Rightarrow \delta$, demnach gilt $\Phi \vdash \sigma \Rightarrow \delta$. Also ist $\sim_\Phi$ eine Äquivalenzrelation.

2.1. Wir zeigen, daß $\sim_\Phi$ kongruent bezüglich der Operation $f_\neg$ ist: es gelte $\sigma \sim_\Phi \gamma$, also $\Phi \vdash \sigma \Rightarrow \gamma$ und $\Phi \vdash \gamma \Rightarrow \sigma$. Wir ergänzen eine Herleitung von $\sigma \Rightarrow \gamma$ aus Φ zunächst durch $(\sigma \Rightarrow \gamma) \Rightarrow (\neg\gamma \Rightarrow \neg\sigma)$ (Axiom (11)). Durch Anwenden des Modus Ponens erhalten wir dann $\neg\gamma \Rightarrow \neg\sigma$. Also gilt $\Phi \vdash \neg\gamma \Rightarrow \neg\sigma$. Symmetrisch folgt $\Phi \vdash \neg\sigma \Rightarrow \neg\gamma$. Damit gilt $\neg\sigma \sim_\Phi \neg\gamma$.

2.2. Wir zeigen, daß $\sim_\Phi$ kongruent bezüglich der Operation $f_\wedge$ ist: es gelte $\sigma \sim_\Phi \sigma'$ und $\gamma \sim_\Phi \gamma'$. Also gibt es eine Herleitungen von $\sigma \Rightarrow \sigma'$ und $\gamma \Rightarrow \gamma'$ aus Φ. Diese Herleitungen können wir durch $(\sigma \wedge \gamma) \Rightarrow \sigma$ und $(\sigma \wedge \gamma) \Rightarrow \gamma$ (Axiome (7)) verlängern. Vermöge Axiom (3) kann man aus den Formeln $(\sigma \wedge \gamma) \Rightarrow \sigma$ und $\sigma \Rightarrow \sigma'$ durch zweimaliges Anwenden des Modus Ponens nun die Formel $(\sigma \wedge \gamma) \Rightarrow \sigma'$, und analog aus $(\sigma \wedge \gamma) \Rightarrow \gamma$ sowie $\gamma \Rightarrow \gamma'$ die Formel $(\sigma \wedge \gamma) \Rightarrow \gamma'$ herleiten. Aus diesen beiden Formeln und Axiom (8) erhält man durch zweimaliges Anwenden des Modus Ponens die Formel $((\sigma \wedge \gamma) \Rightarrow (\sigma' \wedge \gamma'))$. Es gilt also $\Phi \vdash ((\sigma \wedge \gamma) \Rightarrow (\sigma' \wedge \gamma'))$. Symmetrisch folgt $\Phi \vdash ((\sigma' \wedge \gamma') \Rightarrow (\sigma \wedge \gamma))$. Also gilt $(\sigma \wedge \gamma) \sim_\Phi (\sigma' \wedge \gamma')$.

2.3 Der Nachweis, daß $\sim_\Phi$ kongruent bezüglich der Operation $f_\vee$ ist, bleibt dem Leser überlassen.

2.4 Um schließlich nachzuweisen, daß $\sim_\Phi$ eine Kongruenzrelation bezüglich $f_\Rightarrow$ ist, zeigen wir, daß $\sigma \sim_\Phi \sigma'$ und $\gamma \sim_\Phi \gamma'$ stets $\sigma \Rightarrow \gamma \sim_\Phi \sigma' \Rightarrow \gamma'$ implizieren. Dazu verlängern wir eine Herleitung, die die Formeln $\sigma \Rightarrow \sigma'$, $\sigma' \Rightarrow \sigma$, $\gamma' \Rightarrow \gamma$ und $\gamma \Rightarrow \gamma'$ enthält, wie folgt:

$$
\begin{array}{lr}
(\gamma \Rightarrow \gamma') \Rightarrow (\sigma \Rightarrow (\gamma \Rightarrow \gamma')) & \text{aus (2)} \\
\sigma \Rightarrow (\gamma \Rightarrow \gamma') & \text{mit MP} \\
(\sigma \Rightarrow (\gamma \Rightarrow \gamma')) \Rightarrow ((\sigma \Rightarrow \gamma) \Rightarrow (\sigma \Rightarrow \gamma')) & \text{aus (4)} \\
(\sigma \Rightarrow \gamma) \Rightarrow (\sigma \Rightarrow \gamma') & \text{mit MP} \\
(\sigma \Rightarrow \gamma') \Rightarrow (\sigma \Rightarrow \gamma) & \text{erhält man analog}
\end{array}
$$

Also gilt $(\sigma \Rightarrow \gamma) \sim_\Phi (\sigma \Rightarrow \gamma')$. Ferner sind ableitbar

$$
\begin{array}{lr}
(\sigma' \Rightarrow \sigma) \Rightarrow ((\sigma \Rightarrow \gamma') \Rightarrow (\sigma' \Rightarrow \gamma')) & \text{Axiom (3)} \\
(\sigma \Rightarrow \gamma') \Rightarrow (\sigma' \Rightarrow \gamma') & \text{mit MP} \\
(\sigma' \Rightarrow \gamma') \Rightarrow (\sigma \Rightarrow \gamma') & \text{analog}
\end{array}
$$

Also gilt $(\sigma \Rightarrow \gamma') \sim_\Phi (\sigma' \Rightarrow \gamma')$. Mittels der Transitivität von $\sim_\Phi$ erhalten wir nun in der Tat $\sigma \Rightarrow \gamma \sim_\Phi \sigma' \Rightarrow \gamma'$. ■

Lemma 11.6 *Die Quotientenstruktur* $Form(At)/\sim_\Phi$ *ist eine Boolesche Algebra. Die Restklasse* $1 := [A_0 \vee \neg A_0]$ *besteht genau aus den aus* Φ *herleitbaren Formeln.*

Beweis. Auf der Menge $\mathcal{L}_0/\sim_\Phi$ der Äquivalenzklassen definieren wir eine partielle Ordnung wie folgt: $[\sigma] \leq [\gamma]$ genau dann, wenn $\Phi \vdash \sigma \Rightarrow \gamma$. Wir haben zunächst zu zeigen, daß $\leq$ unabhängig von den gewählten Repräsentanten ist. Gilt $\Phi \vdash \sigma \Rightarrow \gamma$, $\sigma \sim_\Phi \sigma'$ und $\gamma \sim_\Phi \gamma'$, so folgt wie oben bei der Transitivität $\Phi \vdash \sigma' \Rightarrow \gamma$, bei nochmaliger Anwendung dann $\Phi \vdash \sigma' \Rightarrow \gamma'$. Der Nachweis, daß $\leq$ eine partielle Ordnung ist, ist nun leicht.

Wir zeigen nun, daß sich auf $\mathcal{L}_0/\sim_\Phi$ Operationen $\sqcup, \sqcap, \sqsubset, -$ einführen lassen, die $\langle \mathcal{L}_0/\sim_\Phi, \sqcup, \sqcap, \sqsubset, - \rangle$ zu einer Booleschen Algebra machen. Bezüglich $\leq$ exisitieren Suprema und Infima über zwei-elementige Mengen: Es folgt aus (7) $[\sigma \wedge \gamma] \leq [\sigma]$ und $[\sigma \wedge \gamma] \leq [\gamma]$. Aus (8) und dem Modus Ponens folgt $[\sigma \wedge \gamma] = \mathit{inf}\{[\sigma], [\gamma]\}$, ähnlich sieht man mit (5) und (6), daß $[\sigma \vee \gamma] = \mathit{inf}\{[\sigma], [\gamma]\}$. Definieren wir $[\sigma] \sqcup [\gamma] = [\sigma \vee \gamma]$ und $[\sigma] \sqcap [\gamma] = [\sigma \wedge \gamma]$, so ist demnach $\langle \mathcal{L}_0/\sim_\Phi, \sqcup, \sqcap \rangle$ ein Verband. Wegen (9) und (10) ist dieser Verband distributiv. Wir setzen $-[\sigma] = [\neg\sigma]$. Aus (11) folgt leicht, daß $-$ wohldefiniert ist. Aus (1), (12) und (6) folgt, daß stets $\Phi \vdash ((\sigma \wedge \neg\sigma) \vee \gamma) \Rightarrow \gamma$, umgekehrt folgt $\Phi \vdash \gamma \Rightarrow ((\sigma \wedge \neg\sigma) \vee \gamma)$ sofort aus (5). Demnach gilt $([\sigma] \sqcap -[\sigma]) \sqcup [\gamma] = [\gamma]$. Ähnlich folgt $([\sigma] \sqcup -[\sigma]) \sqcap [\gamma] = [\gamma]$, d.h. $\langle \mathcal{L}_0/\sim_\Phi, \sqcup, \sqcap, \sqsubset, - \rangle$ ist eine Boolesche Algebra (mit $a \sqsubset b = -a \sqcup b$).

Offenkundig bilden die aus Φ herleitbaren Formeln gerade die Restklasse $1 = [A_0 \vee \neg A_0]$.

Es bleibt noch zu zeigen, daß die Boolesche Algebra $\langle \mathcal{L}_0/{\sim_\Phi}, \sqcup, \sqcap, \sqsubset, - \rangle$ gerade die Quotientenalgebra von $Form(At)$ modulo $\sim_\Phi$ ist. Die Operationen $\sqcup, \sqcap, -$ sind offenkundig gemäß Definition genau die Operationen der Quotientenalgebra von $Form(At)$ modulo $\sim_\Phi$ Weiter ergibt sich aus (14) und (15), daß $[\sigma] \sqsubset [\gamma] = [\gamma] \sqcup -[\sigma] = [\gamma \vee \neg\sigma] = [\sigma \Rightarrow \gamma]$ ist. Also ist $\langle \mathcal{L}_0/{\sim_\Phi}, \sqcup, \sqcap, \sqsubset, - \rangle$ in der Tat die Quotientenalgebra von $Form(At)$ modulo $\sim_\Phi$. ■

Im nachfolgenden Beweis können wir nun Nutzen ziehen aus dem Begriff der allgemeinen Bewertung und Satz 10.32.

Satz 11.7 (Vollständigkeit des Hilbert-Kalküls) *Für alle $\Phi \subseteq \mathcal{L}_0$ und $\varphi \in \mathcal{L}_0$ folgt aus $\Phi \models \varphi$ stets $\Phi \vdash \varphi$.*

Beweis. Falls $\Phi \vdash \varphi$ nicht gilt, so ist nach dem vorangegangenen Lemma der kanonische Homomorphismus von $Form(At)$ nach $Form(At)/{\sim_\Phi}$ eine Bewertung, die alle Formeln aus Φ auf 1, φ jedoch nicht auf 1 abbildet. Daher gilt nach Satz 10.32 auch nicht $\Phi \models \varphi$. ■

Bemerkung 11.8 Um aus der Formelalgebra das Gegenmodell $Form(At)/{\sim_\Phi}$ zu erhalten, hatten wir im wesentlichen Formeln identifiziert, die äquivalent bezüglich der Herleitbarkeit aus Φ waren. Da gemäß Satz 11.7 der syntaktische Herleitungsbegriff mit dem semantischen Folgerungsbegriff zusammenfällt, erhalten wir aber dieselbe Algebra, wenn wir genau die Elemente σ und γ identifizieren, für die $\Phi \models \sigma \Leftrightarrow \gamma$ gilt. Dies zeigt, daß $Form(At)/{\sim_\Phi}$ genau die aus Abschnitt 10.4 bekannte Lindenbaumalgebra zur Formelmenge Φ ist.

Aus der Vollständigkeit des Hilbert-Kalküls erhalten wir nun sofort eine interessante Schlußfolgerung.

Satz 11.9 (Kompaktheitssatz der Aussagenlogik) *Es sei Φ eine möglicherweise unendliche Menge von Formeln und $\varphi \in \mathcal{L}_0$. Gilt $\Phi \models \varphi$, so gibt es eine endliche Teilmenge Φ_0 von Φ mit $\Phi_0 \models \varphi$.*

Beweis. Gilt $\Phi \models \varphi$, so nach Satz 11.7 auch $\Phi \vdash \varphi$ und es existiert eine Herleitung von φ aus Φ. In dieser können nur endlich viele Formeln aus Φ auftreten. Ist Φ_0 die Menge der Formeln aus Φ, die in der Herleitung auftreten, so gilt $\Phi_0 \vdash \varphi$, folglich nach Lemma 11.4 auch $\Phi_0 \vdash \varphi$. ■

Eine etwas andere Formulierung des Kompaktheitssatzes ergibt sich, wenn wir den Erfüllbarkeitsbegriff verwenden.

Satz 11.10 *Es sei Φ eine möglicherweise unendliche Menge von Formeln. Falls jede endliche Teilmenge von Φ erfüllbar ist, so auch Φ.*

Beweis. Angenommen, Φ sei unerfüllbar. Dann gilt $\Phi \models A_0 \wedge \neg A_0$. Folglich existiert nach dem vorangegangenen Satz eine endliche Teilmenge Φ_0 von Φ so daß $\Phi_0 \models A_0 \wedge \neg A_0$. Dies impliziert aber, daß Φ_0 unerfüllbar ist, in Widerspruch zu unserer Annahme. Daher muß auch Φ erfüllbar sein. ■

11.2 Der Sequenzen-Kalkül G'

Der nun zu besprechende Kalkül, der durch seine Klarheit und Symmetrie besticht, ist auch unter dem Namen *Gentzens*[3] *System G'* bekannt. Der Kalkül operiert nicht auf Formeln, sondern auf komplexeren Ausdrücken, den sogenannten Sequenzen.

Sequenzenformat der Aussagenlogik

Definition 11.11 Eine *Sequenz* ist ein Paar $\langle \Gamma, \Delta \rangle$ von endlichen Folgen $\Gamma = \langle \gamma_1, \ldots, \gamma_m \rangle$ und $\Delta = \langle \delta_1, \ldots, \delta_n \rangle$ von aussagenlogischen Formeln. Dabei kann jede der beiden Folgen leer sein, nicht jedoch alle beide. Die Folge Γ heißt das *Antezedens*, die Folge Δ heißt das *Sukzedens* der Sequenz $\langle \Gamma, \Delta \rangle$.

Üblicherweise werden Sequenzen in der Form $\Gamma \rightarrow \Delta$ notiert. Man sollte sich von der mannigfachen Verwendung des Zeichens $\rightarrow$ nicht verwirren lassen—hier sind definitiv keine Abbildungen von Γ nach Δ gemeint. Intuitiv gesehen steht $\Gamma \rightarrow \Delta$ für die Formel

$$(\gamma_1 \wedge \ldots \wedge \gamma_m) \Rightarrow (\delta_1 \vee \ldots \vee \delta_n).$$

[3]Gerhard Gentzen, ...

Man beachte, daß in dieser Formel die Glieder des Antezedens Γ konjunktiv, die Glieder des Sukzedens hingegen *disjunktiv* zu lesen sind! Selbst nach dieser Übersetzung bleibt etwas unklar, wie Sequenzen zu interpretieren sind, wo das Antezedens oder das Sukzedens leer sind. Daher werden wir formal eine logisch äquivalente Formel verwenden, wo sich dieses Interpretationsproblem nicht stellt.

Definition 11.12 Die zu einer Sequenz $\Gamma \to \Delta$ der Form $\gamma_1, \ldots, \gamma_m \to \delta_1, \ldots, \delta_n$ *assoziierte Formel* ist

$$\neg\gamma_1 \vee \ldots \vee \neg\gamma_m \quad \vee \quad \delta_1 \vee \ldots \vee \delta_n.$$

Diese Formel werden wir kurz in der Form $\bigvee \bar{\Gamma} \vee \bigvee \Delta$ notieren.

Definition 11.13 Eine Sequenz $\Gamma \to \Delta$ heißt *gültig*, im Zeichen $\models \Gamma \to \Delta$, genau dann, wenn die assoziierte Formel $\bigvee \bar{\Gamma} \vee \bigvee \Delta$ eine aussagenlogische Tautologie ist.

Der einfache Beweis des nachfolgenden Lemmas bleibt als Übungsaufgabe offen (vgl. Aufgabe 11.5).

Lemma 11.14 *Es sei $\Gamma \to \Delta$ eine Sequenz, wo alle Formeln aus Γ respektive Δ Atomformeln sind. Dann ist $\Gamma \to \Delta$ gültig genau dann, wenn die Folgen Γ und Δ ein gemeinsames Element haben.*

Die in der Einleitung des Kapitels beschriebenen Probleme zu entscheiden, ob eine aussagenlogischge Formel eine Tautologie ist, oder ob sie aus anderen gegebenen Formeln folgt, lassen sich nun in ein Problem der Gültigkeit einer Sequenz übersetzen.

Lemma 11.15 *Eine aussagenlogische Formel α ist genau dann eine aussagenlogische Tautologie, wenn die Sequenz $\to \alpha$ (leeres Antezedens) gültig ist. Ist $\Phi = \{\sigma_1, \ldots, \sigma_n\}$ eine endliche Menge aussagenlogischer Formeln, so gilt $\Phi \models \alpha$ genau dann, wenn die Sequenz $\sigma_1, \ldots, \sigma_n \to \alpha$ gültig ist.*

Beweis. Wir zeigen nur den zweiten Teil. Es gelte $\Phi \models \alpha$. Dann ist nach Korollar 10.25 die Formel $(\sigma_1 \wedge \ldots \wedge \sigma_n) \Rightarrow \alpha$, damit auch die logisch äquivalente Formel $\neg\sigma_1 \vee \ldots \vee \neg\sigma_n \vee \alpha$ eine aussagenlogische Tautologie. Somit ist $\sigma_1, \ldots, \sigma_n \to \alpha$ gültig. Die Umkehrung folgt analog. ■

Kalkülregeln und Herleitbarkeit

Gentzens Kalkül G' hat als *Axiome* alle Sequenzen der Form $\Gamma \to \Delta$, wo Γ und Δ ein gemeinsames Element haben. Die nachfolgende Tabelle faßt die *Ableitungsregeln* des Kalküls zusammen. Hierbei steht etwa Γ, α, Λ für eine Formelfolge, die an irgendeiner Stelle die Formel α enthält.

$$\frac{\Gamma,\alpha_1,\alpha_2,\Delta\to\Lambda}{\Gamma,\alpha_1\wedge\alpha_2,\Delta\to\Lambda}\ (\wedge\text{: links}) \qquad \frac{\Gamma\to\Delta,\beta_1,\Lambda \quad \Gamma\to\Delta,\beta_2,\Lambda}{\Gamma\to\Delta,\beta_1\wedge\beta_2,\Lambda}\ (\wedge\text{: rechts})$$

$$\frac{\Gamma,\beta_1,\Delta\to\Lambda \quad \Gamma,\beta_2,\Delta\to\Lambda}{\Gamma,\beta_1\vee\beta_2,\Delta\to\Lambda}\ (\vee\text{: links}) \qquad \frac{\Gamma\to\Delta,\alpha_1,\alpha_2,\Lambda}{\Gamma\to\Delta,\alpha_1\vee\alpha_2,\Lambda}\ (\vee\text{: rechts})$$

$$\frac{\Gamma,\Delta\to\beta_1,\Lambda \quad \beta_2,\Gamma,\Delta\to\Lambda}{\Gamma,\beta_1\Rightarrow\beta_2,\Delta\to\Lambda}\ (\Rightarrow\text{: links}) \qquad \frac{\alpha_1,\Gamma\to\Delta,\alpha_2,\Lambda}{\Gamma\to\Delta,\alpha_1\Rightarrow\alpha_2,\Lambda}\ (\Rightarrow\text{: rechts})$$

$$\frac{\Gamma,\Delta\to,\alpha,\Lambda}{\Gamma,\neg\alpha,\Delta\to\Lambda}\ (\neg\text{: links}) \qquad \frac{\alpha,\Gamma\to\Delta,\Lambda}{\Gamma\to\Delta,\neg\alpha,\Lambda}\ (\neg\text{: rechts})$$

Jede Regel hat eine oder zwei Prämissen, die über dem Strich notiert sind, und eine Konklusion, die unter dem Strich erscheint. Die Regeln haben Namen, die rechts vom Strich notiert sind. Die Formel, die in der Konklusion neu entsteht, wird *Hauptformel* der Regeln genannt. Die mit der Hauptformel zusammenhängende(n) Formel(n) der Prämisse(n) sind die *Nebenformeln*. Für jeden Junktor gibt es zwei Regeln, die es erlauben, den Junktor auf der linken bzw. auf der rechten Seite einzuführen.

Die nachfolgend einzuführenden Beweisbäume sind Bäume im Sinn von Kapitel 7, wo die Knoten mit Sequenzen dekoriert sind.

Definition 11.16 Die Menge der *Beweisbäume* ist die kleinste Menge von gelabelten Bäumen, die unter den folgenden Bildungsregeln abgeschlossen ist:

1. jeder Baum, der aus einem einzigen gelabelten Blatt besteht, das mit einem Axiom gelabelt ist, ist ein Beweisbaum der Höhe 0.

2. Für jeden Beweisbaum der Höhe n, dessen Wurzel mit einer Sequenz $\Gamma \to \Delta$ gelabelt ist, die die einzige Prämisse einer Ableitungsregel mit Konklusion $\Gamma' \to \Delta'$ darstellt, ist der erweiterte Baum, der entsteht, wenn man eine neue Wurzel $\Gamma' \to \Delta'$ unten anfügt, ein Beweisbaum der Höhe $n + 1$.

3. Sind T_1 und T_2 Beweisbäume der Höhe n_1 und n_2, deren Wurzeln mit den Sequenzen $\Gamma_1 \to \Delta_1$ und $\Gamma_2 \to \Delta_2$ gelabelt sind, die die beiden Prämissen einer Ableitungsregel mit Konklusion $\Gamma' \to \Delta'$ darstellen, so ist derjenige Baum ein Beweisbaum, dessen Wurzel mit $\Gamma' \to \Delta'$ gelabelt ist, und dessen unmittelbare Teilbäume die Bäume T_1 und T_2 sind. Seine Höhe ist $max\{n_1, n_2\} + 1$.

Die Menge der *Ableitungsbäume* ist analog definiert, nur daß wir in 1. als Blätter nicht nur Axiome, sondern beliebige Sequenzen zulassen. Ein Ableitungsbaum T heißt *abgeschlossen*, wenn jedes Blatt von T mit einer Sequenz gelabelt ist, die nur noch atomare Formeln enthält.

Typischerweise wird bei der graphischen Repräsentation von Beweisbäumen die Wurzel als *unterster* Knoten geführt, Statt Kanten zu den Kindern zu zeichnen, werden die Kinder über einer horizontalen Linie aufgeführt, wie bei den Kalkülregeln. Zur Illustration hier ein Beispiel für einen abgeschlossenen Ableitungsbaum.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\longrightarrow A_0, A_0, A_1 \qquad A_1 \longrightarrow A_0, A_1}{(A_0 \Rightarrow A_1) \longrightarrow A_0, A_1}}{(A_0 \Rightarrow A_1) \longrightarrow (A_0 \vee A_1)}}{\longrightarrow (A_0 \vee A_1), \neg(A_0 \Rightarrow A_1)}}{\neg(A_0 \vee A_1) \longrightarrow \neg(A_0 \Rightarrow A_1)} \qquad \dfrac{\dfrac{\dfrac{A_1 \longrightarrow A_0, A_1}{A_1 \longrightarrow (A_0 \vee A_1)}}{\longrightarrow (A_0 \vee A_1), \neg A_1}}{\neg(A_0 \vee A_1) \longrightarrow \neg A_1}}{\neg(A_0 \vee A_1) \longrightarrow \neg(A_0 \Rightarrow A_1) \wedge \neg A_1}$$

Von den beiden unmittelbaren Teilbäumen ist der rechte ein Beweisbaum, der linke hingegen nicht.

Definition 11.17 Eine Sequenz $\Gamma \to \Delta$ heißt *herleitbar*, im Zeichen $\vdash \Gamma \to \Delta$, genau dann, wenn es einen Beweisbaum mit Wurzel $\Gamma \to \Delta$ gibt.

Korrektheit und Vollständigkeit

Um Korrektheit und Vollständigkeit zu beweisen, stellen wir zunächst einige einfache Hilfsbeobachtungen zusammen.

Proposition 11.18 *Für jede Regel des Kalküls G' der Form*

$$\frac{\Gamma \to \Delta}{\Gamma' \to \Delta'}$$

sind die assoziierten Formeln $\bigvee \bar{\Gamma} \vee \bigvee \Delta$ und $\bigvee \bar{\Gamma'} \vee \bigvee \Delta'$ aussagenlogisch äquivalent. Für jede Regel des Kalküls G' der Form

$$\frac{\Gamma_1 \to \Delta_1 \quad \Gamma_2 \to \Delta_2}{\Gamma' \to \Delta'}$$

ist die Konjunktion $(\bigvee \bar{\Gamma_1} \vee \bigvee \Delta_1) \wedge (\bigvee \bar{\Gamma_2} \vee \bigvee \Delta_2)$ zur assoziierten Formel $\bigvee \bar{\Gamma'} \vee \bigvee \Delta'$ aussagenlogisch äquivalent.

Beweis. Dies folgt unter Verwendung des Ersetzungslemmas 10.19 durch einfache Inspektion (Übung). ■

Proposition 11.19 *Alle Axiome des Sequenzen-Kalküls sind gültig. Für jede ein-Prämissenregel gilt: die Prämisse der Regel ist gültig genau dann, wenn die Konklusion der Regel gültig ist. Für jede zwei-Prämissenregel gilt: die beiden Prämissen der Regel sind gültig genau dann, wenn die Konklusion der Regel gültig ist.*

Beweis. Die zu den Axiomen nach Definition 11.12 assoziierten Disjunktionen enthalten stets eine Teildisjunktion der Form $\alpha \vee \neg\alpha$ und sind damit aussagenlogische Tautologien. Also sind alle Axiome nach Definition 11.13 gültige Sequenzen. Die weiteren Behauptungen folgen sofort aus der vorangegangenen Proposition. ■

Theorem 11.20 (Korrektheit und Vollständigkeit des Kalküls G') *Eine Sequenz $\Gamma \to \Delta$ ist gültig genau dann, wenn $\vdash \Gamma \to \Delta$ gilt.*

1. Die Korrektheit des Kalküls G' ist einfach zu zeigen: es sei $\Gamma \to \Delta$ eine herleitbare Sequenz. Dann gibt es einen Beweisbaum T mit Wurzel

$\Gamma \to \Delta$. Aus Proposition 11.19 folgt durch eine einfache Induktion über die Höhe des Beweisbaums, daß $\Gamma \to \Delta$ gültig ist.

2. Wir zeigen die Vollständigkeit des Kalküls G'. Es sei $\Gamma \to \Delta$ eine gültige Sequenz. Es sei m die Anzahl der in der Sequenz $\Gamma \to \Delta$ vorkommenden Junktoren. Wenn wir die Regeln des Kalküls etwas näher anschauen, sieht man, daß Rückwärtsanwendungen der Regeln immer einen Junktor eliminieren. Man wird also, wenn man von einer zu testenden Sequenz $\Gamma \to \Delta$ startet, durch iterierte Rückwärtsanwendungen der Regeln schließlich einen Ableitungsbaum der Höhe $\leq m$ konstruieren, der bei Blättern endet, die mit Sequenzen der Form $\Gamma' \to \Delta'$ gelabelt sind, die auf beiden Seiten nur Aussagenvariablen enthalten. Aus Proposition 11.19 folgt durch eine einfache Induktion, daß all diese Blatt-Sequenzen $\Gamma' \to \Delta'$ gültig sind. Aus Lemma 11.14 folgt nun, daß all diese Sequenzen Axiome sind. Daher ist der konstruierte Ableitungsbaum eine Beweisbaum und es gilt $\vdash \Gamma \to \Delta$. ■

Ein Entscheidungsverfahren

Wir können nun mit Hilfe des Kalküls G' entscheiden, ob eine aussagenlogische Formel φ aus einer *endlichen* Menge von Formeln $\Phi = \{\sigma_1, \ldots, \sigma_n\}$ folgt. Dies ist, wie wir gesehen haben, genau dann der Fall, wenn die Sequenz $\sigma_1, \ldots, \sigma_n \to \varphi$ gültig ist. Um nun die Gültigkeit von $\sigma_1, \ldots, \sigma_n \to \varphi$ zu entscheiden, brauchen wir—wie der Beweis von Theorem 11.20 zeigte—nur beginnend mit $\sigma_1, \ldots, \sigma_n \to \varphi$ die Regeln von G' rückwärts anzuwenden. Falls hierbei gleich mehrere invertierte Regeln auf eine Sequenz anwendbar sind, so wählen wir irgendeine davon aus, um den Ableitungsbaum weiterzuentwickeln. Die Höhe des entstehenden Ableitungsbaums ist durch die Anzahl der Junktoren in $\sigma_1, \ldots, \sigma_n \to \varphi$ begrenzt. Wenn der entstehende Ableitungsbaum abgeschlossen ist, so prüfen wir nach, ob alle Blätter Axiome sind. In diesem Fall, aber auch nur in diesem Fall, ist die Sequenz $\sigma_1, \ldots, \sigma_n \to \varphi$ gültig. In der Tat, falls alle Blätter Axiome sind, so ist der entstandenen Ableitungsbaum ein Beweisbaum, und es folgt mit Theorem 11.20, daß $\sigma_1, \ldots, \sigma_n \to \varphi$ gültig ist. Ist aber nur eines der entstandenen Blätter kein Axiom, so ist es keine gültige Sequenz, wie Lemma 11.14 zeigt. Aus Proposition 11.19 folgt jetzt durch eine einfache Induktion, daß auch $\sigma_1, \ldots, \sigma_n \to \varphi$ nicht gültig ist.

Als Korollar erhalten wir die schon aus dem vorigen Kapitel bekannte

Entscheidbarkeit der klassischen Aussagenlogik.

Theorem 11.21 *Es ist für vorgegebene Formeln $\gamma \in \mathcal{L}_0$ und $\sigma_1, \ldots, \sigma_n \in \mathcal{L}_0$ entscheidbar, ob γ eine aussagenlogische Tautologie ist, beziehungsweise ob γ aus $\{\sigma_1, \ldots, \sigma_n\}$ folgt.*

Der Schnitt

Eine Regel, die es oft erlaubt, Beweisbäume stark zu verkleinern, ohne die Korrektheit des Kalküls zu beeinträchtigen, ist die sogenannte „Schnittregel", oft auch englisch „cut" genannt:

$$\frac{\Gamma \to \Delta, \alpha \qquad \alpha, \Lambda \to \Theta}{\Gamma, \Lambda \to \Delta, \Theta}$$

Es ist leicht nachprüfbar, daß Anwendungen der Schnittregel auf gültige Sequenzen wieder zu einer gültigen Sequenz führen. Damit können wir die Schnittregel zu G' hinzunehmen, ohne daß neue (und damit ungültige) Sequenzen herleitbar würden. Die Bedeutung des Schnitts wir deutlicher, wenn wir eine vereinfachte Variante betrachten, wo die Folge Δ leer ist, und wo Θ aus nur einem Element besteht:

$$\frac{\Gamma \to \alpha \qquad \alpha, \Lambda \to \theta}{\Gamma, \Lambda \to \theta.}$$

Man sieht hier, daß man die *Schnittformel* α als eine Hilfsbehauptung ansehen kann. Falls wir aus Γ die Hilfsbehauptung α beweisen können, und falls aus α und Λ dann θ folgt, so ist damit gezeigt, daß θ aus Γ und Λ folgt.

Es gibt einen einfachen Grund dafür, daß wir die Schnittregel nicht direkt zum Kalkül hinzugefügt haben. Wir können den Cut nicht ohne weiteres invertieren. Das nämlich würde voraussetzen, daß wir die aus dem Beweis „geschnittene" Formel α kennen. Die Schnittformel ist aber gerade nicht aus der Konklusion rekonstruierbar. Beweise mit und ohne Schnitt geben einen wichtigen Unterschied zwischen dem Vorgang des Beweisens beim Menschen und beim Computer wieder. Der Mathematiker wird typischerweise bei komplexen Beweisen eine ganze Reihe geeigneter Zwischenbehauptungen einführen. Zum Auffinden geeigneter Zwischenbehauptungen gehört eine gute Portion Kreativität und Intuition. Es ist bislang auch nicht annähernd gelungen, diese Art von Kreativität so zu formalisieren, daß sie einem Computer in Form eines Programms oder Algorithmus zugänglich gemacht werden kann.

Dies ist ein wichtiger Grund dafür, daß die Möglichkeiten, Beweise vollautomatisch mit dem Computer zu finden, bis heute stark eingeschränkt sind.

Das Problem, im mit dem Schnitt erweiterten Kalkül G' die Schnittformel α zu finden, ist verwandt zu dem Problem, im Hilbert-Kalkül zu gegebener Formel β eine Formel α zu finden, so daß α und $\alpha \Rightarrow \beta$ herleitbar sind, um danach per Modus Ponens β herzuleiten. In beiden Fällen hat man zunächst keinen Anhaltspunkt, wie die gesuchte Formel α aussieht, da sie ja aus der resultierenden Formel verschwunden ist.

11.3 G' als Widerlegungs-Kalkül

In unserem bisherigen Bild des Sequenzen-Kalküls G' haben wir eine Sequenz $\Gamma \to \Delta$ durch die assoziierte Formel $\bigvee \bar{\Gamma} \vee \bigvee \Delta$ interpretiert. Beweisbäume waren damit Schlußfolgen, die uns von einfachsten, tautologisch richtigen Sequenzen an den Blättern hin zu komplexeren Tautologien an der Wurzel führten. Basis hierfür war die Beobachtung, daß die Konklusion einer Sequenzenregel genau dann gültig ist, wenn *alle* Prämissen ebenfalls gültig sind (Proposition 11.19). Wir können das bisherige *Bild des Gentzen-Kalküls G' als affirmativen Kalkül* wie folgt zusammenfassen:

- die Herleitbarkeit einer Sequenz ist äquivalent zur Gültigkeit der assoziierten Formel,
- die Formeln innerhalb einer Sequenz sind disjunktiv zu lesen,
- die Formeln des Antezedens sind hierbei negiert zu denken,
- Aufspaltungen im Ableitungsbaum sind konjunktiv zu lesen (vgl. Prop. 11.18).

Die nun einzuführende alternative Betrachtungsweise soll vorab an einem Beispiel verdeutlicht werden.

Beispiel 11.22 Wir betrachten die Sequenz

$$(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0)) \quad \to \quad A_2.$$

Im bisherigen Bild repräsentiert der Beweisbaum

$$
\dfrac{
\dfrac{A_0 \rightarrow A_1,A_0,A_2 \qquad A_1 \rightarrow A_1,A_0,A_2}{(A_0 \vee A_1) \rightarrow A_1,A_0,A_2}
\qquad
\dfrac{
\dfrac{A_2,A_0 \rightarrow A_0,A_2 \qquad A_2,A_1 \rightarrow A_0,A_2}{A_2,(A_0 \vee A_1) \rightarrow A_0,A_2}
\qquad
\dfrac{A_0,A_0 \rightarrow A_0,A_2 \qquad A_0,A_1 \rightarrow A_0,A_2}{A_0,(A_0 \vee A_1) \rightarrow A_0,A_2}
}{(A_2 \vee A_0),(A_0 \vee A_1) \rightarrow A_0,A_2}
}{
\dfrac{(A_0 \vee A_1),(A_1 \Rightarrow (A_2 \vee A_0)) \rightarrow A_0,A_2}{(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0)) \rightarrow A_2}
}
$$

eine Bestätigung dafür, daß die assoziierte Formel

$$\neg(A_0 \vee A_1) \vee \neg(\neg A_0) \vee \neg(A_1 \Rightarrow (A_2 \vee A_0)) \quad \vee \quad A_2 \qquad (*)$$

eine aussagenlogische Tautologie ist. Die Begründung war, in kurzen Worten, die folgende: die zu den Blättern assoziierten Formeln enthalten allesamt eine Teildisjunktion der Form $A_i \vee \neg A_i$ und sind damit Tautologien. Wegen Proposition 11.19 folgt dann induktiv auch, daß $\neg(A_0 \vee A_1) \vee \neg(\neg A_0) \vee \neg(A_1 \Rightarrow (A_2 \vee A_0)) \vee A_2$ eine Tautologie ist.

Dualisierung. Wir können nun aber auch eine duale Perspektive einnehmen. Hierzu assoziieren wir zu $(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0) \rightarrow A_2)$ die neue Formel

$$(A_0 \vee A_1) \wedge \neg A_0 \wedge (A_1 \Rightarrow (A_2 \vee A_0)) \quad \wedge \quad \neg A_2.$$

Man beachte, daß sich diese Formel wegen der de Morganschen Regel bis auf logische Äquivalenz genau aus der *Negation* der oben betrachteten Formel $(*)$ ergibt. Inhaltlich steht diese Formel für die—nachfolgend zu widerlegende—Annahme, daß alle Formeln im Antezedens wahr sind, alle Formeln des Sukzedens hingegen falsch. Die Formeln einer Sequenz sind nun also konjunktiv zu lesen, wobei die Glieder des Sukzedens zu negieren sind. Die Erfüllbarkeit einer Sequenz in dieser Leseart würde nun voraussetzen, daß zumindest eine Vorgängersequenz erfüllbar ist, wie man sich leicht klarmacht. Induktiv verallgemeinert mußte man dann unter der obigen Annahme eine erfüllbare Formel an den Blättern finden. Wie man jedoch sieht, enthalten die zu den Blättern in dieser Leseart assoziierten Formeln allesamt eine Teilkonjunktion der Form $A_i \wedge \neg A_i$ und sind daher unerfüllbar. Daraus können wir schließen, daß auch die Formel $(A_0 \vee A_1) \wedge \neg A_0 \wedge (A_1 \Rightarrow (A_2 \vee A_0)) \wedge \neg A_2$ unerfüllbar ist.

Um ein allgemeines Bild des Gentzen-Kalküls G' als *Widerlegungs-Kalkül* zu erhalten, führen wir die zu einer Sequenz dual assoziierte Formel ein.

Definition 11.23 Die zu einer Sequenz $\Gamma \to \Delta$ der Form $\gamma_1, \ldots, \gamma_m \to \delta_1, \ldots, \delta_n$ *dual assoziierte Formel* ist die Formel

$$\bigwedge \Gamma \wedge \bigwedge \bar{\Delta} \quad := \quad \gamma_1 \wedge \ldots \wedge \gamma_m \quad \wedge \quad \neg\delta_1 \wedge \ldots \wedge \neg\delta_n.$$

Es gelten folgende duale Versionen der Propositionen 11.18 und 11.19.

Proposition 11.24 *Für jede Regel des Kalküls G' der Form*

$$\frac{\Gamma \to \Delta}{\Gamma' \to \Delta'}$$

sind die dual assoziierten Formeln $\bigwedge \Gamma \wedge \bigwedge \bar{\Delta}$ und $\bigwedge \Gamma' \wedge \bigwedge \bar{\Delta}'$ aussagenlogisch äquivalent. Für jede Regel des Kalküls G' der Form

$$\frac{\Gamma_1 \to \Delta_1 \quad \Gamma_2 \to \Delta_2}{\Gamma' \to \Delta'}$$

ist die Disjunktion $(\bigwedge \Gamma_1 \wedge \bigwedge \bar{\Delta}_1) \vee (\bigwedge \Gamma_2 \wedge \bigwedge \bar{\Delta}_2)$ zur dual assoziierten Formel $\bigwedge \Gamma' \wedge \bigwedge \bar{\Delta}'$ aussagenlogisch äquivalent.

Beweis. Die erste Behauptung ist sehr einfach zu zeigen: nach Proposition 11.18 wissen wir, daß die zu den beiden Sequenzen (disjunktiv) assoziierten Formeln aussagenlogisch äquivalent sind. Da sich die dual assoziierten Formeln aus diesen bis auf logische Äquivalenz durch Negation ergeben, sind sie ebenfalls aussagenlogisch äquivalent. Die zweite Behauptung folgt analog mit den de Morganschen Regeln. ■

Proposition 11.25 *Die zu den Axiomen des Sequenzen-Kalküls G' dual assoziierten Formeln sind unerfüllbar. Für jede ein-Prämissenregel gilt: die Prämisse repräsentiert eine erfüllbare Formel genau dann, wenn dies auch auf die Konklusion der Regel zutrifft. Für jede zwei-Prämissenregel gilt: die Konklusion der Regel repräsentiert eine erfüllbare Formel genau dann, wenn dies auch auf zumindest eine der Prämissen zutrifft.*

Beweis. Dies folgt unmittelbar aus Proposition 11.19. ■

Da wir den Begriff des Beweisbaums völlig unverändert lassen, ergibt sich aus Theorem 11.20 sofort die folgende duale Version.

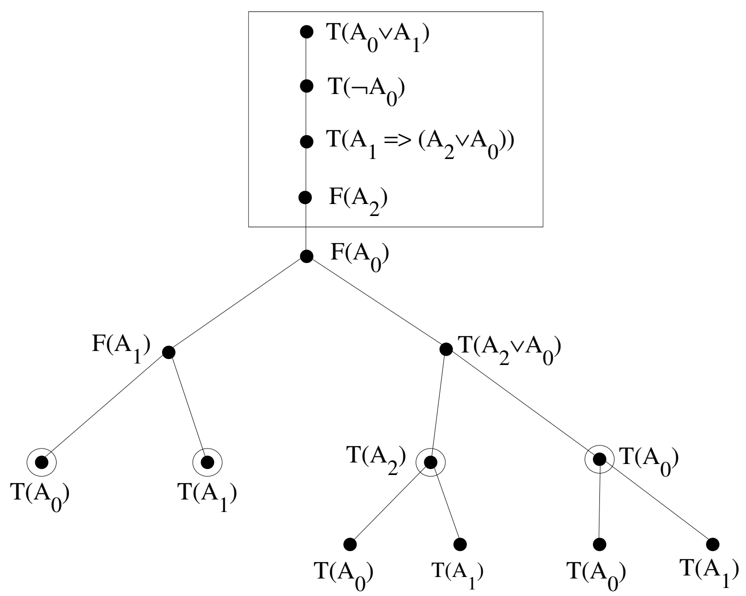
Theorem 11.26 (Korrektheit und Vollständigkeit des Kalküls G') *Eine Sequenz $\Gamma \to \Delta$ ist herleitbar (d.h. $\vdash \Gamma \to \Delta$) genau dann, wenn die dual assoziierte Formel $\bigwedge \Gamma \wedge \bigwedge \bar{\Delta}$ unerfüllbar ist.*

Abschließend können wir das *Bild des Gentzen-Kalküls G' als Widerlegungs-Kalkül* wie folgt zusammenfassen:

- die Herleitbarkeit einer Sequenz ist äquivalent zur Unerfüllbarkeit der dual assoziierten Formel,
- die Formeln innerhalb einer Sequenz sind konjunktiv zu lesen,
- die Formeln des Sukzedens sind hierbei negiert zu lesen,
- Aufspaltungen im Ableitungsbaum sind disjunktiv zu lesen.

11.4 Tableau-Kalküle

Die im vorangegangenen Abschnitt besprochene Beweismethode, die auf dem Invertieren der Regeln des Gentzen-Kalküls beruht, hat Vor- und Nachteile, wenn man an eine Automatisierung denkt. Es ist angenehm, daß man letztlich nur die Blätter des konstruierten Baums inspizieren muß, um festzustellen, ob die zu testende Sequenz gültig ist. Sehr unangenehm hingegen ist, daß man ständig alle nicht-bearbeiteten Formeln von Knoten zu Vorgängerknoten weiterkopieren muß, was augenscheinlich eine beträchtliche Redundanz darstellt. Ein Beispiel hierfür sind Beweise für die Distributivität der Junktoren $\wedge$ und $\vee$, wie wir sie in Beispiel 11.22 in einem speziellen Fall gegeben hatten. Wenn wir in Kauf nehmen, daß wir ganze Zweige inspizieren, kann man sich diese Kopierarbeit weitgehend sparen. Bei der nun zu besprechenden Methode der *signierten Tableaus* notieren wir beim Auflösen einer zusammengesetzten Formel einfach, welche Anteile sich für die linke und rechte Seite der nach oben entstehenden Sequenzen ergeben. Formeln γ, die bei der Umkehrung der Gentzen-Regeln auf der linken (bzw, rechten) Seite einer Sequenz entstehen, werden mit einem entsprechenden Label versehen. Der Interpretation des Gentzen-Kalküls G' als Widerlegungs-Kalkül folgend werden wir alle Formeln γ, die in einem Antezedens (Sukzedens) entstehen, in der Form $T(\gamma)$ (bzw. $F(\gamma)$) markieren. Hierbei steht „T“ für „true“ und „F“ für „false“. Da die in einem bestimmten Schritt nicht bearbeiteten Formeln nun nicht mehr zu den Nachfolgerknoten weiterkopiert werden sollen, kann man sich nicht mehr darauf beschränken, nur diejenige Formel weiteraufzulösen, die den zuletzt erreichten Knoten labelt. Vielmehr müssen wir allgemeiner eine Zugriff auf all diejenigen Formeln erlauben, die im jeweiligen Ast des bisher konstruierten Baumes zu finden sind. Diese Formeln innerhalb eines Asts sind konjunktiv zu lesen, wie die Glieder einer Sequenz des Kalküls G' im Widerlegungsbild. Aus dem in Beispiel 11.22 beschriebenen Beweis für die Sequenz $(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0)) \rightarrow A_2$ im Gentzen-Kalkül G' wird vermöge der so skizzierten Übersetzungstechnik das folgende signierte semantische Tableau:



Das obere Rechteck markiert die signierten Formeln, die sich aus der Übersetzung der Sequenz $(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0)) \rightarrow A_2$ ergeben. Da in Beispiel 11.22 alle Blätter Axiome waren, und damit ein gemeinsames Atom auf der linken und der rechten Seite (Antezedens respektive Sukzedens) hatten, wird es nicht überraschen, daß wir nun entsprechend in jedem Ast des Tableaus ein Paar von Atomformeln mit komplementärer Markierung „T“ beziehungsweise „F“ erhalten (die betreffenden Formeln sind durch Kreise markiert). Da wir—wie im Bild des Gentzen-Kalküls G' als Widerlegungs-Kalkül—die Formeln entlang eines Asts konjunktiv lesen wollen, repräsentieren aber komplementär gelabelte Formeln einen offenkundigen Widerspruch. Die verschiedenen Äste des Tableaus stellen aber eine vollständige Fallunterscheidung dar. Das Tableau *widerlegt* demnach die im markierten Rechteck getroffene Annahme, indem es zeigt, daß es keine 0-1 Bewertungsfunktion geben kann, die $A_0 \vee A_1$, $\neg A_0$ und $A_1 \Rightarrow (A_2 \vee A_0)$ wahr macht, aber A_2 falsch. Wenn wir den Blickwinkel nun wieder umkehren, *beweist* es damit die Gültigkeit der Sequenz $(A_0 \vee A_1), \neg A_0, (A_1 \Rightarrow (A_2 \vee A_0)) \rightarrow A_2$.

Wir wollen nun die so skizzierte Idee näher formalisieren.

Definition 11.27 Eine *signierte Formel* der Aussagenlogik ist ein Ausdruck der Form $T(\alpha)$ oder $F(\alpha)$, wo α eine aussagenlogische Formel ist. Die *zugehörigen unsignierten Formeln* sind α (für $T(\alpha)$) respektive $\neg\alpha$ (für $F(\alpha)$).

Definition 11.28 Eine signierte Formel heißt *erfüllbar* genau dann, wenn die zugehörige unsignierte Formel erfüllbar ist. Eine Menge signierter Formeln heißt *erfüllbar* genau dann, wenn die Menge der zugehörigen unsignierten Formeln erfüllbar ist.

Gemäß allgemeiner Gepflogenheit werden wir wie im obigen Beispiel Tableaus stets *von oben noch unten* entwickeln, und damit die Orientierung im Gentzen-Kalkül G' invertieren. Aus der Übersetzung der Regeln des Kalküls G' erhalten wir dann folgende Regeln.

Definition 11.29 *Erweiterungsregeln* für signierte semantische Tableaus für die klassische Aussagenlogik:

$$\frac{T(\alpha_1 \wedge \alpha_2)}{\begin{array}{c}T(\alpha_1)\\ T(\alpha_2)\end{array}} \qquad \frac{F(\beta_1 \wedge \beta_2)}{F(\beta_1) \qquad F(\beta_2)}$$

$$\frac{T(\beta_1 \vee \beta_2)}{T(\beta_1) \qquad T(\beta_2)} \qquad \frac{F(\alpha_1 \vee \alpha_2)}{\begin{array}{c}F(\alpha_1)\\ F(\alpha_2)\end{array}}$$

$$\frac{T(\beta_1 => \beta_2)}{T(\beta_2) \qquad F(\beta_1)} \qquad \frac{F(\alpha_1 => \alpha_2)}{\begin{array}{c}T(\alpha_1)\\ F(\alpha_2)\end{array}}$$

$$\frac{T(\neg\alpha)}{F(\alpha)} \qquad \frac{F(\neg\alpha)}{T(\alpha)}$$

Die komplexe signierte Formel über dem Strich werden wir die *Hauptformel* der Regel nennen, die signierte(n) Formel(n) unter dem Strich die *Nebenformel(n)* der Regel. Regeln mit Teilformeln β_1, β_2 heißen *verzweigend*, Regeln mit Teilformeln α_1, α_2 oder α heißen *unverzweigend*.

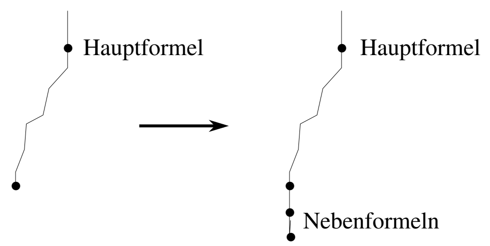
Man sieht in offenkundiger Weise, daß die Tableau-Regeln erschöpfende Fallunterscheidungen darstellen. Formaler können wir diesen Sachverhalt wie folgt wiedergeben:

Proposition 11.30 *Es sei g eine* 0-1 *Bewertung. Es sei α die Hauptformel einer unverzweigenden Tableau-Regel, α' (resp. α_1 sowie α_2) seien die Nebenformel(n) der Regel. Dann gilt $g(\alpha) = 1$ genau dann, wenn $g(\alpha') = 1$ (resp. $g(\alpha_1) = 1$ und $g(\alpha_2) = 1$). Es sei β die Hauptformel einer verzweigenden Tableau-Regel, β_1 sowie β_2 seien die Nebenformeln der Regel. Dann gilt $g(\beta) = 1$ genau dann, wenn $g(\beta_1) = 1$ oder $g(\beta_2) = 1$.*

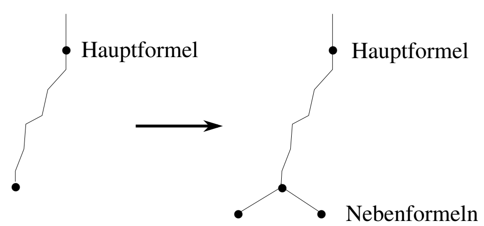
Beweis. Trivial. ■

Es bleibt noch die Art und Weise zu klären, wie die Tableau-Erweiterungsregeln angewandt werden können.

1. Finden wir in einem Ast an einer beliebigen Stelle die Hauptformel einer unverzweigenden Regel, so können wir den Ast an der Spitze linear um einen neuen Knoten für jede Nebenformel erweitern.



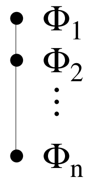
2. Finden wir in einem Ast an einer beliebigen Stelle die Hauptformel einer verzweigenden Regel, so können wir den tiefsten Knoten des Asts um zwei Kinderknoten erweitern. Die beiden Kinderknoten werden mit den zwei Nebenformeln gelabelt.



Dementsprechend repräsentieren verschiedenen Äste Alternativen, während die Formeln auf ein und demselben Ast konjunktiv zu lesen sind. Nun läßt sich die Menge der signierten aussagenlogischen Tableaus wie folgt definieren.

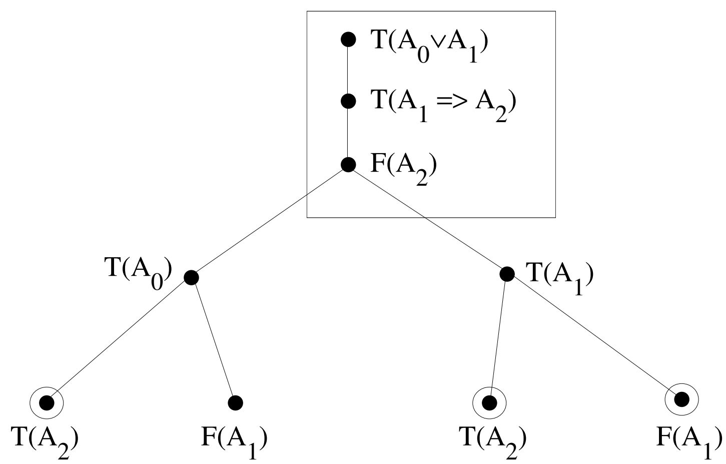
Definition 11.31 Es sei $\{\Phi_1, \ldots, \Phi_n\}$ eine endliche Menge signierter Formeln. Die Menge der *signierten semantischen Tableaus für* $\{\Phi_1, \ldots, \Phi_n\}$ ist die kleinste Menge gelabelter Bäume, die unter den folgenden zwei Konstruktionsregeln abgeschlossen ist:

1. Der folgende Ast ist ein signiertes semantisches Tableau für $\{\Phi_1, \ldots, \Phi_n\}$:



2. Ist $\mathcal{T}_0$ ein signiertes semantisches Tableau für $\{\Phi_1, \ldots, \Phi_n\}$, und ergibt sich $\mathcal{T}_1$ aus $\mathcal{T}_0$ durch Anwendung einer Tableau-Erweiterungsregel, so ist auch $\mathcal{T}_1$ ein signiertes semantisches Tableau für $\{\Phi_1, \ldots, \Phi_n\}$.

Beispiel 11.32 Ein signiertes semantisches Tableau für $\{T(A_0 \vee A_1), T(A_1 \Rightarrow A_2), F(A_2)\}$:



Lemma 11.33 *Es sei $M = \{\Phi_1, \ldots, \Phi_n\}$ eine Menge signierter Formeln. Falls M erfüllbar ist, so enthält jedes signierte semantische Tableau für M zumindest einen Ast, so daß die Menge aller signierter Formeln auf dem Ast erfüllbar ist.*

Beweis. Das signierte Tableau, das aus nur einem Ast besteht, dessen Knoten mit den Formeln aus M gelabelt sind, hat nach Voraussetzung die oben genannte Eigenschaft. Ist nun $\mathcal{T}_0$ ein signiertes Tableau für $\{\Phi_1, \ldots, \Phi_n\}$, das die genannte Eigenschaft besitzt, und ergibt sich $\mathcal{T}_1$ aus $\mathcal{T}_0$ durch Anwendung einer Tableau-Erweiterungsregel, so hat auch $\mathcal{T}_1$ die genannte Eigenschaft. Um dies einzusehen betrachte man einen Ast π_0 von $\mathcal{T}_0$, so daß die Menge aller signierter Formeln auf π erfüllbar ist. Wenn π_0 beim Erweiterungsschritt nicht betroffen war, sind wir fertig. Wenn der Ast π_0 erweitert wurde, so folgt aus Proposition 11.30, daß π_0 Teil eines Asts π_1 von $\mathcal{T}_1$ ist, so daß die Menge aller Formeln auf π_1 erfüllbar ist. Da damit die Menge der signierten semantischen Tableaus für $\{\Phi_1, \ldots, \Phi_n\}$, die die genannte Eigenschaft besitzen, unter den beiden Tableau-Bildungsregeln aus Definition 11.31 abgeschlossen sind, folgt aus Definition 11.31, daß alle signierten semantischen Tableaus für M die genannte Eigenschaft haben. ■

Definition 11.34 (i) Ein Ast eines Tableaus heißt *geschlossen*, wenn er zwei Knoten enthält, die mit komplementären signierten Formeln der Form $T(\alpha)$ und $F(\alpha)$ gelabelt sind. Andernfalls heißt der Ast *offen*.
(ii) Ein Tableau heißt *geschlossen*, wenn jeder seiner Äste geschlossen ist. Andernfalls heißt das Tableau *offen*.

Bei der nachfolgenden Definition halte man sich vor Augen, daß wir den Tableau-Kalkül als *Widerlegungs-Kalkül* beschreiben.

Definition 11.35 Ein *Tableau-Beweis für die Formel* γ ist ein geschlossenes Tableau für $\{F(\gamma)\}$. Ein *Tableau-Beweis von* γ *aus der Formelmenge* $\{\sigma_1, \ldots, \sigma_n\}$ ist ein geschlossenes Tableau für $\{F(\gamma), T(\sigma_1), \ldots, T(\sigma_n)\}$.

In der Einleitung dieses Abschnitts hatten wir an einem Beispiel angedeutet, wie sich Beweisbäume im Sequenzen-Kalkül G' in signierte semantische Tableaus übersetzen lassen. Es sei dem Leser überlassen, eine formale Übersetzung anzugeben, mit der sich ein ganz allgemein ein Beweisbaum im System G' für eine Sequenz $\sigma_1, \ldots, \sigma_n \to \varphi$ in einen Tableau-Beweis von φ aus $\{\sigma_1, \ldots, \sigma_n\}$ transformieren läßt. Aus einer solchen Übersetzung erhält man mit Theorem 11.20, daß es stets einen Tableau-Beweis von φ aus $\{\sigma_1, \ldots, \sigma_n\}$ gibt, wenn φ aus $\{\sigma_1, \ldots, \sigma_n\}$ folgt.

Gibt es umgekehrt einen Tableau-Beweis von φ aus $\{\sigma_1, \ldots, \sigma_n\}$, so kann $\{F(\varphi), T(\sigma_1), \ldots, T(\sigma_n)\}$ nach Lemma 11.33 offenkundig nicht erfüllbar sein. Also folgt in diesem Fall φ aus $\{\sigma_1, \ldots, \sigma_n\}$.

Zusammenfassend erhalten wir die Korrektheit und Vollständigkeit des beschriebenen Tableau-Kalküls.

Satz 11.36 *Es sei φ eine aussagenlogische Formel. Dann besitzt φ einen Tableau-Beweis genau dann, wenn φ eine aussagenlogische Tautologie ist. Es seien weiter $\sigma_1, \ldots, \sigma_n$ aussagenlogische Formeln. Dann besitzt φ einen Tableau-Beweis aus $\{\sigma_1, \ldots, \sigma_n\}$ genau dann, wenn $\Phi \models \varphi$ gilt.*

Tableaus als Entscheidungsverfahren

Wenn wir den Tableau-Kalkül zur Basis eines Entscheidungsverfahrens für die Aussagenlogik machen wollen, so sollten wir eine Vorsichtsmaßnahme treffen. In der bisherigen Formulierung kann ein Tableau für eine endliche Menge signierter Formeln beliebig groß werden. Der Grund hierfür ist, daß wir ein und dieselbe komplexe signierte Formel bislang beliebig oft zum Erweitern des Tableaus verwenden können. Dies ist natürlich bei einer Implementierung nicht zu akzeptieren. Ein naheliegender Ausweg besteht darin, jede komplexe signierte Formel nach ihrer Verwendung als „verbraucht“ zu markieren. Verbrauchte Formeln dürfen nachfolgend nicht mehr zum Erweitern des Tableaus verwendet werden. Ein solches Vorgehen setzt freilich voraus, daß wir die Formel bei ihrer erstmaligen Verwendung gleich zur simultanen Erweiterung all derjenigen Pfade verwenden, die durch den Knoten gehen, der mit der betreffenden Formel gelabelt ist. Andernfalls würde die Vollständigkeit des Tableau-Kalküls verloren gehen. Wenn wir tatsächlich diese Vorsichtsmaßnahme treffen, kann die Vollständigkeit nicht eingeschränkt werden, da ja eine mehrfache Verwendung der Formel in einem Ast nur zu einer Wiederholung bereits vorhandener Formeln führen würde (wir verzichten hier auf einen formalen Beweis). Wenn wir als *Gewicht* eines Asts die Anzahl der Junktoren, die in unmarkierten signierten Formeln des Asts vorkommen, festlegen, so sieht man, daß jeder Erweiterungsschritt das Astgewicht reduziert. Daher kann jeder Ast offenkundig nur maximal k mal erweitert werden, wo k die Anzahl der Junktoren in der Eingabemenge von signierten Formeln ist. Offenkundig hat dann das gesamte Tableau höchstens 2^{k+1} viele Knoten, und das Tableau-Entwicklungsverfahren terminiert.

Gegenmodelle aus offenen Tableaus

Eine attraktive Eigenschaft von Tableaus beruht darauf, daß man aus offenen Tableaus unter bestimmten Voraussetzungen 0-1 Bewertungen erhalten kann, die konkret zeigen, wie die Eingabeformelmenge zu erfüllen ist. Wir nennen ein Tableau *vollständig entwickelt*, wenn jeder seiner Äste die folgenden Bedingung erfüllt: jede komplexe Formel, die sich auf dem Ast befindet, wurde zumindest einmal zum Erweitern des Asts verwendet. Im vorangegangenen Abschnitt hatten wir eine Möglichkeit gezeigt, wie man in systematischer Weise mit endlich vielen Erweiterungen ein vollständig entwickeltes Tableau für eine gegebene endliche Menge signierter Formeln erhalten kann. Wenn nun ein solches vollständig entwickeltes Tableau einen offenen Ast enthält, so können wir dadurch eine partielle 0-1 Belegung erhalten, daß wir jede Atomformel A_i, die mit Signum „T" („F") im Ast auftritt, auf 1 (respektive 0) abbilden. Jede 0-1 Bewertung die diese partielle Bewertung erweitert, bildet die Menge der zu den Eingabeformeln gehörenden unsignierten Formeln auf 1 ab. Auch hier geben wir keinen formalen Beweis, sondern begnügen uns mit einem Beispiel.

Beispiel 11.37 Wir wollen testen, ob die Formeln $A_0 \vee A_1, \neg A_1, A_0 \Rightarrow (A_2 \vee A_3)$ die Formel A_2 impliziert. Da wir eine Widerlegungsstrategie verfolgen, untersuchen wir, ob die Formelmenge $\{A_0 \vee A_1, \neg A_1, A_0 \Rightarrow (A_2 \vee A_3), \neg A_2\}$ erfüllbar ist. Hierzu konstruieren wir das in Abbildung 11.2 dargestellte vollständig entwickelte Tableau für $\{T(A_0 \vee A_1), T(\neg A_1), T(A_0 \Rightarrow (A_2 \vee A_3)), F(A_2)\}$. Wir erhalten einen offenen Ast mit den atomaren signierten Formeln $T(A_0), F(A_1), F(A_2), T(A_3)$. Die 0-1 Bewertung, die A_0 und A_3 auf 1, A_1 und A_2 auf 0 abbildet, zeigt, daß $\{A_0 \vee A_1, \neg A_1, A_0 \Rightarrow (A_2 \vee A_3), \neg A_2\}$ erfüllbar ist.

11.5 Der Resolutions-Kalkül

Die Bedeutung des nun zu besprechenden Resolutions-Kalküls beruht vor allem darauf, daß die später betrachtete Erweiterung auf prädikatenlogische Formeln, die auf dem aussagenlogischen Formalismus aufbaut, eines der wichtigsten Beweisverfahren für die Prädikatenlogik darstellt.

Klauselformat des Aussagenlogik

Ausgangspunkt des Verfahrens ist eine Darstellung aussagenlogischer Formeln im sogenannten Klauselformat. Um diese Format einzuführen, besprechen wir zunächst zwei Normalformen für aussagenlogische Formeln.

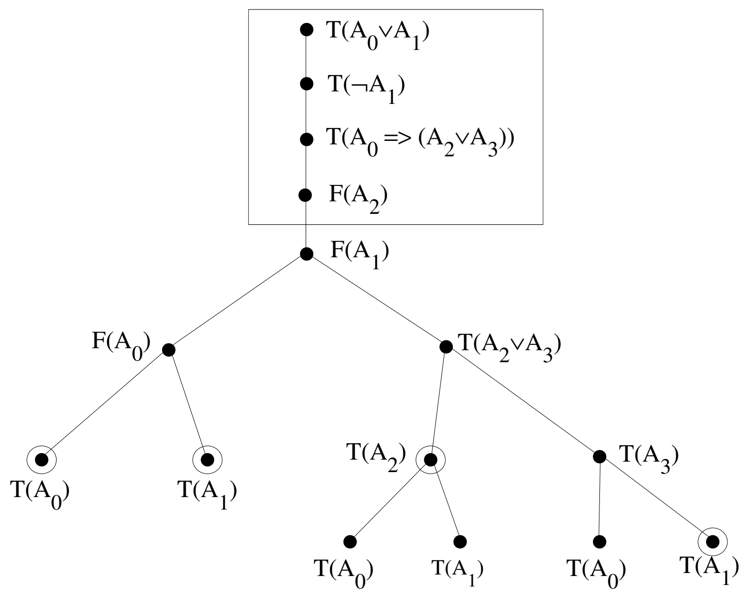


Abbildung 11.2: Ablesen eines Gegenmodells aus einem offenen Tableau

Definition 11.38 Ein *Literal* ist eine Aussagenvariable oder eine negierte Aussagenvariable. Zwei Literale der Form A_i und $\neg A_i$ werden *komplementär* genannt.

Sind L_1 und L_2 komplementäre Literale, so schreiben wir $\bar{L}_1 = L_2$ und $\bar{L}_2 = L_1$.

Definition 11.39 Eine aussagenlogische Formel α ist in *konjunktiver Normalform* falls sie die Form

$$\bigwedge_{i=1}^{n} (\bigvee_{j=1}^{m_i} L_{i,j})$$

hat, wobei die Teilformeln $L_{i,j}$ Literale sind. Eine aussagenlogische Formel α ist in *disjunktiver Normalform*, falls sie die Form

$$\bigvee_{i=1}^{n} (\bigwedge_{j=1}^{m_i} L_{i,j})$$

hat, wobei die Teilformeln $L_{i,j}$ Literale sind.

Satz 11.40 *Zu jeder aussagenlogischen Formel α kann man effektiv eine aussagenlogisch äquivalente Formel β berechnen, die in konjunktiver (disjunktiver) Normalform ist.*

Beweis. Der wohl schönste Beweis ergibt sich unter Verwendung des Gentzen-Kalküls G' wie in den Aufgaben 11.9 und 11.10 angedeutet. Nachfolgend skizzieren wir einen weiteren, direkten Beweis. Wir verwenden, daß die folgenden Formeln stets aussagenlogisch äquivalent sind:

$$\begin{array}{rcl}
\gamma \Rightarrow \delta & \text{und} & \delta \vee \neg\gamma \\
\neg(\gamma \wedge \delta) & \text{und} & \neg\delta \vee \neg\gamma \\
\neg(\gamma \vee \delta) & \text{und} & \neg\delta \wedge \neg\gamma \\
\neg(\neg\delta) & \text{und} & \delta \\
\gamma \vee (\delta_1 \wedge \delta_2) & \text{und} & (\gamma \vee \delta_1) \wedge (\gamma \vee \delta_2) \\
(\delta_1 \wedge \delta_2) \vee \gamma & \text{und} & (\delta_1 \vee \gamma) \wedge (\delta_2 \vee \gamma) \\
\gamma \wedge (\delta_1 \vee \delta_2) & \text{und} & (\gamma \wedge \delta_1) \vee (\gamma \wedge \delta_2) \\
(\delta_1 \vee \delta_2) \wedge \gamma & \text{und} & (\delta_1 \wedge \gamma) \vee (\delta_2 \wedge \gamma).
\end{array}$$

Damit können wir zunächst eine zur gegebenen Formel äquivalente Formel bilden, in der der Junktor „$\Rightarrow$“ nicht mehr vorkommt. Nachfolgend können wir die drei Regeln zur Negation so oft wie möglich in gerichteter Form von links nach rechts anwenden. Man kann sich überlegen, daß dieses Verfahren terminiert. Die resultierende Formel besitzt Negationszeichen „$\neg$“ nur noch direkt vor Atomformeln. Anschließend wenden wir die ersten beiden (resp. die beiden unteren) Distributivitätsregeln so oft wie möglich in gerichteter Form von links nach rechts an. Auch dieses Verfahren terminiert und liefert die konjunktive (resp. disjunktive) Normalform.

Der Resolutions-Kalkül verwendet Formeln in konjunktiver Normalform. Diese werden allerdings leicht verändert, nämlich als Mengen von Literalen dargestellt.

Definition 11.41 Eine *Klausel* ist eine—möglicherweise leere—endliche Menge von Literalen.

Die leere Klausel wird allgemein in der Form „$\Box$“ notiert.

Definition 11.42 Es sei $\varphi = \bigwedge_{i=1}^{n}(\bigvee_{j=1}^{m_i} L_{i,j})$ eine aussagenlogische Formel in konjunktiver Normalform. Die zu φ *assoziierte Klauselmenge* ist $M\varphi := \{\{L_{i,1}, \ldots, L_{i,m_i}\} \mid i = 1, \ldots, n\}$.

Definition 11.43 Die zu einer Klausel $K = \{L_1, \ldots, L_m\}$ *assoziierte Formel* ist die Disjunktion $\bigvee K := L_1 \vee \ldots \vee L_m$. Die zu einer Klauselmenge $M = \{K_1, \ldots, K_n\}$ *assoziierte Formel* ist die Formel in konjunktiver Normalform $\varphi_M := \bigwedge_{i=1}^{n}(\bigvee K_i)$.

Die zu einer Klausel assoziierte Formel ist natürlich nur bis auf eine irrelevante Reihenfolge der Disjunktionsglieder eindeutig. Die zu der leeren Klausel assoziierte „leere Disjunktion" ist bislang nicht formal als Formel eingeführt. Wir können jede unerfüllbare Formel (also etwa $A_0 \wedge \neg A_0$) als assoziierte Formel in diesem Spezialfall verwenden. Während die zur leeren Klausel assoziierte Formel damit unerfüllbar ist, ist für eine nichtleere Klausel K offensichtlich $\bigvee K$ stets erfüllbar. Wir definieren nun den Erfüllbarkeitsbegriff für Klauselmengen.

Definition 11.44 Eine Klauselmenge M heißt *erfüllbar*, falls die Menge der zu Klauseln aus M assoziierten Formeln erfüllbar ist im Sinn von Definition 10.13.

Lemma 11.45 *Eine aussagenlogische Formel in konjunktiver Normalform ist (un)erfüllbar genau dann, wenn die zugehörige Klauselmenge (un)erfüllbar ist.*

Hintergrund

Der Hintergrund des Resolutions-Kalküls läßt sich wie folgt darstellen. Wir interessieren uns eigentlich für das Problem, ob eine aussagenlogische Formel φ aus einer endlichen Menge von Formeln $\Phi = \{\sigma_1, \ldots, \sigma_n\}$ folgt. Dieses Problem wird nun wie folgt übersetzt in ein Problem, wo die Erfüllbarkeit einer endlichen Klauselmenge M zu testen ist:

Zunächst gilt nach Lemma 10.23 $\Phi \models \varphi$ genau dann, wenn $\{\sigma_1, \ldots, \sigma_n, \neg\varphi\}$ beziehungsweise $\sigma_1 \wedge \ldots \wedge \sigma_n \wedge \neg\varphi$ unerfüllbar ist. Es sei nun ψ eine logisch zu $\sigma_1 \wedge \ldots \wedge \sigma_n \wedge \neg\varphi$ äquivalente Formel in konjunktiver

Normalform und M_ψ die assoziierte Klauselmenge. Lemma 11.45 zeigt in Zusammenhang mit obigen Schritten, daß $\Phi \models \varphi$ genau dann gilt, wenn M_ψ unerfüllbar ist. Wir können daher das Ausgangsproblem darauf zurückführen, die (Un)erfüllbarkeit einer endlichen Klauselmenge zu testen. Genau einen solchen Test liefert der Resolutions-Kalkül.

Der Resolutions-Kalkül

Der Resolutions-Kalkül operiert auf Klauselmengen. Es gibt eine einzige Regel, die es erlaubt, neue Klauseln hinzuzufügen. Dies soll zuerst dargestellt werden, bevor wir das Rahmenschema des Kalküls beschreiben.

Definition 11.46 Es seien K_1 und K_2 zwei Klauseln einer Klauselmenge M, die komplementäre Literale $L \in K_1$ und $\bar{L} \in K_2$ enthalten. Dann heißt die Klausel

$$K := (K_1 \setminus \{L\}) \cup (K_2 \setminus \{\bar{L}\})$$

eine *Resolvente* der Klauseln K_1 und K_2. Die Resolvente darf zur Klauselmenge M hinzugefügt werden.

Beispiel 11.47 Es sei $K_1 = \{A_0, \neg A_1, A_3\}$, $K_2 = \{A_0, A_2, \neg A_3\}$. Dann ist $\{A_0, \neg A_1, A_2\}$ eine Resolvente von K_1 und K_2.

Die Eingabe des aussagenlogischen Resolutions-Kalküls ist eine Menge von Klauseln M. In jedem Schritt wird eine Resolvente aus zwei beliebigen Klauseln der aktuellen Klauselmenge gebildet, die zu der Menge der bereits existierenden Klauseln hinzugefügt wird. Das Verfahren kann auf zwei Weisen terminieren.

1. Es wird die leere Klausel „$\Box$“ erzeugt. Dann bricht das Verfahren ab und meldet die Unerfüllbarkeit von M.

2. Obwohl „$\Box$“ kein Element der Klauselmenge ist, kann keine neue Resolvente mehr gebildet werden, die nicht schon in der Menge der aktuell vorhandenen Klauseln läge. Dann bricht das Verfahren ab und meldet die Erfüllbarkeit von M.

Es ist ganz leicht einzusehen, daß das Verfahren für eine endliche Eingabemenge M stets terminiert, da man über der endlichen Menge von Aussagenvariablen, die in M dann vorhanden sind, nur endlich viele verschiedene Klauseln bilden kann.

Beispiel 11.48 Es sei M die Menge mit den Klauseln $K_1 := \{A_0, A_1\}$, $K_2 := \{\neg A_1\}$, $K_3 := \{A_2, \neg A_0\}$ und $K_4 := \{\neg A_2\}$. Dann können wir im Resolutions-Kalkül aus K_1 und K_2 die Klausel $K_5 := \{A_0\}$ resolvieren, anschließend aus K_3 und K_5 die Klausel $K_6 := \{A_2\}$. Schließlich ergibt sich aus K_6 und K_4 die leere Klausel „$\Box$“. Demnach ist die Klauselmenge M unerfüllbar.

Korrektheit und Vollständigkeit

Die Korrektheit des Resolutions-Kalküls ist leicht zu zeigen. Sie basiert auf folgendem Lemma.

Lemma 11.49 *Es sei K_0 eine Resolvente der Klauseln K_1 und K_2 der Klauselmenge M. Dann folgt $\bigvee K_0$ aus $\bigvee K_1 \wedge \bigvee K_2$, und M ist erfüllbar genau dann, wenn auch $M \cup \{K_0\}$ erfüllbar ist.*

Beweis. Selbstverständlich ist M stets erfüllbar, wenn $M \cup \{K\}$ erfüllbar ist. Wir zeigen, daß $\bigvee K_0$ aus $\bigvee K_1$ und $\bigvee K_2$ folgt. Damit folgt insbesondere, daß $M \cup \{K_0\}$ erfüllbar ist, falls M erfüllbar ist.

Es sei g eine 0-1 Bewertung, die $\bigvee K_1$ und $\bigvee K_2$ auf 1 abbildet. Es sei etwa $K_1 = \{L, L_1, \ldots, L_m\}$, $K_2 = \{\bar{L}, L'_1, \ldots, L'_n\}$ und $K_2 = \{L_1, \ldots, L_m, L'_1, \ldots, L'_n\}$. Wir unterscheiden 2 Fälle.

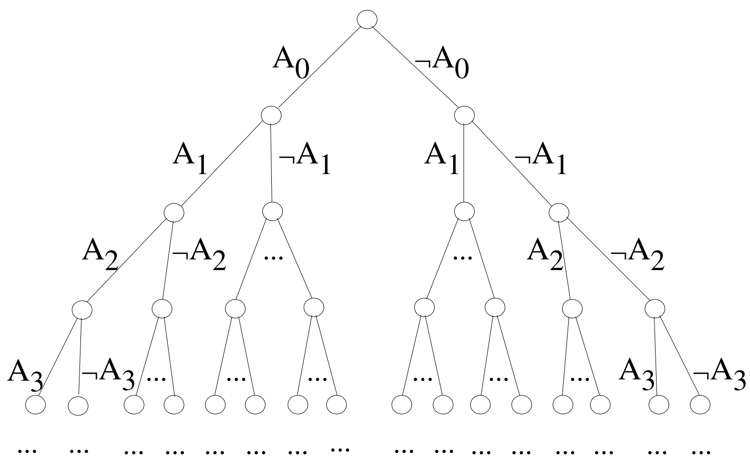
Fall 1: $g(L) = 1$. In diesem Fall folgt $g(\bar{L}) = 0$ und damit $g(L'_1 \vee \ldots \vee L'_n) = 1$. Also gilt $g(\bigvee K_0) = 1$.

Fall 2: $g(L) = 0$. In diesem Fall folgt $g(L_1 \vee \ldots \vee L_m) = 1$ und $g(\bigvee K_0) = 1$. ■

Lemma 11.50 (Korrektheit des Resolutions-Kalküls) *Wenn bei Eingabe von M im Resolutions-Kalkül die leere Klausel herleitbar ist, so ist M unerfüllbar.*

Beweis. Es sei bei Eingabe von M im Resolutions-Kalkül die leere Klausel „$\square$“ herleitbar. Da die leere Klausel unerfüllbar ist, ist dann die nach der Herleitung der leeren Klausel angehäufte Klauselmenge unerfüllbar. Aus Lemma 11.49 folgt mittels Induktion sofort, daß dann auch die Eingabemenge M unerfüllbar ist. ■

Um den Beweis der Vollständigkeit des Resolutions-Kalküls vorzubereiten, betrachten wir den folgenden Baum $\mathcal{T}_{\text{sem}}$, den wir den *semantischen Baum* nennen wollen.



Man beachte, daß die Äste von $\mathcal{T}_{\text{sem}}$ gerade die möglichen 0-1 Bewertungsfunktionen repräsentieren. Jedem Knoten von $\mathcal{T}_{\text{sem}}$ können wir die Menge all derjenigen 0-1 Bewertungen zuordnen, die auf den Pfaden durch diesen Knoten realisiert werden.

Im nachfolgenden bezeichnet M stets eine Klauselmenge.

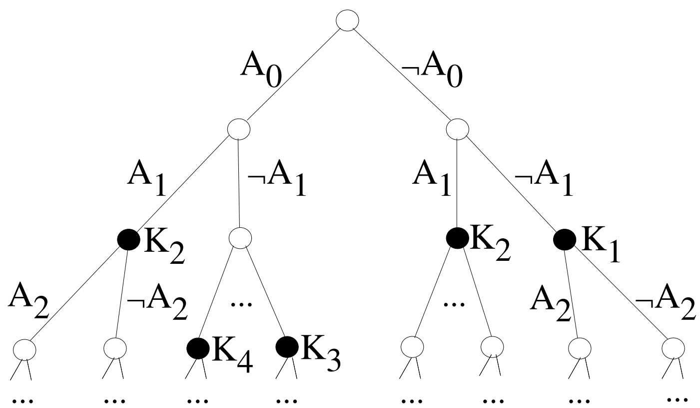
Definition 11.51 Ein Knoten η von $\mathcal{T}_{\text{sem}}$ heißt *Fehlerknoten für M* unter den folgenden Bedingungen.

1. Es gibt eine Klausel $K \in M$ so, daß jede dem Knoten η zugeordnete 0-1 Bewertungsfunktion die Formel $\bigvee K$ auf 0 abbildet,

2. η ist ein minimaler Knoten mit dieser Eigenschaft, das heißt kein Vorgänger von η hat die Eigenschaft 1.

Wir sagen, daß die Klausel K bei η *widerlegt* wird.

Beispiel 11.52 Es sei M wie in Beispiel 11.48 die Menge, die die Klauseln $K_1 := \{A_0, A_1\}$, $K_2 := \{\neg A_1\}$, $K_3 := \{A_2, \neg A_0\}$ und $K_4 := \{\neg A_2\}$ enthält. In der folgendenden Figur sind die Fehlerknoten für M schwarz markiert; die Labels geben die Klauseln wieder, die an dem betreffenden Knoten widerlegt werden.



Definition 11.53 Das *Gewicht von $\mathcal{T}_{sem}$ bezüglich der Klauselmenge M* ist die Anzahl der Knoten von $\mathcal{T}_{\text{sem}}$, die selbst keine Fehlerknoten sind, und die auch nicht unterhalb eines Fehlerknotens für M liegen.

Beispiel 11.54 Es sei M wie in Beispiel 11.52. Dann hat $\mathcal{T}_{\text{sem}}$ bezüglich M das Gewicht 4.

Lemma 11.55 (Königs Lemma) *Ein endlich verzweigender Baum, der keinen unendlichen Pfad hat, ist endlich.*

Beweis. Sei $\mathcal{T}$ ein endlich verzweigender Baum. Angenommen, $\mathcal{T}$ sei unendlich. Wir zeigen, daß in diesem Fall $\mathcal{T}$ zumindest einen unendlichen Pfad haben muß. Diesen Pfad definieren wir induktiv. Als ersten Knoten η_0 des zu konstruierenden Pfades nehmen wir die Wurzel. Nun gibt es unter den endlich vielen direkten Nachfolgerknoten von η_0 einen Knoten η_1, so daß der bei η_1 beginnende Teilbaum unendlich ist. Angenommen, wir haben schon den Knoten η_n der Tiefe n gefunden, so daß der bei η_n beginnende Teilbaum unendlich ist. Dann muß es wiederum unter den endlich vielen direkten

Nachfolgerknoten von η_n einen Knoten η_{n+1}, so daß der bei η_{n+1} beginnende Teilbaum unendlich ist. Durch Iteration erhalten wir offenkundig einen unendlichen Pfad von $\mathcal{T}$. ■

Es sei darauf hingewiesen, daß der vorausgegangene Beweis nicht konstruktiv ist.

Lemma 11.56 *Es sei M eine Klauselmenge. Dann ist das Gewicht von $\mathcal{T}_{sem}$ bezüglich M endlich genau dann, wenn M unerfüllbar ist. Das Gewicht von $\mathcal{T}_{sem}$ bezüglich M ist* 0 *genau dann, wenn M die leere Klausel enthält.*

Beweis. Es sei das Gewicht von $\mathcal{T}_{\text{sem}}$ bezüglich M endlich. Wenn wir eine 0-1 Bewertungsfunktion g als Pfad in $\mathcal{T}_{\text{sem}}$ auffassen, so geht dieser durch einen Fehlerknoten, da ansonsten alle Punkte des unendlichen Pfades von g zum Gewicht einen Beitrag 1 leisten würden. Damit erfüllt aber g die betreffende Klausel von M nicht, die an diesem Fehlerknoten widerlegt wird. Somit erfüllt g auch M nicht.

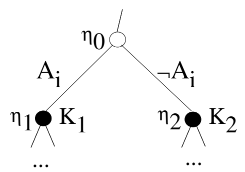
Es sei M unerfüllbar. Da ein Pfad ohne Fehlerknoten eine 0-1 Bewertungsfunktion repräsentieren würde, die jede Klausel aus M und damit ganz M erfüllt, so liegt auf jedem Pfad von $\mathcal{T}_{\text{sem}}$ ein Fehlerknoten. Nach Königs Lemma ist dann derjenige Teil von $\mathcal{T}_{\text{sem}}$, der aus den Knoten oberhalb der Fehlerknoten besteht, endlich. Damit hat $\mathcal{T}_{\text{sem}}$ bezüglich M endliches Gewicht.

Hat das Gewicht von $\mathcal{T}_{\text{sem}}$ bezüglich M den Wert 0, so ist die Wurzel von $\mathcal{T}_{\text{sem}}$ ein Fehlerknoten bezüglich M. Daher muß M eine Klausel enthalten, die von keiner Belegung erfüllt wird. Dies kann nur die leere Klausel sein. Umgekehrt ist die Wurzel von $\mathcal{T}_{\text{sem}}$ bezüglich M ein Fehlerknoten von M, wenn M die leere Klausel enthält. ■

Wir können nun die Vollständigkeit des aussagenlogischen Resolutions-Kalküls beweisen.

Theorem 11.57 (Vollständigkeit des Resolutions-Kalküls) *Bei Eingabe einer unerfüllbaren Klauselmenge M ist im Resolutions-Kalkül stets die leere Klausel herleitbar.*

Beweis. Es sei M eine unerfüllbare Klauselmenge. Dann hat der semantische Baum $\mathcal{T}_{\text{sem}}$ bezüglich M nach Lemma 11.56 endliches Gewicht n. Wir verwenden nun Induktion nach n. Falls $n = 0$ gilt, so enthält M nach Lemma 11.56 die leere Klausel und wir sind fertig. Andernfalls muß es einen Knoten η_0 in $\mathcal{T}_{\text{sem}}$ geben, so daß beide unmittelbaren Nachfolgerknoten η_1 und η_2 Fehlerknoten bezüglich M sind. In der Tat könnte man ansonsten sofort einen unendlichen Pfad finden, der keinen Fehlerknoten bezüglich M enthält. Wir betrachten nun gemäß der nachfolgenden Abbildung die beiden Klauseln K_1 und K_2 aus M, die bei η_1 und η_2 widerlegt werden. Die beiden Wege, die von η_0 zu η_1 und η_2 führen, seien hierbei mit den komplementären Literalen A_i und $\neg A_i$ gelabelt.

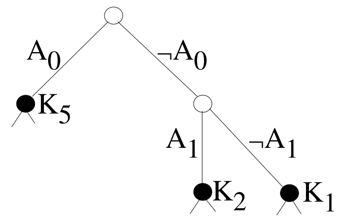


Nun muß aber K_1 das Literal $\neg A_i$ enthalten, andernfalls wäre nämlich K_1 bereits bei η_0 (oder oberhalb) widerlegt, was der Definition des Fehlerknotens widersprechen würde. Außerdem muß die (nicht zu M gehörende) Klausel $K_1 \setminus \{\neg A_i\}$ bereits bei η_0 oder oberhalb widerlegt sein. Analog muß K_2 das Literal A_i enthalten, und $K_2 \setminus \{A_i\}$ muß bereits bei η_0 oder oberhalb widerlegt sein. Wenn wir nun die Resolvente

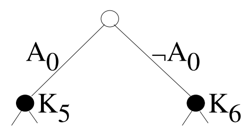
$$K_0 =:= (K_1 \setminus \{\neg A_i\}) \cup (K_2 \setminus \{A_i\})$$

aus K_1 und K_2 bilden, so ist die neue Klausel K_0 bereits bei η_0 oder oberhalb widerlegt. Damit hat $\mathcal{T}_{\text{sem}}$ bezüglich $M \cup \{K_0\}$ kleineres Gewicht als n. Nach der Induktionsvoraussetzung können wir damit aus $M \cup \{K_0\}$ die leere Klausel resolvieren. Damit ist die leere Klausel auch aus M selbst herleitbar.
■

Wir können den im Beweis geschilderten Prozeß der Gewichtsverringerung durch Resolventenbildung anhand von Beispiel 11.54 verfolgen. Wir betrachten wieder die Klauseln $K_1 := \{A_0, A_1\}$, $K_2 := \{\neg A_1\}$, $K_3 := \{A_2, \neg A_0\}$ und $K_4 := \{\neg A_2\}$. Die Klauseln K_4 und K_3 enthalten die komplementären Literale $\neg A_2$ und A_2, wie man bereits der Abbildung in Beispiel 11.54 entnehmen kann. Durch Resolventenbildung erhalten wir die neue Klausel $K_5 := \{\neg A_0\}$. Der semantische Baum hat nunmehr nur noch Gewicht 2 bezüglich der vergrößerten Klauselmenge:



Eine zweite Resolventenbildung mit K_2 und K_1 liefert $K_6 := \{A_0\}$, wodurch der semantische Baum Gewicht 1 bekommt:



Schließlich liefern K_5 und K_6 die leere Klausel K_7 als Resolvente, und der semantische Baum erhält Gewicht 0:

Bemerkung 11.58 Unsere Darstellung, die sich an der üblichen Darstellung des Resolutions-Kalküls orientierte, folgte dem Bild des Resolutionskalküls als Widerlegungskalkül. Die Herleitung der leeren Klausel interpretierten wir als Beweis der Unerfüllbarkeit der Eingabeklauselmenge. Wir hatten bereits der Darstellung des Gentzen-Kalküls gesehen, daß ein und dasselbe Beweissystem unter Umständen wahlweise als Widerlegungskalkül oder als affirmativer Kalkül aufgefaßt werden kann, je nach der Art und Weise, wie man die Ausdrücke des Kalküls als Formeln deutet. Genau diese Art von Dualismus findet man auch beim Resolutions-Kalkül.

Anstatt Klauseln als Disjunktionen zu lesen und eine Klauselmenge als Konjunktion der zu den Klauseln gehörenden Disjunktionen, kann man auch dual Klauseln als Konjunktionen lesen. Jede Klauselmenge wird dann interpretiert als Disjunktion über die zu den Klauseln gehörenden Konjunktionen.

Das Hinzufügen einer neuen Resolvente, das sich der Form nach in keiner Weise verändert, erhält nunmehr gerade die *Gültigkeit* der Klauselmenge beziehungsweise der zugeordneten Formel. Dies kann man leicht wie folgt einsehen. Offenbar entsteht aus einer gültigen Formel – also einer Tautologie — wieder eine Tautologie, wenn sie disjunktiv mit einer neuen Formel verknüpft wird. Also erhält das Hinzufügen einer neuen Klausel trivialerweise die Gültigkeit der zugeordneten Formel. Es sei umgekehrt die nach Hinzufügen einer Resolvente entstandene Klauselmenge in der neuen Lesart eine gültige Formel. Jede 0-1 Bewertungsfunktion, die die Resolvente — beziehungsweise die zugeordnete Konjunktion — auf 1 abbildet, bildet zumindest eine der Elternklauseln auf 1 ab. Hieraus folgt aber sofort, daß auch die Ausgangsklauselmenge gültig ist.

Die leere Klausel, als triviale Konjunktion, ist nunmehr tautologisch wahr. Nach Herleitung der leeren Klausel haben wir eine Disjunktion vorliegen, die eine Tautologie als Disjunktionsglied enthält, die also selbst Tautologie ist. Daher zeigt die Herleitbarkeit der leeren Klausel in diesem Bild gerade die Gültigkeit der Eingabeklauselmenge an.

11.6 Aufgaben zu Kapitel 11

Aufgaben zu Teilkapitel 11.1

Aufgabe 11.1 Es sei Φ eine Formelmenge, σ und γ seien Formeln. Wie immer sei $\sigma \Leftrightarrow \gamma$ eine Abkürzung für die Formel $(\sigma \Rightarrow \gamma) \wedge (\gamma \Rightarrow \sigma)$. Zeigen Sie, daß im Hilbert-Kalkül aus $\Phi \vdash \sigma \Rightarrow \gamma$ und $\Phi \vdash \gamma \Rightarrow \sigma$ stets $\Phi \vdash \sigma \Leftrightarrow \gamma$ folgt und umgekehrt.

Aufgabe 11.2 Welches Axiom müßte man zum Hilbert-Kalkül hinzufügen, um — unter Erhaltung der Korrektheit des Kalküls — möglichst bequem aus Formeln α und β die Konjunktion $\alpha \wedge \beta$ herleiten zu können? Wie sähe die Herleitung aus?

Aufgabe 11.3 Es seien α und β Formeln. Geben Sie eine Herleitung von $\neg(\alpha \vee \beta) \Rightarrow (\neg\alpha \wedge \neg\beta)$ im Hilbert-Kalkül an.

Aufgabe 11.4 Erledigen Sie den Subfall 2.3 im Beweis von Proposition 11.5.

Aufgaben zu Teilkapitel 11.2

Aufgabe 11.5 Es sei $\Gamma \to \Delta$ eine Sequenz, wo alle Formeln aus Γ respektive Δ Atomformeln sind. Zeigen Sie, daß $\Gamma \to \Delta$ gültig ist genau dann, wenn die Folgen Γ und Δ ein gemeinsames Element haben.

Aufgabe 11.6 Zeigen Sie: Für jede Regel des Kalküls G' der Form

$$\frac{\Gamma \to \Delta}{\Gamma' \to \Delta'}$$

sind die assoziierten Formeln $\bigvee \bar{\Gamma} \vee \bigvee \Delta$ und $\bigvee \bar{\Gamma}' \vee \bigvee \Delta'$ aussagenlogisch äquivalent.

Aufgabe 11.7 Beweisen Sie die Gültigkeit der Sequenz

$$\neg(A_0 \vee A_1) \longrightarrow \neg A_0 \wedge \neg A_1,$$

indem sie einen Beweisbaum angeben, dessen Wurzel mit der Sequenz gelabelt ist.

Aufgabe 11.8 Beweisen Sie die Gültigkeit der Sequenz

$$A_0 \vee (A_1 \wedge A_2) \longrightarrow (A_0 \vee A_1) \wedge (A_0 \vee A_2),$$

indem sie einen Beweisbaum angeben, dessen Wurzel mit der Sequenz gelabelt ist.

Aufgabe 11.9 Ein *Literal* ist eine Atomformel oder eine negierte Atomformel. Eine aussagenlogische Formel γ heißt *in konjunktiver Normalform* falls γ eine Konjunktion von Formeln ist, die ihrerseits Disjunktionen von Literalen sind. Zeigen Sie, daß es zu jeder aussagenlogischen Formel α eine logisch äquivalente Formel in konjunktiver Normalform gibt. Anleitung: geben Sie zunächst eine Sequenz an, deren assoziierte Formel zu α äquivalent ist. Konstruieren Sie einen abgeschlossenen Ableitungsbaum für die Sequenz. An diesem Baum kann man eine zu α logisch äquivalente Formel in konjunktiver Normalform ablesen — wie?

Aufgaben zu Teilkapitel 11.3

Aufgabe 11.10 Die nachfolgende Aufgabe kann als eine duale Version von Aufgabe 11.9 betrachtet werden. Eine aussagenlogische Formel γ heißt *in disjunktiver Normalform* falls γ eine Disjunktion von Formeln ist, die ihrerseits Konjunktionen von Literalen sind. Zeigen Sie mit Hilfe des Bildes von G' als Widerlegungs-Kalkül, daß es zu jeder aussagenlogischen Formel α eine logisch äquivalente Formel in disjunktiver Normalform gibt. Tip: Übersetzen Sie die Methode aus Aufgabe 11.9.

Aufgaben zu Teilkapitel 11.4

Aufgabe 11.11 Geben Sie ein geschlossenes Tableau für

$$\{T(\neg(A_0 \vee A_1)), F(\neg A_0 \wedge \neg A_1))\}$$

an, indem Sie den in Aufgabe 11.7 konstruierten Beweisbaum im Gentzen-Kalkül übersetzen.

Aufgabe 11.12 Geben Sie ein geschlossenes Tableau für

$$\{T(A_0 \vee (A_1 \wedge A_2)), F((A_0 \vee A_1) \wedge (A_0 \vee A_2))\}$$

an, indem Sie den in Aufgabe 11.8 konstruierten Beweisbaum im Gentzen-Kalkül übersetzen.

Aufgabe 11.13 Axiom (1) des Hilbert-Kalküls lautet $\alpha \rightarrow \alpha$. Zeigen Sie die Gültigkeit dieses Axioms, indem Sie ein geschlossenes Tableau für

$$\{F(\alpha \rightarrow \alpha)\}$$

konstruieren. Beweisen Sie auf dieselbe Weise die Gültigkeit aller Axiome des Hilbert-Kalküls.

Aufgabe 11.14 Gegeben sei die Formel α der Form

$$(A_0 \rightarrow (A_1 \vee (A_2 \Rightarrow A_0))) \wedge ((A_0 \vee A_2) \Rightarrow (A_1 \vee (A_2 \Rightarrow A_1))).$$

Geben Sie alle Wahrheitswertbelegungen an, die die Formel auf 0 abbilden, indem sie ein vollständig entwickeltes Tableau für $\{\alpha\}$ konstruieren, und dort die offenen Äste betrachten.

Aufgaben zu Teilkapitel 11.5

Aufgabe 11.15 Unter welchen Umständen kann man eine Resolvente zwischen ein und derselben Klausel K bilden? Wie läßt sich die Resolvente beschreiben?

Aufgabe 11.16 Geben Sie eine Abschätzung an, wieviel Klauseln im Resolutions-Kalkül bei Eingabe einer Klauselmenge M, in der n verschiedene Atomformeln auftreten, höchstens erzeugt werden können.

Aufgabe 11.17 Zeigen Sie mit Hilfe des Resolutions-Kalküls die Unerfüllbarkeit der Menge M mit den Klauseln $\{A_0, A_2, \neg A_1\}$, $\{A_0, A_1\}$, $\{A_3, \neg A_0\}$, $\{\neg A_3, \neg A_0\}$, $\{\neg A_2\}$.

Aufgabe 11.18 Zeichnen Sie im semantischen Baum $\mathcal{T}_{\text{sem}}$ die Fehlerknoten bezüglich der Menge M mit den Klauseln $\{A_0, A_2, \neg A_1\}$, $\{A_0, A_1\}$, $\{A_3, \neg A_0\}$, $\{\neg A_3, \neg A_0\}$, $\{\neg A_2\}$ ein. Zeigen Sie, wie Resolutionsschritte zur Gewichtsverringerung beitragen, bis schließlich die leere Klausel hergeleitet ist.

Aufgabe 11.19 Welche Klauseln sind durch iterierte Resolventenbildung aus der Klauselmenge $\{\{A_0, A_1, A_2\}, \{\neg A_0, \neg A_1, \neg A_2\}\}$ herleitbar?

Aufgabe 11.20 Gegeben sei das aussagenlogische (bzw. variablenfreie) Prolog-Programm mit den Klauseln

$$\begin{array}{lcl} A & :- & B, C, E. \\ B & :- & C, D. \\ E & :- & B, C. \\ C. & & \\ D. & & \end{array}$$

und das Goal A. Betrachten Sie die dazugehörige Klauselmenge $\{\{\neg A, B, C, E\}, \{\neg B, C, D\}, \{\neg E, B, C\}, \{\neg C\}, \{\neg D\}, \{A\}\}$. Geben Sie eine Herleitung der leeren Klausel im Resolutionskalkül an, die die Prolog-Abarbeitung imitiert.

Aufgabe 11.21 Geben Sie eine vereinfachte Variante des Beweises von Theorem 11.57 für den Fall, daß die Eingabe-Klauselmenge M endlich ist. Es soll die Verwendung von Königs Lemma eliminiert werden.

11.7 Bibliographische Angaben

To be added.

12

Prädikatenlogik 1. Stufe

Im Rahmen der Aussagenlogik hatten wir eine sehr primitive Formalisierung des Umgangs mit mathematischen Aussagen erreicht. Modellieren konnten wir lediglich die Verknüpfung von Aussagen mit Hilfe von Junktoren. Es war aber nicht möglich, auf die innere Form der atomaren Aussagen selbst Bezug zu nehmen, da wir diese als unstrukturierte Variablen behandelt hatten. Insbesondere blieb die Rolle der Quantoren beim in Kapitel 1 kurz angedeuteten Übergang von Aussageformen zu Aussagen völlig ungeklärt. Damit blieb jedoch für viele offenkundig richtige Schlußfolgerungen eine formale Analyse unmöglich. Wir können dies mit einem oft zitierten Beispiel verdeutlichen: aus den beiden Aussagen „Alle Menschen sind sterblich" und „Sokrates ist ein Mensch" folgt die Aussage „Sokrates ist sterblich". Dies ist aber keine aussagenlogische Folgerung.

In der *Prädikatenlogik* bemüht man sich um eine weitergehende Präzisierung des Begriffs der „mathematischen Aussage", der „Folgerungsbeziehung" und des „Beweises". Charakteristisch ist die Verwendung der Quantoren „$\forall$" (lies: für alle) und „$\exists$" (lies: es existiert ein) als Zeichen der formalen Sprache. Allerdings ist die im nachfolgenden betrachtete *Prädikatenlogik erster Stufe* beim Gebrauch dieser Quantoren deutlichen Einschränkungen unterworfen. Betrachtet werden lediglich Aussagen, wo Quantifizierungen mittels „$\exists$" und „$\forall$" sich nur auf Elemente beziehungsweise Individuen beziehen.[1] Unter der so gegebenen Einschränkung werden wir nachfolgend ei-

[1] Im Gegensatz hierzu ist in der *Logik höherer Stufe*, auf die wir hier nicht eingehen, auch eine Quantifizierung über Teilmengen und komplexere Objekte möglich.

ne partielle Formalisierung der oben genannten Begriffe erreichen. Zunächst müssen die sprachlichen Ausdrucksmittel genau beschrieben werden. Wie immer ist es ein wichtiges Ziel der einzuführenden formalen Sprache, daß alle Ambiguitäten vermieden werden.

12.1 Syntax der Prädikatenlogik

Die Prädikatenlogik erster Stufe basiert auf einer Familie von Sprachen. Jede dieser Sprachen wird verwendet zur logischen Beschreibung von Strukturen einer festen Signatur Σ. Demzufolge ist die Signatur gerade der Parameter, der die konkrete Sprache festlegt. Nachfolgend sei Σ eine fest gewählte Signatur, die das Gleichheitssymbol „$\equiv$" nicht enthalten soll. Wie in Kapitel 6 bezeichnen wir mit $\Sigma_{\mathcal{F}}$ die Menge der Funktionssymbole aus Σ, mit $\Sigma_{\mathcal{R}}$ die Menge der Relationssymbole aus Σ und mit $\Sigma_{\mathcal{E}}$ die Menge der Individuenkonstanten aus Σ.

Die Beschreibung der Sprache $\mathcal{L}_\Sigma$ der Prädikatenlogik erster Stufe zur Signatur Σ vollzieht sich in drei Schritten. Wir werden zunächst den Zeichenvorrat angeben, danach die syntaktischen Ausdrücke, mittels derer wir auf Individien einer Struktur referieren, danach die Formeln der Sprache definieren, die als Aussagen und Aussageformen aufzufassen sind.

Zeichenvorrat

Zunächst hier also das *Alphabet* der Sprache $\mathcal{L}_\Sigma$ der Prädikatenlogik erster Stufe zur Signatur Σ.

Definition 12.1 Das Alphabet der Sprache $\mathcal{L}_\Sigma$ umfaßt die folgenden Symbole:

(1) eine abzählbar unendliche Menge von (Individuen-) *Variablen* $X = \{v_i \mid i \in \mathbb{N}\}$,

(2) aussagenlogischen Junktoren $\neg, \wedge, \vee, \Rightarrow$,

(3) die *Quantoren* $\forall$ (Allquantor; lies: „für alle") und $\exists$ (Existenzquantor; lies: „es existiert ein"),

(4) die Klammern "(" und ")" sowie das Komma "," als Hilfszeichen,

(5) die Symbole der Signatur Σ.

Das Gleichheitssymbol „$\equiv$" kann zusätzlich in das Alphabet mitaufgenommen werden, man spricht dann gegebenenfalls von *Logik mit Gleichheit.* Die unter (1)-(4) erwähnten Symbole sind also in jedem logischen Alphabet vorhanden. Sie, oder zumindest die Symbole in (2) und (3), werden daher auch als „logische Symbole" bezeichnet. Der Unterschied zwischen verschiedenen „Dialekten" beziehungsweisen Alphabeten beruht also auf der Auswahl der „nicht-logischen" Symbole aus der Signatur Σ. Weiterhin ist zwischen Logik mit Gleichheit und Logik ohne Gleichheit zu unterscheiden.

Im Zusammenhang mit logischen Alphabeten verwenden wir folgende Mitteilungszeichen:

- Symbole f, g, h, gegebenenfalls mit Indizes, für Funktionssymbole,
- Symbole R, S, T, gegebenenfalls mit Indizes, für Relationssymbole,
- Symbole c, d, e, gegebenenfalls mit Indizes, für Individuenkonstanten.

Diese „anonymen" Symbole werden genauer eingesetzt, um über nicht näher bestimmte Bereiche zu reden. In spezielleren Kontexten werden assoziationsträchtigere Funktions- und Relationssymbole sowie Individuenkonstanten verwendet. Beispiele folgen unten. Schließlich verwenden wir Symbole x, y, z, u, v, w, gegebenenfalls mit Indizes, für Individuenvariablen.

Terme

Im zweiten Schritt führt man diejenigen Ausdrücke der Sprache $\mathcal{L}_\Sigma$ ein, mit denen auf *Individuen* oder *Elemente* eines Bereichs referiert wird. Dies führt auf die bereits aus Abschnitt 12.1 bekannte Definition des Terms, die wir hier bequemlichkeitshalber wiederholen. Anstelle von Termen über Σ und X werden wie nachfolgend häufig von $\mathcal{L}_\Sigma$-Termen reden.

Definition 12.2 Die Menge der $\mathcal{L}_\Sigma$-*Terme* 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 $\mathcal{L}_{\Sigma}$-Term.
2. Sind $t_1, \ldots, t_n$ jeweils $\mathcal{L}_{\Sigma}$-Terme, und ist $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol, so ist $f(t_1, \ldots, t_n)$ ein $\mathcal{L}_{\Sigma}$-Term.

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

Man beachte, daß die hier betrachteten *Individuen*variablen eine andere Rolle spielen als die in der Aussagenlogik verwendeten *Aussagen*variablen. Mit $Var(t)$ bezeichnen wir die Menge der im $\mathcal{L}_{\Sigma}$-Term t vorkommenden Variablen, wir ersparen uns die offenkundige Definition.

Wir werden in der nachfolgenden Diskussion einige Begriffe (Teilterme, Tiefe eines Terms, Baumdarstellung, Grundterm) und Ergebnisse (Eindeutige Lesbarkeit von Termen) aus Abschnitt 12.1 verwenden. Der Leser möge sich den Inhalt gegebenenfalls nochmals vergegenwärtigen.

Beispiel 12.3 Es sei $\Sigma_{\mathcal{F}}$ die Teilsignatur mit den zweistelligen Funktionssymbolen „+“ und „·“ und $\Sigma_{\mathcal{E}} := \{0, 1\}$, es enthalte $\Sigma_{\mathcal{R}}$ das zweistellige Relationssymbol „≤“. Die Signatur $\Sigma := \Sigma_{\mathcal{F}} \cup \Sigma_{\mathcal{R}} \cup \Sigma_{\mathcal{E}}$ werden wir verwenden, um über die gewohnten Zahlbereiche zu reden, wobei wir die üblichen Schreibkonventionen verwenden. Es sind dann x, $x + x$, $x \cdot (3 + y \cdot z)$ sowie $3 \cdot x^2 + 3 \cdot x \cdot y + 3 \cdot z$ Beispiele für Terme. Hierbei ist x^2 (resp. 3) eine Abkürzung für $x \cdot x$ (resp. $1 + 1 + 1$). Die Terme dieser Signatur erreichen beliebige Tiefe.

Beispiel 12.4 Es sei $\Sigma_{\mathcal{R}} := \{kennt, sieht, liebt, Freund, Feind, Frau, Mann\}$ und $\Sigma_{\mathcal{E}} = \{max, maria\}$. Mit der Signatur $\Sigma := \Sigma_{\mathcal{R}} \cup \Sigma_{\mathcal{E}}$ werden wir über Strukturen reden, die geeignete Äußerungskontexte für natürlichsprachliche Sätze formalisieren. Neben den Individuenvariablen sind hier die Individuenkonstanten „*max*“ und „*maria*“ die einzigen Terme.

Beispiel 12.5 Um über Schachkonfigurationen zu reden, führen wir eine Signatur Σ mit den Individuenkonstanten „*weißer_König*“, „*weiße_Dame*“, „*schwarzer_König*“, „*schwarze_Dame*“ ohne Funktionssymbole ein. Die einzigen Terme sind dann die Individuenvariablen und die vier Individuenkonstanten.

Beispiel 12.6 Um über Konkatenation von Wörtern zu reden, verwenden wir die Signatur Σ, die als einziges Symbol das zweistellige Funktionssymbol „*concat*" enthält. Es sind x, $concat(x, y)$, $concat(concat(x, y), z)$ und $concat(concat(x, x), concat(y, y))$ Beispiele für Terme. Es gibt neben den Variablen unendlich viele Terme, die beliebige Tiefe annehmen können.

Formeln

Im dritten Schritt können wir nun diejenigen Ausdrücke der Sprache $\mathcal{L}_\Sigma$ einführen, die es erlauben, Eigenschaften von Elementen und Beziehungen zwischen Elementen auszudrücken.

Definition 12.7 Die Menge der *$\mathcal{L}_\Sigma$-Formeln* ist die kleinste Menge von Symbolfolgen, die unter den folgenden Bildungsregeln abgeschlossen ist:

1. sind t_1 und t_2 $\mathcal{L}_\Sigma$-Terme, so ist $(t_1 \equiv t_2)$ eine *atomare* $\mathcal{L}_\Sigma$-Formel,
2. ist $R \in \Sigma_\mathcal{R}$ ein k-stelliges Relationssymbol und sind $t_1, \ldots, t_k$ $\mathcal{L}_\Sigma$-Terme, so ist $R(t_1, \ldots, t_k)$ eine *atomare* $\mathcal{L}_\Sigma$-Formel,
3. alle aussagenlogischen Verknüpfungen von $\mathcal{L}_\Sigma$-Formeln mit den Junktoren $\neg, \wedge, \vee, \Rightarrow$ sind $\mathcal{L}_\Sigma$-Formeln,
4. ist φ eine $\mathcal{L}_\Sigma$-Formel und $x \in X$, so ist $(\forall x \varphi)$ eine $\mathcal{L}_\Sigma$-Formel,
5. ist φ eine $\mathcal{L}_\Sigma$-Formel und $x \in X$, so ist $(\exists x \varphi)$ eine $\mathcal{L}_\Sigma$-Formel,

Für Logik ohne Gleichheit entfallen die Atomformeln vom Typ 1 natürlich. Mit $\mathcal{L}_\Sigma$ bezeichnen wir die Menge der $\mathcal{L}_\Sigma$-Formeln. Für diese verwenden wir Symbole wie φ, ψ, γ.

Die Begriffe der *Teilformel* und des *Vorkommens einer Teilformel* sind in der offenkundigen Weise definiert. Enthält φ ein Vorkommen einer Teilformel $(\forall x \psi)$ oder $(\exists x \psi)$, so ist (das entsprechende Vorkommen von) ψ der *Skopus* der betrachteten Bindung $\forall x$ respektive $\exists x$. Es ist dann x eine gebundene Variable der Formel φ.

Auch für $\mathcal{L}_\Sigma$-Formeln gilt das *Prinzip der eindeutigen Lesbarkeit*, wir verzichten auf die explizite Formulierung und auf den Beweis. Wir werden im nachfolgenden in Formeln Klammern weglassen, wenn die eindeutige Lesbarkeit erhalten bleibt.

Beispiel 12.8 Es sei Σ die Signatur aus Beispiel 12.3. In der Logik mit Gleichheit sind $\forall x\ (x \equiv 0)$, $0 + 0 \leq x \cdot (y + z)$ und $\exists x\ (x + 0 \equiv x)$ Formeln der betrachteten Sprache.

Beispiel 12.9 Es sei Σ die Signatur aus Beispiel 12.4. Die nachfolgenden Beispiele von $\mathcal{L}_\Sigma$-Formeln lassen sich als natürlichsprachliche Aussagen und Aussageformen wiedergeben:

- $\mathit{Mann}(z)$ (bzw. z ist ein Mann),
- $\mathit{liebt}(x, \mathit{maria})$ (bzw. x liebt Maria),
- $\mathit{kennt}(\mathit{max}, \mathit{maria})$ (bzw. Max kennt Maria),
- $\forall x(\mathit{Mann}(x) \vee \mathit{Frau}(x))$ (bzw. jedes Individuum ist ein Mann oder eine Frau),
- $\forall x \forall y((\mathit{Freund}(x, \mathit{maria}) \wedge \mathit{Feind}(y, x)) \Rightarrow \mathit{Feind}(y, \mathit{maria}))$ (bzw. die Feinde von Marias Freunden sind Marias Feinde) sowie
- $\forall x(\mathit{Mann}(x) \Rightarrow \exists y(\mathit{kennt}(\mathit{maria}, y) \wedge \mathit{Frau}(y) \wedge \mathit{liebt}(x, y)))$ (bzw. jeder Mann liebt irgendeine Bekannte von Maria).

Es sei darauf hingewiesen, daß sich bei der umgekehrten Übersetzung von natürlichsprachlichen Sätzen in Prädikatenlogik verschiedene Probleme ergeben können. So sind einerseits viele Sätze selbst bei passender Wahl der Signatur gar nicht in Prädikatenlogik erster Stufe ausdrückbar. Andererseits ergeben sich aufgrund der Ambiguität der natürlichen Sprache teilweise mehrere Übersetzungen.

Beispiel 12.10 Dem Alphabet aus Beispiel 12.5 mit den Individuenkonstanten *weißer_König*, *weiße_Dame*, *schwarzer_König*, *schwarze_Dame* fügen wir die einstelligen Relationszeichen *schwarz*, *weiß*, *Läufer*, *Springer*, *Turm*, *Bauer* und die zweistelligen Relationssymbol *deckt* und *bedroht* hinzu. Dann sind

- $\forall x \forall y \forall z ((\mathit{Springer}(x) \wedge \mathit{Bauer}(y)) \Rightarrow \mathit{deckt}(x, y))$,
- $\neg \mathit{bedroht}(\mathit{weiße\_Dame}, \mathit{schwarzer\_König})$ sowie
- $\exists x \ (\mathit{weiß}(x) \wedge \forall y \ (\mathit{schwarz}(y) \Rightarrow \mathit{bedroht}(y, x)))$

Formeln der betrachteten Sprache.

Beispiel 12.11 Wir betrachten Logik mit Gleichheit. Das Alphabet enthalte das zweistellige Funktionssymbol f. Dann ist

$$\forall x \forall y \forall z \ (f(x, f(y, z)) \equiv f(f(x, y), z)))$$

eine Formel.

Definition 12.12 Die *freien Variablen-Vorkommen in einer $\mathcal{L}_\Sigma$-Formel φ* sind wie folgt erklärt.

- die freien Variablen-Vorkommen in einer atomaren Formel φ sind die Vorkommen von Variablen in den Termen der atomaren Formel,
- die freien Variablen-Vorkommen in einer booleschen Verknüpfung der Formeln $\varphi_1, \ldots, \varphi_k$ sind die freien Variablen-Vorkommen in den Teilformeln $\varphi_1, \ldots, \varphi_k$,
- die freien Variablen-Vorkommen in einer quantifizierten Formel $\forall x \varphi$ oder $\exists x \varphi$ sind die freien Variablen-Vorkommen in φ der Variablen $y \neq x$.

Die *Menge der in der Formel φ frei vorkommenden Variablen* bezeichnen wir mit $\mathit{fr}(\varphi)$. Die Schreibweise $\varphi(x_1, \ldots, x_n)$ deutet an, daß $\mathit{fr}(\varphi) \subseteq \{x_1, \ldots, x_n\}$.

Definition 12.13 Eine $\mathcal{L}_\Sigma$-Formel φ heißt *geschlossen* genau dann, falls $\mathit{fr}(\varphi) = \emptyset$ gilt, andernfalls heißt φ *offen*. Geschlossene Formeln werden auch als *Sätze* bezeichnet.

Während offene Formeln als Aussageformen aufzufassen sind, stellen lediglich die Sätze Aussagen im eigentlichen Sinn dar.

12.2 Semantik der Prädikatenlogik

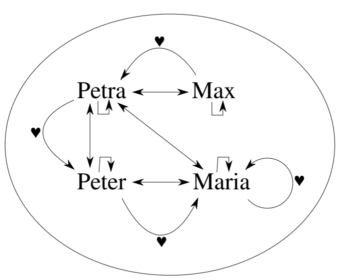
Nachfolgend sei $\mathcal{L}_\Sigma$ eine fest gewählte prädikatenlogische Sprache wie oben beschrieben. In Kapitel 6 hatten wir bereits den Begriff der Σ-Struktur eingeführt. Zur Erinnerung hier nochmals die Definition.

Definition 12.14 Eine Σ-*Struktur* ist ein Paar $\mathcal{A} = \langle A, I \rangle$ wo A eine nichtleere Menge ist und I eine *Interpretationsfunktion*, das heißt eine Funktion mit Definitionsbereich Σ, die folgende Bedingungen erfüllt:

1. für jedes Funktionssymbol $f \in \Sigma_\mathcal{F}$ der Stelligkeit n ist $I(f)$ eine n-stellige Funktion auf A,
2. für jedes Relationssymbol $R \in \Sigma_\mathcal{R}$ der Stelligkeit m ist $I(R)$ eine m-stellige Relation auf A,
3. für jedes $e \in \Sigma_\mathcal{E}$ ist $I(e)$ ein Element von A.

Statt $I(f)$ (resp. $I(R)$, $I(e)$ schreiben wir auch $f_\mathcal{A}$ (resp. $R_\mathcal{A}$, $e_\mathcal{A}$) und reden von der *Interpretation von* f (resp. R, e) *in* $\mathcal{A}$. Im übrigen übernehmen wir die im Kapitel 6 eingeführten Konventionen bei der Darstellung von Strukturen.

Beispiel 12.15 Es sei Σ die Signatur aus Beispiel 12.4 bzw. 12.9. Die nachfolgende Abbildung deutet eine Σ-Struktur $\mathcal{A}$ mit den vier Elementen Max, Maria, Peter und Petra an. Gerade Pfeile deuten die „kennt"-Relation $kennt_\mathcal{A}$ an, die mit einem Herz markierten Pfeile die „liebt"-Relation $liebt_\mathcal{A}$. Zu $Mann_\mathcal{A}$ gehören Max und Peter, zu $Frau_\mathcal{A}$ Maria und Petra. Die Interpretation der anderen Relationssymbole aus Σ sei jeweils leer.



In einer gegebenen Σ-Struktur ist auch durch die Interpretationsfunktion den Variablen noch kein Wert zugewiesen. Variablen werden ja gerade darum eingeführt, um auf beliebige Elemente eines Bereichs referieren zu können. Zur Formalisierung benötigen wir hier den folgenden Begriff.

Definition 12.16 Eine *Variablenbelegung* in eine Σ-Struktur $\mathcal{A} = \langle A, I \rangle$ ist eine Abbildung ν: $X \to A$. Ist ν eine Variablenbelegung in $\mathcal{A}$, ist $x \in X$ und $a \in A$, so bezeichnet $\nu_{x/a}$ die Variablenbelegung mit

$$\nu_{x/a}(y) := \begin{cases} \nu(y) & \text{falls } x \neq y \\ a & \text{falls } y = x. \end{cases}$$

Diese Variablenbelegung wird die „x-Variante von ν, die x auf a abbildet“, genannt.

Unter einer festen Variablenbelegung bezeichnet nun jeder Term ein eindeutige bestimmtes Element der gegebenen Struktur.

Definition 12.17 Die *Auswertung eines $\mathcal{L}_\Sigma$-Terms t in einer Σ-Struktur $\mathcal{A}$ unter der Variablenbelegung ν* ist das folgende Element $\nu_\mathcal{A}(t) \in A$:

1. $\nu_\mathcal{A}(x) := \nu(x)$ für $x \in X$,
2. $\nu_\mathcal{A}(c) := c_\mathcal{A}$ für $c \in \Sigma_\mathcal{E}$,
3. $\nu_\mathcal{A}(f(t_1, \ldots, t_k) := f_\mathcal{A}(\nu_\mathcal{A}(t_1), \ldots, \nu_\mathcal{A}(t_k))$ für $f \in \Sigma_\mathcal{F}$, k-stellig.

Beispiel 12.18 Es sei Σ die in Beispiel 12.3 eingeführte Signatur. Es bezeichne $\mathcal{R}$ die Struktur der reellen Zahlen, die wir als Σ-Struktur auffassen. Die Variablen $x_1, x_2, x_3 \ldots$ sollen durch die Variablenbelegung ν auf die Zahlen $1, 2, 3, \ldots$ abgebildet werden. Dann ist $\nu_\mathcal{R}(x_1 + x_2) = 3$ und $\nu_\mathcal{R}((0 \cdot (x_1 + x_3)) + x_2^2) = 4$.

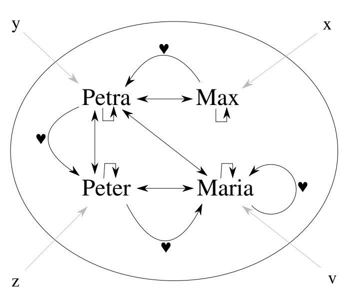
Beispiel 12.19 Es sei Σ die in Beispiel 12.6 eingeführte Signatur. Es bezeichne $\mathcal{A}$ die Menge aller Wörter über dem Alphabet mit den Symbolen a und b, zusammen mit der Konkatenation als Interpretation von „*concat*“. Die Variablen x, y, z sollen durch die Variablenbelegung ν auf die Wörter aaa, bb und $abab$ abgebildet werden. Dann ist $\nu_\mathcal{R}(concat(x, y)) = aaabb$ und $\nu_\mathcal{R}(concat(concat(y, y), concat(x, z))) = bbbbaaaabab$.

Nachdem wir die Auswertung von Termen formalisiert haben, können wir im zweiten Schritt nun den $\mathcal{L}_\Sigma$-Formeln eine Bedeutung zuweisen.

Definition 12.20 [A. Tarski[2]] Die *Auswertung einer $\mathcal{L}_\Sigma$-Formel φ in einer Σ-Struktur $\mathcal{A}$ unter der Variablenbelegung ν* ist der folgende Wahrheitswert $|\varphi|^{\mathcal{A},\nu} \in \{0,1\}$:

1. $|t_1 \equiv t_2|^{\mathcal{A},\nu} = 1$ genau dann, wenn $\nu_\mathcal{A}(t_1) = \nu_\mathcal{A}(t_2)$,
2. $|R(t_1 \ldots, t_m)|^{\mathcal{A},\nu} = 1$ genau dann, wenn $\langle \nu_\mathcal{A}(t_1), \ldots, \nu_\mathcal{A}(t_m)\rangle \in R_\mathcal{A}$,
3. $|\varphi_1 \wedge \varphi_2|^{\mathcal{A},\nu} = 1$ genau dann, wenn $|\varphi_1|^{\mathcal{A},\nu} = 1$ und $|\varphi_2|^{\mathcal{A},\nu} = 1$,
   $|\varphi_1 \vee \varphi_2|^{\mathcal{A},\nu} = 1$ genau dann, wenn $|\varphi_1|^{\mathcal{A},\nu} = 1$ oder $|\varphi_2|^{\mathcal{A},\nu} = 1$,
   $|\varphi_1 \Rightarrow \varphi_2|^{\mathcal{A},\nu} = 1$ genau dann, wenn $|\varphi_1|^{\mathcal{A},\nu} = 1$ impliziert $|\varphi_2|^{\mathcal{A},\nu} = 1$,
   $|\neg\varphi|^{\mathcal{A},\nu} = 1$ genau dann, wenn $|\varphi|^{\mathcal{A},\nu} = 0$,
4. $|\forall\varphi|^{\mathcal{A},\nu} = 1$ genau dann, wenn für alle $a \in A$ stets $|\varphi|^{\mathcal{A},\nu_{x/a}} = 1$ gilt,
   $|\exists\varphi|^{\mathcal{A},\nu} = 1$ genau dann, wenn $|\varphi|^{\mathcal{A},\nu_{x/a}} = 1$ für zumindest ein $a \in A$ gilt.

Beispiel 12.21 Es sei Σ die Signatur aus Beispiel 12.4 bzw. 12.9 und $\mathcal{A}$ die Σ-Struktur aus Beispiel 12.15. Die Variablenbelegung ν bilde die Variablen x, y, z, v wie in der Abbildung angegeben ab.



Dann gilt zum Beispiel

- $|Mann(z) \wedge Frau(y)|^{\mathcal{A},\nu} = 1$,

[2]Alfred Tarski,

- $|liebt(x, maria)|^{\mathcal{A},\nu} = 0$,
- $|kennt(max, maria)|^{\mathcal{A},\nu} = 0$,
- $|\forall x(Mann(x) \vee Frau(x))|^{\mathcal{A},\nu} = 1$,
- $|\exists y(kennt(maria, y) \wedge Frau(y) \wedge liebt(x, y))|^{\mathcal{A},\nu} = 1$.

Beispiel 12.22 Es sei Σ die in Beispiel 12.3 eingeführte Signatur. Es bezeichne $\mathcal{R}$ die Struktur der reellen Zahlen, die wir als Σ-Struktur auffassen. Die Variablen $x_1, x_2, x_3 \ldots$ sollen durch die Variablenbelegung ν auf die Zahlen $1, 2, 3, \ldots$ abgebildet werden. Dann gilt

$$\begin{aligned}
|x_1 + x_2 \equiv x_3|^{\mathcal{R},\nu} &= 1 \\
|x_1 + x_2 \equiv x_3 + x_1|^{\mathcal{R},\nu} &= 0 \\
|\exists z\ (0 \leq z \wedge x_3 \equiv x_1 + z)|^{\mathcal{R},\nu} &= 1. \\
|\forall z(0 \leq z \Rightarrow \exists y(z \equiv y \cdot y))|^{\mathcal{R},\nu} &= 1.
\end{aligned}$$

Wir hatten bei der Auswertung aussagenlogischer Formeln gesehen, daß der Wahrheitswert einer aussagenlogischen Formel φ unter einer 0-1 Bewertung g nur von den Wahrheitswerten der in φ auftretenden Atomformeln abhängt (vgl. Lemma 10.12). In ähnlicher Weise kann zeigen, daß der Wert eines $\mathcal{L}_\Sigma$-Terms t unter einer Variablenbelegung ν nur von den Werten der Variablen in $Var(t)$ anhängt, der Wert einer $\mathcal{L}_\Sigma$-Formel φ nur von den Werten der freien Variablen von φ. Diese einfache Beobachtung ist ein wichtiger Schlüssel bei vielen einfachen Beweisen in der Prädikatenlogik.

Theorem 12.23 (Koinzidenztheorem) *Es sei $\mathcal{A}$ eine Σ-Struktur. Dann gilt für alle Variablenbelegungen ν, ν' in A, für jeden $\mathcal{L}_\Sigma$-Term t und für jede $\mathcal{L}_\Sigma$-Formel φ:*

1. *falls ν und ν' auf $\mathrm{Var}(t)$ übereinstimmen, so $\nu_{\mathcal{A}}(t) = \nu'_{\mathcal{A}}(t)$,*
2. *falls ν und ν' auf $\mathrm{fr}(\varphi)$ übereinstimmen, so $|\varphi|^{\mathcal{A},\nu} = |\varphi|^{\mathcal{A},\nu'}$.*

Beweis. Teil 1 beweist man per Induktion über die Tiefe des Terms t. Dies bleibt als Übung offen. Wir zeigen nun Teil 2 per Induktion über den Formelaufbau von φ.

Fall 1: φ ist eine Atomformel der Form $t_1 \equiv t_2$. Dann ist $Var(t_1) \cup Var(t_2) = fr(\varphi)$ und es folgt mit Teil 1

$$\begin{aligned} |t_1 \equiv t_2|^{\mathcal{A},\nu} = 1 \quad & \text{iff} \quad \nu_{\mathcal{A}}(t_1) = \nu_{\mathcal{A}}(t_2) \\ & \text{iff} \quad \nu'_{\mathcal{A}}(t_1) = \nu'_{\mathcal{A}}(t_2) \\ & \text{iff} \quad |t_1 \equiv t_2|^{\mathcal{A},\nu'} = 1. \end{aligned}$$

Fall 2: φ ist eine Atomformel der Form $R(t_1, \dots, t_n)$. Dann ist $\bigcup_{i=1}^{n} \mathrm{Var}(t_i) = fr(\varphi)$ und es folgt mit Teil 1

$$\begin{aligned} |R(t_1, \dots, t_n)|^{\mathcal{A},\nu} = 1 \quad & \text{iff} \quad \langle \nu_{\mathcal{A}}(t_1), \dots, \nu_{\mathcal{A}}(t_n) \rangle \in R^{\mathcal{A}} \\ & \text{iff} \quad \langle \nu'_{\mathcal{A}}(t_1), \dots, \nu'_{\mathcal{A}}(t_n) \rangle \in R^{\mathcal{A}} \\ & \text{iff} \quad |R(t_1, \dots, t_n)|^{\mathcal{A},\nu'} = 1. \end{aligned}$$

Fall 3: φ ist eine Boolesche Verknüpfung von Teilformeln. Dieser Fall ist einfach. Exemplarisch behandeln wir den Fall, wo φ die Form $\varphi_1 \vee \varphi_2$ hat. Dann ist $\bigcup_{i=1}^{2} fr(\varphi_i) = fr(\varphi)$ und es folgt nach Induktionsvoraussetzung

$$\begin{aligned} |\varphi_1 \vee \varphi_2|^{\mathcal{A},\nu} = 1 \quad & \text{iff} \quad |\varphi_1|^{\mathcal{A},\nu} = 1 \text{ oder } |\varphi_2|^{\mathcal{A},\nu} = 1 \\ & \text{iff} \quad |\varphi_1|^{\mathcal{A},\nu'} = 1 \text{ oder } |\varphi_2|^{\mathcal{A},\nu'} = 1 \\ & \text{iff} \quad |\varphi_1 \vee \varphi_2|^{\mathcal{A},\nu'} = 1. \end{aligned}$$

Fall 4: φ hat die Form $\exists x\psi$. Dann ist $fr(\varphi) = fr(\psi) \setminus \{x\}$ und es gilt nach Induktionsvoraussetzung $|\psi|^{\mathcal{A},\nu'_1} = |\psi|^{\mathcal{A},\nu'_2}$ für alle Variablenbelegungen ν'_1, ν'_2, die auf $fr(\psi)$ übereinstimmen. Demnach gilt

$$\begin{aligned} |\exists x\psi|^{\mathcal{A},\nu} = 1 \quad & \text{iff} \quad \text{es ex. } a \in A \text{ mit } |\psi|^{\mathcal{A},\nu_{x/a}} = 1 \\ & \text{iff} \quad \text{es ex. } a \in A \text{ mit } |\psi|^{\mathcal{A},\nu'_{x/a}} = 1 \\ & \text{iff} \quad |\exists x\psi|^{\mathcal{A},\nu'} = 1. \end{aligned}$$

Fall 5: φ hat die Form $\forall x\psi$. Dann ist $fr(\varphi) = fr(\psi) \setminus \{x\}$ und es gilt nach Induktionsvoraussetzung $|\psi|^{\mathcal{A},\nu'_1} = |\psi|^{\mathcal{A},\nu'_2}$ für alle Variablenbelegungen ν'_1, ν'_2, die auf $fr(\psi)$ übereinstimmen. Demnach gilt

$$\begin{aligned} |\forall x\psi|^{\mathcal{A},\nu} = 1 \quad & \text{iff} \quad \text{für alle } a \in A \text{ gilt } |\psi|^{\mathcal{A},\nu_{x/a}} = 1 \\ & \text{iff} \quad \text{für alle } a \in A \text{ gilt } |\psi|^{\mathcal{A},\nu'_{x/a}} = 1 \\ & \text{iff} \quad |\forall x\psi|^{\mathcal{A},\nu'} = 1. \end{aligned}$$

■

Korollar 12.24 *Es seien φ ein $\mathcal{L}_\Sigma$-Satz, $\mathcal{A}$ eine Σ-Struktur und ν, ν' Variablenbelegungen in A. Dann gilt $|\varphi|^{\mathcal{A},\nu} = |\varphi|^{\mathcal{A},\nu'}$.*

12.3 Gültigkeit in Strukturen

Bei der Auswertung von prädikatenlogischen Formeln sind zwei Situationen zu unterscheiden. Wir können uns für die Auswertung in einer fest vorgegebenen Σ-Struktur interessieren, oder aber für die Auswertung in beliebigen Σ-Strukturen. Wir stellen in diesem Abschnitt die wichtigsten Eigenschaften von Formeln zusammen, die die Auswertung in einer *fest vorgegebenen* Struktur betreffen.

Definition 12.25 Es sei $\mathcal{A}$ eine Σ-Struktur. Die $\mathcal{L}_\Sigma$-Formel φ *gilt in* $\mathcal{A}$, im Zeichen $\mathcal{A} \models \varphi$, und $\mathcal{A}$ heißt *Modell von* φ genau dann, wenn $|\phi|^{\mathcal{A},\nu} = 1$ für jede Variablenbelegung ν in $\mathcal{A}$.

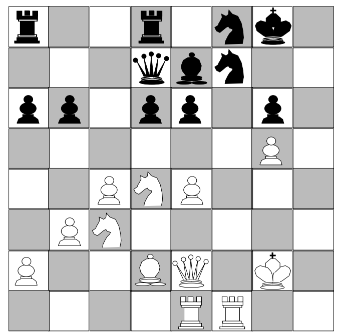
Ist φ eine offene $\mathcal{L}_\Sigma$-Formel mit den freien Variablen $x_1, \ldots, x_n$, so heißt $\forall x_1 \ldots \forall x_n \varphi$ der *universelle Abschluß* von φ.

Bemerkung 12.26 Sei $\varphi = \varphi(x_1, \ldots, x_k)$ eine $\mathcal{L}_\Sigma$-Formel mit den freien Variablen $x_1, \ldots, x_n$. Dann gilt $\mathcal{A} \models \varphi$ genau dann, wenn $\mathcal{A} \models \forall x_1 \ldots \forall x_n \varphi$. Damit sind Formeln mit freien Variablen mit Hinblick auf ihre Gültigkeit in einer gegebenen Struktur äquivalent zu ihrem universellen Abschluß.

Beweis. Es gelte $\mathcal{A} \models \varphi$. Es sei ν eine beliebige Variablenbelegung in $\mathcal{A}$. Wir wollen $|\forall x_1 \ldots \forall x_n \varphi|^{\mathcal{A},\nu} = 1$ zeigen. Es seien $a_1, \ldots, a_n$ beliebige gewählte Elemente aus A. Wegen $\mathcal{A} \models \varphi$ gilt $|\varphi|^{\mathcal{A},\nu_{x_1/a_1 \cdots x_n/a_n}} = 1$. Hieraus folgt aber $|\forall x_1 \ldots \forall x_n \varphi|^{\mathcal{A},\nu} = 1$. Die Umkehrung bleibt dem Leser überlassen. ■

Definition 12.27 Sei Φ eine Menge von $\mathcal{L}_\Sigma$-Formeln. Eine Σ-Struktur $\mathcal{A}$ heißt *Modell von* Φ genau dann, wenn $\mathcal{A} \models \varphi$ für alle $\varphi \in \Phi$ gilt. Wir schreiben kurz $\mathcal{A} \models \Phi$.

Beispiel 12.28 Wir betrachten das Schach-Alphabet aus Beispiel 12.5, es sei $\mathcal{A}$ die Struktur, die durch folgende Schachkonfiguration beschrieben wird

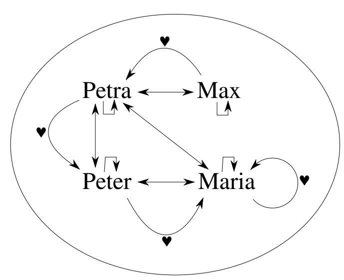


Dann gilt

$$\begin{aligned} &\mathcal{A} \models \forall x\ ((\mathit{weiß}(x) \wedge \neg \mathit{Springer}(x)) \Rightarrow \exists y\ \mathit{deckt}(y,x)),\\ &\mathcal{A} \models \exists x \exists y\ (\mathit{schwarz}(x) \wedge \mathit{bedroht}(y,x) \wedge \\ &\qquad \neg \exists z\ (\mathit{weiß}(z) \wedge \mathit{bedroht}(x,z))), \end{aligned}$$

hingegen nicht $\mathcal{A} \models \exists x \exists y\ (\mathit{bedroht}(x,y) \wedge \mathit{bedroht}(y,x))$.

Beispiel 12.29 Es sei Σ die Signatur aus Beispiel 12.4 bzw. 12.9 und $\mathcal{A}$ die in Beispiel 12.15 erläuterte Σ-Struktur aus der nachfolgenden Abbildung.



Dann gilt

$$\begin{aligned} &\mathcal{A} \models \neg \mathit{liebt}(\mathit{max}, \mathit{maria}),\\ &\mathcal{A} \models \forall x(\mathit{Mann}(x) \Rightarrow (\exists y(\mathit{kennt}(\mathit{maria}, y) \wedge \mathit{Frau}(y) \wedge \mathit{liebt}(x,y)))),\\ &\mathcal{A} \models \exists x \forall y(\mathit{liebt}(x,y) \Rightarrow x = y) \end{aligned}$$

hingegen nicht $\mathcal{A} \models \exists x \exists y\ ((\neg x \equiv y) \wedge \mathit{liebt}(x,y) \wedge \mathit{liebt}(y,x))$.

12.4 Tautologien, Folgerungsbegriff und Äquivalenz

Die nachfolgenden Definitionen stellen nun die wichtigsten Eigenschaften von Formeln und Formelklassen zusammen, die die Auswertung in *beliebigen* Σ-Strukturen betreffen.

Definition 12.30 Sei $\Phi \cup \{\varphi\}$ eine Menge von $\mathcal{L}_\Sigma$-Sätzen. Man sagt φ *folgt aus* Φ, im Zeichen $\Phi \models \varphi$, genau dann, wenn jedes Modell von Φ stets auch Modell von $\{\varphi\}$ ist.

In der obigen Definition haben wir uns bewußt auf $\mathcal{L}_\Sigma$-Sätze beschränkt. Man findet in der Literatur auch allgemeinere Definitionen des Folgerungsbegriffs für beliebige Formeln, allerdings in zwei unterschiedlichen Varianten. Falls $\Phi = \emptyset$, so schreiben wir $\models \varphi$ statt $\emptyset \models \varphi$. Man beachte, daß das Zeichen „$\models$“ nunmehr zwei unterschiedliche Verwendungsweisen hat. Es bezeichnet einerseits die Gültigkeit in einer Σ-Struktur ($\mathcal{A} \models \varphi$), andererseits die Folgerungsbeziehung ($\Phi \models \varphi$).

Definition 12.31 Eine $\mathcal{L}_\Sigma$-Formel φ heißt *logisch allgemeingültig* oder *prädikatenlogische Tautologie* genau dann, wenn φ in jeder Σ-Struktur $\mathcal{A}$ gilt. Eine $\mathcal{L}_\Sigma$-Formel φ heißt *erfüllbar* genau dann, wenn es eine Σ-Struktur $\mathcal{A}$ und eine Variablenbelegung ν in $\mathcal{A}$ gibt mit $|\varphi|^{\mathcal{A},\nu} = 1$. Eine $\mathcal{L}_\Sigma$-Formelmenge Φ heißt *erfüllbar* genau dann, wenn es eine Σ-Struktur $\mathcal{A}$ und eine Variablenbelegung ν in $\mathcal{A}$ gibt mit $|\varphi|^{\mathcal{A},\nu} = 1$ für alle $\varphi \in \Phi$.

Beispiel 12.32 Es enthalte Σ als einziges Zeichen das zweistellige Funktionssymbol „$\circ$“. Dann ist der $\mathcal{L}_\Sigma$-Satz $\forall x \forall y (x \circ y = y \circ x)$ erfüllbar, aber keine Tautologie. In der Tat können wir die reellen Zahlen als Σ-Struktur $\mathcal{R}$ auffassen, und es gilt dann $\mathcal{R} \models \forall x \forall y (x \circ y = y \circ x)$. Andererseits bildet auch die Menge der Wörter über dem Alphabet $\{a, b\}$ zusammen mit der Konkatenation eine Σ-Struktur $\mathcal{A}$. Hier gilt *nicht* $\mathcal{A} \models \forall x \forall y (x \circ y = y \circ x)$.

Beispiel 12.33 In der Logik mit Gleichheit ist die Formelmenge $\{x \equiv y, y \equiv z\}$ erfüllbar, aber keine Tautologie. Es ist $\{x \equiv y, y \equiv z, \neg z \equiv x\}$ unerfüllbar.

Das folgende Lemma enthält eine harmlose, aber häufig nützliche Beobachtung.

Lemma 12.34 *Es sei φ ein $\mathcal{L}_\Sigma$-Satz. Dann ist φ eine Tautologie genau dann, wenn $\neg\varphi$ unerfüllbar ist. Es sei $\Phi \cup \{\varphi\}$ eine Menge von $\mathcal{L}_\Sigma$-Sätzen. Dann gilt $\Phi \models \varphi$ genau dann, wenn $\Phi \cup \{\neg\varphi\}$ unerfüllbar ist.*

Beweis. Wir beweisen die zweite Aussage. Es gelte $\Phi \models \varphi$. Es sei nun $\mathcal{A}$ eine beliebige Σ-Struktur und ν eine Variablenbelegung in $\mathcal{A}$. Falls $|\psi|^{\mathcal{A},\nu} = 1$ für alle $\psi \in \Phi$ gilt, so ist nach Korollar 12.24 $\mathcal{A}$ ein Modell der Satzmenge Φ. Aus der Voraussetzung ergibt sich, daß $\mathcal{A}$ ein Modell von φ ist. Insbesondere folgt $|\varphi|^{\mathcal{A},\nu} = 1$, damit $|\neg\varphi|^{\mathcal{A},\nu} = 0$. Daher ist $\Phi \cup \{\neg\varphi\}$ unerfüllbar. Es sei umgekehrt $\Phi \cup \{\neg\varphi\}$ unerfüllbar. Ist $\mathcal{A}$ ein Modell von Φ, so kann $\mathcal{A}$ kein Modell von $\{\neg\varphi\}$ sein. Es gibt daher eine Variablenbelegung ν in $\mathcal{A}$ mit $|\neg\varphi|^{\mathcal{A},\nu} = 0$ beziehungsweise $|\varphi|^{\mathcal{A},\nu} = 1$. Da φ ein Satz ist, folgt nach Korollar 12.24, daß $\mathcal{A}$ ein Modell von φ ist. Es ist damit jedes Modell von Φ ein Modell von φ, also gilt $\Phi \models \varphi$. ■

Oft möchte man eine gegebene Klasse von Σ-Strukturen durch die Angabe einiger $\mathcal{L}_\Sigma$-Sätze, die in allen Strukturen der Klasse gelten, eindeutig charakterisieren. Hierzu dient die folgende Begriffsbildung.

Definition 12.35 Es sei Σ eine Signatur und $\mathcal{K}$ eine Klasse von Σ-Strukturen. Eine Menge von $\mathcal{L}_\Sigma$-Sätzen Φ heißt *Axiomensystem* oder *Axiomatisierung* von $\mathcal{K}$ genau dann, wenn für jede Σ-Struktur $\mathcal{A}$ gilt: $\mathcal{A} \in \mathcal{K}$ genau dann, wenn $\mathcal{A} \models \Phi$.

Es sei angemerkt, daß es nicht immer möglich ist, für eine gegebene Klasse von Σ-Strukturen eine Axiomatisierung anzugeben. Für eine Reihe spezieller Klassen von Algebren (Halbgruppen, Monoide, Gruppen, Ringe, Körper) hatten wir in Abschnitt 6.1.1 bereits eine Axiomatisierung angegeben.

Beispiel 12.36 Es enthalte Σ als einziges Zeichen das zweistellige Funktionssymbol „$\cdot$". Die Klasse der Halbgruppen (vgl. Def. 6.4) wird durch den Satz $\forall x \forall y \forall z\ (x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ axiomatisiert.

Definition 12.37 Zwei $\mathcal{L}_\Sigma$-Formeln φ_1 und φ_2 heißen *logisch äquivalent* genau dann, wenn $\varphi_1 \Leftrightarrow \varphi_2$ eine Tautologie ist.

Lemma 12.38 *Für beliebige $\mathcal{L}_\Sigma$-Formeln φ und ψ sind die folgenden Formeln sind stets logisch äquivalent:*

1. $\exists x(\varphi \vee \psi)$ *und* $(\exists x\varphi) \vee (\exists x\psi)$,
2. $\forall x(\varphi \wedge \psi)$ *und* $(\forall x\varphi) \wedge (\forall x\psi)$,
3. $\forall x\forall y\varphi$ *und* $\forall y\forall x\varphi$,
4. $\exists x\exists y\varphi$ *und* $\exists y\exists x\varphi$,
5. $\exists x\varphi$ *und* $\neg\forall x\neg\varphi$,
6. $\forall x\varphi$ *und* $\neg\exists x\neg\varphi$,
7. $\neg\exists x\varphi$ *und* $\forall x\neg\varphi$,
8. $\neg\forall x\varphi$ *und* $\exists x\neg\varphi$.

Beweis. Ein Beweis ergibt sich jeweils unmittelbar aus der Anwendung der Regeln von Definition 12.20. Wir zeigen dies für das erste Formelpaar und überlassen die weiteren Fälle dem Leser. Es reicht offenbar zu zeigen, daß die beiden genannten Formeln in jeder Σ-Struktur $\mathcal{A}$ und unter jeder Variablenbelegung ν identisch ausgewertet werden. Dies folgt jedoch aus der Äquivalenz der nachfolgenden Aussagen.

(1) $|\exists x(\varphi \vee \psi)|^{\mathcal{A},\nu} = 1$
(2) es ex. $a \in A$ mit$|\varphi \vee \psi|^{\mathcal{A},\nu_{x/a}} = 1$
(3) es ex. $a \in A$ mit$|\varphi|^{\mathcal{A},\nu_{x/a}} = 1$ oder $|\psi|^{\mathcal{A},\nu_{x/a}} = 1$
(4) es ex. $a \in A$ mit$|\varphi|^{\mathcal{A},\nu_{x/a}} = 1$ oder es ex. $a \in A$ mit$|\psi|^{\mathcal{A},\nu_{x/a}} = 1$
(5) $|\exists x\varphi|^{\mathcal{A},\nu} = 1$ oder $|\exists x\psi|^{\mathcal{A},\nu} = 1$
(6) $|(\exists x\varphi) \vee (\exists x\psi)|^{\mathcal{A},\nu} = 1$

■

Es bleibt dem Leser als Übung überlassen, sich Beispiele von Strukturen auszudenken, die zeigen, daß folgende Formeln hingegen *nicht* logisch äquivalent sind.

1. $\exists x(\varphi \wedge \psi)$ und $(\exists x\varphi) \wedge (\exists x\psi)$,

2. $\forall x(\varphi \vee \psi)$ und $(\forall x\varphi) \vee (\forall x\psi)$,

3. $\forall x \exists y \varphi$ und $\exists y \forall x \varphi$.

Ohne Beweis sei das folgende Theorem erwähnt, von dem wir in vorausgegangenen Beweisen implizit schon oft Gebrauch gemacht haben. Auch in der Folge werden wir dieses Resultat, das eine Verallgemeinerung des aussagenlogischen Ersetzungslemmas 10.19 darstellt, ohne besondere Erwähnung verwenden.

Theorem 12.39 *Ersetzt man in einer $\mathcal{L}_\Sigma$-Formel φ einige Vorkommen der Teilformel ψ durch eine zu ψ logisch äquivalente Formel ψ', so ist die resultierende Formel φ' zur Ausgangsformel φ logisch äquivalent.*

Durch Verbindung mit dem nachfolgenden Lemma sehen wir, daß man stets Teilformeln gegebener Formeln durch aussagenlogisch äquivalente Teilformeln ersetzen darf, ohne die logische Äquivalenz zu verletzen.

Lemma 12.40 *Es sei α eine aussagenlogische Tautologie, die die Aussagenvariablen $A_1, \ldots, A_n$ enthält. Sind $\varphi_1, \ldots, \varphi_n$ beliebige $\mathcal{L}_\Sigma$-Formeln, so ist diejenige Formel γ, die man aus α erhält, indem man jedes Vorkommen von A_i durch φ_i ersetzt ($1 \leq i \leq n$) eine Tautologie der Prädikatenlogik.*

Beweis. Es sei $\mathcal{A}$ eine Σ-Struktur und ν eine Variablenbelegung in $\mathcal{A}$. Gemäß der Auswertung der Booleschen Junktoren in der Prädikatenlogik entspricht $|\gamma|^{\mathcal{A},\nu}$ genau dem Wert von α unter einer 0-1 Bewertung g mit $g(A_i) := |\varphi_i|^{\mathcal{A},\nu}$ ($1 \leq i \leq n$). Da α aussagenlogische Tautologie ist, ist $g(\alpha) = 1$, damit auch $|\gamma|^{\mathcal{A},\nu}$. Es folgt, daß γ eine Tautologie der Prädikatenlogik ist.
■

Korollar 12.41 *Jede $\mathcal{L}_\Sigma$-Formel φ kann bis auf logische Äquivalenz nur mit den Junktoren „$\neg$“ und „$\wedge$“ (das heißt ohne die Junktoren „$\Rightarrow$“ und „$\vee$“) formuliert werden.*

12.5 Normalformen für prädikatenlogische Formeln

Lemma 12.42 *Es seien φ und ψ zwei beliebige $\mathcal{L}_\Sigma$-Formeln. Falls $x \notin$ fr(ψ), so sind*

$$\begin{array}{llcl} (a) & (\exists x\varphi) \wedge \psi & \text{und} & \exists x(\varphi \wedge \psi) \\ (b) & (\forall x\varphi) \wedge \psi & \text{und} & \forall x(\varphi \wedge \psi) \\ (c) & (\exists x\varphi) \vee \psi & \text{und} & \exists x(\varphi \vee \psi) \\ (d) & (\forall x\varphi) \vee \psi & \text{und} & \forall x(\varphi \vee \psi) \end{array}$$

logisch äquivalent.

Beweis. (a) Es sei $\mathcal{A}$ eine Σ-Struktur und ν eine Variablenbelegung in A. Dann gilt $|(\exists x\varphi) \wedge \psi|^{\mathcal{A},\nu} = 1$ genau dann, wenn es ein a in A gibt mit $|\varphi|^{\mathcal{A},\nu_{x/a}} = 1$ und $|\psi|^{\mathcal{A},\nu} = 1$. Da $x \notin \mathrm{fr}(\psi)$ ist nach dem Koinzidenztheorem $|\psi|^{\mathcal{A},\nu} = |\psi|^{\mathcal{A},\nu_{x/a}}$. Also ist die obige Aussage äquivalent zu $|\exists x(\varphi \wedge \psi)|^{\mathcal{A},\nu} = 1$. Damit ist Teil (a) bewiesen, die anderen Teile folgen analog. ■

Formeln, wo gebundene Variablen auch frei vorkommen, oder verschiedene Quantorenvorkommen dieselbe Variable binden, kann man als pathologisch betrachten. Es ist einfach zu zeigen, daß man zu jeder vorgegebenen Formel φ eine logisch äquivalente Formel φ' konstruieren kann, wo freie und gebundene Variablen stets voneinander verschieden sind, und wo zwei Quantorenvorkommen nie dieselbe Variable binden. Im nachfolgenden nehmen wir stets an, daß die betrachteten Formeln von dieser Gestalt sind.

Definition 12.43 Eine $\mathcal{L}_\Sigma$-Formel φ heißt in *pränexer Normalform* genau dann, wenn φ die Form $Q_1x_1 \ldots Q_nx_n\psi$ hat, wo die Q_i Quantoren aus $\{\forall, \exists\}$ sind, und wo die „Matrix" ψ quantorenfrei ist. Hat darüberhinaus die Matrix ψ die Form $\bigwedge_{i=1}^{r}(\bigvee_{j=1}^{s_i} \gamma_{i,j})$, wo die Formeln $\gamma_{i,j}$ negierte oder unnegierte Atomformeln sind, so heißt φ in *konjunktiver pränexer Normalform*. Dual heißt φ in *disjunktiver pränexer Normalform* genau dann, wenn die Matrix ψ die Form $\bigvee_{i=1}^{r}(\bigwedge_{j=1}^{s_i} \gamma_{i,j})$ hat, wo wieder die Formeln $\gamma_{i,j}$ negierte oder unnegierte Atomformeln sind.

Satz 12.44 *Zu jeder vorgegebenen $\mathcal{L}_\Sigma$-Formel φ kann man effektiv eine logisch äquivalente Formeln φ' in pränexer (konjunktiver, oder disjunktiver) Normalform berechnen.*

Beweis. (Skizze) Wir können mit Korollar 12.41 annehmen, daß φ nur die Junktoren „$\neg$" und „$\wedge$" enthält. Wir setzen außerdem voraus, daß φ nicht pathologisch im oben erklärten Sinn ist. Durch Anwendung der Äquivalenzen 7 und 8 aus Lemma 12.38 und der de Morganschen Regeln kann man eine logisch äquivalente Formel φ_1 erreichen, die nur die Junktoren „$\vee$", „$\wedge$" und „$\neg$" verwendet, und wo Negationszeichen nur direkt vor Atomformeln vorkommen. Außerdem ist auch φ_1 nicht pathologisch im oben erklärten Sinn. Man kann daher nach Lemma 12.42 die Quantoren in φ_1 stets vor Konjunktionen und Disjunktionen ziehen. Hierdurch wird eine Formel φ_2 in pränexer Normalform erreicht, die zu φ logisch äquivalent ist. Nun kann man die Matrix offenkundig durch rein aussagenlogische Umformungen in die bei der konjunktiven (bzw. disjunktiven) pränexen Normalform geforderten Form bringen (vgl. Satz 11.40 und Theorem 12.39). ■

12.6 Ergänzung: Natürliche Sprache und Prädikatenlogik

Die Quantoren „ein" und „alle". Standardsemantik. "Ein" als „jeder". Skopusambiguitäten. De re, de-dicto Lesarten.

Weitere NL Quantoren. Zwei etc.

Nichtausdrückbarkeit.

12.7 Aufgaben zu Kapitel 12

Aufgaben zu Teilkapitel 12.1

Aufgabe 12.1 Geben Sie eine exakte Charakterisierung derjenigen endlichen Signaturen Σ, für die es $\mathcal{L}_\Sigma$-Terme beliebiger Tiefe gibt.

Aufgabe 12.2 Geben Sie für die nachfolgenden Sätze eine natürlichsprachliche Formulierung.

$$\exists y \mathit{Frau}(y) \wedge \mathit{liebt}(\mathit{max}, y)$$
$$\forall x(\mathit{Mann}(x) \Rightarrow (\exists y \mathit{Frau}(y) \wedge \mathit{liebt}(x, y)))$$

$$\forall x \forall y (\mathit{Freund}(x,p) \wedge \mathit{Feind}(y,x)) \Rightarrow \mathit{Feind}(y,p))$$
$$\forall x (\mathit{Mann}(x) \wedge \forall y ((\mathit{Frau}(y) \wedge \mathit{kennt}(x,y)) \Rightarrow \mathit{liebt}(x,y)) \Rightarrow (\neg \exists z (\mathit{Frau}(z) \wedge \mathit{liebt}(z,x))))$$

Aufgabe 12.3 Geben Sie eine geeignete Signatur Σ an, die verwendet werden kann, um über Gruppen (vgl. Def. 6.9) zu reden. Geben Sie dann einige Terme und Formeln dieser Sprache an.

Aufgabe 12.4 Markieren Sie die freien Variablen-Vorkommen in den folgenden Formeln. Geben Sie die Menge der freien Variablen der Formeln an.

(1) $(\forall x\ ((\exists y\ R(x,y)) \wedge S(y))) \wedge R(x,u)$

(2) $(\forall u\ (\forall x \forall y\ ((\exists z\ R(x,y)) \wedge S(z))) \wedge R(x,u))$

Aufgaben zu Teilkapitel 12.2

Aufgabe 12.5 Es sei Σ die in Beispiel 12.3 eingeführte Signatur. Es bezeichne $\mathcal{R}$ die Struktur der reellen Zahlen, die wir als Σ-Struktur auffassen. Zeigen Sie: eine reelle Zahl r tritt genau dann als Auswertung eines variablenfreien $\mathcal{L}_\Sigma$-Terms auf (unter irgendeiner Variablenbelegung in $\mathcal{R}$) auf, wenn r eine natürliche Zahl ist. Erweitern Sie nun die Signatur und geben Sie eine dazu passende Interpretation der neuen Symbole derart an, daß alle ganzen Zahlen (rationalen Zahlen) als Auswertung eines variablenfreien Terms auftreten. Auf welche Schwierigkeit stößt man, wenn man ein einstelliges Funktionssymbol „$^{-1}$" einführt, das die Inversenbildung $x \mapsto 1/x$ beschreibt?

Aufgabe 12.6 Man kann die erste Klausel von Definition 12.20 wie folgt paraphrasieren: eine Gleichung $t_1 \equiv t_2$ gilt in der Struktur $\mathcal{A}$ unter der Variablenbelegung ν genau dann, wenn die Auswertung der Terme t_1 und t_2 in $\mathcal{A}$ unter ν dasselbe Element ergibt. Geben Sie entsprechende Paraphrasierungen für die anderen Klauseln von Definition 12.20 an.

Aufgabe 12.7 Welche Auswertungsergebnisse in Beispiel 12.21 würde sich durch Wahl einer anderen Variablenbelegung verändern lassen? Wie sehe eine passende Variablenbelegung jeweils aus? Welche Auswertungsergebnisse in Beispiel 12.22 würde sich durch Wahl einer anderen Variablenbelegung verändern lassen? Wie sehe eine passende Variablenbelegung jeweils aus?

Aufgabe 12.8 Beweisen Sie Teil 1 des Koinzidenztheorems.

Aufgaben zu Teilkapitel 12.3

Aufgabe 12.9 Es sei φ ein $\mathcal{L}_\Sigma$-Satz und $\mathcal{A}$ eine Σ-Struktur. Zeigen Sie: es gilt entweder $\mathcal{A} \models \varphi$ oder $\mathcal{A} \models \neg\varphi$. Geben Sie ein Beispiel an, das zeigt, daß dies nicht in entsprechender Weise für $\mathcal{L}_\Sigma$-Formeln φ gilt.

Aufgabe 12.10 Erweitern Sie Beispiel 12.28, indem Sie mindestens drei weitere Sätze angeben, die in $\mathcal{A}$ gelten. Jeder Satz soll zumindest zwei Quantoren haben. Geben Sie auch drei weitere Sätze an, die in $\mathcal{A}$ nicht gelten.

Aufgabe 12.11 Modifizieren Sie Beispiel 12.29, indem Sie eine andere Σ-Struktur $\mathcal{B}$ angeben, wo die drei angegebenen Sätze, die in $\mathcal{A}$ gelten, nun falsch sind. Hingegen soll der angegebene Satz, der in $\mathcal{A}$ nicht gilt, nun in $\mathcal{B}$ gelten.

Aufgabe 12.12 Geben Sie in der Logik mit Gleichheit einen Satz an, der in einer Struktur $\mathcal{A}$ gilt genau dann, wenn A genau (resp. mindestens) zwei Elemente hat. Es sei nun R ein einstelliges Relationssymbol. Schreiben Sie einen Satz, der ausdrückt, daß es genau zwei Elemente mit der Eigenschaft R gibt.

Aufgaben zu Teilkapitel 12.4

Aufgabe 12.13 Ist der Satz $\exists x\ (x \equiv x)$ eine Tautologie? Es sei c eine Individuenkonstante. Ist der Satz $\forall x \forall y\ ((x \equiv c) \wedge (y \equiv c)) \Rightarrow (x \equiv y))$ eine Tautologie? Geben Sie jeweils einen Beweis oder ein Gegenbeispiel.

Aufgabe 12.14 Es sei R ein einstelliges Relationssymbol der Signatur Σ. Zeigen Sie, daß der Satz $\exists x\ (R(x) \Rightarrow (\forall y\ R(y)))$ eine Tautologie ist.

Aufgabe 12.15 Es sei f ein einstelliges Funktionssymbol der Signatur Σ. Zeigen Sie, daß der Satz $(\forall x\ (x \equiv f(x))) \Rightarrow (\forall x\ (x \equiv f(f(x))))$ eine Tautologie ist.

Aufgabe 12.16 Eine mögliche Verallgemeinerung von Definition 12.30 auf $\mathcal{L}_\Sigma$-Formeln besteht darin, daß man die Definition in der angegebenen Form auf $\mathcal{L}_\Sigma$-Formeln erweitert. Was ist hierbei problemetisch, und wie könnte eine alternative Erweiterung des Folgerungsbegriffs auf Formeln aussehen?

Aufgabe 12.17 Beweisen Sie die Teile 2–8 von Lemma 12.38.

Aufgabe 12.18 Es seien $R, S \in \Sigma_\mathcal{R}$ einstellige Relationssymbole. Geben Sie Beispiele von Σ-Strukturen an, die zeigen, daß folgende Formeln nicht logisch äquivalent sind.

1. $\exists x\ (R(x) \wedge S(x))$ und $(\exists x\ R(x)) \wedge (\exists x\ S(X))$,
2. $\forall x\ (R(x) \vee S(x))$ und $(\forall x\ R(x)) \vee (\forall x\ S(x))$,
3. $\forall x \exists y\ R(x)$ und $\exists y \forall x\ R(x)$.

Aufgaben zu Teilkapitel 12.5

Aufgabe 12.19 Es sei φ eine beliebige Formel und ψ eine Formel der „pathologischen“ Gestalt $\forall x \exists x \forall y \exists y\ \varphi$. Welche der angegebenen vier Quantifizierungen können weggelassen werden unter Erhaltung der logischen Äquivalenz?

Aufgabe 12.20 Geben Sie einen Satz in konjunktiver pränexer Normalform an, der zu

$$\forall x(\mathit{Mann}(x) \Rightarrow (\exists y(\mathit{kennt}(\mathit{maria}, y) \wedge \mathit{Frau}(y) \wedge \mathit{liebt}(x, y))))$$

äquivalent ist.

Aufgabe 12.21 Geben Sie ein Beispiel an, daß zeigt, daß zwei Formeln $(\exists x \varphi) \wedge \psi$ und $\exists x(\varphi \wedge \psi)$ nicht notwendig logisch äquivalent sind.

12.8 Bibliographische Angaben

Tiefergehende Darstellungen der Prädikatenlogik finden sich in praktisch allen Lehrbüchern zur mathematischen Logik wie zum Beispiel [BM77, Her78, Kle67, Sho67, BM77]. Unter anspruchsvollen weiterführenden Werken seien zwei erwähnt. Einen Überblick über viele der damaligen Forschungsrichtungen und -Fragen innerhalb der mathematischen Logik bietet [(Ed77]. Das Teilgebiet der Modelltheorie ist ausführlich in [CK73] dargestellt.

14

Lösungen zu einigen Übungsaufgaben

Lösungen zu Aufgaben aus Kapitel 1

1.2 Teil 1. Setze $\delta_1 := (\alpha \Rightarrow \beta)$, $\delta_2 := (\alpha \wedge \beta) \Rightarrow \gamma$, $\delta_3 := \alpha \Rightarrow \gamma$. Die erwähnte Aussage ist damit $\delta_4 := (\delta_1 \wedge \delta_2) \Rightarrow \delta_3$. Bei der Berechnung der Wahrheitswerte der Teilformeln haben wir 8 Fälle zu unterscheiden.

α	β	γ	δ_1	$\alpha \wedge \beta$	δ_2	$\delta_1 \wedge \delta_2$	δ_3	δ_4
0	0	0	1	0	1	1	1	1
0	0	1	1	0	1	1	1	1
0	1	0	1	0	1	1	1	1
1	0	0	0	0	1	0	0	1
0	1	1	1	0	1	1	1	1
1	0	1	0	0	1	0	1	1
1	1	0	1	1	0	0	0	1
1	1	1	1	1	1	1	1	1

Die Formel $((\alpha \Rightarrow \beta) \wedge ((\alpha \wedge \beta) \Rightarrow \gamma)) \Rightarrow (\alpha \Rightarrow \gamma)$ kann wie folgt interpretiert werden: Gelingt es, aus einer Prämisse α eine Zwischenbehauptung β zu beweisen, und kann unter Zuhilfenahme der Zwischenbehauptung eine Aussage γ bewiesen werden, so stellt dies einen gültigen Beweis dar, daß aus α die Aussage γ folgt. Die anderen Teile der Aufgabe bleiben dem Leser überlassen.

1.4 Teil 1. Es gibt eine natürliche Zahl, die größergleich jeder weiteren (nicht notwendig verschiedenen!) natürlichen Zahl ist. Die Aussage ist falsch.

Teil 2. Es gibt eine natürliche Zahl, die kleinergleich jeder weiteren (nicht notwendig verschiedenen!) natürlichen Zahl ist. Die Aussage ist richtig, die Zahl Null hat die angegebene Eigenschaft.
Teil 3. Jede natürliche Zahl ist gleich Null, oder sie ist Nachfolgerzahl einer anderen natürlichen Zahl. Die Aussage ist richtig.
Teil 4. Zu jeder natürlichen Zahl x gibt es eine größere Zahl y, so daß jede natürliche Zahl zwischen x und y mit x oder mit y übereinstimmt. Die Aussage ist richtig, $x := y + 1$.
Teil 5. Es gibt eine von Null verschiedene Zahl x, so daß für jede natürliche Zahl y das Produkt von x und y kleiner ist als die Summe von x und y. Die Aussage ist richtig, die Zahl 1 hat die Eigenschaft.
Teil 6. Zu je zwei Zahlen x und y gibt es eine Zahl z, so daß jede Zahl u, die sowohl x als auch y teilt, auch u teilt. Die Aussage ist richtig. Das Produkt (oder auch das kleinste gemeinsame Vielfache) von x und y hat diese Eigenschaft.

1.6 Die Wahrheitswert-Tabelle für „+“ ist

α	β	$\neg\alpha$	$\neg\beta$	$\alpha \wedge \neg\beta$	$\beta \wedge \neg\alpha$	$\alpha + \beta$
0	0	1	1	0	0	0
0	1	1	0	0	1	1
1	0	0	1	1	0	1
1	1	0	0	0	0	0

Der Junktor „+“ repräsentiert das „ausschließende oder“: Eine Aussage $\alpha + \beta$ ist wahr genau dann, wenn *genau eine* der beiden Teilaussagen α und β wahr sind. Der Nachweis, daß die angegebenen Aussagen Tautologien sind, erfolgt wie üblich mit Wahrheitswerttabellen unter direkter Verwendung der obigen Tabelle und bleibt dem Leser überlassen.

Lösungen zu Aufgaben aus Kapitel 2

2.1 Eine mögliche Darstellung der ersten Menge ist

$$\{n \in \mathbb{N} \mid \exists k, l \in \mathbb{N}\colon n = k^2 + l^2\},$$

eine mögliche Darstellung der zweiten Menge ist

$$\{n \in \mathbb{N} \mid \forall k, l \in \mathbb{N}\colon (k \cdot l = n \to k < 1000)\}.$$

Andere Darstellungen sind möglich.

2.2 $\{\{n\} \mid n \in \mathbb{N}\}$ (und *nicht etwa* $\{n \mid n \in \mathbb{N}\}$).

2.5 $\{\{\{a\}, \{\{a\}\}\}\}$.

2.6 Wir brauchen nur den Fall zu betrachten wo in der Tat $M \subseteq N$ und $N \subseteq M$. Nach der Definition der Teilmengenbeziehung in 2.4 bedeutet dies $\forall x\colon x \in M \Rightarrow x \in N$ und umgekehrt $\forall x\colon x \in N \Rightarrow x \in M$. Sei x gegeben. Es folgt $x \in M \Rightarrow x \in N \wedge x \in N \Rightarrow x \in M$. Mit Tautologie 16 aus Satz 1.1 ergibt sich $x \in M \Leftrightarrow x \in N$. Da x beliebig war, haben wir damit $\forall x\colon x \in M \Leftrightarrow x \in N$ gezeigt. Nach Definition 2.2 bedeutet dies $M = N$. Der Beweis wird hier nur so detailliert geführt, um nochmals die Verwendung der Prinzipien aus Kapitel 1 deutlich zu machen. Sobald die Verwendungsweise dieser Prinzipien klar ist, sollte eine kürzere Darstellung gewählt werden.

2.7 Ist C eine beliebige Teilmenge von $\{1,2,3\}$, so ist $\{1,2,3\} \cup (C \cup \{4,5\}) = \{1,2,3,4,5,6\}$. Damit ergeben sich folgende Mengen B: $\{4,5,6\}$, $\{1,4,5,6\}$, $\{2,4,5,6\}$, $\{3,4,5,6\}$, $\{1,2,4,5,6\}$, $\{1,3,4,5,6\}$, $\{2,3,4,5,6\}$, $\{1,2,3,4,5,6\}$.

2.16 Es sei M eine beliebige Menge, $A := M$ und $B := M \cup \{1\}$ (oder $B := M$ und $A := M \cup \{1\}$). Dann ist $A \bowtie B = \{1\}$, wie man leicht nachrechnet. Sind umgekehrt A, B Mengen mit $A \bowtie B = \{1\}$, so können wir ohne Beschränkung der Allgemeinheit $1 \in B$ annehmen. Es kann 1 dann nicht in A sein. Jedes von 1 verschiedene Element $b \in B$ muß aber in A liegen, da b ansonsten in $A \bowtie B$ enthalten wäre. Damit gilt $A = B \setminus \{1\}$, das heißt A und B haben auch stets die oben angegebene Form.

2.17 In der symmetrischen Differenz von $n \geq 2$ Mengen liegen genau diejenigen Elemente, die in einer ungeraden Zahl von beteiligten Mengen als Element auftreten. Eine Aussage $\alpha_1 + \cdots + \alpha_n$ ist wahr genau dann, wenn eine ungerade Zahl von Teilaussagen wahr ist. Man könnte naiv annehmen, daß genau eine Teilaussage wahr sein muß.

2.18 Wir betrachten die Aussagen

$$
\begin{array}{ll}
(1) & x \in -(A_1 \cup A_2) \\
(2) & (x \in B) \wedge \neg((x \in A_1) \vee (x \in A_2)) \\
(3) & (x \in B) \wedge (\neg(x \in A_1) \wedge \neg(x \in A_2)) \\
(4) & ((x \in B) \wedge \neg(x \in A_1)) \wedge ((x \in B) \wedge \neg(x \in A_2)) \\
(5) & x \in (-A_1) \cap (-A_2).
\end{array}
$$

(1) und (2) sind aufgrund der Definition des Komplements äquivalent. (2) und (3) sind nach Tautologie 10 aus Satz 1.1 äquivalent. Die Äquivalenz von (3) und (4), wo wir im wesentlichen die Teilkonjunktion $x \in B$ verdoppeln, folgt aus den Tautologien 2, 3, 5. (4) und (5) sind aufgrund der Definition des Komplements und des Mengendurchschnitts äquivalent. Damit sind (1) und (5) äquivalent und es gilt nach dem Extensionalitätsprinzip 2.2 $-(A_1 \cup A_2) = (-A_1) \cap (-A_2)$.

2.21 Es gelte zunächst $A \subseteq B$. Sei x gegeben. Dann gilt $x \in A \Rightarrow x \in B$. Nach Tautologie 12 aus Satz 1.1 gilt dann $x \in A \vee x \in B$ genau dann, wenn $x \in B$. Es ist also $x \in A \cup B$ äquivalent zu $x \in B$. Da x beliebig war, folgt nach dem

Extensionalitätsprinzip 2.2 $A \cup B = B$. Es gelte nun umgekehrt $A \cup B = B$. Sei x gegeben. Falls $x \in A$ so auch $x \in A \cup B = B$. Da x beliebig war, gilt $A \subseteq B$. Damit ist die erste Äquivalenz bewiesen. Die zweite folgt analog unter Verwendung von Tautologie 13 aus Satz 1.1.

2.23 Offenkundig impliziert $M = N$ auch $\mathcal{P}(M) = \mathcal{P}(N)$. Es gelte nun umgekehrt $\mathcal{P}(M) = \mathcal{P}(N)$. Sei $m \in M$. Dann gilt $\{m\} \in \mathcal{P}(M) = \mathcal{P}(N)$, nach Definition der Potenzmenge folgt $\{m\} \subseteq N$ und nach Definition der Teilmengenbeziehung $m \in N$. Symmetrisch impliziert $m \in N$ auch $m \in M$. Nach dem Extensionalitätsprinzip 2.2 folgt $M = N$.

2.24 $\mathcal{P}(\mathcal{P}(\{1,2\})) = \mathcal{P}(\{\emptyset, \{1\}, \{2\}, \{1,2\}\})$ enthält genau die 16 Elemente $\emptyset$, $\{\emptyset\}$, $\{\{1\}\}$, $\{\{2\}\}$, $\{\{1,2\}\}$, $\{\emptyset, \{\{1\}\}$, $\{\emptyset, \{\{2\}\}$, $\{\emptyset, \{\{1,2\}\}$, $\{\{\{1\}, \{2\}\}$, $\{\{\{1\}, \{1,2\}\}$, $\{\{\{2\}, \{1,2\}\}$, $\{\emptyset, \{1\}, \{2\}\}$, $\{\emptyset, \{1\}, \{1,2\}\}$, $\{\emptyset, \{2\}, \{1,2\}\}$, $\{\{1\}, \{2\}, \{1,2\}\}$ und $\{\emptyset, \{1\}, \{2\}, \{1,2\}\}$

2.27 Es ist $\bigcup A = \{\{1,2,3\}, \{1,2,4\}, \{1,4\}, \{1,2,5\}\}$ und $\bigcap A = \{\{1,2,3\}, \{1,4\}\}$, damit $\bigcup(\bigcup A) = \{1,2,3,4,5\}, \bigcap(\bigcup A) = \{1\}, \bigcap(\bigcap A) = \{1\}, \bigcup(\bigcap A) = \{1,2,3,4\}$.

2.31 Sei M eine beliebige Menge. Dann gilt $\bigcup \mathcal{P}(M) = M$ und damit $\mathcal{P}(\bigcup \mathcal{P}(M)) = \mathcal{P}(M)$. Gilt umgekehrt $\mathcal{P}(\bigcup A) = A$, so ist A die Potenzmenge von $M := \bigcup A$.

2.33 Die leere Menge hat genau eine Partition, nämlich $\emptyset$. Jede einelementige Menge $\{a\}$ hat genau eine Partition, nämlich $\{\{a\}\}$. Hat M genau zwei Partitionen, so hat M genau zwei Elemente. Hat umgekehrt $M = \{a, b\}$ genau zwei Elemente, so sind $\{\{a,b\}\}$ und $\{\{a\}, \{b\}\}$ die einzigen Partitionen von M. Es sei $M = \{a, b, c\} \cup N$ eine Menge mit zumindest drei Elementen a, b, c. Dann sind $\{M\}$, $\{\{m\} \mid m \in M\}$, $\{\{a,b\}\} \cup \{\{m\} \mid m \in M, a \neq m \neq b\}$, $\{\{b,c\}\} \cup \{\{m\} \mid m \in M, b \neq m \neq c\}$ vier (unter anderen möglichen) Partitionen. Es kann daher keine Menge mit genau drei Partitionen geben.

2.36 Beide Annahmen führen auf einen Widerspruch: ist „heterologisch" autologisch, so besitzt es die durch „heterologisch" ausgedrückte Eigenschaft und ist heterologisch. „heterologisch" heterologisch, so besitzt es die durch „heterologisch" ausgedrückte Eigenschaft nicht und ist autologisch.

2.37 Angenommen in $\bigcap A = \{x \mid \forall M \in A\colon x \in M\}$ liegt ein Element N von A. Dann muß $N \in N$ gelten. Wir erhalten eine unendlich absteigende Kette $\ldots \in N \in N \in N$ von Elementbeziehungen. Dies ist ausgeschlossen, wenn — wie meist üblich — das Prinzip der Fundiertheit von Mengen gefordert wird.

Lösungen zu Aufgaben aus Kapitel 3

3.1 Es sei $b \neq d$. Mit $a := \{d\}$ und $c := \{b\}$ folgt $\{a, \{b\}\} = \{\{d\}, \{b\}\} = \{\{b\}, \{d\}\} = \{c, \{d\}\}$, es gilt jedoch $a \neq c$ und $b \neq d$. Das Lemma würde also falsch.

3.2 $\prod_{i=1,\ldots,n} A_i = \emptyset$.

3.4 Die Menge A muß die Elemente a, b, c enthalten, die als vordere Komponenten in den genannten Tupeln auftreten. Die Menge B muß die Elemente a, c, d enthalten. Da $A \times B$ nach Voraussetzung $5+4 = 9$ Elemente hat, muß $A = \{a, b, c\}$ und $\{a, c, d\}$ gelten. Die herausgenommenen Elemente sind $\langle a, d\rangle$, $\langle b, c\rangle$, $\langle b, d\rangle$ und $\langle c, a\rangle$.

3.5 Es habe A genau m, B genau n Elemente. Es sei k das Maximum der Zahlen. Dann gibt es eine Menge $K = \{\langle a_1, b_1\rangle, \ldots, \langle a_k, b_k\rangle\} \subseteq A \times B$ wo $A \subseteq \{a_1, \ldots, a_k\}$ und $B \subseteq \{b_1, \ldots, b_k\}$. Man kann damit an K alle Elemente von A und B ablesen.

3.7. Ist $x \in A \times (B \setminus C)$, so hat x die Form $x = \langle a, b\rangle$ wo $a \in A$, $b \in B$, $b \notin C$ gilt. Es ist dann $x = \langle a, b\rangle \in A \times B$ und $x = \langle a, b\rangle \notin A \times C$, also $x \in (A \times B) \setminus (A \times C)$. Ist $x \in (A \times B) \setminus (A \times C)$, so ist $x \in A \times B$ und $x \notin A \times C$. Aus der ersten Aussage folgt, daß x die Form $x = \langle a, b\rangle$ mit $a \in A$, $b \in B$ hat. Wegen der zweiten Aussage gilt $b \notin C$. Damit ergibt sich $b \in B \setminus C$ und $x \in A \times (B \setminus C)$. Nach dem Extensionalitätsprinzip 2.2 folgt $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.

3.8 Ein Beispiel ist $\{a^n b^n c^n \mid n \in \mathbb{N}\}$.

3.9 Es sind alle Teilmengen von $\{1,2\} \times \{1,2\} = \{\langle 1,1\rangle, \langle 1,2\rangle, \langle 2,1\rangle, \langle 2,2\rangle\}$ anzugeben. Es gibt 16 solche Teilmengen, nämlich $\emptyset$, $\{\langle 1,1\rangle\}$, $\{\langle 1,2\rangle\}$, $\{\langle 2,1\rangle\}$, $\{\langle 2,2\rangle\}$, $\{\langle 1,1\rangle, \langle 1,2\rangle\}$, $\{\langle 1,1\rangle, \langle 2,1\rangle\}$, $\{\langle 1,1\rangle, \langle 2,2\rangle\}$, $\{\langle 1,2\rangle, \langle 2,1\rangle\}$, $\{\langle 1,2\rangle, \langle 2,2\rangle\}$, $\{\langle 2,1\rangle, \langle 2,2\rangle\}$, $\{\langle 1,1\rangle, \langle 1,2\rangle, \langle 2,1\rangle\}$, $\{\langle 1,1\rangle, \langle 1,2\rangle, \langle 2,2\rangle\}$, $\{\langle 1,1\rangle, \langle 2,1\rangle, \langle 2,2\rangle\}$, $\{\langle 1,2\rangle, \langle 2,1\rangle, \langle 2,2\rangle\}$, $\{\langle 1,1\rangle, \langle 1,2\rangle, \langle 2,1\rangle, \langle 2,2\rangle\}$.

3.10 Es hat A^m genau n^m Elemente. Jede Teilmenge ist eine m-stellige Relationen auf A. Nach Lemma 2.27 gibt es 2^{n^m} viele.

3.14 $R^{-1} = \{\langle 2,0\rangle, \langle 4,2\rangle, \langle 6,4\rangle\}$, $S^{-1} := \{\langle 3,1\rangle, \langle 5,3\rangle, \langle 4,3\rangle, \langle 0,3\rangle\}$, $R \circ S = \emptyset$, $S \circ R = \{\langle 3,2\rangle, \langle 3,6\rangle\}$ und $R^{-1} \circ S^{-1} = \{\langle 2,3\rangle, \langle 6,3\rangle\}$. Es ist $(S \circ R)^{-1} = R^{-1} \circ S^{-1}$ (diese Gleichung gilt für beliebige zweistellige Relationen $R \subseteq A \times B$ und $S \subseteq B \times C$).

3.16 Teil 1: Es sei $x \in (R^{-1})^{-1}$. Dann hat x die Form $\langle y, z\rangle$ wo $\langle z, y\rangle \in R^{-1}$, also $x = \langle y, z\rangle \in R$ ist. Ist umgekehrt $x \in R$, so hat x die Form $\langle y, z\rangle$ wo $\langle z, y\rangle \in R^{-1}$, also $x = \langle y, z\rangle \in (R^{-1})^{-1}$ ist. Also folgt $(R^{-1})^{-1} = R$.
Teil 2: Es sei $x \in (R \cup S)^{-1}$. Dann hat x die Form $\langle y, z\rangle$ wo $\langle z, y\rangle \in R \cup S$. Falls $\langle z, y\rangle \in R$ folgt $x = \langle y, z\rangle \in R^{-1} \subseteq R^{-1} \cup S^{-1}$. Analog folgt $x \in R^{-1} \cup S^{-1}$ im Fall $\langle z, y\rangle \in S$. Wir haben $(R \cup S)^{-1} \subseteq R^{-1} \cup S^{-1}$ gezeigt. Umgekehrt ist offenkundig

$R^{-1} \subseteq (R \cup S)^{-1}$ und analog $S^{-1} \subseteq (R \cup S)^{-1}$, also auch $R^{-1} \cup S^{-1} \subseteq (R \cup S)^{-1}$. Mit Teil 2 von Lemma 2.6 folgt $R^{-1} \cup S^{-1} = (R \cup S)^{-1}$.

3.18 Es bezeichne $l \in \{0, 1, \ldots, k\}$ die Zahl derjenigen Elemente aus A, die als zweite Komponente a_2 eines Elements $\langle a_1, a_2 \rangle$ aus R auftreten. Es bezeichne m die Zahl derjenigen Elemente aus A, die als erste Komponente b_1 eines Elements $\langle b_1, b_2 \rangle$ aus S auftreten. Da $R \circ S = \emptyset$ gelten soll, kann kein Element zugleich als zweite Komponente eines Elements von R und als erste Komponente eines Elements von S auftreten. Damit folgt $l + m \leq k$. Offenkundig ist die Zahl der Elemente von R (resp. S) kleinergleich $k \cdot l$ (resp. $m \cdot k$). Damit ist die Gesamtzahl der Elemente von $R \cup S$ durch $k \cdot l + m \cdot k = (l + m) \cdot k \leq k^2$ beschränkt. Die Wahl $R := A \times A$ und $S := \emptyset$ zeigt, daß $R \cup S$ andererseits k^2 Elemente haben kann.

3.21 Sei M eine Menge, seien Funktionen $f_i : M \to A_i$ für $1 \leq i \leq n$ vorgegeben. Definieren wir $g(m) := \langle f_1(m), \ldots, f_n(m) \rangle$, so hat $g : M \to A_1 \times \ldots A_n$ die angegebene Eigenschaft: Für jedes $i \in \{1, \ldots, n\}$ haben die Funktionen f_i und $g \circ \pi_i$ denselben Definitionsbereich M und bilden m auf $f_i(m)$ ab. Damit gilt $f_i = g \circ \pi_i$ nach Lemma 3.39. Es bleibt die Eindeutigkeit dieser Abbildung zu verifizieren. Hat $g' : M \to \Pi_{i=1}^n A_i$ die angegebenen Eigenschaften und ist $m \in M$, so muß $f_i(m) = (g \circ \pi_i)(m) = \pi_i(g(m))$ die i-te Komponente von $g(m)$ sein $(1 \leq i \leq n)$. Damit gilt $g'(m) = \langle f_1(m), \ldots, f_n(m) \rangle = g(m)$. Nach Lemma 3.39 folgt $g = g'$.

3.23 $f(\emptyset) = \emptyset$, $f(\{1, 2\}) = \{2, 4\}$, $f^{-1}(\{2, 6, 8\}) = \{1, 3\}$, $f\lceil\{1, 3\} = \{\langle 1, 2 \rangle, \langle 3, 6 \rangle\}$, $f\lceil\{1, 3\}^{-1} = \{\langle 2, 1 \rangle, \langle 6, 3 \rangle\}$, $(f\lceil\{1, 3\})^{-1}(\{2, 4, 8\}) = \{1\}$.

3.25 Es sei $A := \{x_1, x_2\}$, $B := \{y_1, y_2\}$, die Abbildung $f : A \to B$ bilde x_1 und x_2 auf y_1 ab. Mit $X := \{x_1\}$ ist X eine echte Teilmenge von $f^{-1}(f(X)) = \{x_1, x_2\}$. Mit $Y := \{y_1, y_2\}$ ist $f(f^{-1}(Y)) = \{y_1\}$ eine echte Teilmenge von Y.

3.26 Es seien $f : A \to B$ und $g : B \to A$ Funktionen, es gelte $f \circ g = Id_A$ und $g \circ f = Id_B$. Wir zeigen zunächst, daß f injektiv ist: es seien $a, a' \in A$ und $f(a) = f(a')$. Dann folgt $a = (f \circ g)(a) = g(f(a)) = g(f(a')) = (f \circ g)(a') = a'$. Damit ist f injektiv. Wir zeigen nun, daß f surjektiv ist: mit $b \in B$ gilt $b = (g \circ f)(b) = f(g(b))$, also tritt b als Bild unter f auf. Somit ist f surjektiv. Zusammen folgt, daß f bijektiv ist. Symmetrisch folgt, daß auch g bijektiv ist. Wir zeigen $f \subseteq g^{-1}$: ist $\langle a, b \rangle \in f$, so gilt $f(a) = b = f(g(b))$. Da f injektiv ist, folgt $a = g(b)$, $\langle b, a \rangle \in g$ und $\langle a, b \rangle \in g^{-1}$. Wir zeigen $g^{-1} \subseteq f$: ist $\langle a, b \rangle \in g^{-1}$, so ist $\langle b, a \rangle \in g$ und $g(b) = a = g(f(a))$. Da g injektiv ist, folgt $b = f(a)$, also $\langle a, b \rangle \in f$. Zusammen folgt also $f = g^{-1}$. Symmetrisch folgt $g = f^{-1}$.

3.27 Es sei $n \geq 2$ die Zahl der Personen in P. Die Abbildung $f : P \to \{0, 1, \ldots, n-1\}$ ordne jeder Person die Zahl von anderen Personen aus P zu, die diese kennt. Falls es eine Person a gibt mit $f(a) = n - 1$, so kann 0 nicht als Wert unter f auftreten. Andernfalls tritt $n - 1$ nicht als Wert unter f auf. Damit hat $f(P)$ höchstens $n - 1$ Elemente und f kann nicht injektiv sein.

3.29 Sei $f : A \to B$. Falls f injektiv ist und $\langle b, a \rangle$, $\langle b, a' \rangle$ zu $f^{-1} \subseteq B \times A$ gehören, so gilt $\langle a, b \rangle, \langle a', b \rangle \in f$ und damit $f(a) = b = f(a')$. Aufgrund der Injektivität von f gilt dann $a = a'$. Dies zeigt, daß f^{-1} eine Funktion ist. Gilt $f^{-1}(b) = f^{-1}(b') =: a$, so folgt $f(a) = b$ und $f(a) = b'$. Da f Funktion ist, muß dann $b = b'$ gelten. Dies zeigt, daß f^{-1} injektiv ist. Ist andererseits f^{-1} eine Funktion und gilt $f(a) = f(a') =: b$, so folgt $\langle a, b \rangle, \langle a', b \rangle \in f$, also $\langle b, a \rangle, \langle b, a' \rangle \in f^{-1}$. Da f^{-1} Funktion ist, folgt $a = a'$. Dies zeigt, daß f injektiv ist, falls f^{-1} Funktion ist. Es sei nun f injektiv und surjektiv. Dann gilt offenkundig $f^{-1} : B \to A$. Für $a \in A$ folgt aus $\langle a, f(a) \rangle \in f$ sofort $\langle f(a), a \rangle \in f^{-1}$, also $f^{-1}(f(a)) = a$. Somit gilt $f \circ f^{-1} = Id_A$ und f^{-1} ist surjektiv. Da f^{-1} stets injektiv ist, ist f^{-1} in diesem Fall eine Bijektion. Ist $b \in B$ so gilt $\langle b, f^{-1}(b) \rangle \in f^{-1}$, $\langle f^{-1}(b), b \rangle \in f$, damit $f(f^{-1}(b)) = b$, also ist $f^{-1} \circ f = Id_B$.

3.32 Wir hatten geordnete Paare und kartesische Produkte eingeführt, um den Funktionsbegriff überhaupt erst definieren zu können.

Lösungen zu Aufgaben aus Kapitel 4

4.1 Links oben: reflexiv, transitiv, symmetisch. Rechts oben: reflexiv, transitiv aber nicht symmetrisch. Links unten: nicht reflexiv, symmetrisch, nicht transitiv. Rechts unten: reflexiv, transitiv, symmetrisch.

4.3 Sei $f : A \to A$ nilpotent, d.h. $f \circ f = Id_A$. Mit $\langle a, b \rangle \in f$ gilt $f(a) = b$. Aus $f(f(a)) = a$ ergibt sich $f(b) = a$, also $\langle b, a \rangle \in f$. Dies zeigt, daß f symmetrisch ist. Es sei nun umgekehrt f symmetrisch, $a \in A$ und $b := f(a)$. Dann folgt $\langle a, b \rangle \in f$, wegen der Symmetrie $\langle b, a \rangle \in f$ und somit $f(f(a)) = f(b) = a$. Also gilt $f \circ f = Id_A$.

Sei $f : A \to A$ idempotent, d.h. $f \circ f = f$. Für alle $a, b, c \in A$ gilt nun: Aus $\langle a, b \rangle \in f$ und $\langle b, c \rangle \in f$ folgt $c = f(f(a) = f(a) = b$, somit $\langle a, c \rangle \in f$. Dies zeigt, daß f transitiv ist. Es sei nun umgekehrt f transitiv, $a \in A$ und $b := f(a)$ sowie $c := f(b)$. Dann gilt $\langle a, b \rangle \in f$ und $\langle b, c \rangle \in f$, wegen der Transitivität also $\langle a, c \rangle \in f$ oder $f(a) = c$. Da f Funktion ist, muß nun $b = c$ bzw. $f(f(a) = f(a)$ gelten. Da a beliebig war folgt $f \circ f = f$.

Um die verbleibende Äquivalenz zu verifizieren, verwenden wir einen Ringschluß.

Aus Eigenschaft 3 folgt Eigenschaft 2: Offenkundig gilt $Id_A \circ Id_A = Id_A$. Daher ist Id_A sowohl idempotent als auch nilpotent.

Aus Eigenschaft 2 folgt Eigenschaft 1: Ist f idempotent und nilpotent, so gilt $Id_A = f \circ f = f$. Offenkundig ist Id_A reflexiv.

Aus Eigenschaft 1 folgt Eigenschaft 3: Ist f reflexiv, so ist $\langle a, a \rangle \in f$ und

$f(a) = a$ für jedes $a \in A$. Folglich gilt $f = Id_A$.

4.5 Es seien R und S zwei zweistellige Relationen auf A. Sind R und S *reflexiv* und ist $a \in A$, so gilt $\langle a, a\rangle \in R \subseteq R \cup S$, damit ist $R \cup S$ reflexiv. Dies gilt bereits, wenn eine der beiden Relationen R und S reflexiv ist. Sind R und S *symmetrisch* und ist $\langle a, b\rangle \in R \cup S$, so folgt $R(a, b)$ oder $S(a, b)$. Da R und S symmetrisch sind, folgt $R(b, a)$ oder $S(b, a)$, also $\langle b, a\rangle \in R \cup S$. Somit ist $R \cup S$ symmetrisch. Hingegen ist die Vereinigung zweier *antisymmetrischer* (oder *transitiver*) Relationen auf einer Menge A nicht notwendig wieder antisymmetrisch (resp. transitiv). Als Gegenbeispiel betrachten wir $A : \{1, 2\}$, $R := \{\langle 1, 2\rangle\}$, $S := \{\langle 2, 1\rangle\}$. R und S sind antisymmetrisch und transitiv, $R \cup S$ ist weder antisymmetrisch noch transitiv.

4.8 Sei A eine Menge. Offenkundig ist Id_A reflexiv. Aus $Id_A(a, b)$ folgt $a = b$ und $Id_A(b, a)$. Also ist Id_A symmetrisch und antisymmetrisch. Aus $Id_A(a, b)$ und $Id_A(b, c)$ folgt $a = b = c$ und $Id_A(a, c)$, also ist Id_A transitiv. Ist $R \subseteq A \times A$ reflexiv, symmetrisch, antisymmetrisch und transitiv, so folgt wegen der Reflexivität sofort $Id_A \subseteq R$. Enthielte R ein Paar $R(c, d)$ mit $c \neq d$, so folgte aus der Symmetrie $R(d, c)$, damit wäre R nicht antisymmetrisch, also gilt $R = Id_A$.

4.9 Sei zunächst $R \subseteq A \times A$ reflexiv und euklidisch. Seien $a, b, c \in A$ beliebig. Gilt $R(a, b)$, so folgt wegen der Reflexivität von R auch $R(a, a)$, wegen der der Euklidizität dann $R(b, a)$. Somit ist R symmetrisch. Gilt $R(a, b)$ und $R(b, c)$, so muß zunächst $R(b, a)$ gelten, wie wir gesehen haben. Aus $R(b, a)$ und $R(b, c)$ folgt nun $R(a, c)$. Damit ist R transitiv. Sei nun $R \subseteq A \times A$ reflexiv und anti-euklidisch. Gilt $R(a, b)$, so folgt wegen der Reflexivität von R auch $R(b, b)$. Aus $R(b, b)$ und $R(a, b)$ folgt mit Anti-Euklidizität $R(b, a)$. Somit ist R symmetrisch. Gilt $R(a, b)$ und $R(b, c)$, so muß zunächst $R(c, b)$ gelten, wie wir gesehen haben. Aus $R(a, b)$ und $R(c, b)$ folgt wegen der Anti-Euklidizität $R(a, c)$. Damit ist R transitiv.

4.10 $R := \{\langle a, b\rangle, \langle b, a\rangle\}$. Offenkundig ist R nicht reflexiv, nicht transitiv und nicht antisymmetrisch. Es gilt $R(a, b)$ und $R(a, b)$, jedoch weder $R(b, b)$ noch $R(a, a)$. Somit ist R nicht euklidisch und nicht anti-euklidisch.

4.11 $R := \{\langle a, b\rangle, \langle a, c\rangle, \langle c, a\rangle, \langle c, c\rangle\}$.

4.14 Es seien A, A^*, „$\circ$", „$\sim_L^r$", „$\sim_L^l$" und „$\sim_L$" wie in der Aufgabe angegeben.
1. Wir zeigen, daß „$\sim_L^r$" eine Äquivalenzrelation auf A^* ist (der Beweis, daß auch „$\sim_L^l$" und „$\sim_L$" Äquivalenzrelationen sind, ist ähnlich). *Reflexivität.* Sei $v \in A^*$. Trivialerweise gilt $\forall w \in A^*\colon v \circ w \in L \Leftrightarrow v \circ w \in L)$. *Symmetrie.* Gilt per definitionem, da „$\Leftrightarrow$" symmetrisch ist. *Transitivität.* Seien $v_1, v_2, v_3 \in A^*$. Aus $\forall w \in A^*\colon v_1 \circ w \in L \Leftrightarrow v_2 \circ w \in L$ und $\forall w \in A^*\colon v_2 \circ w \in L \Leftrightarrow v_3 \circ w \in L$ folgt offenkundig $\forall w \in A^*\colon v_1 \circ w \in L \Leftrightarrow v_3 \circ w \in L$.
2. Es gelte $v \sim_L^r v'$, d.h. $\forall w \in A^*\colon v \circ w \in L \Leftrightarrow v' \circ w \in L$. Für alle $w' \in A^*$ gilt dann $v \circ (w \circ w') = (v \circ w) \circ w' \in L \Leftrightarrow v' \circ (w \circ w') = (v' \circ w) \circ w' \in L$. Damit folgt $v \circ w \sim_L^r v' \circ w$.
3. Ähnlich folgen Teile 3 und 4.

5. Es gelte $v_1 \sim_L v_1'$ und $v_2 \sim_L v_2'$. Seien $u, w \in A^*$ gegeben. Dann folgt

$$\begin{aligned} u \circ (v_1 \circ v_2) \circ w \in L &\Leftrightarrow u \circ v_1 \circ (v_2 \circ w) \in L \\ &\Leftrightarrow u \circ v_1' \circ (v_2 \circ w) \in L \\ &\Leftrightarrow (u \circ v_1') \circ v_2 \circ w \in L \\ &\Leftrightarrow (u \circ v_1') \circ v_2' \circ w \in L \\ &\Leftrightarrow u \circ (v_1' \circ v_2') \circ w \in L. \end{aligned}$$

Dies zeigt $v_1 \circ v_2 \sim_L v_1' \circ v_2'$.

4.16 Es sei für jeden Index i der nichtleeren Indexmenge I stets R_i eine Äquivalenzrelation auf der Menge A. Sei $a \in A$. Dann gilt für alle $b \in A$

$$\begin{aligned} b \in [a]_{\bigcap_{i \in I} R_i} &\Leftrightarrow (\bigcap_{i \in I} R_i)(a, b) \\ &\Leftrightarrow R_i(a, b) \text{ für alle } i \in I \\ &\Leftrightarrow b \in [a]_{R_i} \text{ für alle } i \in I \\ &\Leftrightarrow b \in \bigcap_{i \in I} [a]_{R_i}. \end{aligned}$$

Es folgt $[a]_{\bigcap_{i \in I} R_i} = \bigcap_{i \in I} [a]_{R_i}$.

4.20 Die direkten Nachfolger einer Zahl n haben die Form $n \cdot p$ wo p Primzahl ist. Die direkten Vorgänger einer Zahl n erhält man, indem man n durch eine Primzahl p teilt, die ihrerseits n teilt.

4.23 Offenkundig folgt aus der Gleichheit der linearen Ordnungen „$\leq_1$“ und „$\leq_2$“, daß der Durchschnitt wieder eine lineare Ordnung ist. Es sei nun umgekehrt „$\leq_{1,2}$“ eine lineare Ordnung. Seien a, b zwei verschiedene Elemente aus M. Gilt $a \leq_1 b$, so folgt $a \leq_{1,2} b$: ansonsten würde aufgrund der Linearität $b \leq_{1,2} a$ gelten, damit aber auch $b \leq_1 a$, was der Antisymmetrie von „$\leq_1$“ widersprechen würde. Aus $a \leq_{1,2} b$ folgt aber $a \leq_2 b$. Damit ist „$\leq_1$“ Teilrelation von „$\leq_2$“. Symmetrisch folgt, daß auch „$\leq_2$“ Teilrelation von „$\leq_1$“ ist.

4.25 Es sei $H\colon \mathcal{P}(A \times A) \to \mathcal{P}(A \times A)$ die Abbildung, die einer binären Relation R auf A ihre reflexive Hülle $H(R) := \bigcap\{S \subseteq A \times A \mid R \subseteq S, S \text{ reflexiv}\}$ zuordnet. Offenkundig gilt $R \subseteq H(R)$, also hat H die Erweiterungseigenschaft. Da $A \times A$ eine R umfassende reflexive Relation ist, ist der Durchschnitt nichtleer. Nach Lemma 4.4 Teil 1 ist $H(R)$ reflexiv. Daraus folgt sofort, daß $H(H(R)) = H(R)$ gilt. Also ist H idempotent. Gilt $R_1 \subseteq R_2$, so enthält jede R_2 umfassende reflexive Relation auch R_1. Es folgt

$$\bigcap\{S \subseteq A \times A \mid R_1 \subseteq S, S \text{ reflexiv}\} \subseteq \bigcap\{S \subseteq A \times A \mid R_2 \subseteq S, S \text{ reflexiv}\}$$

und H hat die Monotonie-Eigenschaft. Also ist H eine Hüllenabbildung. Analog zeigt man, daß die Abbildungen, die einer Relation ihre symmetrische, transitive oder reflexiv-transitive Hülle zuweisen, Hüllenabbildungen sind.

4.30 Für jedes Paar $\langle a, b\rangle$ von R^+ existieren Elemente a' und b' mit $R(a, a')$ und $R(b', b)$. Nach Voraussetzung kann es damit maximal n verschiedene Elemente der Form a bzw. b geben. Damit ist die Zahl der Paare von R^+ durch n^2 beschränkt.

4.34 Mit $R = \{\langle n, n+1\rangle \mid n \in \mathbb{N}\}$ gilt $R^* = \mathbb{N} \times \mathbb{N}$.

4.36 Man zeigt zunächst mittels einer Induktion über die Zahl der Elemente von M, daß für je zwei verschiedene Elemente m, n mit $m \leq n$ stets eine Kette der Form $m = m_0 \leq m_1 \leq \cdots \leq m_k = n$ existiert ($k \geq 1$), wo m_{i+1} ein direkter Nachfolger von m_i ist ($0 \leq i < k$). Nach Lemma 3.29 gilt in dieser Situation $\langle m, n\rangle \in R^k$. Da $R^* = \bigcup_{n \geq 0} R^n$ (vgl. Bem. 4.43) folgt $\langle m, n\rangle \in R^*$. Umgekehrt folgt aus $\langle m, n\rangle \in R^* = \bigcup_{n \geq 0} R^n$ sofort $m \leq n$, da „$\leq$" reflexiv und transitiv ist. Also stimmen R^* und „$\leq$" überein. Da R kein Paar der Form $\langle a, a\rangle$ enthält, stimmen dann auch „$<$" und R^+ überein. Bei partiellen Ordnungen auf unendlichen Mengen stimmen die genannten Relationen im allgemeinen nicht überein. Wenn wir eine lineare Ordnung betrachten, bei der zwei Kopien der ganzen Zahlen so hintereinandergelegt werden, daß alle Zahlen der ersten Kopie kleiner sind als alle Zahlen der zweiten Kopie, so sind in R^* bzw. R^+ nur Paare, wo beide Komponenten aus derselben Kopie kommen.

Lösungen zu Aufgaben aus Kapitel 5

5.1 Gilt $|A| \leq |B|$ und $|B| \leq |C|$, so gibt es injektive Funktionen $f\colon A \to B$ und $g\colon B \to C$. Dann ist nach Lemma 3.56 auch $f \circ g\colon A \to C$ injektiv. Also gilt $|A| \leq |C|$.

5.2 Jede rationale Zahl ist als Bruch m/n darstellbar, wo m eine ganze Zahl ist und $n \in \mathbb{N}$, $n \neq 0$. In der nachfolgenden unendlichen Tabelle tritt daher jede rationale Zahl zumindest einmal auf (sogar unendlich oft, dies spielt jedoch nachfolgend keine Rolle!).

...	$\frac{-6}{1}$	$\frac{-5}{1}$	$\frac{-4}{1}$	$\frac{-3}{1}$	$\frac{-2}{1}$	$\frac{-1}{1}$	$\boxed{\frac{0}{1}}$	$\frac{1}{1}$	$\frac{2}{1}$	$\frac{3}{1}$	$\frac{4}{1}$	$\frac{5}{1}$	$\frac{6}{1}$	...
...	$\frac{-6}{2}$	$\frac{-5}{2}$	$\frac{-4}{2}$	$\frac{-3}{2}$	$\frac{-2}{2}$	$\frac{-1}{2}$	$\frac{0}{2}$	$\frac{1}{2}$	$\frac{2}{2}$	$\frac{3}{2}$	$\frac{4}{2}$	$\frac{5}{2}$	$\frac{6}{2}$	...
...	$\frac{-6}{3}$	$\frac{-5}{3}$	$\frac{-4}{3}$	$\frac{-3}{3}$	$\frac{-2}{3}$	$\frac{-1}{3}$	$\frac{0}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{3}{3}$	$\frac{4}{3}$	$\frac{5}{3}$	$\frac{6}{3}$	...
...	$\frac{-6}{4}$	$\frac{-5}{4}$	$\frac{-4}{4}$	$\frac{-3}{4}$	$\frac{-2}{4}$	$\frac{-1}{4}$	$\frac{0}{4}$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	$\frac{4}{4}$	$\frac{5}{4}$	$\frac{6}{4}$	...
...	$\frac{-6}{5}$	$\frac{-5}{5}$	$\frac{-4}{5}$	$\frac{-3}{5}$	$\frac{-2}{5}$	$\frac{-1}{5}$	$\frac{0}{5}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{5}{5}$	$\frac{6}{5}$	...
...	$\frac{-6}{6}$	$\frac{-5}{6}$	$\frac{-4}{6}$	$\frac{-3}{6}$	$\frac{-2}{6}$	$\frac{-1}{6}$	$\frac{0}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$	...
	...						...						...	

Wir können nun nacheinander alle Tabelleneinträge besuchen, indem wir zunächst den Eintrag des innersten Rechtecks, danach die Einträge des nächstgrößeren Rechtecks durchlaufen, etc. Wir nummerieren beim Besuch die Zahlen mit natürlichen Zahlen $0, 1, 2, \ldots$ durch, wobei wir eine rationale Zahl, die schon eine Nummer erhalten hat, jeweils bei der Nummerierung überspringen. Offenkundig erhalten wir dadurch eine Bijektion zwischen $\mathbb{N}$ und der Menge der rationalen Zahlen.

Lösungen zu Aufgaben aus Kapitel 6

6.1 Offenkundig ist die Konkatenation von Wörtern assoziativ. Sind L_1, L_2, L_3 Sprachen über dem Alphabet A, so gilt damit

$$\begin{aligned}(L_1 \odot L_2) \circ L_3 &= \{u \circ v \mid u \in L_1, v \in L_2\} \circ L_3\\ &= \{(u \circ v) \circ w \mid u \in L_1, v \in L_2, w \in L_3\}\\ &= \{u \circ (v \circ w) \mid u \in L_1, v \in L_2, w \in L_3\}\\ &= L_1 \circ \{v \circ w \mid v \in L_2, w \in L_3\}\\ &= L_1 \circ (L_2 \circ L_3)\end{aligned}$$

6.3 Es bildet $\mathbb{N} \setminus \{0\}$ zusammen mit der Addition eine Halbgruppe ohne Neutralelement.

6.4 Die Multiplikationstabelle hat die Form

$\cdot$	1	d	dd	k	kd	kdd
1	1	d	dd	k	kd	kdd
d	d	dd	1	kdd	k	kd
dd	dd	1	d	kd	kdd	k
k	k	kd	kdd	1	d	dd
kd	kd	kdd	k	dd	1	d
kdd	kdd	k	kd	d	dd	1

Das Inverse eines Elements x ist das eindeutig bestimmte Element y wo $x \cdot y = 1$. Es ist leicht aus der Tabelle abzulesen.

6.6 Es sei 0 das Nullelement das Rings R. Wir setzen $0_F \colon M \to R; m \mapsto 0$. Für $f \in F$ definieren wir $-_F(f)$ als die Abbildung $m \mapsto -f(m)$. Wir zeigen zunächst, daß $\langle F, +_F, O_F, -_F \rangle$ eine kommutative Gruppe ist. Die Assoziativität von „$+_F$“ ergibt sich unter Verwendung der Assoziativität von „$+$“ wie folgt: für beliebige $f, g, h \in F$ und $m \in M$ gilt

$$\begin{aligned}((f +_F g) +_F h)(m) &= (f +_F g)(m) + h(m) = (f(m) + g(m)) + h(m)\\ &= f(m) + (g(m) + h(m)) = f(m) + (g +_F h)(m)\\ &= (f +_F (g +_F h))(m)\end{aligned}$$

Da die Abbildungen $(f +_F g) +_F h$ und $f +_F (g +_F h)$ Definitionbereich M haben und stets dieselben Werte annehmen, sind sie nach dem Extensionalitätsprinzip für Funktionen identisch. Die Kommutativität von „$+_F$“ folgt nach demselben Muster sofort aus der Kommutativität von „$+$“. Aus

$$(f +_F 0_F)(m) = f(m) + 0 = f(m) = 0 + f(m) = (0_F +_F f)(m)$$

folgt, daß 0_F ein Neutralelement bezüglich „$+_F$“ ist. Weiter gilt stets

$$(f +_F (-_F f))(m) = f(m) + (-f(m)) = 0 = 0_F(m) = ((-_F f) + f)(m),$$

was zeigt, daß $-_F$ die Regeln der Inversenbildung in Gruppen erfüllt.

In völlig analoger Weise folgt, daß $\langle F, \cdot_F \rangle$ eine Halbgruppe ist. Es bleibt zu zeigen, daß die Distributivgesetze gelten. Da in $\mathcal{R}$ die Distributivgesetze gelten, ist stets

$$\begin{aligned} (f \cdot_F (g +_F h))(m) &= f(m) \cdot ((g +_F h)(m)) = f(m) \cdot (g(m) + h(m)) \\ &= (f(m) \cdot g(m)) + (f(m) \cdot h(m)) = \dots \\ &= ((f \cdot_F g) +_F (f \cdot_F h))(m) \end{aligned}$$

Hieraus folgt eines der beiden Distributivgesetze. Das andere ergibt sich symmetrisch.

6.9 Die wir in $\mathcal{A}$ keine ausgezeichneten Elemente oder Funktionen haben, ist nach Lemma 6.27 jede nichtleere Teilmenge von A der Grundbereich einer geeigneten Teilstruktur von $\mathcal{A}$. Es gibt demnach $2^4 - 1 = 15$ Teilstrukturen. Die jeweiligen Interpretationen der Relationssymbole *liebt* und *Geiger* ergeben sich durch Einschränkung von $liebt_{\mathcal{A}}$ und $Geiger_{\mathcal{A}}$ auf die betreffende Teilmenge.

6.11 Da $B_k \subseteq A$ für alle $k \in K$, gilt $B \subseteq A$. Also ist B eine nichtleere Teilmenge von A. Wir verifizieren zunächst, daß $\mathcal{B}$ Bedingung (ii) aus Definition 6.25 erfüllt. Sei $R \in \Sigma_{\mathcal{R}}$ mit Stelligkeit m vorgegegen und $b_1, \dots, b_m \in B$. Aus $R_{\mathcal{A}}(b_1, \dots, b_m)$ bzw. $\langle b_1, \dots, b_m \rangle \in I(R)$ folgt für $k \in K$ nach Definition 6.25 stets $\langle b_1, \dots, b_m \rangle \in I_k(R)$, da nach Voraussetzung $\mathcal{B}_k$ Teilstruktur von $\mathcal{A}$ ist. Daher gilt $\langle b_1, \dots, b_m \rangle \in J(R)$. Gilt umgekehrt $\langle b_1, \dots, b_m \rangle \in J(R)$, so wählen wir einen Index $k \in K$. Es ist $\langle b_1, \dots, b_m \rangle \in I_k(R)$ nach Definition von J. Da $\mathcal{B}_k$ Teilstruktur von $\mathcal{A}$ ist, folgt $R_{\mathcal{A}}(b_1, \dots, b_m)$. Damit ist Bedingung (ii) nachgewiesen. Es sei nun $f \in \Sigma_{\mathcal{F}}$ ein Funktionssymbol der Stelligkeit n und $b_1, \dots, b_n \in B$. Wir setzen $b := f_{\mathcal{A}}(b_1, \dots, b_n)$. Für jedes $k \in K$ ist nun $b \in B_k$. Dies folgt nach Bedingung (iii) aus Definition 6.25, da ja $\mathcal{B}_K$ Teilstruktur von $\mathcal{A}$ ist. Also ist $b \in B$. Da Funktionen spezielle Relationen sind, folgt aufgrund des oben Bewiesenen aus $\langle b_1, \dots, b_n, b \rangle \in I(f)$ nun $\langle b_1, \dots, b_n, b \rangle \in J(f)$. Damit ist Eigenschaft (iii) aus Definition 6.25 für $\mathcal{B}$ nachgewiesen. Es folgt unmittelbar aus der Definition von J, daß auch Eigenschaft (iv) gilt.

6.12 Die Mengen der geraden Zahlen mit der Addition bildet eine Teilstruktur $\mathcal{B}_1$ von $\langle \mathbb{N}, + \rangle$, ebenso die Teilmenge der durch 3 teilbaren Zahlen. Die Vereinigung ist aber keine Teilstruktur, da sie nicht unter Addition abgeschlossen ist.

6.15 Wir nehmen auf die Beschreibung des Erzeugnisses am Ende von Abschnitt 6.3 Bezug. Die Basisregel besagt, daß neben 4 und 6 auch 0 zum Erzeugnis von $\{4,6\}$ gehört. Wenn wir durch interierte Anwendung der induktiven Regeln die Menge $\{0.4.6\}$ unter Inversenbildung und Addition abschließen, erhalten wir die Menge der durch 2 teilbaren ganzen Zahlen.

6.16 Es kommen die vier Abbildungen

$$F_1 := \begin{bmatrix} 0 & \mapsto 0 \\ 1 & \mapsto 0 \end{bmatrix},\ F_2 := \begin{bmatrix} 0 & \mapsto 1 \\ 1 & \mapsto 1 \end{bmatrix},\ F_3 := \begin{bmatrix} 0 & \mapsto 0 \\ 1 & \mapsto 1 \end{bmatrix},\ F_4 := \begin{bmatrix} 0 & \mapsto 1 \\ 1 & \mapsto 0 \end{bmatrix}$$

in Frage. Da

$$\begin{aligned}
F_1(f_\wedge(0,0)) &= F_1(0) = 0 = f_\wedge(0,0) = f_\wedge(F_1(0), F_1(0)) \\
F_1(f_\wedge(0,1)) &= F_1(0) = 0 = f_\wedge(0,0) = f_\wedge(F_1(0), F_1(1)) \\
F_1(f_\wedge(1,0)) &= F_1(0) = 0 = f_\wedge(0,0) = f_\wedge(F_1(1), F_1(0)) \\
F_1(f_\wedge(1,1)) &= F_1(1) = 0 = f_\wedge(0,0) = f_\wedge(F_1(1), F_1(1))
\end{aligned}$$

ist F_1 ein Endomorphismus. Völlig analog folgt, daß F_2 und F_3 Homomorphismen sind. Hingegen zeigt $F_4(f_\wedge(0,1)) = F_4(0) = 1 \neq 0 = f_\wedge(1,0) = f_\wedge(F_4(0), F_4(1))$, daß F_4 kein Homomorphismus ist.

6.18 Wir beschränken uns darauf, die Homomorphieeigenschaft bezüglich der ausgezeichneten Funktionen nachzuweisen. Sei $f \in \Sigma_{\mathcal{F}}$ ein n-stelliges Funktionssymbol und $a_1, \ldots, a_n \in A$. Unter Ausnützung der Homomorphie von h_1 und h_2 erhalten wir wie gewünscht

$$\begin{aligned}
(h_1 \circ h_2)(f_{\mathcal{A}}(a_1, \ldots, a_n)) &= h_2(h_1(f_{\mathcal{A}}(a_1, \ldots, a_n))) \\
&= h_2(f_{\mathcal{B}}(h_1(a_1), \ldots, h_1(a_n))) \\
&= f_{\mathcal{C}}(h_2(h_1(a_1)), \ldots, h_2(h_1(a_n))) \\
&= f_{\mathcal{C}}((h_1 \circ h_2)(a_1), \ldots, (h_1 \circ h_2)(a_n)).
\end{aligned}$$

6.20 Angenommen h ist ein Epimorphismus (also ein surjektiver Homomorphismus) von $\langle \mathbb{N}, \leq \rangle$ auf $\langle \mathcal{Z}, \leq \rangle$. Dann muß die Zahl $z := h(0) - 1$ als Bild unter h auftreten. Es muß also ein $n > 0$ geben mit $h(n) = z$. Nun gilt $0 \leq n$, aber nicht $h(0) \leq h(n)$. Dies liefert einen Widerspruch.

6.24 Wir zeigen Teil 1: Da h ein Automorphismus ist, ist h bijektiv. Für jede Teilmengen $N \subseteq M$ gilt aufgrund der Homomorphie von h stets $h(N) = h(\emptyset \cup N) = h(\emptyset) \cup h(N)$ und somit $h(\emptyset) \subseteq h(N)$. Aufgrund der Surjektivität von h ist demnach $h(\emptyset)$ in jeder Teilmenge von M enthalten. Es folgt $h(\emptyset) = \emptyset$. Wir zeigen Teil 2: Sei $m \in M$. Wir wissen nach Teil 1 aufgrund der Bijektivität von h, daß $h(\{m\}) \neq \emptyset$. Hätte $h(m)$ zwei verschiedene Elemente n_1 und n_2, so könnten wir $h(m)$ in der Form $\{n_1\} \cup \{n_2\} \cup N$ darstellen. Da nach Lemma 6.43 mit h auch h^{-1}

ein Automorphismus ist, erhalten wir

$$\begin{aligned}\{m\} &= h^{-1}(h(\{m\})) = h^{-1}(\{n_1\} \cup \{n_2\} \cup N) \\ &= h^{-1}(\{n_1\}) \cup h^{-1}(\{n_2\}) \cup h^{-1}(N).\end{aligned}$$

Nach Teil 1 sind aber $h^{-1}(\{n_1\})$ und $h^{-1}(\{n_2\})$ beide nichtleer, da ja h^{-1} bijektiv ist. Es folgt nun $h^{-1}(\{n_1\}) = h^{-1}(\{n_2\}) = \{m\}$. Dies widerspricht der Injektivität von h^{-1}. Also kann $h(m)$ nur einelementig sein.

6.26 Da wir in $\langle \mathbb{N}, +, \cdot \rangle$ keine ausgezeichneten Relationen haben, müssen wir Bedingung 2 aus Definition 6.45 für die Funktionen „+“, und „·“ verifizieren.

„+“: Es gelte $a \sim a'$ und $b \sim b'$. Dann sind $a - a'$ und $b - b'$ ganzzahlige Vielfache von 5. Es folgt, daß auch $(a+b)-(a'+b') = (a-a')+(b-b')$ ein ganzzahliges Vielfaches von 5 ist. Damit folgt wie erwünscht $a + b \sim a' + b'$.

„·“: Es gelte $a \sim a'$ und $b \sim b'$. Damit gibt es ganze Zahlen k und l mit $a = a'+k\cdot 5$ und $b = b' + l \cdot 5$. Es folgt

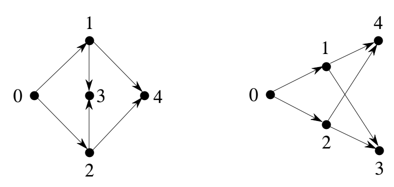
$$\begin{aligned}a \cdot b &= (a' + k \cdot 5) \cdot (b' + l \cdot 5) \\ &= a' \cdot b' + 5 \cdot (a' \cdot l + k \cdot b' + k \cdot l \cdot 5).\end{aligned}$$

Folglich ist $a \cdot b - a' \cdot b'$ ein ganzzahliges Vielfaches von 5 und $a \cdot b \sim a' \cdot b'$.

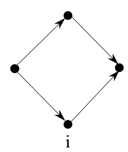
6.31 Die Aussage würde falsch. Es bilden die ganzen Zahlen mit der üblichen Ordnung eine abzählbare diskret geordnete, linear geordnete Menge ohne Endpunkte. Ein weiteres Beispiel einer solchen Menge erhalten wir, wenn wir zwei Kopien der ganzen Zahlen hintereinanderlegen, wobei also jede Zahl aus Kopie 1 kleiner ist als jede Zahl von Kopie 2. In der erstgenannten Struktur (ganze Zahlen) liegen zwischen zwei gegebenen Elementen immer nur endlich viele. Dies gilt in der zweiten Struktur nicht mehr. Hieraus folgt leicht, daß beide Strukturen nicht isomorph sein können.

Lösungen zu Aufgaben aus Kapitel 7

7.1 Bei der hier angedeuteten Benennung der Elemente repräsentiert die identische Abbildung einen Isomorphismus.

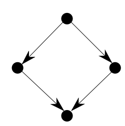


7.2 Es gibt 2^k Pfade der Länge $2k$. Um den Pfad der Länge $2k$ mit dem geringsten Gewicht zu finden, stellen wir für jedes „Karo"



lokal fest, ob der obere Kantenzug oder der untere geringeres Gewicht hat. Wir kombinieren dann die lokal konstruierten Kantenzüge. Das Verfahren kann als eine primitive Form der sogenannten „dynamischen Programmierung" aufgefaßt werden.

7.6 Da in einer Menge M ein Element immer genau einmal auftritt, gibt es keine Menge, deren DAG-Repräsentation die Form



hat.

7.7 Es sei $\mathcal{G} = \langle N, E \rangle$ zusammenhängend. Sind $m, n \in N$ zwei verschiedene Knoten, so gibt es damit einen Pfad von m nach n. Nach Lemma 7.6 gilt dann $\langle m, n \rangle \in E^+$, somit auch $\langle m, n \rangle \in E^*$. Damit gilt $[m]_{E^*} = [n]_{E^*}$ nach Lemma 4.10. Es folgt, daß N/E^* genau ein Element hat. Hat umgekehrt N/E^* genau ein Element, so sind nach Lemma 4.10 je zwei Elemente m, n von N stets bezüglich E^* äquivalent. Sind m und n verschiedene Punkte von N, so gilt demnach $\langle m, n \rangle \in E^+$. Nach Lemma 7.6 gibt es dann einen Pfad von m nach n. Dies zeigt, daß $\mathcal{G}$ zusammenhängend ist.

7.9 Wir zeigen durch Induktion über k, daß die Zahl der Blätter durch n^k beschränkt ist. Für $k = 0$ besteht $\mathcal{T}$ nur aus der Wurzel, die in diesem Fall auch das einzige Blatt ist. Es sei nun $k = m + 1$ für ein $m \geq 0$. Die unmittelbaren Teilbäume haben nach Induktionsvoraussetzung jeweils höchstens n^m Blätter. Es gibt maximal n unmittelbare Teilbäume, da die Wurzel höchstens Verzweigungsgrad n hat. Damit ist die Zahl der Blätter von $\mathcal{T}$ durch $n \cdot n^m = n^k$ beschränkt. Gilt $n = 2$, so kann man die Gesamtzahl aller Knoten durch $1 + 2 + 2^2 + \cdots + 2^k = 2^{k+1} - 1$ abschätzen.

7.10 Es sei f die Abbildung, die jedem von der Wurzel verschiedenen Knoten h die Kante $\langle k, h \rangle$ zuordnet, wo k der nach Lemma 7.28 eindeutig bestimmte Vorgänger von h ist. Die Abbildung ist injektiv, da für $h \neq h'$ die Bildkanten von h und h' verschiedene Endknoten (nämlich h bzw. h') haben. Offenkundig tritt jede Kante $\langle k, h \rangle$ von $\mathcal{T}$ als Bild unter f auf. Daher ist f eine Bijektion zwischen der Menge der von der Wurzel verschiedenen Knoten auf die Menge der Kanten. Mit Lemma 3.50 folgt nun $n = m - 1$.

7.12 Es sei $\mathcal{T}$ ein rationaler Baum. Nach Definition ist $\mathcal{T}$ dann endlich verzweigend. Angenommen, $\mathcal{T}$ hat nicht uniform endlichen Verzweigungsgrad. Dann gibt es eine unendliche Folge von Knoten $n_1, n_2, \ldots$, wo stets n_{i+1} höheren Verzweigungsgrad hat als n_i. Die durch diese Knoten bestimmten Teilbäume $\mathcal{T}_1, \mathcal{T}_2, \ldots$ haben an der Wurzel paarweise unterschiedlichen Verzweigungsgrad. Es ist leicht zu sehen, daß sie damit nicht isomorph sein können. Damit kann aber $\mathcal{T}$ nicht rational sein, wir erhalten einen Widerspruch.

7.15 Ansonsten wäre die Menge der Σ-Bäume leer, da bei einem Σ-Baum die Blätter stets mit Individuenkonstanten gelabelt sind.

7.17 *Basisregel*: Für jede Individuenkonstante $e \in \Sigma_{\mathcal{E}}$ ist stets $\langle \{\langle\ \rangle\}, D \rangle$ ein Σ-Baum, wo D die leere Folge $\langle\ \rangle$ auf e abbildet.
Induktive Regel: Ist $f \in \Sigma_{\mathcal{F}}$ n-stellig, und ist für $1 \leq i \leq n$ jeweils T_i ein Σ-Baum der Form $\langle S_i, D_i \rangle$, so auch $\langle S, D \rangle$ ein Σ-Baum, wo S und D wie im Beweis von Lemma 7.47 erklärt sind.
Es bleibt zu zeigen, daß durch diese induktive Definition gerade die Menge der Σ-Bäume definiert wird. Der Beweis, daß jedes Element der induktiv definierten Menge M ein Σ-Baum ist, erfolgt per Induktion über die Zahl k der Anwendungen der induktiven Regel. Der Fall $k = 0$ ist trivial, da $\{\langle\ \rangle\}$ ein Baumskelett ist. Es ergebe sich nun $T \in M$ durch Anwendung der induktiven Regel auf die Elemente $T_1, \ldots, T_n \in M$. Nach Induktionsvoraussetzung können wir dann $T_i \in B(\Sigma)$ annehmen. Da wir bei der Anwendung der induktiven Regel genau wie im Beweis von Lemma 7.47 vorgehen, zeigt das Lemma, daß auch T zu $B(\Sigma)$ gehört. Umgekehrt kann man durch Induktion über die Baumhöhe mit Hilfe von Lemma 7.48 leicht zeigen, daß sich jeder Σ-Baum mit Hilfe der oben angegebenen Basisregel durch eine endliche Zahl von Anwendungen der induktiven Regel gewinnen läßt. Die Details werden hier dem Leser überlassen.

7.19 *Fall 1. Σ enthält kein Funktionssymbol der Stelligkeit ≥ 1.* Ist n ein Knoten des Σ-Baums T, so muss n mit einer Individuenkonstante gelabelt sein. Daher kann n kein Kind haben. Es besteht T daher nur aus der Wurzel. Dies zeigt, daß die Zahl der Σ-Bäume genau der Zahl der Individuenkonstanten in Σ entspricht und endlich ist.
Fall 2. Σ enthält zumindest ein Funktionssymbol der Stelligkeit ≥ 1. Wir zeigen mittels vollständiger Induktion nach n, daß die Menge B_n aller Σ-Bäume der Höhe n für alle $n \in \mathbb{N}$ endlich und nichtleer ist. Die Behauptung folgt dann mit Aufgabe 5.3.
$n = 0$. Es ist $\{\langle\ \rangle\}$ das einzige Baumskelett der Höhe 0. Es kann mit jeder der endlich vielen Individienkonstanten aus Σ gelabelt werden. Da Σ nach Voraussetzung

zumindest eine Individuenkonstante enthält, folgt die Behauptung.
$n = k+1 \geq 1$. Nach Induktionsvoraussetzung ist für $0 \leq i \leq k$ die Menge B_i endlich und nichtleer. Damit ist auch $\bigcup_{i=1}^{k} B_i$ endlich und nichtleer. Es ist aufgrund der Voraussetzung von Fall 2 trivial zu zeigen, daß es einen Σ-Baum der Höhe n gibt. Für jeden Σ-Baum der Höhe n liegen die unmittelbaren Teilbäume nach Lemma 7.48 stets in $\bigcup_{i=1}^{k} B_i$. Es gibt also unter allen Σ-Bäumen der Höhe n nach Induktionsvoraussetzung nur endlich viele unterschiedliche unmittelbare Teilbäume. Da Σ endlich ist, gibt es nur endlich viele mögliche Wurzeldekorationen. Daher folgt nun aus der Eindeutigkeitsaussage von Lemma 7.47, daß es nur endlich viele Σ-Bäume der Höhe n gibt.

7.22 Ist $T \in B(\Sigma)$ ein Baum der Höhe 0, so stellt er die Interpretation von *Eins* in $T \in \mathcal{B}(\Sigma)$ dar und wird auf die Interpretation der *Eins* in $\langle \mathbb{N}, 1, + \rangle$, also auf 1 abgebildet. Wir nehmen an, daß jeder Baum der Höhe $i \leq k$ auf die Zahl b seiner Blätter abgebildet wird. Ist nun $T \in B(\Sigma)$ ein Baum der Höhe $n = k + 1$, so hat er Wurzeldekoration Add, die unmittelbaren Teilbäume haben nach Lemma 7.48 Höhe $\leq k$. Jeder dieser Teilbäume wird nach Induktionsvoraussetzung unter h auf die Zahl seiner Blätter abgebildet. Da Add in $\langle \mathbb{N}, 1, + \rangle$ als Addition interpretiert wird, wird T auf die Summe der Zahlen der Blätter der unmittelbaren Teilbäume — also auf die Zahl der Blätter von T selbst — abgebildet.

7.23 Da $Id_A \circ Id_A = Id_A$ folgt mit einer trivialen Induktion über die Baumhöhe, daß unter h jeder Baum in $B(\Sigma)$ auf Id_A abgebildet wird.

7.27 Die Terme $f(f(a,x),x)$ und $f(y, f(y,a))$ können nicht unifizierbar sein: wäre σ ein Unifikator, so müßte $\sigma(y) = \sigma(f(a,x)) = f(a, \sigma(x)$ den Term $\sigma(x)$ echt enthalten. Umgekehrt müßte auch $\sigma(x) = \sigma(f(y,a)) = f(\sigma(y), a)$ den Term $\sigma(y)$ echt enthalten. Da Terme endliche Zeichenfolgen sind, ist dies ausgeschlossen.
Die Terme $f(f(a,x),y)$ und $f(z, f(a,z))$ sind unifizierbar. Wenn wir den Robinsonschen Algorithmus schrittweise anwenden, wird zunächst z auf $f(a,x)$, anschließend y auf $f(a, f(a,x))$ abgebildet. Es ist $\langle z/f(a,x), y/f(a, f(a,x)) \rangle$ ein allgemeinster Unifikator.

Lösungen zu Aufgaben aus Kapitel 8

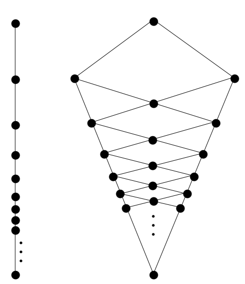
8.1 Alle bis auf das dritte Beispiel sind Verbände. Im dritten Beispiel treten Elemente auf, die zwei unvergleichbare minimale obere Schranken (bzw, maximale untere Schranken) besitzen.

8.2 Verbände mit 4 Elementen gibt es bis auf Isomorphie lediglich zwei: sind alle vier Elemente paarweise bezüglich der Verbandsordnung $\leq$ vergleichbar, ergibt sich eine lineare Ordnung mit vier Elementen. Im anderen Fall gibt es zwei unvergleichbare Elemente a und b. Die beiden anderen Elemente müssen dann $sup\ \{a, b\}$ und

inf $\{a,b\}$ sein, wodurch sich die Form eines Karos ergibt.
Verbände mit 5 Elementen gibt es 6. Sind alle Elemente paarweise bezüglich der Verbandsordnung $\leq$ vergleichbar, ergibt sich eine lineare Ordnung mit fünf Elementen. Im anderen Fall muß es zwei unvergleichbare Elemente a, b geben. Zusammen mit *sup* $\{a,b\}$ und *inf* $\{a,b\}$ ergibt sich ein Karo. Das fünfte Element kann (1) echt oberhalb *sup* $\{a,b\}$, (2) echt zwischen a und *sup* $\{a,b\}$, (3) echt zwischen a und *inf* $\{a,b\}$, (4) echt unterhalb *inf* $\{a,b\}$ liegen. Durch Austausch der Rolle von a und b ergeben sich isomorphe Verbände. Im letzten verbleibenden Fall ist das fünfte Element zu a und b unvergleichbar. Es kann weder die 1 noch die 0 des Verbands darstellen und liegt daher zwischen *inf* $\{a,b\}$ und *sup* $\{a,b\}$.

8.5 Sind A und B zwei (co-) endliche Teilmengen von $\mathbb{N}$, so sind auch $A \cup B$ und $A\cap$ (co-) endlich. Ist A endlich und B co-endlich, so ist $A \cup B$ co-endlich und $A \cap B$ endlich. Demnach ist $\mathcal{K}$ unter Vereinigung und Durchschnitt abgeschlossen. Offenkundig folgt aus $A \subseteq M$ und $B \subseteq M$ stets $A \cup B \subseteq M$, daher ist $A \cup B$ das Supremum von A und $B \in \mathcal{K}$ bezüglich $\subseteq$. Analog folgt, daß $A \cap B = \mathit{inf}\ \{A, B\}$. Das Supremum über alle Elemente von $\mathcal{K}$ der Form $\{0\}, \{2\}, \{4\}, \ldots$ existiert nicht: Dies müßte offenkundig eine co-endliche Menge sein. Durch Wegnehmen einer geeigneten ungeraden Zahl ergäbe sich eine kleinere Menge, die oberhalb aller Elemente von $\mathcal{K}$ liegt. Also ist der Verband nicht vollständig.

8.6 In der folgenden Abbildung sind zwei Beispiele in Form eines Hasse-Diagramms angegeben.



Durch Verallgemeinerung erhält man zahlreiche weitere Beispiele. Hierunter fallen auch die linearen Ordnungen der ganzen, rationalen und reellen Zahlen.

8.8 Aus $a \leq b$ folgt $a \leq \mathit{sup}\ \{b,c\} = b \sqcup c$. Zusammen mit Also ist $b \sqcup c$ eine obere Schranke für $\{a,c\}$ und es folgt $\mathit{sup}\ \{a,c\} = a \sqcup c \leq b \sqcup c$. Symmetrisch zeigt man, daß $a \sqcap c$ untere Schranke für $\{b,c\}$ ist, woraus die zweite Ungleichung folgt.

8.9 Wir verwenden die Bezeichnung $c := \emptyset, a := \{0\}, b := \{1\}, d := \{0,1\}$. Es kann keinen starken Homomorphismus zwischen den Ordnungsstrukturen geben. Beweis: angemommen, h ist ein starker Homomorphismus. Dieser müßte zumindest zwei der Elemente aus $\{a,b,c,d\}$ — wir nennen diese x, y— auf dasselbe Element z

aus $\{0,1\}$ abbilden. Aus $z = h(x) \leq h(y) = z$ und $z = h(y) \leq h(x) = z$ folgte nun $x \subseteq y$ und $y \subseteq x$, also $x = y$, was der Wahl von x, y widerspricht. Die schwachen Homomorphismen können eindeutig durch das Urbild von 0 gekennzeichnet werden. Dieses kann sein: (1) $\emptyset$, (2) $\{c\}$, (3) $\{c, a\}$, (4) $\{c, b\}$ (5) $\{a, b, c, d\}$. Mehr Möglichkeiten gibt es nicht, da mit jedem der Elemente a, b, c, d, das auf 0 abgebildet wird, aufgrund der Homomorphieeigenschaft auch jede Teilmenge auf 0 abgebildet werden muß. Durch Nachrechnen folgt, daß die angegebenen Abbildungen schwache Homomorphismen sind. Die Abbildung, die c, a und b auf 0 abbildet, ist kein algebraischer Homomorphismus, da das Supremum von a und b (also d) nicht auf das Supremum der Bilder (also 0) abgebildet wird. In derselben Weise zeigt man, daß die Abbildung, die nur c auf 0 abbildet, kein ein algebraischer Homomorphismus ist, da das Infimum von a und b (also c) nicht auf das Infimum der Bilder (also 1) abgebildet wird. Die beiden anderen genannten schwachen Homomorphismen sind auch Homomorphismen im algebraischen Sinn, wie man leicht nachrechnen kann.

8.11 Sei $\mathcal{V} = \langle V, \sqcup, \sqcap, \rangle$. Nachzuweisen ist die Äquivalenz der Bedingungen

$$\begin{array}{ll} (1) & \forall a, b, c \in V\colon (a \sqcap b) \sqcup (a \sqcap c) = a \sqcap (b \sqcup (a \sqcap c)) \\ (2) & \forall a, b, c \in V\colon a \leq b \Rightarrow (a \sqcup (b \sqcap c) = b \sqcap (a \sqcup c)) \end{array}$$

„(1) $\to$ (2)“: (1) sei gültig, es seien $a' \leq b'$ und $c' \in V$. Dann folgt

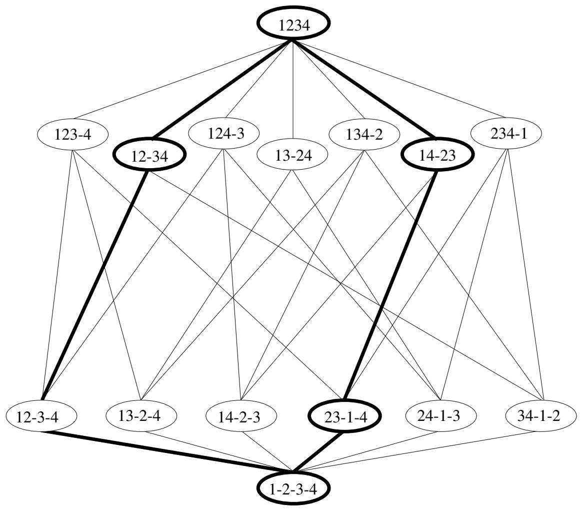
$$\begin{aligned} a' \sqcup (b' \sqcap c') &\overset{(i)}{=} (b' \sqcap a') \sqcup (b' \sqcap c') \\ &\overset{(ii)}{=} (b' \sqcap c') \sqcup (b' \sqcap a') \\ &\overset{(iii)}{=} (b' \sqcap a') \sqcup (b' \sqcap c') \\ &\overset{(iv)}{=} b' \sqcap (c' \sqcup (b' \sqcap a')) \\ &\overset{(v)}{=} b' \sqcap (c' \sqcup a') \\ &\overset{(vi)}{=} b' \sqcap (a' \sqcup c') \end{aligned}$$

Gleichung (i) folgt, da $a' = a' \sqcap b'$. Gleichungen (ii), (vi) gelten wegen der Kommutativität von „$\sqcup$“, Gleichungen (iii),(v) folgen aus $a' = a' \sqcap b' = b' \sqcap a'$, Gleichung (iii) wendet (1) an mit $a := b'$, $b := c'$ und $c := a'$.
„(2) $\to$ (1)“: (2) sei gültig, es seien a', b' und $c' \in V$. Es ist $a' \sqcap c' \leq a'$. Hieraus erhält man

$$\begin{aligned} (a' \sqcap b') \sqcup (a' \sqcap c') &\overset{(i)}{=} (a' \sqcap c') \sqcup (a' \sqcap b') \\ &\overset{(ii)}{=} a' \sqcap ((a' \sqcap c') \sqcup b') \\ &\overset{(ii)}{=} a' \sqcap (b' \sqcup (a' \sqcap c')) \end{aligned}$$

Gleichungen (i), (iii) gelten wegen der Kommutativität von „$\sqcup$“, Gleichung (ii) wendet (2) an mit $a := a' \sqcap c'$, $b := a'$, $c := b'$.

8.16 Am nachfolgenden Hasse-Diagramm läßt sich ablesen, daß der Verband einen Teilverband enthält, der zum Pentagon isomorph ist.



Daher ist der Verband nicht modular. Da jeder distributive Verband modular ist, ist der gegebene Verband auch nicht distributiv.

8.19 Die maximalen eigentlichen Ideale haben die Form $\{N \subseteq M \backslash \{m\}\}$ für $m \in M$. Die maximalen eigentlichen Filter haben die Form $\{N \subseteq M \mid m \in N\}$ für $m \in M$.

8.21 Offenkundig ist das kleinste gemeinsame Präfix zweier Wörter U und V gerade das Infimum über $\{U, V\}$. Sobald des Alphabet zwei Buchstaben enthält, gibt es aber Wörter U und V, wo weder U Präfix von V ist noch umgekehrt. Dann kann es aber kein Wort geben, das U und V als Präfixe hat. Demzufolge existieren für zweielementige Teilmengen i.a. nicht einmal obere Schranken. Hat aber das Alphabet nur ein Symbol, so entspricht die Präfixordnung gerade der $\leq$-Ordnung der Wortlängen. Es ergibt sich ein Verband.

Lösungen zu Aufgaben aus Kapitel 9

9.3 Es sei M eine dreielementige Menge und $\mathbf{P_3} = \langle \mathcal{P}(M), \cup, \cap, - \rangle$. Jede Teilmenge K von M bestimmt ein Ideal, die Menge $\{N \subseteq M \mid N \subseteq K\}$ aller Untermengen von K. Es gibt demzufolge genau 8 Ideale. Nachfolgend verwenden wir die Elementbezeichnungen aus Abbildung 9.1. Dort ist ein Ideal $I = \{0, r, t, z\}$ angeben. Die zugehörige Kongruenzrelation hat die Äquivalenzklassen I und J, die Quotientenstruktur ist isomorph zu $\mathbf{P_1}$, wie in Beispiel 9.21 gezeigt wurde. Symmetrisch ergeben sich zwei weiter Ideale I' mit vier Elementen. Die zugehörigen Kongruenzrelationen haben die Äquivalenzklassen I' und $\mathcal{P}(M) \setminus I'$, die Quotientenstruktur ist in beiden Fällen isomorph zu $\mathbf{P_1}$. Darüberhinaus gibt es drei Ideale mit jeweils zwei Elementen. Wir betrachten $\{0, r\}$. Die Äquivalenzklassen der zugeordneten Kongruenzrelation sind $\{0, r\}, \{t, z\}, \{s, x\}, \{y, 1\}$. Die Quotientenstruktur ist isomorph zu $\mathbf{P_2}$, wobei $\{0, r\}$ das Nullelement, $\{y, 1\}$ die Eins darstellen. Weiter stellt auch $\{0\}$ ein Ideal dar. Die Äquivalenzklassen sind die einelementigen Teilmengen von $\mathbf{P_3}$ und die Quotientenstruktur ist isomorph zu $\mathbf{P_3}$. Schließlich ist auch $\mathcal{P}(M)$ selbst ein (uneigentliches) Ideal. Die Kongruenrelation enthält $\mathcal{P}(M)$ als einzige Äquivalenzklasse, die Quotientenstruktur ist die triviale Boolesche Algebra, die aus einem Element besteht.

9.5 Teil 1: Offenkundig ist I endlich. Da Ideale unter Supremumsbildung über zweielementige Mengen abgeschlossen sind, folgt mit einer trivialen Induktion, daß *sup* I ein Element von I ist. Da es eine obere Schranke der Element von I ist, ist es das größte Element von I. Teil 2: Offenkundig folgt aus $a \equiv_I b$ sofort $a \sim_I b$, for alle $a, b \in B$. Gilt umgekehrt $a \sim_I b$, so existieren nach Definition von "$\sim_I$" Elemente $x, y \in I$ mit $a \sqcup x = b \sqcup y$. Hieraus folgt $a \sqcup \mathit{sup}\ I = a \sqcup x \sqcup \mathit{sup}\ I = b \sqcup y \sqcup \mathit{sup}\ I = b \sqcup \mathit{sup}\ I$, also gilt $a \equiv_I b$. Teil 3: Es sei $b \in B$ und $b' \in [b]_{\sim_I}$. Aus $b \sqcup \mathit{sup}\ I = b' \sqcup \mathit{sup}\ I$ folgt einerseits $(b \sqcap b') \sqcup \mathit{sup}\ I = (b \sqcup \mathit{sup}\ I) \sqcap (b' \sqcup \mathit{sup}\ I) = b \sqcup \mathit{sup}\ I$, also gilt $b \sqcap b' \in [b]_{\sim_I}$. Analog ergibt sich $(b \sqcup b') \sqcup \mathit{sup}\ I = (b \sqcup \mathit{sup}\ I) \sqcup (b' \sqcup \mathit{sup}\ I) = b \sqcup \mathit{sup}\ I$, also gilt $b \sqcup b' \in [b]_{\sim_I}$. Dies zeigt, daß die Kongruenzklassen $[b]_{\sim_I}$ unter Infima und Suprema abgeschlossen sind.

9.7 Die modifizierte Abbildung g wäre im allgemeinen nicht mehr homomorph bezüglich "$\sqcap$", da die Gleichung $g(a) \cap g(b) = g(a \sqcap b)$ ungültig sein kann. Ist $a \sqcap b$ Element des Primideals P, so folgt im allgemeinen nicht $a \in P$ *und* $b \in P$. Beispiel: bezeichnen wir mit $0, u, v, 1$ die Elemente von $\mathbf{P_2}$, so sind $P_u := \{0, u\}$ und $P_v := \{0, v\}$ die einzigen Primideale. Es gilt $g(u) = \{P_u\}, g(v) = \{P_v\}$ also $g(u) \cap g(v) = \emptyset$, jedoch $g(u \sqcap v) = g(0) = \{P_u, P_v\}$.

9.8 Die Atome sind genau die Bitvektoren mit genau einem Eintrag 1. Nachfolgend seien $a = \langle a_1, \dots, a_n \rangle$ und $b = \langle b_1, \dots, b_n \rangle$ beliebige Elemente von $\mathbf{2}^n$. Es ist $a \leq b$ zu $a \sqcap b = a$ äquivalent. Diese Bedingung bedeutet gerade $a_i \sqcap b_i = a_i$ bzw. $a_i \leq b_i$ für $1 \leq i \leq n$. Damit sind die Atome von $\mathbf{2}^n$ genau die Bitvektoren mit genau einem Eintrag 1. Sei $M = \{1, \dots, n\}$ als Menge mit n Elementen gewählt. Dann ist die Abbildung h, die jedem Element $a = \langle a_1, \dots, a_n \rangle$ die Menge $\{i \in \{1, \dots, n\} \mid a_i = 1\}$ zuordnet, ein Isomorphismus von $\mathbf{2}^n$ auf $\mathbf{P_n}$. Es ist trivial zu zeigen, daß h bijektiv ist. Man vergleiche hierzu die Diskussion in Abschnitt 3.6.2. Wir zeigen die Homomorphie bezüglich „$\sqcup$“: $h(a) \cup h(b) = \{i \in \{1, \dots, n\} \mid a_i = 1\} \cup \{i \in \{1, \dots, n\} \mid b_i = 1\} = \{i \in \{1, \dots, n\} \mid a_i \sqcup b_i = 1\} = h(a) \sqcup h(b)$. Analog folgen auch die anderen Homomorphieeigenschaften.

Lösungen zu Aufgaben aus Kapitel 10

10.1 Die Teilformeln von $((A_0 \vee A_1) \wedge A_2) \Rightarrow (A_0 \wedge \neg A_1)$ sind genau die Formeln $((A_0 \vee A_1) \wedge A_2) \Rightarrow (A_0 \wedge \neg A_1)$, $((A_0 \vee A_1) \wedge A_2)$, $(A_0 \wedge \neg A_1)$, $(A_0 \vee A_1)$, $\neg A_1$, A_0, A_1 und A_2.

10.3 Für $k \in \mathbb{N}$ bezeichne $\mathcal{L}_0^{(k)}$ die Menge der aussagenlogischen Formeln des Gewichts k. Es gilt $\mathcal{L}_0^{(0)} = At$. Diese Menge ist nach Voraussetzung abzählbar unendlich. Es sei nun $n \geq 0$, für alle $0 \leq k \leq n$ sei gezeigt, daß $\mathcal{L}_0^{(k)}$ abzählbar unendlich ist. Nach Lemma 5.7 ist $\mathcal{L}_0^{\leq n} := \bigcup_{k=0}^{n} \mathcal{L}_0^{(k)}$ abzählbar unendlich. Es ist offenkundig $\{(\alpha \wedge \beta) \mid w(\alpha) = n, w(\beta) \leq n\}$ mit $\mathcal{L}_0^{(n)} \times \mathcal{L}_0^{\leq n}$ gleichmächtig, damit nach Lemma 5.8 abzählbar unendlich. Dasselbe gilt für $\{(\alpha \wedge \beta) \mid w(\alpha) \leq n, w(\beta) = n\}$. Daraus folgt nun mit Lemma 5.6, daß die Menge der Formeln der Form $(\alpha \wedge \beta)$ aus $\mathcal{L}_0^{(n+1)}$ abzählbar unendlich ist. Analog folgt dies für Formeln aus $\mathcal{L}_0^{(n+1)}$ der Form $(\alpha \vee \beta)$ (resp. $(\alpha \Rightarrow \beta)$, $\neg\alpha$). Hieraus ergibt sich mit Lemma 5.7, daß $\mathcal{L}_0^{(n+1)}$ als Vereinigung der obigen Mengen abzählbar unendlich ist. Wir haben induktiv gezeigt, daß für jedes $n \geq 0$ die Menge $\mathcal{L}_0^{(n)}$ abzählbar unendlich ist. Es ist damit $\mathcal{L}_0 = \bigcup_{n=0}^{\infty} \mathcal{L}_0^{(n)}$ nach Lemma 5.10 abzählbar unendlich.

10.4 Wir verwenden Lemma 10.14. Sei g eine 0-1 Bewertung. Falls $g(\alpha) = 0$ gilt, folgt wegen $g(\alpha \wedge \beta) = 0$ sofort $g(\alpha \Rightarrow (\beta \Rightarrow \gamma)) = 1 = g((\alpha \wedge \beta) \Rightarrow \gamma)$. Wir können daher $g(\alpha) = 1$ annehmen. Falls $g(\beta) = 0$ gilt, folgt nun wegen $g(\alpha \wedge \beta) = 0$ und $g(\beta \Rightarrow \gamma) = 1$ sofort $g(\alpha \Rightarrow (\beta \Rightarrow \gamma)) = 1 = g((\alpha \wedge \beta) \Rightarrow \gamma)$. Daher dürfen wir auch $g(\beta) = 1$ annehmen. Offenkundig erhalten wir dann aber $g(\alpha \Rightarrow (\beta \Rightarrow \gamma)) = g((\alpha \wedge \beta) \Rightarrow \gamma)$, wobei der Wert mit dieser Formeln mit dem von γ übereinstimmt.

10.5 Wir verwenden Lemma 10.14. Die möglichen Klammerungen führen auf die Formeln

$$\begin{aligned}
\alpha &:= A_0 \Rightarrow (A_1 \Rightarrow (A_2 \Rightarrow A_3)), \\
\beta &:= (A_0 \Rightarrow A_1) \Rightarrow (A_2 \Rightarrow A_3), \\
\gamma &:= ((A_0 \Rightarrow A_1) \Rightarrow A_2) \Rightarrow A_3
\end{aligned}$$

Wir zeigen, daß alle drei Formeln paarweise nichtäquivalent sind. Es sei g_1 eine 0-1 Bewertung mit $g_1(A_0) = g_1(A_1) = g_1(A_3) = 0$, $g_1(A_2) = 1$. Dann folgt $g_1(\alpha) = 1$ aber $g_1(\beta) = 0 = g_1(\gamma)$. Somit ist α weder zu β noch zu γ äquivalent. Um noch die Nichtäquivalenz von β und γ zu zeigen, verwendet man analog eine 0-1 Bewertung g_2 mit $g_2(A_1) = g_2(A_2) = g_2(A_3) = 0$ und $g_2(A_0) = 1$. Es ist dann $g_2(\beta) = 1$ und $g_2(\gamma) = 0$.

10.8 Es sind $A_0 \wedge \neg A_0$, $A_0 \vee \neg A_0$, A_0 und $\neg A_0$ vier nichtäquivalente Formeln der angegebenen Art. Aufgrund des Koinzidenzlemmas muß für jede Formel γ der angegebenen Art einer der folgenden vier Fälle gelten:

1. Der Wahrheitswert von γ ist unter jeder 0-1-Belegung g stets 0, unabhängig vom Wert $g(A_0)$,
2. Der Wahrheitswert von γ ist unter jeder 0-1-Belegung g stets 1, unabhängig vom Wert $g(A_0)$,
3. Der Wahrheitswert von γ unter einer 0-1-Belegung g stimmt mit $g(A_1)$ überein,
4. Der Wahrheitswert von γ unter einer 0-1-Belegung g stimmt mit $1 - g(A_1)$ überein.

Es kann daher nicht mehr als vier nichtäquivalente Formeln der angegebenen Art geben.

10.10 Wir nennen zwei 0-1 Bewertungsfunktionen äquivalent, wenn sie allen Atomformeln aus $\{A_1, \ldots, A_n\}$ denselben Wert zuordnen. Damit gibt es genau 2^n viele paarweise nicht äquivalente 0-1 Bewertungsfunktionen. Wir wählen Vertreter $g_1, \ldots, g_{2^n}$ der Äquivalenzklassen. Sei M die Menge der Formeln, deren Atomformeln in $\{A_1, \ldots, A_n\}$ liegen. Sind $\alpha, \beta \in M$ und gilt $g_i(\alpha) = g_i(\beta)$ für alle $1 \leq i \leq 2^n$, so sind nach dem Koinzidenztheorem α und β aussagenlogisch äquivalent. Jeder Formel $\alpha \in M$ ordnen wir die Menge $G(\alpha) := \{g_i \mid 1 \leq i \leq 2^n, g_i(\alpha) = 1\}$ zu. Aus $G(\alpha) = G(\beta)$ folgt also, daß α und β äquivalent sind. Dies zeigt, daß die Anzahl nichtäquivalenter Formeln aus M durch die Zahl der Teilmengen von $\{1, \ldots, 2^n\}$ beschränkt ist. Diese beträgt 2^{2^n}.

10.15 Die Eigenschaft $\{\gamma\} \models \neg\gamma$ gilt genau dann, wenn γ unerfüllbar ist. Ist nämlich γ unerfüllbar, so auch $\{\gamma\}$, es folgt sofort $\{\gamma\} \models \alpha$ für beliebiges $\alpha \in L$. Ist andererseits γ erfüllbar, etwa durch die 0-1 Bewertung g, so gilt $g(\neg\gamma) = 0$, damit nicht $\{\gamma\} \models \neg\gamma$.
Es ist γ zu $\gamma \vee \neg\gamma$ aussagenlogisch äquivalent genau dann, wenn γ eine aussagenlogische Tautologie ist. Dies ergibt sich daraus, daß offenkundig $\gamma \vee \neg\gamma$ eine aussagenlogische Tautologie ist.

Lösungen zu Aufgaben aus Kapitel 11

11.1 Es steht τ für irgendein Axiom, etwa $A_0 \Rightarrow A_0$. Wenn wir eine Herleitung von $\sigma \Rightarrow \gamma$ und von $\gamma \Rightarrow \sigma$ aus Σ haben, können wir diese mittels Regel (2) und Anwendungen des Modus Ponens verlängern, um die Formeln $\alpha := \tau \Rightarrow (\sigma \Rightarrow \gamma)$ und $\beta := \tau \Rightarrow (\gamma \Rightarrow \sigma)$ herzuleiten. Es ist nun die Formel

$$\alpha \Rightarrow (\beta \Rightarrow (\tau \to ((\sigma \Rightarrow \gamma) \wedge (\gamma \Rightarrow \sigma))))$$

eine Instanz von Axiom 8. Aus ihr erhalten wir mittels α, β und τ und drei Anwendungen des Moduls Ponens die Formel $(\sigma \Rightarrow \gamma) \wedge (\gamma \Rightarrow \sigma)$. Ist umgekehrt die zuletztgenannte Formel aus Σ herleitbar, so kann man die Axiome (7) und den Modus Ponens verwenden, um die Formeln $\sigma \Rightarrow \gamma$ und $\gamma \Rightarrow \sigma$ herzuleiten.

11.2 Mit dem neuen Axiom $\alpha \Rightarrow (\beta \Rightarrow (\alpha \wedge \beta))$ erhielte man zunächst die Herleitung mit den Gliedern α, β, $\alpha \Rightarrow (\beta \Rightarrow (\alpha \wedge \beta))$. Zwei Anwendungen des Modus Ponens ergeben nun $\alpha \wedge \beta$.

11.3 Axiom (5) ist $\alpha \Rightarrow (\alpha \vee \beta)$. Axiom (11) ergibt $(\alpha \Rightarrow (\alpha \vee \beta)) \Rightarrow (\neg(\alpha \vee \beta) \Rightarrow \neg\alpha)$. Durch Modus Ponens erhält man $\neg(\alpha \vee \beta) \Rightarrow \neg\alpha)$. Analog folgt $\neg(\alpha \vee \beta) \Rightarrow \neg\beta)$ durch Verwendung der zweiten Variante von Axiom (5). Der Rest ergibt sich einfach durch Verwendung des Axioms (8) und zweimaliger Anwendung des Moduls Ponens.

11.6 Wir betrachten exemplarisch die Regel ($\Rightarrow$: rechts). Die betreffenden assoziierten Formeln sind

$$\begin{array}{ll} (1) & \neg\alpha_1 \vee \bigvee \neg\Gamma \vee \bigvee \Delta \vee \alpha_2 \vee \bigvee \Lambda \\ (2) & \bigvee \neg\Gamma \vee \bigvee \Delta \vee \alpha_1 \Rightarrow \alpha_2 \vee \bigvee \Lambda \end{array}$$

Wenn wir von der Reihenfolge der Disjunktionsglieder abstrahieren, geht (2) aus (1) hervor, indem wir die Disjunktion $\neg\alpha_1 \vee \alpha_2$ durch $\alpha_1 \Rightarrow \alpha_2$ ersetzen. Es ist trivial (etwa mit Wahrheitswert-Tabellen) zu zeigen, daß die beiden zuletztgenannten Formeln logisch äquivalent sind. Damit sind nach dem Ersetzungslemma 10.19 auch (1) und (2) logisch äquivalent.

11.9 Die zur Sequenz $\rightarrow \alpha$ assoziierte Formel ist α. Wir wenden auf $\rightarrow \alpha$ Regeln des Kalküls G' in invertierter Form an. Da jede Rückwärtsanwendung einen Junktor eliminiert, stoppt das Verfahren nach endlich viele Schritten, wobei alle Formeln der Blattsequenzen Atomformeln sind. Die assoziierten Formeln sind daher Disjunktionen von Literalen. Es folgt nun aus Proposition 11.18 durch Induktion über die Zahl der invertierten Regelanwendungen, daß die Konjunktion über all diese Disjunktionen von Literalen zu α logisch äquivalent ist.

11.16 Jede erzeugte Klausel ist eine Teilmenge der $2n$ Literale, die über den Atomformeln gebildet werden können. Daher kann die Zahl der Klauseln erzeugten nicht größer sein als 2^{2n}.

11.17 Aus $\{A_3, \neg A_0\}$ und $\{\neg A_3, \neg A_0\}$ ergibt sich per Resolventenbildung die neue Klausel $\{\neg A_0\}$. Aus $\{A_0, A_1\}$ und $\{\neg A_0\}$ ergibt sich per Resolventenbildung die neue Klausel $\{A_1\}$. Aus $\{A_1\}$ und $\{A_0, A_2, \neg A_1\}$ ergibt sich $\{A_0, A_2\}$. Aus $\{A_0, A_2\}$ und $\{\neg A_0\}$ ergibt sich $\{A_2\}$. Aus $\{A_2\}$ und $\{\neg A_2\}$ ergibt sich die leere Klausel. Also ist M unerfüllbar.

11.21 Es sei M eine endliche Klauselmenge und n der größte Index einer in M als Teilformel auftretenden Atomformel. Sei M unerfüllbar. Angenommen, der semantische Baum $\mathcal{T}_{\text{sem}}$ hat unendliches Gewicht bezüglich M. Dann muß auch ein Knoten η in Tiefe $> n$ zum Gewicht beitragen. Bei η ist dann noch keine der Klauseln aus

M widerlegt. Nun wäre aber nach dem aussagenlogischen Koinzidenztheorem jede Wahrheitswertbelegung, die einem Pfad durch η zugeordnet ist, eine Belegung, die M erfüllt.

Lösungen zu Aufgaben aus Kapitel 12

12.1 Es gibt $\mathcal{L}_\Sigma$-Terme beliebiger Tiefe genau dann, wenn Σ zumindest ein Funktionssymbol enthält. Zu beachten ist hierbei, daß wir in Definition 6.19 festgelegt haben, daß Funktionssymbole stets positive Stelligkeit haben. Ist f ein Funktionssymbol der Stelligkeit n, so erreichen die Terme der Folge $f(x, x, \dots, x)$, $f(f(x, x, \dots, x), x, \dots, x)$, $f(f(f(x, x, \dots, x), x, \dots, x), x, \dots, x), \dots$ beliebige Tiefe. Falls Σ hingegen kein Funktionssymbol hat, so sind die Variablen und die Individuenkonstanten die einzigen Terme.

12.4 Die freien Variablenvorkommen sind nachfolgend unterstrichen.

$$\begin{array}{ll} (1) & (\forall x\ ((\exists y\ R(x,y)) \wedge S(\underline{y}))) \wedge R(\underline{x}, \underline{u}) \\ (2) & (\forall u\ (\forall x \forall y\ ((\exists z\ R(x,y,z)) \wedge S(\underline{z}))) \wedge R(\underline{x}, u)) \end{array}$$

Die Mengen der freien Variablen sind also $\{x, y, u\}$ (1) beziehungsweise $\{x, z\}$ (2).

12.5 Alle variablenfreien Terme enthalten als atomare Teilterme höchstens die Individienkonstanten 0 und 1, als Funktionsszeichen höchstens „+“ und „$\cdot$“. Ist n die Auswertung eines solchen Terms in $\mathcal{R}$, so ergibt sich n vermöge Summation bzw. Multiplition aus den Auswertungen dieser Individuenkonstanten. Offenkundig entstehen durch Summation bzw. Multiplition aus Null und Eins nur natürliche Zahlen. Ist n eine vorgegebene natürliche Zahl, so ergibt sich n als Auswertung des variablenfreien Terms $(\dots(1 + 1) + \dots + 1)$ mit n Vorkommen von 1. Wir erweitern nun die Signatur mit den einstelligen Funktionszeichen „$-$“ und „$^{-1}$“. Das Zeichen „$-$“ interpretieren wir als Abbildung $x \mapsto -x$, das Zeichen „$^{-1}$“ als Inversenbildung $x \mapsto 1/x$. Hierbei ergibt sich das Problem, daß $1/0$ nicht definiert ist. Die Interpretation eines Funktionssymbols muß aber eine *totale* Funktion sein. Als einen Ausweg definieren wir willkürlich z.B. $1/0 := 0$. Offenkundig ergibt sich nun jede ganze bzw. rationale Zahl als Auswertung eines variablenfreien Terms. Dies folgt sofort daraus, daß jede negative ganze Zahl darstellbar ist in der Form $-n$ mit $n \in \mathbb{N}$, jede rationale Zahl als Bruch zweier ganzer Zahlen.

12.7 Der Wert der ersten Formel wird zu 0, wenn die Variablenbelegung z auf Petra oder Maria abbildet, oder wenn sie y auf Max oder Peter abbildet. Der Wert der zweiten Formel wird 1, wenn die Variablenbelegung x auf Peter abbildet. Der Wert der letzten Formel wird zu 0, wenn die Variablenbelegung x auf Petra abbildet. Der Wert der dritten und vierten Formel ist unabhängig von der gewählten Variablenbelegung. Beides sind Sätze. Vergleiche hierzu Korollar 12.24.

12.9 Angenommen es gilt nicht $\mathcal{A} \models \varphi$. Dann gibt es eine Variablenbelegung ν in $\mathcal{A}$ mit $|\varphi|^{\mathcal{A},\nu} = 0$ bzw. $|\neg\varphi|^{\mathcal{A},\nu} = 1$. Da $\neg\varphi$ ein Satz ist, folgt mit Korollar 12.24 nun $|\neg\varphi|^{\mathcal{A},\nu'} = 1$ für jede Variablenbelegung ν' in $\mathcal{A}$. Damit gilt $\mathcal{A} \models \neg\varphi$. Um zu zeigen, daß derselbe Zusammenhang für Formeln nicht notwendig gilt, nehmen wir an, die Signatur enthalte ein einstelliges Relationssymbol R. Falls in $\mathcal{A}$ Elemente a, b existieren, wo $R_{\mathcal{A}}(a)$ aber nicht $R_{\mathcal{A}}(b)$ gilt, so folgt, daß weder $\mathcal{A} \models R(x)$ gilt, noch $\mathcal{A} \models \neg R(x)$. Dies folgt u.a. auch sofort aus Bemerkung 12.26.

12.12 Der Satz $\exists x \exists y\ (\neg x \equiv y)$ gilt in einer Struktur $\mathcal{A}$ genau dann, wenn A zumindest zwei Elemente hat. Der Satz $\exists x \exists y\ (\neg x \equiv y \wedge \forall z\ (z = x \vee z = y))$ gilt in einer Struktur $\mathcal{A}$ genau dann, wenn A genau zwei Elemente hat. Der Satz $\exists x \exists y\ (\neg x \equiv y \wedge R(x) \wedge R(y) \wedge \forall z\ (R(z) \Rightarrow (z = x \vee z = y)))$ gilt in einer Struktur $\mathcal{A}$ genau dann, $R_{\mathcal{A}}$ genau zwei Elemente hat.

12.13 Beides sind Tautologien. Dies ergibt sich für den ersten Satz sofort daraus, daß der Grundbereich einer Σ-Struktur stets nichtleer ist. Im zweiten Fall folgt dies aus der Transitivität der Gleichheit.

12.14 Es sei $\mathcal{A}$ eine Σ-Struktur. Wir führen eine Fallunterscheidung durch. Falls es in A ein Element a gibt, das nicht zu $R_{\mathcal{A}}$ gehört, so folgt für jede Variablenbelegung ν in $\mathcal{A}$ sofort $|R(x) \Rightarrow (\forall y\ R(y))|^{\mathcal{A},\nu_{x/a}} = 1$ wegen $|R(x)|^{\mathcal{A},\nu_{x/a}} = 0$. Wir erhalten $|\exists x\ (R(x) \Rightarrow (\forall y\ R(y)))|^{\mathcal{A},\nu} = 1$. Es folgt $\mathcal{A} \models \exists x\ (R(x) \Rightarrow (\forall y\ R(y)))$. Im andern Fall gehören alle Elemente aus A zu $R_{\mathcal{A}}$. Hieraus folgt $|\forall y\ R(y)|^{\mathcal{A},\nu} = 1$. Hieraus ergibt sich leicht $\mathcal{A} \models \exists x\ (R(x) \Rightarrow (\forall y\ R(y)))$.

Index

Literaturverzeichnis

[Acz88] P. Aczel. Non-well-founded sets. CSLI Lecture Notes 14, Stanford University, 1988.

[Big89] Norman L. Biggs. *Discrete Mathematics.* Oxford Science Publications, 1989.

[Bir84] Garrett Birkhoff. *Lattice Theory.* Americam Mathematical Society, Providence, Rhode Island, 1984.

[BM77] J.L. Bell and M. Machover. *A Course in Mathematical Logic.* North-Holland, Amsterdam, 1977.

[BN98] Franz Baader and Tobias Nipkow. *Term Rewriting and All That.* Cambridge University Press, 1998.

[Bra88] James Bradley. *Introduction to Discrete Mathematics.* Addision-Wesley, Reading, Massachusetts, 1988.

[Büc98] J. Richard Büchi. *Finite Automata, Their Algebras and Grammars.* Springer Verlag, Berlin Heidelberg, 1998. Editiert von Dirk Siefkes.

[Can55] Georg Cantor. *Contributions to the Founding of the Theory of Transfinite Numbers.* Dover Publications, New York, 1955. Übersetzt von P. Jourdain.

[CK73] C.C. Chang and H.J. Keisler. *Model Theory.* North-Holland, New York, Amsterdam, 1973.

[DM79] Nachum Dershowitz and Zohar Manna. Proving termination with multiset orderings. *Communications of the ACM*, 22(8):465–476, 1979.

[dW67] W. d. Waerden. *Algebra I, II.* Springer, 1966/67.

[(Ed77] J. Barwise (Ed.). *Handbook of Mathematical Logic.* North-Holland, Amsterdam, 1977.

[Fel79] Walter Felscher. *Naive Mengen und Abstrakte Zahlen Bd. III, Transfinite Methoden.* BI Wissenschaftsverlag, Bibliograpisches Institut AG, Zürich, 1979.

[FS91] P. A. Fejer and D. A. Simovici. *Mathematical Foundations of Computer Science, Vol. I: Sets, Relations, and Induction.* Springer Verlag, Berlin Heidelberg, 1991.

[Ger72] Hans-Dieter Gerster. *Aussagenlogik, Mengen, Relationen.* Studienbücher Mathematik. Herder KG, Freiburg, 1972.

[Grä68] George Grätzer. *Universal Algebra.* Van Nostrand, Princeton, New York, 1968.

[GW99] Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis - Mathematical Foundations.* Springer Verlag, Berlin Heidelberg, 1999.

[Hal87] P. R. Halmos. *Naive Set Theory.* Springer Verlag, Berlin Heidelberg New York, 1987.

[Her78] H. Hermes. *Aufzählbarkeit, Entscheidbarkeit, Berechenbarkeit.* Springer, Berlin Heidelberg, 1978.

[Hof79] Douglas R. Hofstadter. *Gödel, Escher, Bach; an Eternal Golden Braid.* Basic Books, New York, 1979. Deutsche Übersetzung: Gödel, Escher, Bach ein Endloses Geflochtenes Band, Klett-Cotta, Stuttgart 1985.

[HU79] J.E. Hopcroft and J.D. Ullman. *Introduction to Automata Theory, Languages, and Computation.* Addison-Wesley, Reading, MA, 1979.

[Jac85] Nathan Jacobson. *Basic Algebra I.* W.H. Freeman and Company, New York, 1985.

[Jun99] Dieter Jungnickel. *Graphs, Networks and Algorithms.* Springer Verlag, 1999.

[Kle67] S.C. Kleene. *Mathematical Logic.* Wiley, New York, 1967.

[Kö06] J. König. Sur la théorie des ensembles. *C.R.Paris*, 143:110–112, 1906.

[Koz97] Dexter C. Kozen. *Automata and Computability.* Springer, New York, Berlin, 1997.

[Lan65] S. Lang. *Algebra.* Addison-Wesley, 1965.

[Lev79] A. Levy. *Basic Set Theory.* Springer Verlag, Berlin Heidelberg New York, 1979.

[LMM87] Jean-Louis Lassez, Michael Maher, and Kim Mariott. Unification revisited. In J. Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 587–625. Morgan Kaufman, 1987.

[Mey75] K. Meyberg. *Algebra Teil I, II.* Carl Hanser Verlag, München Wien, 1975.

[MT92] Karl Meinke and John Tucker. Universal algebra. In S. Abramsky, D.M. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science, Vol. 1*, pages 189–411. Oxford University Press, 1992.

[Pen90] Roger Penrose. *The Emperors New Mind.* Oxford University Press, New York, Vintage Edition, 1990.

[PtMW90] Barbara H. Partee, Alice ter Meulen, and Robert E. Wall. *Mathematical Methods in Linguistics.* Kluwer Academic Publishers, Dordrecht, 1990.

[Rob65] J.A. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 27:23–41, 1965.

[Roi90] J. Roitman. *Introduction to Modern Set Theory.* John Wiley & Sons, New York, 1990. Übersetzt von P. Jourdain.

[RSV69] H.J. Reiffen, G. Scheja, and U. Vetter. *Algebra.* BI Mannheim, 1969.

[RW92] Kenneth A. Ross and Charles R.B. Wright. *Discrete Mathematics.* Prentice-Hall International Editions, 1992.

[SE00] Gerd Stumme and Rudolf Wille (Editoren). *Begriffliche Wissensverarbeitung.* Springer Verlag, Berlin Heidelberg, 2000.

[Sho67] J.R. Shoenfield. *Mathematical Logic.* Addison-Wesley, Reading, M.A., 1967.

[Sik70] R. Sikorski. *Boolean Algebras.* Springer Verlag, Berlin Heidelberg, 1970.

[SS93] Gunther Schmidt and Thomas Ströhlein. *Relations and Graphs.* Springer Verlag, Berlin Heidelberg, 1993.

[uLW74] Harald Scheid und Lutz Warlich. *Mathematik für Lehramtskandidaten Band I: Mengen, Relationen, Abbildungen.* Akademische Verlagsgesellschaft, Frankfurt am Main, 1974.