Vom Gefüge des Denkens

Logische Systeme zur Grundlegung der Mathematik, der Informatik und der Philosophie

Vom Gefüge des Denkens

Logische Systeme zur Grundlegung der Mathematik, der Informatik und der Philosophie

Juni 2024

Inhaltsverzeichnis

1. Logisches Schließen

1.1. Grundbegriffe

1.1.1. Schlussregeln

Logisches Schließen findet in einzelnen Schritten statt. Ein Schritt stellt hierbei immer die Ableitung einer Schlussfolgerung aus einer oder mehreren Voraussetzungen dar. Die Voraussetzungen heißen *Prämissen*, die Schlussfolgerung *Konklusion*. Darstellen wollen wir den Schritt durch eine waagerechte Linie, wobei die Prämissen oberhalb befindlich sein sollen, und die Konklusion unterhalb. Der Schritt

$$\frac{\text{Wenn es regnet, wird die Straße nass} \qquad \text{Es regnet}}{\text{Die Straße wird nass}}$$

beschreibt beispielsweise, dass aus den Prämissen »Wenn es regnet, wird die Straße nass« und »Es regnet« die Konklusion »Die Straße wird nass« gefolgert wird.

Schlüsse wie der Obige treten in der Mathematik ständig auf. Ihnen allen liegt ein bestimmtes Muster zugrunde, welches sich durch eine als *Modus ponens* oder *Abtrennungsregel* bezeichnete schematische *Schlussregel* beschreiben lässt. Es bezeichne hierzu $A \Rightarrow B$ die Implikation »wenn A, dann B«. Es dürfen nun in

$$\frac{A \Rightarrow B \qquad A}{B}$$

für A, B beliebige Aussagen eingesetzt werden. So darf »Es regnet« für A und »Die Straße wird nass« für B eingesetzt werden.

1.1.2. Sequenzen

Das Schließen von Aussagen allein genügt nicht. Um freier argumentieren zu können, würden wir gerne den Umstand beschreiben können, dass eine Aussage unter bestimmten Annahmen abgeleitet werden konnte. Diese Annahmen A_k sind selbst Aussagen. Wir fassen sie zu einer endlichen Ansammlung

$$\Gamma := [A_1, A_2, \ldots, A_n]$$

zusammen, worunter wir eine endliche Liste, oder auch eine endliche Menge verstehen wollen, denn man soll mit dieser Liste umgehen können wie mit einer Menge. Das heißt, es ist nicht von Bedeutung, wie oft eine Aussage vorkommt oder in welcher Reihenfolge die Aussagen stehen. Wir bezeichnen die Symbolik

$$\Gamma \vdash A$$

als *Sequenz*. Sie drückt das *Urteil* aus, dass die Aussage A aus den Annahmen vermittels Schlussregeln ableitbar ist. Darin nennt man A die *Hinterformel* oder *Sukzedenz*. Man bezeichnet Γ als die *Vorderformeln*, die *Antezedenz*, oder die Liste der *Antezedenzen*. Es wird Γ auch der *Kontext* oder die *Umgebung* genannt, das sind auf die Typentheorie zurückzuführende Sprechweisen, die einen ganz ähnlichen Formalismus besitzt. Der Modus ponens wird nun in der allgemeinen Form

$$\frac{\Gamma \vdash A \Rightarrow B \qquad \Gamma \vdash A}{\Gamma \vdash B}$$

beschrieben. Wir argumentieren beim Schließen also ab jetzt nicht mehr mit den Aussagen selbst, sondern mit den Sequenzen. Dies hat einen wichtigen Grund, nämlich dass die Berücksichtigung der Abhängigkeit von Annahmen expliziter Teil des Schließens wird.

Ein Kontext kann auch eine leere Liste sein. Besitzt eine vermittels Schlussregeln ableitbare Sequenz einen leeren Kontext, so bezeichnet man die Sukzedenz als ein *Theorem* im engeren Sinne. Ein Theorem ist also eine Aussage, die für sich allein gilt, ohne dass dafür irgendwelche Annahmen getroffen werden müssen.

Für Sequenzen gilt die *Abschwächungsregel*. Sie besagt, dass falls die Aussage A bereits aus Γ ableitbar ist, diese Aussage erst recht ableitbar ist, wenn zu Γ weitere Annahmen Γ' hinzugefügt werden. Kurzum gilt die Regel

$$\frac{\Gamma \vdash A}{\Gamma, \Gamma' \vdash A}.$$

Hierbei bedeutet Γ, Γ' die Konkatenation der Listen Γ und Γ', also im Wesentlichen dasselbe wie die Vereinigung $\Gamma \cup \Gamma'$, insofern man die Kontexte als Mengen betrachtet.

Der Umstand, dass mit dem Kontext umgegangen werden darf, wie mit einer Menge, drückt sich in zwei weiteren Regeln aus. Es gilt die *Kontraktionsregel*

$$\frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}, \quad \text{oder allgemein} \quad \frac{\Gamma, \Gamma', \Gamma' \vdash B}{\Gamma, \Gamma' \vdash B}$$

und die *Vertauschungsregel*

$$\frac{\Gamma, A, B, \Gamma' \vdash C}{\Gamma, B, A, \Gamma' \vdash C},$$

wobei Γ, Γ' leer sein dürfen. Schließlich gelten für Mengen die Termumformungen $\Gamma \cup \Gamma' \cup \Gamma' = \Gamma \cup \Gamma'$ und

$$\Gamma \cup \{A\} \cup \{B\} \cup \Gamma' = \Gamma \cup \{B\} \cup \{A\} \cup \Gamma'.$$

1.1.3. Zulässige Schlussregeln

Wiewohl die Regeln des Schließens den Mechanismus zum Beweis von Aussagen bilden, ist ihre Rolle sogar noch ein wenig tiefgreifender. Wir können sie nämlich ebenfalls zur Ableitung *weiterer Regeln* nutzen. Das heißt, wir können sie dazu nutzen, den logischen Kalkül selbst zu erweitern. Erweiterungen dieser Art nennen wir *zulässige Schlussregeln*.

Mit den bisherigen Regeln ist bereits die zulässige Regel

$$\frac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash B}$$

ableitbar, die eine allgemeinere Form des Modus ponens darstellt. Man erhält sie kurzerhand, indem den Prämissen des Modus ponens jeweils die Abschwächungsregel vorgesetzt wird:

$$\frac{\dfrac{\Gamma \vdash A \Rightarrow B}{\Gamma, \Gamma' \vdash A \Rightarrow B} \qquad \dfrac{\Gamma' \vdash A}{\Gamma, \Gamma' \vdash A}}{\Gamma, \Gamma' \vdash B}$$

Die einfache Form des Modus ponens erhält man mit $\Gamma' := \Gamma$ als Spezialfall unter Anwendung der Kontraktionsregel.

1.1.4. Implikationseinführung

Ich möchte mich nun der Frage zuwenden, wie eine Implikation $A \Rightarrow B$ bewiesen wird. Intuitiv ist hierzu aus der Annahme A die Aussage B zu folgern. Das heißt, es genügt die Ableitung der Sequenz $A \vdash B$. Ein weiteres Mal gilt es noch zu berücksichtigen, dass ein Beweis auch auf einen vorausgesetzten Kontext Γ beschränkt sein dürfen soll. Reflektiert man darüber eine Weile, dürfte es der Überlegung nach wohl genügen, dass A einfach dem Kontext Γ hinzugefügt wird, woraus B zu folgern ist. Man gelangt zur Regel

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B}.$$

Wer diese Regel nicht so leicht fassbar findet, insbesondere nicht direkt plausibel, ob sie bedenkenlos eingesetzt werden darf, der ist nicht allein. Es gibt auch logische Kalküle, die diese Regel nicht explizit enthalten. Sie tritt dennoch als *Deduktionstheorem* in Erscheinung, ein metalogisches Theorem, dessen Beweis erst erbracht werden muss. Ich möchte diesen Weg allerdings aus einem bestimmten Grund nicht gehen. Nämlich ist beim Beweis eigentlich natürliches Schließen auf der metalogischen Ebene zu verwenden, wenn dies auch in informaler Weise stattfinden mag. Aber nicht jeder Leser weiß zu diesem Zeitpunkt, wie akkurates logisches Schließen geht. Der Leser benötigt am Anfang etwas, um sich an den eigenen Haaren aus dem Sumpf zu ziehen.

1.1.5. Axiome

Zur Komplettierung des Kalküls gesellen sich schließlich auch noch *Axiome* hinzu, das sind gemachte Grundannahmen, die nicht weiter bewiesen werden müssen. Sie sollten daher möglichst plausibel, oder besser noch zweifelsfrei einsichtig sein. Für die Logik selbst genügt das Axiom

$$A \vdash A.$$

Der Kalkül funktioniert dergestalt, dass für A eine beliebige Aussage eingesetzt werden darf, worunter auch zusammengesetzte Aussagen fallen. Eine gern gewählter Weg der Definition des logischen Kalküls sieht A als eine metasprachliche Variable, für die eine beliebige Formel eingesetzt werden darf. Unter dieser Sichtweise spricht man von einem *Axiomenschema*. Wie eine Schablone produziert es für jede Einsetzung einer konkreten logischen Formel ein eigenes Axiom.

Statt A, B, C sind für metasprachliche Variablen auch die griechischen Buchstaben φ, ψ, χ geläufig. Man muss sie von atomaren logischen Variablen unterscheiden, für die ich in diesem Buch, um Missverständnissen aus dem Weg zu gehen, kleine Buchstaben a, b, c oder p, q, r verwenden werde. Sprachlich suggestiv steht φ für *Formel* oder *formula*, a für *Aussage* und p für *proposition*.

Streng genommen notiert man nur konkrete atomare Variablen als p, q, r, wogegen P, Q, R metasprachliche Variablen sind, für die konkrete atomare Variablen eingesetzt werden dürfen. Hier ergibt sich allerdings eine Überschneidung mit der Prädikatenlogik, wo P, Q, R Prädikate bezeichnen. Das ist aber eigentlich nicht sonderlich schlimm, da atomare Variablen nichts anderes als nullstellige Prädikate sind. Statt von atomaren Variablen wird daher auch von logischen Konstanten gesprochen, um sie von den Individuenvariablen zu unterscheiden, die nicht für logische Aussagen, sondern für Objekte wie Zahlen oder Mengen stehen.

In diesem Sinne sind auch die Schlussregeln Schemata. Sofern man Schlussregeln mit null Prämissen gestattet, lässt sich das Axiomenschema auch als Regel

$$\frac{}{A \vdash A}$$

auffassen. In dieser Weise wollen wir die Anwendung von Axiomen in den Beweisbäumen darstellen.

Axiome in der Form von Sequenzen heißen auch *Grundsequenzen*.

Wir haben nun die Mittel in der Hand, um erste Theoreme beweisen zu können. Es ist $A \Rightarrow A$ ein Theorem. Der Beweisbaum ist:

$$\dfrac{\dfrac{}{A \vdash A}\,\text{Axiom}}{\vdash A \Rightarrow A}\,\text{Subjunktionseinführung}$$

Unter der Lesung, dass A eine Metavariable ist, handelt es eigentlich nicht nur um ein Theorem, sondern um ein Schema von Theoremen. Setzt man für A bspw. die konkrete Formel $p \Rightarrow q$ ein, bekommt man das konkrete Theorem

$$(p \Rightarrow q) \Rightarrow (p \Rightarrow q).$$

1.1.6. Junktoren

Bisher traten zusammengesetzte Aussagen alleinig in Form einer Implikation auf. Will man logische Zusammenhänge beschreiben können, muss die logische Sprache um weitere Junktoren bereichert werden. Unter einem *Junktor* versteht man einen logischen Operator, der Aussagen zu einer zusammengesetzten Aussage verknüpft. Von Belang sind zunächst fünf Stück.

Wir werden einen Junktor durch *Einführungsregeln* und *Beseitigungsregeln* charakterisieren. Die Regeln der Implikation wurden bereits beschrieben; die Einführung geschieht per Implikationseinführung, die Beseitigung per Modus ponens. Für die restlichen Junktoren der Aussagenlogik lassen sich die Regeln wahlweise in Form von Axiomenschemata oder in Form von Schlussregeln darstellen. Ich möchte das per Schemata machen, weil diese ein wenig kompakter sind, was sie hoffentlich ein wenig leichter einsichtig macht. Die entsprechenden Schlussregeln leiten wir anschließend als zulässige Regeln ab.

Die Konjunktion $A \wedge B$, auch Und-Verknüpfung genannt, sprich »A und B«, ist charakterisiert durch die Sequenzen

$$A, B \vdash A \wedge B; \qquad A \wedge B \vdash A; \qquad A \wedge B \vdash B.$$

Aus dem Fall von sowohl Regen als auch Schnee ist der Fall von Schneeregen ableitbar. Aus dem Fall von Schneeregen ist der Fall von Regen ableitbar. Aus dem

Fall von Schneeregen ist der Fall von Schnee ableitbar. So sind diese Sequenzen zu verstehen.

Die Einführung der Konjunktion geschieht mit der Regel

$$\frac{\Gamma \vdash A \qquad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A \wedge B}.$$

Denn es findet sich der Beweisbaum:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{A, B \vdash A \wedge B}\,\text{Axiom}}{A \vdash B \Rightarrow A \wedge B}\,\text{Subj-Einf.}}{\vdash A \Rightarrow (B \Rightarrow A \wedge B)}\,\text{Subj-Einf.} \qquad \Gamma \vdash A}{\dfrac{\Gamma \vdash B \Rightarrow A \wedge B \qquad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A \wedge B}\,\text{MP}}\,\text{MP}$$

Es steht MP als Abkürzung für Modus ponens, und Sub-Einf. für Subjunktionseinführung. Man schreibt alternativ auch das Kürzel ⇒E statt Subj-Einf. und ⇒B statt MP. Hierbei steht E offenkundig für *Einführung* und B für *Beseitigung*. Aber Vorsicht, in der englischsprachigen Literatur sind das I für *introduction* und E für *elimination*. Unmissverständlich wären subj-intro und subj-elim.

Die beiden Regeln zur Beseitigung der Konjunktion sind

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}, \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}.$$

Denn es findet sich:

$$\dfrac{\dfrac{\dfrac{}{A \wedge B \vdash A}\,\text{Axiom}}{\vdash A \wedge B \Rightarrow A}\,\text{Subj-Einf.} \qquad \Gamma \vdash A \wedge B}{\Gamma \vdash A}\,\text{MP}$$

Die Disjunktion $A \vee B$, auch Oder-Verknüpfung genannt, sprich »A oder B«, ist charakterisiert durch die Sequenzen

$$A \vdash A \vee B; \qquad B \vdash A \vee B; \qquad A \vee B, (A \Rightarrow C), (B \Rightarrow C) \vdash C.$$

So ist »Die Erde des Beetes ist nass« ableitbar aus »Es hat geregnet oder das Beet wurde gegossen«. Denn sowohl »Es hat geregnet« als auch »Das Beet wurde gegossen« impliziert »Die Erde des Beetes ist nass«.

Die beiden Regeln zur Einführung der Disjunktion sind

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B}, \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B}.$$

Die Regel zur Beseitigung der Disjunktion ist

$$\frac{\Gamma \vdash A \vee B \qquad \Gamma', A \vdash C \qquad \Gamma'', B \vdash C}{\Gamma, \Gamma', \Gamma'' \vdash C}.$$

Die Beweise dieser Regeln seien dem Leser überlassen.

Eine Aussage wie »Bertram wird seine Hausaufgaben nicht machen« formuliert man gern in der Form »Wenn Bertram seine Hausaufgaben macht, färbt sich der Mond grün«. In gleichartiger Weise lässt sich die Verneinung auch in der formalen Logik definieren. Hierzu legt man als Hilfsbegriff zunächst $\bot$ als die *Kontradiktion* fest, sie steht für eine widersprüchliche Formel.

Die Negation $\neg A$, auch Verneinung genannt, sprich »nicht A«, definiert man als identisch mit $A \Rightarrow \bot$. Hierdurch sind die Regeln zu ihrer Einführung und Beseitigung auf die der Implikation zurückführbar. Es ergibt sich

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}, \qquad \frac{\Gamma \vdash \neg A \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash \bot}.$$

Alternativ ließe sich die Negation durch die Sequenzen

$$(A \Rightarrow \bot) \vdash \neg A; \qquad A, \neg A \vdash \bot$$

charakterisieren. Man überzeuge sich, dass dies aufs selbe hinausläuft.

Die Äquivalenz $A \Leftrightarrow B$, sprich »A genau dann, wenn B«, definiert man als identisch mit $(A \Rightarrow B) \wedge (B \Rightarrow A)$. Insofern sind die Regeln zu ihrer Einführung und Beseitigung auf die der Konjunktion zurückführbar. Es ergibt sich

$$\frac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash B \Rightarrow A}{\Gamma, \Gamma' \vdash A \Leftrightarrow B}, \qquad \frac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash A \Rightarrow B}, \qquad \frac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash B \Rightarrow A}.$$

Die entsprechenden charakterisierenden Sequenzen sind

$$(A \Rightarrow B), (B \Rightarrow A) \vdash A \Leftrightarrow B; \qquad (A \Leftrightarrow B), A \vdash B; \qquad (A \Leftrightarrow B), B \vdash A.$$

1.1.7. Quantoren

Eine logische Sprache, die der freien Formulierung mathematischer Zusammenhänge dienlich sein soll, muss hinreichend reichhaltig sein. Ebenfalls schrieb der Philosoph Ludwig Wittgenstein in seinem *Tractatus* den ähnlichen Gedanke »Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.« Bislang fehlt das wichtige Konzept der Quantifizierung, das die Aussagenlogik zur Prädikatenlogik erweitert.

In der Prädikatenlogik treten *Aussageformen* auf. Das sind Formeln, die *freie Variablen* enthalten. Erst wenn jede der freien Variablen mit einem Wert belegt wird, entsteht eine Aussage. Außerdem treten Quantoren auf. Ein *Quantor* bindet eine freie Variable, und macht eine Aussageform dabei ebenfalls zu einer bestimmten Aussage.

Es sei $A(x)$ eine Aussageform mit der freien Variable x. Anstelle von $A(x)$ schreibt man auch kurzum A. Anstelle von $A(t)$ schreibt man auch $A[x := t]$ oder $A[t/x]$, womit die Ersetzung jedes Vorkommens von x durch den Term t gemeint ist. Genauer gesagt die Ersetzung jedes *freien* Vorkommens, wobei man außerdem einer möglichen Überschattung einer der in t enthaltenen Variablen aus dem Weg gehen muss. Diese Spitzfindigkeiten tauchen allerdings erst auf, wenn man mit Verschachtelungen von Quantoren hantiert. Ich will später näher darauf eingehen.

Die wesentlichen beiden Quantoren sind der *Allquantor* $\forall$ und der *Existenzquantor* $\exists$. Man ließt $\forall x\colon A(x)$ als »für alle x gilt $A(x)$« oder »jedes x hat die Eigenschaft $A(x)$«. Man ließt $\exists x\colon A(x)$ als »es gibt mindestens ein x, für das $A(x)$ gilt« oder »mindestens ein x hat die Eigenschaft $A(x)$«.

Quantifiziert wird immer über ein bestimmtes *Diskursuniversum*. Darunter versteht man die Gesamtheit der Objekte, auf die sich »für alle« und »es gibt« bezieht. Um bestimmten Komplikationen aus dem Weg zu gehen, muss es nichtleer sein. Zur Veranschaulichung des Übergangs von der Aussagenlogik zur Prädikatenlogik wählen wir ein endliches, das lediglich die Zahlen von eins bis vier enthält. Die Aussage $\forall x\colon A(x)$ bedeutet nun insofern dasselbe wie

$$A(1) \wedge A(2) \wedge A(3) \wedge A(4).$$

Diese schlichte Konjunktion gibt Anlass zu der Schlussregel

$$\frac{\Gamma \vdash \forall x\colon A(x)}{\Gamma \vdash A(t)}. \qquad (t \text{ muss eine der Zahlen von eins bis vier sein})$$

Die Beseitigung der Allquantifizierung darf insofern als Analogon zur Beseitigung der Konjunktion verstanden werden.

Im Fortgang soll $\Gamma \vdash A(a)$ bedeuten, dass die Aussageform $A(a)$ aus dem Kontext Γ ableitbar ist, wobei a beliebig gelassen wird. Man leitet die vier Sequenzen

$$\Gamma \vdash A(1); \quad \Gamma \vdash A(2); \quad \Gamma \vdash A(3); \quad \Gamma \vdash A(4)$$

sozusagen in einen Zug ab. Es stellt sich nun die Frage

$$\frac{\Gamma \vdash A(a)}{\Gamma \vdash \forall x\colon A(x)}?$$

Betrachten wir dazu $a = 1 \vdash A(a)$. Mit der bedenklichen Regel erhielte man aus ihr $a = 1 \vdash \forall x\colon A(x)$. Diese trifft insbesondere im Fall $a := 1$ zu. Nun braucht man $1 = 1$ nicht vorauszusetzen, womit man $\vdash \forall x\colon A(x)$ erhält. Den Quantor beseitigen wir nun noch mit $x := a$, und erhalten $\vdash A(a)$. Die Annahme wurde also aus der Sequenz herausgemogelt. Um dies zu unterbinden, legen wir fest, dass a keine freie Variable einer der Antezedenzen sein darf.

Die Regel zur Einführung ist demnach zu formulieren als

$$\frac{\Gamma \vdash A(a)}{\Gamma \vdash \forall x\colon A(x)}(a \notin \mathrm{FV}(\Gamma, \forall x\colon A(x))).$$

Hierbei steht die Symbolik $\mathrm{FV}(\Gamma)$ für die Menge der freien Variablen von Γ. Mit elementarer Mengenlehre definiert man sie präzise als Rekursion über den Formelaufbau. Man legt fest

$$\begin{aligned}
&\mathrm{FV}(A \wedge B) = \mathrm{FV}(A \vee B) = \mathrm{FV}(A \Rightarrow B) = \mathrm{FV}(A \Leftrightarrow B) = \mathrm{FV}(A) \cup \mathrm{FV}(B),\\
&\mathrm{FV}(\neg A) = \mathrm{FV}(A), \qquad \mathrm{FV}(\forall x\colon A) = \mathrm{FV}(\exists x\colon A) = \mathrm{FV}(A) \setminus \{x\},\\
&\mathrm{FV}(\bot) = \mathrm{FV}(\top) = \emptyset, \quad \mathrm{FV}(P(t_1, \ldots, t_n)) = \mathrm{FV}(t_1) \cup \ldots \cup \mathrm{FV}(t_n).
\end{aligned}$$

Hierbei steht P für ein n-stelliges Prädikat. Und es ist $\mathrm{FV}(t)$ die Menge der Variablen des Terms t. Man legt sie abermals rekursiv fest als

$$\begin{aligned}
&\mathrm{FV}(t_1 + t_2) = \mathrm{FV}(t_1 - t_2) = \mathrm{FV}(t_1 \cdot t_2) = \mathrm{FV}(t_1) \cup \mathrm{FV}(t_2),\\
&\mathrm{FV}(-t) = \mathrm{FV}(t), \quad \mathrm{FV}(v) = \{v\}, \quad \mathrm{FV}(c) = \emptyset,
\end{aligned}$$

wobei v für eine Variable und c für eine Konstante steht.

Die Aussage $\exists x\colon A(x)$ ist gleichwertig mit

$$A(1) \vee A(2) \vee A(3) \vee A(4).$$

Diese Perspektive gibt Anlass zur Einführungsregel

$$\frac{\Gamma \vdash A(t)}{\Gamma \vdash \exists x\colon A(x)}. \quad (t \text{ muss eine der Zahlen von eins bis vier sein})$$

Bei der Beseitigung müssen wir nun gewissermaßen eine Fallunterscheidung in die vier Fälle vornehmen und bestätigen, dass jeder Fall dieselbe Aussage B impliziert. Dies soll allerdings parametrisch in einer einzigen Ableitung stattfinden. Man gelangt zu

$$\frac{\Gamma \vdash \exists x\colon A(x) \qquad \Gamma', A(u) \vdash B}{\Gamma, \Gamma' \vdash B}(u \notin \mathrm{FV}(\Gamma, \Gamma', B, \exists x\colon A(x))).$$

Diese Regel wird wie folgt interpretiert. Mit der Existenzaussage $\exists x\colon A(x)$ liegt ein Zeuge u mit $A(u)$ vor. Unter Verwendung von $A(u)$ wird nun eine Aussage B abgeleitet, in der u nicht frei vorkommt. Somit gilt B unabhängig vom gewählten Zeugen, was notwendig ist, da unbekannt bleibt, welche der Zahlen von eins bis vier als Zeuge vorliegt.

Ohne die Bedingung an u ließe sich leicht Schabernack vollführen. Man könnte beispielsweise kurzerhand eine Existenzaussage zu einer Allaussage machen:

$$\dfrac{\dfrac{\vdash \exists x\colon A(x) \qquad \overline{A(u) \vdash A(u)}}{\vdash A(u)}}{\vdash \forall x\colon A(x)}$$

Abschließend verbleibt noch näher zu erläutern, wie Substitution vonstatten geht. Ersetzt wird nur jedes freie Vorkommen einer Variablen. Ein durch einen Quantor gebundenes Vorkommen der Variable bleibt dagegen erhalten. So resultiert die Substitution

$$(P(x) \land \forall x : Q(x))[x := y] \quad \text{in} \quad P(y) \land \forall x\colon Q(x).$$

Außerdem darf es bei einer Substitution nicht zu einer Überschattung durch eine Variablenbindung kommen, engl. *capture-avoiding substitution*, was bedeuten soll, dass bei $A[x := t]$ keine der freien Variablen des Terms t durch eine in A befindliche Variablenbindung eingefangen wird. Die Substitution

$$(\forall y\colon P(x) \land Q(y))[x := y]$$

darf beispielsweise nicht direkt ausgeführt werden. Man geht der Überschattung aus dem Weg, indem die gebundene Variable y zuerst in eine frische, nehmen wir z, umbenannt wird. Das Resultat der Substitution ist also

$$\forall z\colon P(y) \land Q(z), \quad \text{nicht} \quad \forall y\colon P(y) \land Q(y).$$

1.1.8. Substitution

Es ist noch zu präzisieren, wie Substitution genau vonstatten geht. Substituiert werden können sowohl atomare logische Variablen gegen Formeln als auch Individuenvariablen gegen Terme. Betrachten wir zunächst die logischen.

Allgemein definiert man ihre Substitution

$$A[P_1 := C_1, \ldots P_n := C_n], \text{ kurz } A[S] \text{ oder } S(A)$$

rekursiv über den Formelaufbau als

$$\begin{aligned}(\neg A)[S] &:= \neg(A[S]), & (\forall x\colon A)[S] &:= (\forall x\colon (A[S])),\\ (A \circ B)[S] &:= ((A[S]) \circ (B[S])), & (\exists x\colon A)[S] &:= (\exists x\colon (A[S])).\end{aligned}$$

wobei $\circ$ jeder der zweistelligen Junktoren $\wedge, \vee, \Rightarrow, \Leftrightarrow$ ist. Die Basisfälle sind für die atomaren Variablen $P_1, \ldots, P_n$ und Q definiert gemäß

$$Q[P_1 := C_1, \ldots, P_n := C_n] := \begin{cases} C_k, & \text{wenn sich } k \text{ mit } P_k = Q \text{ findet,}\\ Q, & \text{sonst.}\end{cases}$$

Für $n \geq 2$ spricht man von *simultaner Substitution.*

Für $n = 1$ vereinfacht sich die Substitution zu

$$Q[P := C] := \begin{cases} C, & \text{wenn } P = Q,\\ Q, & \text{wenn } P \neq Q.\end{cases}$$

Beispielsweise ist

$$(a \wedge b \Rightarrow a)[a := a \vee b] = ((a \vee b) \wedge b \Rightarrow (a \vee b)).$$

Simultane Substitution darf im Allgemeinen nicht schrittweise durchgeführt werden, weil dadurch ein anderes Resultat entstehen kann. Zum Beispiel ist

$$\begin{aligned}(a \wedge b)[a := b, b := c] &= (b \wedge c),\\ (a \wedge b)[a := b][b := c] &= (c \wedge c).\end{aligned}$$

Implementiert man die Substitution als Computerprogramm, bildet sie üblicherweise abstrakte Syntaxbäume auf abstrakte Syntaxbäume ab.

Die Substitution von Individuenvariablen gegen Terme definiert man ganz analog. Hier ist allerdings hinsichtlich der Quantoren zu beachten, dass nur freie Variablen substituiert werden, und man eine unter Umständen entstehende Überschattung durch Umbenennung der gebundenen Variablen umgehen muss.

1.1.9. Zur Syntax

So wie »Punktrechnung vor Strichrechnung« gilt, legt man für jeden Junktor zur Einsparung von Klammern eine Stufe der Priorität fest. In absteigender Rangfolge sind dies $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$. So wird die Formel

$$\neg A \wedge B \vee C \Rightarrow D \quad \text{gelesen als} \quad (((\neg A) \wedge B) \vee C) \Rightarrow D.$$

Weiterhin legt man die Implikation als rechtsassoziativ fest. So wird

$$A \Rightarrow B \Rightarrow C \quad \text{gelesen als} \quad A \Rightarrow (B \Rightarrow C).$$

Wie den Junktoren kommt auch den Quantoren eine Rangstufe zu. Weil diese aber präfix sind, ist bei ihnen lediglich die rechte Seite zu berücksichtigen. Hier sind zwei Varianten verbreitet. In der Schreibweise $(\forall x)A(x)$ oder kurz $\forall x A(x)$ haben sie wie die Verneinung die höchste Rangstufe. In der Schreibweise $\forall x\colon A(x)$ oder $\forall x.\, A(x)$ dagegen die niedrigste, also eine Stufe niedriger als das Bikonditional, so dass alles hinter dem Doppelpunkt in den Wirkungsbereich des Quantors fällt. So wird

$$\forall x\colon A(x) \wedge B \Rightarrow C \quad \text{gelesen als} \quad \forall x\colon ((A(x) \wedge B) \Rightarrow C).$$

Manche Schüler haben Schwierigkeiten, die Struktur von Termen zu durchschauen. Infolge kann es bei ihnen zu Flüchtigkeitsfehlern bei der Ersetzung von Variablen durch Terme kommen. Sie vergessen, dass ein Term vor der Einfügung zunächst geklammert werden muss. Erst die Operatorrangfolge gewährt es, die Klammern unter Umständen nachträglich entfallen zu lassen. Für diese Schüler mag es förderlich sein, einen Term als *abstrakten Syntaxbaum* darzustellen. Gleichermaßen verhält es sich mit der Programmiersprache Lisp, die Terme als Schachtelung von Listen darstellt, deren Klammern obligatorisch sind. Die Aussage $A \wedge B \Rightarrow C$ ist beispielsweise beschreibbar als

```
(implies (and A B) C).
```

Im Wesentlichen veranschaulicht diese Schachtelung nichts anderes als den abstrakten Syntaxbaum. Man kann gewissermaßen sagen, dass Lisp eine Programmiersprache ohne Syntax ist. Fast ohne, im höheren Sinne ohne.

Um sich unmissverständlich auszudrücken, formalisieren Logiker die logische Sprache gern. Es wird hierzu eine *formale Sprache* spezifiziert, was vermittels sogenannter *Produktionsregeln* gemacht werden kann. Insofern Produktionsregeln etwas kryptisch anmuten mögen, beschreiben Logiker die syntaktische Struktur auch in Worten. Für die Aussagenlogik üblicherweise folgendermaßen.

1. Die atomaren Variablen a, b, c usw. sind Formeln.
2. Die Symbole $\bot, \top$ sind Formeln.
3. Ist A eine Formel, so ist auch $(\neg A)$ eine.
4. Sind A, B Formeln, so ist auch $(A \wedge B)$ eine.

5. Sind A, B Formeln, so ist auch $(A \vee B)$ eine.
6. Sind A, B Formeln, so ist auch $(A \Rightarrow B)$ eine.
7. Sind A, B Formeln, so ist auch $(A \Leftrightarrow B)$ eine.
8. Nichts anderes ist eine Formel.

Hierbei bleibt allerdings die Rangfolge und Assoziativität der Junktoren unberücksichtigt. Um sie festzulegen, kann man eine Grammatik der Form PEG – dies steht für *parsing expression grammar* – mit unterschiedlichen Nichtterminalsymbolen aufstellen. Die obige Regelung ist diesbezüglich als spezielle Grammatik betrachtbar, in der *Formel* das einzige Nichtterminalsymbol darstellt. Daraufhin kann mit der Technik des rekursiven Abstiegs ein Parser programmiert werden, der Formeln in abstrakte Syntaxbäume umwandelt. Zur Vollendung kann der Parser ggf. anschließend in einen effizienteren Bottom-up-Parser transformiert werden. Ich will hier nicht darauf eingehen, zumal dies Grundkenntnisse in der Programmierung voraussetzt. Ausführliche Erklärungen zu formalen Grammatiken und Parsern finden sich in Büchern über Compilerbau.

Schreibt man viele logische Formeln auf, drängt es, zumindest bei privaten Notizen und Rechnungen, nach Kurzschreibweisen. In der Logik ist für das Konditional $A \Rightarrow B$ auch die Schreibweise $A \to B$ gebräuchlich, für das Bikonditional $A \Leftrightarrow B$ entsprechend $A \leftrightarrow B$. Insbesondere in der Schaltalgebra schreibt man auch $\overline{A}$ anstelle von $\neg A$, und AB anstelle von $A \wedge B$ sowie $A + B$ anstelle von $A \vee B$. Hierbei darf die Disjunktion $A + B$ allerdings nicht mit der Kontravalenz $A \oplus B$ verwechselt werden. Für die Quantifizierung $\forall x\colon A(x)$ bietet sich $\forall_x A_x$ oder $\forall x.\, A_x$ als kurzschriftliche Form an.

1.2. Natürliches Schließen

1.2.1. Darstellungsformen

Abgeleitet werden soll das Theorem

$$\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A).$$

Meine favorisierte und in diesem Buch genutzte Form der Darstellung des natürlichen Schließens fügt die aus den Schlussregeln erhaltenen Schlüsse wie Legosteine zu einem Baum zusammen, dem *Beweisbaum* oder *Herleitungsbaum*. Im Eigentlichen stehen in den Blättern die Grundsequenzen, und in der Wurzel das Theorem. Wie wir es bereits getan haben, arbeitet man allerdings auch mit Exemplaren, die irgendwelche Sequenzen in irgendwelche Sequenzen überführen, womit man zulässige Schlussregeln erhält. Allgemeiner ginge ferner die Formulierung als gerichteter azyklischer Graph, die bei einigen Beweisen ein wenig den Schreibaufwand reduzieren würde.

Der Beweisbaum des genannten Theorems ist:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\neg B \vdash \neg B}\,\text{Axiom} \quad \dfrac{\dfrac{}{A \Rightarrow B \vdash A \Rightarrow B}\,\text{Axiom} \quad \dfrac{}{A \vdash A}\,\text{Axiom}}{A \Rightarrow B, A \vdash B}\,\Rightarrow\text{B}}{A \Rightarrow B, \neg B, A \vdash \bot}\,\neg\text{B}}{A \Rightarrow B, \neg B \vdash \neg A}\,\neg\text{E}}{A \Rightarrow B \vdash \neg B \Rightarrow \neg A}\,\Rightarrow\text{E}}{\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)}\,\Rightarrow\text{E}$$

Die Ausformulierung der Sequenzen verlangt langwieriges erneutes Aufschreiben der Antezedenzen. Sobald man das Prozedere einmal verstanden hat, erscheint es überausführlich. Man kann sich daher verkürzte Darstellungen der Beweisbäume überlegen:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{1 \equiv \neg B} \quad \dfrac{\dfrac{}{2 \equiv A \Rightarrow B} \quad \dfrac{}{3 \equiv A}}{2, 3 \vdash B}}{1, 2, 3 \vdash \bot}}{1, 2 \vdash \neg A}}{2 \vdash \neg B \Rightarrow \neg A}}{\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\neg B}\,1 \quad \dfrac{\dfrac{}{A \Rightarrow B}\,2 \quad \dfrac{}{A}\,3}{B}}{\bot}}{\neg A}\sim 3}{\neg B \Rightarrow \neg A}\sim 1}{(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)}\sim 2$$

Die linke Form kürzt die Antezedenzen durch Nummern ab. In der rechten Form entfallen die Antezedenzen vollständig. Stattdessen tauchen sie als in den Blättern

Tabelle 1.1.: Beweis in Form einer Liste von Tabellenzeilen

Abh.	Nr.	Aussage	Regel	auf
1	1	$\neg B$	Axiom	
2	2	$A \Rightarrow B$	Axiom	
3	3	A	Axiom	
2, 3	4	B	$\Rightarrow$B	2, 3
1, 2, 3	5	$\bot$	$\neg$B	1, 4
1, 2	6	$\neg A$	$\neg$E	5
2	7	$\neg B \Rightarrow \neg A$	$\Rightarrow$E	6
$\emptyset$	8	$(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$	$\Rightarrow$E	7

gemachte nummerierte *Annahmen* auf, die im Fortgang zur Wurzel irgendwann zu tilgen sind. Ihre Tilgung erscheint nun als Randnotiz.

Eine weitere, sehr systematische Darstellung setzt den Beweis aus einer Liste von Tabellenzeilen zusammen. Allerdings ist sie ein wenig mühevoll zu lesen. Jede Zeile enthält eine Aussage und dahinter zusätzlich die Information, wie und woraus die Aussage abgeleitet wurde. Jede der Aussagen bekommt eine Nummer, siehe Tabelle 1.1. Die Nummerierung der Abhängigkeiten ist in derselben Reihenfolge wie zuvor bei den Bäumen angegeben. Wer die Liste genauer betrachtet, erkennt, dass die jeweilige Zeile nichts anderes als die Sequenz Abh. $\vdash$ Nr. darstellt.

Unabhängig von Gentzen entwickelte Stanisław Jaśkowski das natürliche Schließen einige Jahre zuvor. Während Gentzen Beweise als Bäume darstellte, nutzte Jaśkowski zunächst eine grafische Darstellung, die später von Frederic Brenton Fitch adaptiert wurde und in dieser Form nun als *Fitch-Style* bekannt ist. Die Abhängigkeit von einer Annahme wird hier kenntlich gemacht, indem die aus der Annahme abgeleiteten Aussagen hinter einer senkrechten Linie eingerückt stehen. Die Annahme selbst steht am Anfang der Einrückung, und zwar bereits innerhalb, weil sie ja von sich selbst abhängig ist. Siehe Tabelle 1.2.

Zu guter Letzt muss die klassische Darstellung der Beweisführung aufgeführt werden. Die in Worten. Sie zeichnet sich durch die Auslassung mühseliger technischer Details und blumige Formulierungen aus, soll aber genug Information enthalten, dass der Leser im Zweifel eine Formalisierung des Beweises erstellen und verifizieren kann.

Satz 1.1. Es gilt $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$.

Beweis. Aus der Annahme von sowohl $A \Rightarrow B$ als auch $\neg B$ als auch A ist ein

Tabelle 1.2.: Beweis im Fitch-Style

1	$A \Rightarrow B$	
2	$\quad \neg B$	
3	$\quad\quad A$	
4	$\quad\quad B$	$\Rightarrow$B, 1, 3
5	$\quad\quad \bot$	$\neg$B, 2, 4
6	$\quad \neg A$	$\neg$E, 5
7	$\neg B \Rightarrow \neg A$	$\Rightarrow$E, 6
8	$(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$	$\Rightarrow$E, 7

Widerspruch abzuleiten. Man erhält B zunächst per Modus ponens aus $A \Rightarrow B$ und A. Nun steht $\neg B$ bereits im Widerspruch zu B. □

Als komfortablen Bonus erhält man mit dem Theorem nun im Anschluss kurzerhand eine weitere zulässige Regel, die *Kontrapositionsregel*

$$\frac{\Gamma \vdash A \Rightarrow B}{\Gamma \vdash \neg B \Rightarrow \neg A}, \quad \text{denn} \quad \frac{\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A) \quad \Gamma \vdash A \Rightarrow B}{\Gamma \vdash \neg B \Rightarrow \neg A}.$$

Fügt man ihr den Modus ponens an, findet sich der *Modus tollens*

$$\frac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash \neg B}{\Gamma, \Gamma' \vdash \neg A}.$$

1.2.2. Theoreme der Prädikatenlogik

In einer Formelsammlung zur Prädikatenlogik findet man eine Reihe von Äquivalenzen und Implikationen vor, von denen wir einige beweisen wollen.

Satz 1.2. Es ist $A \wedge \forall x\colon B(x)$ äquivalent zu $\forall x\colon A \wedge B(x)$, sofern die Variable x nicht frei in A vorkommt.

Beweis. Zur Implikation von links nach rechts findet sich der Baum:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \wedge \forall x\colon B(x)}\,1}{A} \quad \dfrac{\dfrac{\dfrac{}{A \wedge \forall x\colon B(x)}\,1}{\forall x\colon B(x)}}{B(x)}}{A \wedge B(x)}}{\forall x\colon A \wedge B(x)}}{(A \wedge \forall x\colon B(x)) \Rightarrow (\forall x\colon A \wedge B(x))}\,{\sim}1 \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \wedge \forall x\colon B(x)}\,1}{A} \quad \dfrac{\dfrac{\dfrac{}{A \wedge \forall x\colon B(x)}\,1}{\forall x\colon B(x)}}{B(u)}}{A \wedge B(u)}}{\forall x\colon A \wedge B(x)}}{(A \wedge \forall x\colon B(x)) \Rightarrow (\forall x\colon A \wedge B(x))}\,{\sim}1$$

Beide Formen sind zulässig. Die rechte Form gibt dem Parameter den eigenen Bezeichner u, die linke nennt diesen ebenfalls so wie die gebundene Variable. In Worten würde man diesen Beweis wie folgt führen. Es sei u fest, aber beliebig, woraus $A \wedge B(u)$ abzuleiten ist. Laut Voraussetzung gilt sowohl A als auch $\forall x\colon B(x)$. Spezialisierung $x := u$ liefert $B(u)$. Ergo gilt sowohl A als auch $B(u)$, und somit $A \wedge B(u)$.

Zur Implikation von rechts nach links findet sich:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\forall x\colon A \wedge B(x)}\,1}{A \wedge B(x)}}{A} \quad \dfrac{\dfrac{\dfrac{}{\forall x\colon A \wedge B(x)}\,1}{A \wedge B(x)}}{B(x)}}{A \wedge B(x)}}{\forall x\colon B(x)}}{A \wedge \forall x\colon B(x)}}{(\forall x\colon A \wedge B(x)) \Rightarrow (A \wedge \forall x\colon B(x))}\,{\sim}1 \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\forall x\colon A \wedge B(x)}\,1}{A \wedge B(u)}}{A} \quad \dfrac{\dfrac{\dfrac{}{\forall x\colon A \wedge B(x)}\,1}{A \wedge B(u)}}{B(u)}}{A \wedge B(x)}}{\forall x\colon B(x)}}{A \wedge \forall x\colon B(x)}}{(\forall x\colon A \wedge B(x)) \Rightarrow (A \wedge \forall x\colon B(x))}\,{\sim}1$$

In Worten: Es ist $A \wedge \forall x\colon B(x)$ abzuleiten, also sowohl A als auch $\forall x\colon B(x)$, was sich mit $B(u)$ für ein festes, aber beliebiges u bestätigt. Laut Voraussetzung erhält man $A \wedge B(u)$ per Spezialisierung $x := u$. Ergo gilt sowohl A als auch $B(u)$. □

Die ausführliche Besprechung verdeutlicht nochmals, wie die Floskel *fest, aber beliebig* einen Parameter, über den die Argumentation verläuft, einführt. Der Parameter steht für ein *festes* Objekt, insofern er während der Argumentation für nur ein Objekt steht. Weil das Objekt *beliebig* ist, also keine Einschränkungen an dessen Beschaffenheit gemacht wurden, darf anschließend ohne Weiteres die Einführung einer Allquantifizierung vorgenommen werden. Es besteht im logischen System, wie es in diesem Buch dargelegt wurde, kein formaler Unterschied zwischen einem Parameter und einer Variable. Ein Parameter verhält sich als freie Variable.

1.2.3. Bezug zum Sequenzenkalkül

Einige Regeln des natürlichen Schließens bieten bei der Lesung eines Beweisbaums von der Wurzel aus zu den Blättern hin routinemäßige Zurückführung eines Ziels

Tabelle 1.3.: Die Regeln des Sequenzenkalküls

	Linke Regel	**Rechte Regel**	
$\dfrac{\Gamma \vdash \Delta}{\Gamma, \Gamma' \vdash \Delta}$	$\dfrac{\Gamma, A, B \vdash \Delta}{\Gamma, A \wedge B \vdash \Delta}$	$\dfrac{\Gamma \vdash A, \Delta \qquad \Gamma' \vdash B, \Delta'}{\Gamma, \Gamma' \vdash A \wedge B, \Delta, \Delta'}$	$\dfrac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \Delta'}$
$\dfrac{\Gamma, A, A \vdash \Delta}{\Gamma, A \vdash \Delta}$	$\dfrac{\Gamma, A \vdash \Delta \qquad \Gamma', B \vdash \Delta'}{\Gamma, \Gamma', A \vee B \vdash \Delta, \Delta'}$	$\dfrac{\Gamma \vdash A, B, \Delta}{\Gamma \vdash A \vee B, \Delta}$	$\dfrac{\Gamma \vdash B, B, \Delta}{\Gamma \vdash B, \Delta}$
$\dfrac{\Gamma, \top \vdash \Delta}{\Gamma \vdash \Delta}$	$\dfrac{\Gamma \vdash A, \Delta \qquad \Gamma', B \vdash \Delta'}{\Gamma, \Gamma', A \Rightarrow B \vdash \Delta, \Delta'}$	$\dfrac{\Gamma, A \vdash B, \Delta}{\Gamma \vdash A \Rightarrow B, \Delta}$	$\dfrac{\Gamma \vdash \Delta, \bot}{\Gamma \vdash \Delta}$
$\dfrac{}{\Gamma, \bot \vdash \Delta}$	$\dfrac{\Gamma \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}$	$\dfrac{\Gamma, A \vdash \Delta}{\Gamma \vdash \neg A, \Delta}$	$\dfrac{}{\Gamma \vdash \top, \Delta}$

auf Unterziele. Jedoch nicht jede der Regeln, womit das Schließen dennoch zum Denksport wird. Aufgrund dessen gestaltet sich auch die Auffindung eines Algorithmus' zur automatischen Erzeugung eines Beweisbaums als schwierig.

Der Sequenzenkalkül macht das Schließen nun gänzlich zur Routine. Mithin ist zu diesem Kalkül ein Algorithmus zur Erzeugung von Beweisbäumen vergleichsweise leicht zu finden.

Man darf als dienlich erachten, dass die Darstellung des natürlichen Schließens, wie sie in diesem Buch dargelegt wurde, mit dem Sequenzenkalkül kompatibel ist. In ihm dürfen Sequenzen von der allgemeineren Form

$$A_1 \ldots, A_m \vdash B_1, \ldots, B_n$$

sein, die für die Aussage

$$A_1 \wedge \ldots \wedge A_m \Rightarrow B_1 \vee \ldots \vee B_n$$

steht. Um die Regeln des Sequenzenkalküls vermittels natürlichem Schließen als zulässige Regeln herzuleiten, wird man daher zunächst die Übersetzungsregeln

$$\frac{\Gamma \vdash B_1 \vee \ldots \vee B_n}{\Gamma \vdash B_1, \ldots, B_n}, \qquad \frac{\Gamma \vdash \bot}{\Gamma \vdash}, \qquad \frac{\Gamma \vdash B_1, \ldots, B_n}{\Gamma \vdash B_1 \vee \ldots \vee B_n}, \qquad \frac{\Gamma \vdash}{\Gamma \vdash \bot}.$$

fordern. Eine Auflistung der wesentlichen Regeln zeigt die Tabelle 1.3. Sie untergliedern sich in linke und rechte Regeln. Die linken gestatten es dabei, Ziele ebenfalls bezüglich Antezedenzen auf Unterziele zurückzuführen. Jede der Ansammlungen

$\Gamma, \Gamma', \Delta, \Delta'$ darf leer sein. Eine algorithmische Umsetzung der Beweissuche mag $\Gamma = \Gamma'$ und $\Delta = \Delta'$ setzen, für den Menschen entstünde dadurch aber umständlicher Schreibaufwand. Exemplarisch soll die linke Regel zur Disjunktion als zulässig bestätigt werden. Für sie findet sich der Baum:

$$\dfrac{\dfrac{}{A \vee B \vdash A \vee B} \quad \dfrac{\dfrac{\Gamma, A \vdash \Delta}{\Gamma, A \vdash \Delta, \Delta'}}{\Gamma, A \vdash C} \quad \dfrac{\dfrac{\Gamma', B \vdash \Delta'}{\Gamma', B \vdash \Delta, \Delta'}}{\Gamma', B \vdash C}}{\dfrac{\Gamma, \Gamma', A \vee B \vdash C}{\Gamma, \Gamma', A \vee B \vdash \Delta, \Delta'}}$$

Hierbei soll C die Disjunktion der Aussagen von Δ, Δ' sein.

Es folgt am Beispiel des Theoremschemas zur Kontraposition, wie das Schließen vonstatten geht. Die Beweisbäume wären von unten nach oben zu lesen, weil das, was weiter oben steht, eigentlich noch unbekannt ist:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \vdash A}}{\vdash A, \neg A} \quad \dfrac{\dfrac{}{B \vdash B}}{B, \neg B \vdash}}{A \Rightarrow B, \neg B \vdash \neg A}}{A \Rightarrow B \vdash \neg B \Rightarrow \neg A}}{\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{B \vdash B}}{\vdash \neg B, B} \quad \dfrac{\dfrac{}{A \vdash A}}{\neg A, A \vdash}}{\neg B \Rightarrow \neg A, A \vdash B}}{\neg B \Rightarrow \neg A \vdash A \Rightarrow B}}{\vdash (\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B)}$$

Der Kalkül muss ein klassischer sein, sonst wäre die Kontraposition nicht rückgängig zu machen. Noch drastischere Einsicht diesbezüglich schaffen die Bäume:

$$\dfrac{\dfrac{\dfrac{}{A \vdash A}}{\vdash A, \neg A}\,\neg\mathrm{R}}{\vdash A \vee \neg A}\,\vee\mathrm{L} \qquad \dfrac{\dfrac{\dfrac{}{A \vdash A}}{\vdash \neg A, A}\,\neg\mathrm{L}}{\neg\neg A \vdash A}\,\neg\mathrm{R}$$

Es gibt auch Varianten des Sequenzenkalküls, die ausschließlich die Ableitung von Theoremen der intuitionistischen Logik gestatten. Sie sind allerdings umständlicher zu verwenden, da es sich um Einschränkungen des klassischen Kalküls handelt. Bereits Gentzen beschrieb so einen als LJ in [1]. Eine sorgfältige Diskussion findet man in [12], [7].

Alternative Formen der Regeln zur Implikation sind

$$\dfrac{\Gamma, \neg A \vdash \Delta \qquad \Gamma, B \vdash \Delta}{\Gamma, A \Rightarrow B \vdash \Delta}, \qquad \dfrac{\Gamma \vdash \neg A, B, \Delta}{\Gamma \vdash A \Rightarrow B, \Delta}.$$

Die Regeln der Allquantifizierung sind

$$\dfrac{\Gamma, A(t) \vdash \Delta}{\Gamma, \forall x\colon A(x) \vdash \Delta}, \qquad \dfrac{\Gamma \vdash A(u), \Delta}{\Gamma \vdash (\forall x\colon A(x)), \Delta}\,(u \notin \mathrm{FV}(\Gamma, \Delta, A(x))),$$

die der Existenzquantifizierung sind

$$\frac{\Gamma, A(u) \vdash \Delta}{\Gamma, \exists x\colon A(x) \vdash \Delta}(u \notin \mathrm{FV}(\Gamma, \Delta, A(x))), \qquad \frac{\Gamma \vdash A(t), \Delta}{\Gamma \vdash (\exists x\colon A(x)), \Delta}.$$

Eine besondere Bedeutung besitzt die Schnittregel

$$\frac{\Gamma \vdash A, \Delta \qquad \Gamma', A \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}.$$

Laut Gentzens Hauptsatz findet sich zu jedem Beweis einer Sequenz, in dem die Schnittregel zur Anwendung kommt, ein alternativer Beweis, der auf sie verzichtet. Kurzum ist sie zulässig, aber redundant. Dieses Resultat ist von großer Bedeutung für die Beweistheorie.

1.2.4. Bezug zum Tableaukalkül

Der *Tableaukalkül* ist ein systematisches Beweisverfahren, bei dem durch abermalige Zurückführung einer Formel auf kleinere Formeln ein Baum entsteht. In einer geläufigen Form des Verfahrens kommt der Beweis einer Aussage zustande, indem ihre Verneinung widerlegt wird. Die Widerlegung stellt sich dadurch her, dass jeder Pfad eine Formel enthält, deren Verneinung bereits auf dem direkten Pfad zur Wurzel vorkam, was auch als *Schließung* des Pfades bezeichnet wird. Der Baum verzweigt sich nicht bei jeder Formel. Man unterscheidet zwischen Formeln vom Typ einer Konjunktion und Formeln vom Typ einer Disjunktion. Lediglich bei den Formeln vom Typ einer Disjunktion findet eine Verzweigung statt.

Es stellt sich im Fortgang heraus, dass der Tableaukalkül in der Logik nicht abgeschieden steht. Ganz im Gegenteil lässt sich ein enger Bezug zum Schließen von Sequenzen herstellen. Genauer gesagt gehört zu jeder Regel des Tableaukalküls eine zulässige Regel des Schließens von Sequenzen, womit sich dieser als ein Teilsystem des allgemeinen Sequenzenkalküls erweist.

Laut der Reductio ad absurdum ist $\Gamma \vdash A$ auf $\Gamma, \neg A \vdash \bot$ zurückführbar. Im Weiteren wird $\Gamma \vdash$ als Abkürzung für $\Gamma \vdash \bot$ geschrieben. Man ruft sich nun die Äquivalenz der Aussagen $A \Rightarrow B$ und $\neg A \vee B$ in Erinnerung. Weiterhin befindet man mit den de morganschen Gesetzen die Aussage $\neg(A \wedge B)$ zu $\neg A \vee \neg B$ äquivalent, sowie $\neg(A \vee B)$ zu $\neg A \wedge \neg B$. Aus diesen Überlegungen heraus gelangt man zu den unverzweigenden zulässigen Regeln

$$\frac{\Gamma, A, B \vdash}{\Gamma, A \wedge B \vdash}, \qquad \frac{\Gamma, \neg A, \neg B \vdash}{\Gamma, \neg(A \vee B) \vdash}, \qquad \frac{\Gamma, A, \neg B \vdash}{\Gamma, \neg(A \Rightarrow B) \vdash},$$

und den verzweigenden zulässigen Regeln

$$\frac{\Gamma, A \vdash \quad \Gamma, B \vdash}{\Gamma, A \vee B \vdash}, \qquad \frac{\Gamma, \neg A \vdash \quad \Gamma, \neg B \vdash}{\Gamma, \neg(A \wedge B) \vdash}, \qquad \frac{\Gamma, \neg A \vdash \quad \Gamma, B \vdash}{\Gamma, (A \Rightarrow B) \vdash}.$$

Schließlich wäre noch festzustellen, dass $\Gamma, A, \neg A \vdash$ mit $\Gamma, A \vdash A$ gleichbedeutend ist, und somit die Rolle einer Grundsequenz einnehmen darf.

Ein Beispiel. Mit den aufgestellten Regeln ergibt sich für die Kontraposition der folgende Beweisbaum:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\neg A, \neg B, \neg\neg A \vdash} \quad \dfrac{}{B, \neg B, \neg\neg A \vdash}}{A \Rightarrow B, \neg B, \neg\neg A \vdash}}{A \Rightarrow B, \neg(\neg B \Rightarrow \neg A) \vdash}}{\neg((A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)) \vdash}}{\vdash (A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)} \qquad \begin{array}{c} \neg((A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)) \\ A \Rightarrow B \\ \neg(\neg B \Rightarrow \neg A) \\ \neg B \\ \neg\neg A \\ \hline \neg A \checkmark \qquad B \checkmark \end{array}$$

Wird dieser Baum nun auf den Kopf gestellt und die Notation dergestalt verkürzt, dass jede der Formeln nur einmal aufgeschrieben werden braucht, findet sich die übliche Notation des Tableaukalküls wieder. Die rechte Darstellung zeigt das Resultat dieser Umgestaltung. Man mag das Schließen von Sequenzen insofern als vielseitig bewerten. Die dazugewonnene Sichtweise schafft überdies ein tiefergründiges Verständnis des Tableaukalküls.

1.3. Logik mit Gleichheit

1.3.1. Axiome der Gleichheit

Im Folgenden Abschnitt geht es um allgemeingültige Erwägungen zur Gleichheit. Das wären Gesetzmäßigkeiten, die die Gleichheit immer erfüllen soll, ungeachtet, ob sie zwischen zwei Zahlen, zwei Mengen, oder zwei wie auch immer gearteten Objekten besteht.

Moderne Formulierungen charakterisieren die Gleichheit durch die Axiome

$$\vdash \forall x\colon x = x, \qquad \text{(Reflexivität)}$$
$$\vdash \forall x\colon \forall y\colon x = y \Rightarrow A(x) \Rightarrow A(y). \qquad \text{(Ersetzung)}$$

Das zweite Axiom stellt eigentlich ein Schema dar, weil dieses für jede Formel A gilt. Mit $A(u)$ sei hierbei gemeint, dass in A die ungenannte Variable u frei vorkommen darf, wobei $A(t)$ als $A(u)[u := t]$ für einen Term t, einschließlich $t = x$ und $t = y$, zu verstehen sein soll. Man gewinnt aus dem Schema kurzerhand die Regel

$$\frac{\Gamma \vdash t = t' \qquad \Gamma' \vdash A(t)}{\Gamma, \Gamma' \vdash A(t')}. \qquad (t, t' \text{ sind beliebige Terme})$$

Das erste Axiom charakterisiert insofern die Regel zur Einführung der Gleichheit, das zweite die Regel zur Beseitigung. Die Symmetrie der Gleichheit lässt sich aus den beiden Regeln ableiten. Sei hierzu $A(u) :\Leftrightarrow (u = x)$. Nun ist $A(x) \Leftrightarrow (x = x)$ und $A(y) \Leftrightarrow (y = x)$. Man setze $t := x$ und $t' := y$. Es findet sich:

$$\frac{\overline{x = y \vdash x = y} \qquad \overline{\vdash x = x}}{x = y \vdash y = x}$$

Aus dieser Sequenz erhält man anschließend

$$\vdash \forall x\colon \forall y\colon x = y \Rightarrow y = x.$$

Bezüglich $A(u) :\Leftrightarrow (f(x) = f(u))$ führt die Ausübung der soeben gemachten Vorgehensweise auf

$$\vdash \forall x\colon \forall y\colon x = y \Rightarrow f(x) = f(y).$$

Es induziert die Ersetzungsregel für Funktionen,

$$\frac{\Gamma \vdash t = t'}{\Gamma \vdash f(t) = f(t')}.$$

Zu beachten wäre allerdings, dass f hierfür auf dem gesamten Diskursuniversum definiert sein muss. Würde das Symbol f mit einer Funktion belegt, die für eine bestimmte Belegung von x nicht definiert ist, was soll $f(x)$ dann sein?

Mehrmalige Anwendung des Ersetzungsaxioms ermöglicht darüber hinaus mehrstellige Ersetzungen wie

$$\vdash \forall x, x', y, y' \colon x = x' \land y = y' \Rightarrow A(x, y) \Rightarrow A(x', y'),$$
$$\vdash \forall x, x', y, y' \colon x = x' \land y = y' \Rightarrow f(x, y) = f(x', y').$$

Ferner implizieren die Axiome das Transitivgesetz

$$\vdash \forall x, y, z \colon x = y \land y = z \Rightarrow x = z.$$

Es steht $\forall x, y, z \colon A$ als Abkürzung für $\forall x \colon \forall y \colon \forall z \colon A$.

Ganz allgemein gilt

$$\vdash \forall x \colon \forall y \colon x = y \Rightarrow s(x) = s(y)$$

für jeden Term s. Dies bestätigt sich unschwer folgendermaßen. Sei h eine frische Hilfsvariable, die nicht frei in s vorkommt und

$$A :\Leftrightarrow (s[u := x] = s[u := h]),$$

wobei $s(x) = s[u := x]$ und $s(y) = s[u := y]$ ist. Der Schluss

$$\frac{\Gamma \vdash x = y \qquad \vdash A[h := x]}{\Gamma \vdash A[h := y]}$$

vereinfacht sich nun zu

$$\frac{\Gamma \vdash x = y \qquad \dfrac{}{\vdash s[u := x] = s[u := x]}}{\Gamma \vdash s[u := x] = s[u := y]}, \quad \text{kurz} \quad \frac{\Gamma \vdash x = y \qquad \dfrac{}{\vdash s(x) = s(x)}}{\Gamma \vdash s(x) = s(y)}.$$

Es genügt übrigens, das Ersetzungsaxiom für atomare Aussageformen zu fordern. Seien hierzu P, Q Prädikate. Sei $A(x)$ zum Beispiel die Formel $P(x) \land Q(x)$. Dann ist die Regel

$$\frac{\Gamma \vdash x = y \qquad \Gamma' \vdash A(x)}{\Gamma, \Gamma' \vdash A(y)}$$

zulässig, denn:

$$\dfrac{\dfrac{\Gamma \vdash x = y \qquad \dfrac{\Gamma' \vdash P(x) \land Q(x)}{\Gamma' \vdash P(x)}}{\Gamma, \Gamma' \vdash P(y)} \qquad \dfrac{\Gamma \vdash x = y \qquad \dfrac{\Gamma' \vdash P(x) \land Q(x)}{\Gamma' \vdash Q(x)}}{\Gamma, \Gamma' \vdash Q(y)}}{\Gamma, \Gamma' \vdash P(y) \land Q(y)}$$

Für die anderen Junktoren klappt es analog. Per struktureller Induktion über den Formelaufbau bestätigt sich die Regel daraufhin in allgemeiner Weise als zulässig.

1.3.2. Von der Identität des Ununterscheidbaren

Dem Gleichheitsbegriff wohnt das *Principium identitatis indiscernibilium* inne, das *Prinzip der Identität des Ununterscheidbaren*, englisch *Identity of indiscernibles*. Es besagt, dass man keine zwei ungleichen Objekte finden kann, die in allen ihren Eigenschaften übereinstimmen. Man nennt es auch *Gleichheit nach Leibniz*, weil Wilhelm Gottfried Leibniz dieses im 27. Kaptiel von *Nouveaux Essais sur L'entendement humain II* im Bezug zum Kosmos beschrieb. Am Ende findet man das Wesentliche nochmals in fasslicher, bildhafter Form,

> »Ich erinnere mich, dass eine große Prinzessin, die von erhabenem Geist ist, einmal sagte, als sie in ihrem Garten spazieren ging, dass sie nicht glaube, dass es zwei vollkommen ähnliche Blätter gebe. Ein geistreicher Herr, der mit auf dem Spaziergang war, glaubte, dass es leicht sein würde, solche zu finden; aber obwohl er viel suchte, wurde er durch seine Augen davon überzeugt, dass man immer einen Unterschied bemerken könne.«

Formalisierung erfährt das Prinzip durch die Aussage

$$(\forall P\colon P(x) \Leftrightarrow P(y)) \Rightarrow x = y.$$

Man sollte bedenken, dass diese Formulierung die Prädikatenlogik zweiter Stufe erfordert, da hier über Prädikate quantifiziert wird. Die Umkehrung

$$x = y \Rightarrow (\forall P\colon P(x) \Leftrightarrow P(y)).$$

wird als unbedenklich angesehen, so dass man das Prinzip auch als Äquivalenz formuliert. Die Umkehrung ist fast trivial aus den Axiomen ableitbar, weil es sich bereits um eine gewisse Form des Ersetzungsaxioms handelt. Umgekehrt können wir aus der Äquivalenz die beiden Axiome zurückgewinnen. Reflexivität besteht offenkundig, weil $P(x)$ immer äquivalent zu $P(x)$ ist. Und zum Ersetzungsaxiom wurde bereits ausgeführt, dass es genügt, dieses für Prädikate zu fordern.

Es verbleibt zu untersuchen, ob das Prinzip aus den Axiomen herleitbar ist. Hierzu wird $P(u) := (x = u)$ als Prädikat gewählt, womit $x = x$ als äquivalent zu $x = y$ vorausgesetzt wird. Weil $x = x$ gemäß Reflexivität vorliegt, erhält man wie gewünscht $x = y$.

1.3.3. Eindeutige Existenz

Gelegentlich tritt in der Mathematik eine Aussage der eindeutigen Existenz auf, dass heißt, eine Aussage, laut der das Objekt mit der geforderten Eigenschaft existiert und zudem eindeutig festgelegt ist. Beispielsweise existiert bei einer Funktion zu jedem Argument genau ein Wert. Weiterhin tritt bei sogenannten universellen Eigenschaften die Forderung eindeutiger Existenz auf. Sie führen in die Kategorientheorie, und der Leser mag daraus schließen, auch wenn er noch nie von ihr gehört hat, dass diese anscheinend die Prädikatenlogik mit Gleichheit als logisches System enthalten muss. Zumindest brauchen wir sie für die folgenden Ausführungen.

Formal ist die eindeutige Existenz fassbar als der Quantor

$$(\exists! x\colon A(x)) \;:\Leftrightarrow\; (\exists x\colon A(x) \wedge \forall y\colon A(y) \Rightarrow x = y).$$

Diese Aussage ist in Existenz und Eindeutigkeit zerlegbar gemäß

Satz 1.3. Es gilt die Äquivalenz

$$(\exists! x\colon A(x)) \;\Leftrightarrow\; (\exists x\colon A(x)) \wedge (\forall x, y\colon A(x) \wedge A(y) \Rightarrow x = y).$$

Beweis. Die linke Seite gelte. Dann liegt ein u mit $A(u)$ und $\forall y\colon A(y) \Rightarrow u = y$ vor. Der Existenzaussage wird mit $x := u$ genügt. Die Eindeutigkeit geht via

$$(\forall y\colon A(y) \Rightarrow u = y), A(x), A(y) \vdash x = y.$$

Wir spezialisieren die gegebene Allaussage einmal mit $y := x$ und einmal mit $y := y$. Mit $A(x)$ und $A(y)$ bekommt man daraufhin $u = x$ und $u = y$. Ergo folgt $x = y$ per Symmetrie und Transitivität der Gleichheit.

Die rechte Seite gelte. Dann liegt ein u mit $A(u)$ vor. Wir wählen $x := u$ für die Existenzaussage. Die Aussage $A(u)$ liegt bereits vor. Verbleibt $\forall y\colon A(y) \Rightarrow u = y$ zu bestätigen. Wir nehmen also $A(y)$ an. Aus der Spezialisierung der gegebenen Allaussage mit $x := u$ und $y := y$ wird $u = y$ via $A(u) \wedge A(y)$ abgetrennt. □

Vorsorglich erwähnen möchte ich

$$(\forall x, y\colon A(x) \wedge A(y) \Rightarrow x = y) \Leftrightarrow (\forall x\colon A(x) \Rightarrow \forall y\colon A(y) \Rightarrow x = y).$$

Der Beweis sollte nicht viel Mühe bereiten.

1.4. Induktion

1.4.1. Einfache Induktion

In der Philosophie bezeichnet man als *Induktion* eine Art von Schlussfolgerung, die da ist der Schluss vom Speziellen auf das Allgemeine. Folgendes Beispiel verdeutlicht die Idee der Überlegung. Ein Gegenstand wird einmal fallen gelassen, man beobachtet wie dieser zu Boden fällt. Wiederholung des Experiments führt abermals zum selben Resultat. Induktiv schließt man daraus, dass dieses Resultat *immer* eintreten wird. Jedoch kann Induktion zu Fehlschlüssen führen, weshalb es nicht als mathematisches Beweisverfahren brauchbar ist. Nur weil bereits dreimal eine Toastbrotscheibe auf die Marmeladenseite gefallen ist, heißt das nicht, dass dieses Resultat immer eintreten müsse. In der Mathematik bieten die Borwein-Integrale ein prägnantes Beispiel, wo leichtfertige Argumentation verfänglich wäre.

Mit der bedenklichen Induktion in der Philosophie teilt sich die *vollständige Induktion* den Namen. Sie ist allerdings unfehlbar. Das Verfahren ist für die moderne Mathematik und Informatik von wesentlicher Bedeutung.

Die vollständige Induktion wird vermittelt durch das Axiomenschema

$$\vdash A(0) \wedge (\forall n \in \mathbb{N}\colon A(n) \Rightarrow A(n+1)) \Rightarrow (\forall n \in \mathbb{N}\colon A(n)).$$

Es induziert die Regel

$$\frac{\Gamma \vdash A(0) \qquad \Gamma', n \in \mathbb{N}, A(n) \vdash A(n+1)}{\Gamma, \Gamma' \vdash \forall n \in \mathbb{N}\colon A(n)} (n \notin \mathrm{FV}(\Gamma')).$$

Zur Veranschaulichung wird gerne eine endlose Dominoreihe herangezogen. Fällt der erste Dominostein um, und ist sicher, dass mit einem Dominostein ebenso dessen Nachfolger umfällt, muss *jeder* Dominostein irgendwann umfallen.

Man bezeichnet $A(0)$ als den *Induktionsanfang*. Die Ableitung von $A(n+1)$ aus $A(n)$ heißt *Induktionsschritt*, wobei $A(n)$ darin die Bezeichnung *Induktionsvoraussetzung* trägt.

Ein erstes Beispiel. Man definiert die Potenz einer Zahl a rekursiv als

$$a^0 := 1, \qquad a^{n+1} := aa^n.$$

Zu beweisen sei das Potenzgesetz

$$A(n) \;:\Leftrightarrow\; (ab)^n = a^n b^n.$$

Der Anfang $A(0)$ bestätigt sich via

$$(ab)^0 = 1 = 1 \cdot 1 = a^0 b^0.$$

Den Induktionsschritt $(A(n) \Rightarrow A(n+1))$ bestätigt die Rechnung

$$(ab)^{n+1} \overset{(1)}{=} (ab)(ab)^n \overset{\text{IV}}{=} aba^n b^n = aa^n bb^n \overset{(2)}{=} a^{n+1}b^{n+1}.$$

Die Stelle, wo $A(n)$ zur Anwendung kam, wurde mit IV annotiert, was für *Induktionsvoraussetzung* steht. Die Umformungen (1), (2) gelten laut Definition.

Bislang trat die Induktion so auf, dass der Anfang in der Zahl Null liegt. Aus der Regel folgt aber bereits, dass man den Anfang in jede beliebige natürliche Zahl legen kann.

Satz 1.4. Es gilt für $n, n_0 \in \mathbb{N}$ das allgemeine Schema

$$\vdash A(n_0) \wedge (\forall n \geq n_0 \colon A(n) \Rightarrow A(n+1)) \Rightarrow (\forall n \geq n_0 \colon A(n)).$$

Beweis. Die Prämisse wird spezialisiert mit $n := u + n_0$, wobei u beliebig sein darf, da $u + n_0 \geq n_0$ für jedes $u \in \mathbb{N}$ erfüllt ist. Es sei nun $B(u) :\Leftrightarrow A(u+n_0)$. Demnach ist $A(n_0)$ äquivalent zu $B(0)$. Und es ist $A(u+n_0+1)$ äquivalent zu $B(u+1)$. Insgesamt erhält man aus der Prämisse also

$$B(0) \wedge (\forall u \colon B(u) \Rightarrow B(u+1)).$$

Per herkömmlicher Induktion gilt also $B(u)$ bzw. $A(u+n_0)$ für jedes natürliche u. Wird dies nun mit der Resubstitution $u := n - n_0$ mit $n \geq n_0$ spezialisiert, erhält man $A(n)$, und somit schließlich die gesuchte Konklusion $(\forall n \geq n_0 \colon A(n))$. □

Spezialisieren ist hier formal zu verstehen. Das Resultat der Substitution ist eigentlich nicht weniger allgemein als die ursprüngliche Aussage.

Befasst man sich im Weiteren mit der Mengenlehre, wo gemeinhin mit Mengen argumentiert wird, bietet sich an, auch die Induktion bezüglich einer Menge zu formulieren. Sei hierzu $M \subseteq \mathbb{N}$ definiert als die Aussonderung

$$M := \{n \in \mathbb{N} \mid A(n)\}.$$

Mit $A(n) \Leftrightarrow n \in M$ nimmt das Schema der Induktion daraufhin die Form

$$\vdash 0 \in M \wedge (\forall n \in \mathbb{N} \colon n \in M \Rightarrow n+1 \in M) \Rightarrow M = \mathbb{N}$$

an. Man verwendet diese Variante vorwiegend, um die axiomatische Charakterisierung der natürlichen Zahlen in Bezug zur Mengenlehre zu bringen. Sie wurde Ende des 19. Jahrhunderts von Richard Dedekind und Giuseppe Peano erdacht. Das Axiom der Induktion ist darin als letzte und komplizierteste. Es lautet

$$\vdash \forall P \colon P(0) \wedge (\forall n \colon P(n) \Rightarrow P(S(n))) \Rightarrow (\forall n \colon P(n)).$$

Hiermit wird allerdings die Prädikatenlogik zweiter Stufe vorausgesetzt, weil über alle Prädikate P quantifiziert wird. Ein Verzicht auf die Logik zweiter Stufe ist aber möglich, indem das Axiom wie zuvor als Schema formuliert wird. Das so in die Prädikatenlogik erster Stufe gebrachte System nennt man die *Peano-Arithmetik*.

1.4.2. Starke Induktion

Die starke Induktion verstärkt die Induktionsvoraussetzung dahingehend, dass sie nicht nur für den unmittelbaren Vorgänger, sondern für sämtliche Vorgänger vorausgesetzt werden darf. Damit wird der Induktionsbeweis unter Umständen erleichtert. Die starke muss nicht extra axiomatisch gefordert werden, sie ist aus der herkömmlichen ableitbar.

Satz 1.5 (Starke Induktion). Es gilt das Schema

$$\vdash A(0) \land (\forall n \in \mathbb{N}\colon (\forall k \leq n\colon A(k)) \Rightarrow A(n+1)) \Rightarrow (\forall n \in \mathbb{N}\colon A(n))$$

Beweis. Sei hierzu

$$C(n) :\Leftrightarrow (\forall k \leq n\colon A(k)).$$

Die Prämissen seien angenommen. Es ist $A(0)$ gleichbedeutend mit $C(0)$. Wir überzeugen uns nun von der Folgerung

$$(\forall n \in \mathbb{N}\colon C(n) \Rightarrow A(n+1)) \Rightarrow (\forall n \in \mathbb{N}\colon C(n) \Rightarrow C(n+1)).$$

Sei also $n \in \mathbb{N}$ und $C(n)$ angenommen, dann haben wir auch $A(n+1)$ und somit $C(n) \land A(n+1)$, was äquivalent zu $C(n+1)$ ist. Per herkömmlicher Induktion gilt also $C(n)$ für jedes natürliche n. Aus $C(n)$ folgt aber $A(n)$, womit erst recht $A(n)$ für jedes natürliche n gilt. □

Ein wenig eleganter formuliert sich das Schema auch in der Form

$$\vdash (\forall n \in \mathbb{N}\colon (\forall k < n\colon A(k)) \Rightarrow A(n)) \Rightarrow (\forall n \in \mathbb{N}\colon A(n)).$$

Weil keine natürliche Zahl kleiner als null existiert, greift hier das Prinzip der leeren Wahrheit, womit $A(0)$ trotzdem zu bestätigen ist.

Aus der starken Induktion leitet sich ein Sachverhalt ab, der von der Anschauung her evident erscheinen mag.

Satz 1.6 (Wohlordnungsprinzip).
Jede nichtleere Teilmenge von $\mathbb{N}$ besitzt ein kleinstes Element.

Beweis. Gegenstand der Betrachtung sei die Menge $M \subseteq \mathbb{N}$. Mit $A(n) :\Leftrightarrow n \notin M$ ergibt sich aus dem Schema der starken Induktion die Aussage

$$(\forall n \in \mathbb{N}\colon \underbrace{(\forall k < n\colon k \notin M) \Rightarrow n \notin M}_{\neg(\forall k<n\colon k\notin M)\vee n\notin M}) \Rightarrow \underbrace{(\forall n \in \mathbb{N}\colon n \notin M)}_{M=\emptyset}.$$

Deren Kontraposition ist

$$M \neq \emptyset \Rightarrow (\exists n \in \mathbb{N}\colon (\forall k < n\colon k \notin M) \wedge n \in M).$$

Mit $M \subseteq \mathbb{N}$ ergibt sich in der Konklusion die Verkürzung

$$n \in \mathbb{N} \wedge n \in M \iff n \in \mathbb{N} \cap M = M.$$

Unternimmt man des Weiteren die Umformung

$$(\forall k < n\colon k \notin M) \iff (\forall k \in M\colon n \leq k),$$

gelangt man schließlich zur Behauptung

$$M \neq \emptyset \Rightarrow (\exists n \in M\colon \forall k \in M\colon n \leq k). \square$$

Insofern im Beweis ausschließlich Äquivalenzumformungen unternommen wurden, stellt sich heraus, dass das Wohlordnungsprinzip zur starken Induktion gleichwertig ist. Da in die Aussageform $A(n)$ per se nur natürliche Zahlen n eingesetzt werden, erhält man nämlich mit der Festlegung

$$M := \{n \in \mathbb{N} \mid \neg A(n)\}$$

das allgemeine Schema zurück.

Wird zum Wohlordnungsprinzip neuerlich die Kontraposition gebildet, jedoch ohne weitergehende Umformungen vorzunehmen, gelangt man zu

$$(\forall n \in M\colon \exists k \in M\colon k < n) \Rightarrow M = \emptyset.$$

Dieses Argument kann alternativ zur starken Induktion verwendet werden. In Worten betrachtet man die hypothetische Menge M der Elemente, die die Aussageform $A(n)$ nicht erfüllen. Kann gezeigt werden, dass zu jedem $n \in M$ ein $k \in M$ mit $k < n$ existieren würde, muss M leer sein, da sich andernfalls ein Widerspruch ergäbe, da M als Teilmenge der natürlichen Zahlen ein kleinstes Element besäße. Ist M leer, gilt $A(n)$ für jede natürliche Zahl n.

Statt vom Wohlordnungsprinzip sprechen manche vom *Prinzip des kleinsten Elements*, um eine Verwechslung mit dem *Wohlordnungssatz* zu vermeiden, einem Satz der Mengenlehre, nach welchem jede Menge wohlgeordnet werden kann.

1.4.3. Strukturelle Induktion

Die herkömmliche Induktion verläuft über die natürlichen Zahlen. Sie beginnt ohne Beschränkung der Allgemeinheit in der Null, und setzt sich dann schrittweise auf den Nachfolger fort. Es wird also $A(0)$ bestätigt, und $A(n+1)$ unter Annahme von $A(n)$.

Strukturelle Induktion verläuft allgemeiner über Knoten. Es gibt ein oder mehrere Anfangsknoten v_0, für die jeweils der Induktionsanfang $A(v_0)$ zu bestätigen ist. Außerdem liegen Regeln vor, wie aus bereits vorhandenen Knoten neue Knoten abzuleiten sind. Die neuen Knoten nehmen die Rolle der Nachfolger ein. So entsteht ein Baum oder ein gerichteter Graph. Leitet sich v aus v_1 bis v_n ab, so besteht der zugehörige Induktionsschritt darin, dass $A(v)$ unter Annahme von $A(v_1)$ bis $A(v_n)$ gezeigt wird. Bei einem Baum verläuft die Induktion von den Blättern aus zu einer Wurzel hin. Die gewöhnliche Induktion verlief von der Null zu einer Zahl hin. Wurde alles gezeigt, ist die Induktion also *vollständig*, darf die Wurzel wie die Zahl beliebig sein.

Ich will versuchen, das Prinzip so mit Dominosteinen zu veranschaulichen, wie die herkömmliche Induktion als Fallen einer Dominoreihe verständlich wird. Man kann sich eine Regel als ein Plättchen bestimmter Art vorstellen. Sofern alle notwendigen Dominosteine das Plättchen erreichen, fällt es um, womit auch der auf das Plättchen folgende Dominostein umfällt. Abermals erreichen notwendige Steine, worunter auch oder nur Nachfolger zu finden sind, ein weiteres Plättchen, worauf auch dieses fällt. Das läuft zumindest solange, bis man das gewünschte Fallen des begehrten Wurzelsteins beobachtet. Als dynamischer Prozess gesehen, hat man hier eine kausale Struktur, bei der Wirkungen nur unter einer oder mehreren passenden Ursachen auftreten.

Die herkömmliche Induktion stellt einen Spezialfall der strukturellen Induktion dar, bei der die Bäume endliche Listen sind. Bei der strukturellen Induktion über den Formelaufbau verläuft die Induktion über Formeln. Sie beginnt in den atomaren Formeln. Die Regeln zur Ableitung sind die Produktionsregeln. Bei der strukturellen Induktion über die Konstruktion eines Beweises verläuft die Induktion über Sequenzen. Sie beginnt in den Grundsequenzen. Die Regeln zur Ableitung sind die Schlussregeln.

Ein weiteres Beispiel bietet die Induktion über die Punkte des Gitters $\mathbb{Z}_{\geq 0} \times Z_{\geq 0}$. Sagen wir, der einzige Anfang ist $A(0,0)$. Die erste Schrittweise bestehe in $A(m+1,n)$ unter Annahme $A(m,n)$. Die zweite Schrittweise bestehe in $A(m,n+1)$ unter Annahme $A(m,n)$. Diese Art von struktureller Induktion ist gleichwohl gegen die herkömmliche ersetzbar. Der Beweis untergliedert sich dabei in zwei Induktionsbeweise. Es wird zunächst $A(m,0)$ für jedes m bestätigt und $A(m,0)$ anschließend

als Anfang für $A(m, n)$ benutzt. Der zweite Induktionsbeweis ist so gesehen durch m parametrisiert, was nichts Schlimmes ist, denn parametrisierte Argumentation kennzeichnet ja das übliche Vorgehen zur Bestätigung allquantifizierter Aussagen.

1.5. Modallogik

1.5.1. Das System K

Die Modallogik handelt von den Weisen, wie eine Aussage ausgeprägt sein kann, was durch modalisierende Operatoren ausgedrückt wird. Es gibt unterschiedliche Systeme der Modallogik. In vielen Systemen finden sich zwei Modalitäten, die *Notwendigkeit* der Aussage und die *Möglichkeit* der Aussage. Allerdings artikulieren diese Sprechweisen nur die *alethische* Deutung der Modalitäten. Je nach System und Anwendung sind unterschiedliche Deutungen der Modalitäten zuträglich.

Ist A die Aussage »Es regnet«, drückt $\Box A$ die Aussage »Es regnet notwendigerweise« und $\Diamond A$ die Aussage »Es regnet möglicherweise« aus.

Die Modallogik scheint wichtiger für Philosophen als für Mathematiker. Indessen kamen mit der Zeit Anwendungen in der Mathematik und der Informatik zum Vorschein. Dem Basiswissen besonders dienlich ist meines Erachtens die *dynamische Logik* mit ihrer engen Beziehung zum Hoare-Kalkül bzw. zum dijkstraschen wlp-Kalkül. Diese Kalküle geben uns die Mittel in die Hand, Algorithmen auf ihre Korrektheit hin zu untersuchen. Die Funktion, die der Algorithmus darstellt, bekommt hierzu eine *Spezifikation*, bestehend aus einer *Vorbedingung* und einer *Nachbedingung*. Man zeigt nun, dass der Algorithmus auch das tut, was er soll, dergestalt dass der Wert der Funktion immer die Nachbedingung erfüllt, sofern ihre Argumente die Vorbedingung erfüllen.

Der Modallogik kommt somit auch eine gewisse Bedeutung für die praktische Arbeit zu, nicht nur für höhere mathematische respektive philosophische Erwägungen und Betrachtungen.

Wir wollen uns zunächst mit dem *System K* beschäftigen, dem Grundsystem der *normalen Modallogiken*. Spezifischere Systeme entstehen durch Hinzunahme weiterer Axiome. Das System K enthält sämtliche Regeln und Axiome der klassischen Aussagenlogik. Hinzu kommt die Regel

$$\frac{\vdash A}{\vdash \Box A} \qquad \text{(Nezessisierungsregel)}$$

und das Schema

$$\vdash \Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B). \qquad \text{(Axiomenschema K)}$$

Wie gehabt induziert das Schema sogleich eine zulässige Regel,

$$\frac{\Gamma \vdash \Box(A \Rightarrow B)}{\Gamma \vdash \Box A \Rightarrow \Box B}. \qquad \text{(Regel K)}$$

Wichtig ist, dass nur Theoreme Nezessisierung erfahren dürfen. Dagegen ist

$$\frac{a \vdash a}{a \vdash \Box a} \qquad \text{(verboten)}$$

ein unzulässiger Schluss. So ist $a \Rightarrow \Box a$ kein Theorem. Man wird mit der Semantik für das System K leicht ein Gegenmodell dieser Formel finden.

Man definiert $\Diamond A$ als äquivalent zu $\neg\Box\neg A$.

Statt der einfachen Nezessisierung und Schema K kann man auch eine einzige allgemeine Nezessisierungsregel voraussetzen. Mit $n \geq 0$ lautet sie

$$\frac{A_1, \ldots, A_n \vdash B}{\Box A_1, \ldots, \Box A_n \vdash \Box B}.$$

Man beweist die Regel per Induktion über n als zulässig. Im Anfang $n = 0$ nimmt sie schlicht die Form der Nezessisierungsregel an. Der Induktionsschritt wird durch den Beweisbaum

$$\dfrac{\dfrac{\dfrac{\dfrac{A_1, \ldots, A_n, A_{n+1} \vdash B}{A_1, \ldots, A_n \vdash A_{n+1} \Rightarrow B}}{\Box A_1, \ldots, \Box A_n \vdash \Box(A_{n+1} \Rightarrow B)}\,\text{IV}}{\Box A_1, \ldots, \Box A_n \vdash \Box A_{n+1} \Rightarrow \Box B}\,\text{K} \qquad \dfrac{}{\Box A_{n+1} \vdash \Box A_{n+1}}}{\Box A_1, \ldots, \Box A_n, \Box A_{n+1} \vdash \Box B}\,\text{Modus ponens}$$

bestätigt.

Auch dem Fitch-Style wurde eine Darstellung der Schlüsse der Modallogik hinzugefügt. Sie ist meines Erachtens etwas schwieriger zu durchschauen als das Schließen von Sequenzen. Ich will in diesem Buch nicht näher darauf eingehen.

1.5.2. Das System S4

Das System S4 formt sich aus den Schemata KT4, siehe Tabelle 1.4.

Die Übersetzung nach Gödel-McKinsey-Tarski ist

$$\begin{aligned} P' &= \Box P, & (A \wedge B)' &= A' \wedge B', & (A \Rightarrow B)' &= \Box(A' \Rightarrow B'), \\ \bot' &= \bot, & (A \vee B)' &= A' \vee B', & (\neg A)' &= \Box\neg A'. \end{aligned}$$

Man deutet $\Box A$ als »A ist beweisbar«. Es ist A genau dann ein Theorem der intuitionistischen Logik, wenn A' ein Theorem im System S4 ist.

Tabelle 1.4.: Übersicht über zusätzliche Schemata

	Schema	Relationen	Formel
K	$\Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B)$	sämtliche	keine Einschränkung
T	$\Box A \Rightarrow A$	reflexive	$\forall x\colon R_{xx}$
B	$A \Rightarrow \Box\Diamond A$	symmetrische	$\forall x, y\colon R_{xy} \Rightarrow R_{yx}$
D	$\Box A \Rightarrow \Diamond A$	serielle	$\forall x\colon \exists y\colon R_{xy}$
4	$\Box A \Rightarrow \Box\Box A$	transitive	$\forall x, y, z\colon R_{xy} \wedge R_{yz} \Rightarrow R_{xz}$
5	$\Diamond A \Rightarrow \Box\Diamond A$	euklidische	$\forall x, y, z\colon R_{xy} \wedge R_{xz} \Rightarrow R_{yz}$

1.6. Beweistheoretische Überlegungen

1.6.1. Ableitbarkeit

Auf das Urteil $\Gamma \vdash A$ blickt man aus zwei Sichtweisen. Zum einen ein stellt es sich als syntaktisches Konstrukt dar, das Gegenstand eines formalen Systems ist. Zum anderen stellt es sich als die metasprachliche Aussage dar, dass A aus Γ ableitbar ist. Zur Schaffung von Klarheit wollen wir bei beweistheoretischen Betrachtungen näher zwischen den beiden Sichtweisen unterscheiden. Dazu notieren wir $\Gamma \rhd A$ für die Sequenz als syntaktisches Konstrukt, und im Unterschied dazu $\Gamma \vdash A$ für die metasprachliche Aussage.

Definition 1.1 (Ableitbarkeit).
Man nennt A aus der Formelmenge Γ ableitbar, kurz $\Gamma \vdash A$, wenn ein endliche Menge $\Gamma_0 \subseteq \Gamma$ existiert, so dass die Sequenz $\Gamma_0 \rhd A$ ableitbar ist.

Zu einer endlichen Formelmenge Γ_0 gilt $\Gamma_0 \vdash A$ demzufolge genau dann, wenn die Sequenz $\Gamma_0 \rhd A$ ableitbar ist, symbolisch

$$(\Gamma_0 \vdash A) \;\Leftrightarrow\; (\vdash \Gamma_0 \rhd A).$$

Die Endlichkeit ist genau deshalb erforderlich, weil Sequenzen nur für endliche Kontexte erklärt sind. Für die Ableitbarkeit besteht damit verbunden Kompaktheit, dergestalt dass $\Gamma \vdash A$ genau dann gilt, wenn ein endliches $\Gamma_0 \subseteq \Gamma$ mit $\Gamma_0 \vdash A$ existiert.

Weiterhin sind die Schlussregeln nun als metasprachliche Aussagen fassbar. So ist die Beseitigung der Subjunktion beschrieben durch

$$(\Gamma \vdash A \Rightarrow B) \wedge (\Gamma' \vdash A) \Rightarrow (\Gamma \cup \Gamma' \vdash B).$$

Anders als in der Regel dürfen Γ, Γ' hierbei auch unendlich sein. Eleganter finde ich hier aber, die Regel in der exportierten bzw. geschönfinkelten Form

$$(\Gamma \vdash A \Rightarrow B) \Rightarrow (\Gamma' \vdash A) \Rightarrow (\Gamma \cup \Gamma' \vdash B)$$

zu formulieren. Zur Anwendung einer Regel genügt daraufhin nämlich innerhalb des metalogischen Systems allein der Modus ponens. Außerdem kann eine Regel nun auch teilweise angewendet werden. So bestätigt der Schluss

$$\frac{\overline{(\emptyset \vdash A \Rightarrow B) \Rightarrow (\Gamma' \vdash A) \Rightarrow (\emptyset \cup \Gamma' \vdash B)} \qquad \emptyset \vdash A \Rightarrow B}{(\Gamma' \vdash A) \Rightarrow (\Gamma' \vdash B)}\,\mathrm{MP}$$

die Regel, laut der $\Gamma' \rhd B$ aus $\Gamma' \rhd A$ abgeleitet werden darf, als zulässig, sofern denn die Sequenz $\rhd A \Rightarrow B$ ableitbar ist. Spinnen wir die Vorgehensweise weiter fort, hält uns nichts davon ab, natürliches Schließen innerhalb der Metalogik selbst zu verwenden. Die Subjunktionseinführung

$$\frac{(\Gamma \vdash A) \rhd_{\mathrm{meta}} (\Gamma \vdash B)}{\rhd_{\mathrm{meta}} (\Gamma \vdash A) \Rightarrow (\Gamma \vdash B)}$$

präzisiert nun den Gedankengang, dass die Regel

$$\frac{\Gamma \rhd A}{\Gamma \rhd B}$$

als zulässig erkannt wird, indem $\Gamma \rhd B$ unter Annahme von $\Gamma \rhd A$ abgeleitet wird. Damit fallen die metalogischen Überlegungen selbst unter die Formalisierung. Dies wird man spätestens dann nicht mehr leugnen können, wenn die Metalogik durch einen Beweisassistenten bereitgestellt wird.

Insoweit dem Kalkül und der Metalogik dieselbe Logik zugrunde liegen soll, dürfen Schlussregeln nach ihrer Herleitung auch in der Metalogik Verwendung finden. Es scheint, als würde sich die Abgrenzung zwischen der Logik und der Metalogik ein Stück weit auflösen, was plausibel sein mag, wenn es sich denn eigentlich um dieselbe Logik handelt. Dieser Gedankengang macht allerdings eine Aussage über die Metalogik, muss also in der Metametalogik stattfinden.

Ich will kurz einen logischen Kalkül aufzeigen, der seine eigene Metalogik umfasst und mit vergleichsweise wenig Aufwand im Computer implementiert werden kann. Zum besseren Verständnis wird das Urteil $\Gamma \vdash A$ als modale Aussage $\Box_\Gamma A$ gedeutet. Die einzigen beiden Schlussregeln sind der Modus ponens und die Nezessisierung,

$$\frac{\vdash A \Rightarrow B \qquad \vdash A}{\vdash B}, \qquad \frac{\vdash A}{\vdash \Box_\Gamma A}.$$

Auf der obersten Ebene arbeitet der Kalkül dementsprechend allein mit Theoremen, gewährt nicht die Abhängigkeit von Annahmen. Die Regeln des natürlichen Schließens werden nun als Axiome oder aus den Axiomen herleitbare Theoreme formuliert. Die diesbezüglichen Axiome sind

$$\Box_\Gamma(A \Rightarrow B) \Rightarrow \Box_\Gamma A \Rightarrow \Box_\Gamma B, \qquad A \Rightarrow B \Rightarrow A \wedge B,$$
$$\Box_{\Gamma\cup\{A\}} B \Rightarrow \Box_\Gamma(A \Rightarrow B), \qquad A \wedge B \Rightarrow A$$

und so weiter. Die Einführung von Grundsequenzen gestattet das Axiom $\Box_{\{A\}}A$. Das zusätzliche Axiom $\Box_\emptyset A \Rightarrow A$ drückt den Umstand aus, dass Theoreme in die metalogische Ebene verschoben werden dürfen. Letztlich braucht es zur Kodierung der Abschwächungsregel für $\Gamma \subseteq \Delta$ noch das Axiom $\Box_\Gamma A \Rightarrow \Box_\Delta A$. Insofern Γ, Δ als Listen formuliert werden, gelten analoge Axiome auch für die Fälle, dass Δ als Umordnung oder Kontraktion aus Γ hervorgeht. Möchte man nun Sequenzen unter Annahme von Sequenzen herleiten, muss auf der dritten Ebene gearbeitet werden. Dafür sind zunächst als Regeln fungierende Theoreme wie

$$\Box_\Delta\Box_\Gamma(A \wedge B) \Rightarrow \Box_\Delta\Box_\Gamma A$$

herzuleiten.

Vermittels der Einführung von Grundsequenzen, der Abschwächungsregel und des Modus ponens gelangt man zu den drei Einsichten

$$A \in \Gamma \Rightarrow (\Gamma \vdash A), \qquad \text{(Extensivität)}$$
$$\Gamma \subseteq \Gamma' \wedge (\Gamma \vdash A) \Rightarrow (\Gamma' \vdash A), \qquad \text{(Monotonie)}$$
$$(\Gamma \cup \{A\} \vdash B) \wedge (\Gamma \vdash A) \Rightarrow (\Gamma \vdash B). \qquad \text{(Schnittregel)}$$

Eine Darstellung aus der Sichtweise der Mengenlehre fördert die Einordnung dieser Sachverhalte. Als Hilfsmittel dient hierbei der *tarskische Konsequenzoperator*

$$\mathrm{Cn}(\Gamma) := \{A \mid \Gamma \vdash A\},$$

der Γ die Menge aller Formeln zuordnet, die aus Γ ableitbar sind. Man spricht auch vom *Inferenzoperator* oder vom *deduktiven Abschluss*. Dieser genügt den drei Axiomen

$$\Gamma \subseteq \mathrm{Cn}(\Gamma), \qquad \text{(Extensivität)}$$
$$\Gamma \subseteq \Gamma' \Rightarrow \mathrm{Cn}(\Gamma) \subseteq \mathrm{Cn}(\Gamma'), \qquad \text{(Monotonie)}$$
$$\mathrm{Cn}(\mathrm{Cn}(\Gamma)) \subseteq \mathrm{Cn}(\Gamma). \qquad \text{(Abgeschlossenheit)}$$

Das heißt, es handelt sich bei Cn um einen Hüllenoperator. Aus der Anfügung der Monotonie an die Extensivität folgt auch die Umkehrung der Abgeschlossenheit. Somit gilt mehr noch die Idempotenz

$$\mathrm{Cn}(\mathrm{Cn}(\Gamma)) = \mathrm{Cn}(\Gamma).$$

Die vormals vorhandene Schnittregel leitet sich aus den drei Axiomen ab. Sie wird zunächst in die Form

$$A \in \mathrm{Cn}(\Gamma) \Rightarrow \mathrm{Cn}(\Gamma \cup \{A\}) \subseteq \mathrm{Cn}(\Gamma)$$

gebracht. Es gilt $\Gamma \cup \{A\} \subseteq \mathrm{Cn}(\Gamma)$, denn $\Gamma \subseteq \mathrm{Cn}(\Gamma)$ gilt gemäß Extensivität, und $\{A\} \subseteq \mathrm{Cn}(\Gamma)$ gilt gemäß der Prämisse. Laut Monotonie und Idempotenz folgt

$$\mathrm{Cn}(\Gamma \cup \{A\}) \subseteq \mathrm{Cn}(\mathrm{Cn}(\Gamma)) = \mathrm{Cn}(\Gamma).$$

Obgleich die Abgeschlossenheit intuitiv klar erscheint, ist sie umgekehrt aus der Schnittregel zu gewinnen. Prämisse sei also $A \in \mathrm{Cn}(\mathrm{Cn}(\Gamma)$, womit $\mathrm{Cn}(\Gamma) \vdash A$. Das heißt aber, es existiert eine endliche Teilmenge

$$\{A_1, \ldots, A_n\} \subseteq \mathrm{Cn}(\Gamma) \text{ mit } \{A_1, \ldots, A_n\} \vdash A.$$

Mithin gilt $\Gamma \vdash A_k$ zu jedem k. Nun unternimmt man der Reihe nach den Schnitt der Form

$$\frac{\Gamma \cup \{A_1, \ldots, A_k\} \vdash A \qquad \Gamma \vdash A_k}{\Gamma \cup \{A_1, \ldots, A_{k-1}\} \vdash A}$$

zu den k von $k = n$ bis $k = 1$. Man gelangt damit schließlich zu $\Gamma \vdash A$, also zu der gewünschten Feststellung $A \in \mathrm{Cn}(\Gamma)$.

1.6.2. Die zulässige Ersetzungsregel

Mit der Äquivalenz zweier Aussagen verhält es sich in gewisser Weise wie mit einer Gleichheit. Und zwar vermittelt die *zulässige Ersetzungsregel*, eine Teilformel gegen eine zu ihr äquivalente Formel ersetzen zu dürfen, analog wie bei der Termumformung eines Teilterms. Sie ermöglicht die bequeme Äquivalenzumformung von Aussagen, womit man sie als besonders nützlich erachten darf.

Satz 1.7 (Zulässige Ersetzungsregel).
Es gelten die beiden gleichwertigen Regeln

$$\frac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash C(A) \Leftrightarrow C(B)}, \qquad \frac{\Gamma \vdash A \Leftrightarrow B \qquad \Gamma' \vdash C(A)}{\Gamma, \Gamma' \vdash C(B)}.$$

Beweis. Zunächst zur Gleichwertigkeit der beiden Regeln:

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash C(A) \Leftrightarrow C(B)}}{\Gamma \vdash C(A) \Rightarrow C(B)} \quad \Gamma' \vdash C(A)}{\Gamma, \Gamma' \vdash C(B)} \qquad \dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow B \quad \overline{C(A) \vdash C(A)}}{\Gamma, C(A) \vdash C(B)}}{\Gamma \vdash C(A) \Rightarrow C(B)} \quad \dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow B \quad \overline{C(B) \vdash C(B)}}{\Gamma, C(B) \vdash C(A)}}{\Gamma \vdash C(B) \Rightarrow C(A)}}{\Gamma \vdash C(A) \Leftrightarrow C(B)}$$

Wir führen nun eine strukturelle Induktion über den Formelaufbau durch. Die Behauptung wird hierbei in der Form

$$\frac{\Gamma \vdash F \Leftrightarrow F'}{\Gamma, C(F) \vdash C(F')}, \quad C(F) := C[P := F], \quad C(F') := C[P := F']$$

geschrieben. Weil F, F' vertauscht werden dürfen, erhält man somit auch die umgekehrte Folgerung, so dass die Äquivalenz von $C(F)$ und $C(F')$ hergestellt wird. Wir definieren die Abkürzungen

$$A := C(F), \quad A' := C(F'), \quad B := D(F), \quad B' := D(F').$$

Zunächst die Basisfälle. Die Formeln $C := \bot$, $C := \top$ und $C := Q$ mit atomarer Variable $Q \neq P$ bleiben von der Substitution unbetroffen und sind offenkundig zu sich selbst äquivalent. Für $C = P$ erhält man schlicht die Prämisse.

Zum Induktionsschritt. Man hat nun

$$(C \wedge D)[P := F] \iff C[P := F] \wedge D[P := F] \iff A \wedge B$$

usw. Induktionsvoraussetzung sei also $\Gamma \vdash A \Leftrightarrow A'$ und $\Gamma \vdash B \Leftrightarrow B'$. Zu zeigen ist

$$\Gamma, A \wedge B \vdash A' \wedge B', \qquad \Gamma, A \Rightarrow B \vdash A' \Rightarrow B', \qquad \Gamma, \forall x\colon A \vdash \forall x\colon A',$$
$$\Gamma, A \vee B \vdash A' \vee B', \qquad \Gamma, A \Leftrightarrow B \vdash A' \Leftrightarrow B', \qquad \Gamma, \exists x\colon A \vdash \exists x\colon A'$$

und $\Gamma, \neg A \vdash \neg A'$. Es findet sich:

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow A'}{\Gamma, A' \vdash A} \quad \dfrac{}{\neg A \vdash \neg A}}{\Gamma, \neg A, A' \vdash \bot}}{\Gamma, \neg A \vdash \neg A'} \qquad \dfrac{\dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow A'}{\Gamma \vdash A \Rightarrow A'} \quad \dfrac{\overline{A \wedge B \vdash A \wedge B}}{A \wedge B \vdash A}}{\Gamma, A \wedge B \vdash A'} \quad \dfrac{\dfrac{\Gamma \vdash B \Leftrightarrow B'}{\Gamma \vdash B \Rightarrow B'} \quad \dfrac{\overline{A \wedge B \vdash A \wedge B}}{A \wedge B \vdash B}}{\Gamma, A \wedge B \vdash B'}}{\Gamma, A \wedge B \vdash A' \wedge B'}$$

Die Erstellung der restlichen Bäume sei dem Leser überlassen.

Für die Quantoren findet sich:

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow A'}{\Gamma \vdash A \Rightarrow A'} \quad \dfrac{\overline{\forall x\colon A \vdash \forall x\colon A}}{\forall x\colon A \vdash A}}{\Gamma, \forall x\colon A \vdash A'}}{\Gamma, \forall x\colon A \vdash \forall x\colon A'}\ x \notin \mathrm{FV}(\Gamma) \qquad \dfrac{\overline{\exists x\colon A \vdash \exists x\colon A} \quad \dfrac{\dfrac{\dfrac{\Gamma \vdash A \Leftrightarrow A'}{\Gamma \vdash A \Rightarrow A'} \quad \dfrac{}{A \vdash A}}{\Gamma, A \vdash A'}}{\Gamma, A \vdash \exists x\colon A'}}{\Gamma, \exists x\colon A \vdash \exists x\colon A'}\ x \notin \mathrm{FV}(\Gamma)$$

Es darf hierbei $x \notin \mathrm{FV}(\Gamma)$ vorausgesetzt werden, weil die gebundene Variable andernfalls ja vor der Betrachtung in eine frische umbenannt werden kann. □

Es wäre noch zu erwähnen, dass die Regel bei den modalisierenden Operatoren für gewöhnlich unzulässig ist. So ist in den modallogischen Systemen K, T, B, D, S4, S5 keine der Formeln

$$(a \Leftrightarrow b) \Rightarrow (\Box a \Leftrightarrow \Box b),$$
$$(a \Leftrightarrow b) \Rightarrow (\Diamond a \Leftrightarrow \Diamond b)$$

ein Theorem. Ein Gegenmodell ist jeweils schnell gefunden, dafür bedarf es nicht mehr als zwei Welten.

Es besteht hier aber ein wesentlicher Unterschied zwischen relativer Äquivalenz $\Gamma \vdash A \Leftrightarrow B$ und absoluter Äquivalenz $\vdash A \Leftrightarrow B$. Zu einer absoluten Äquivalenz gilt die Ersetzungsregel selbst in der Modallogik. Der Induktionsbeweis wird hierzu für leeres Γ erweitert um:

$$\dfrac{\dfrac{\dfrac{\vdash A \Leftrightarrow A'}{\vdash A \Rightarrow A'}}{\vdash \Box(A \Rightarrow A')}}{\vdash \Box A \Rightarrow \Box A'} \quad \dfrac{\dfrac{\dfrac{\vdash A \Leftrightarrow A'}{\vdash A' \Rightarrow A}}{\vdash \Box(A' \Rightarrow A)}}{\vdash \Box A' \Rightarrow \Box A}}{\vdash \Box A \Leftrightarrow \Box A'} \qquad \dfrac{\dfrac{\dfrac{\vdash A \Leftrightarrow A'}{\vdash \neg A \Leftrightarrow \neg A'}}{\vdash \Box \neg A \Leftrightarrow \Box \neg A'}}{\vdash \neg \Box \neg A \Leftrightarrow \neg \Box \neg A'}}{\vdash \Diamond A \Leftrightarrow \Diamond A'}$$

Der Leser möge diesbezüglich zunächst Abschnitt 1.5 studieren.

1.7. Zur Beweisführung

1.7.1. Widerspruchsbeweise

Beim *Beweis durch Widerspruch* widerlegt man eine Aussage, indem gezeigt wird, dass die Annahme der Aussage zu einem logischen Widerspruch führt. In manchen Situationen bietet diese Art der Argumentation eine große Hilfe. So schreibt der britische Mathematiker Godfrey Harold Hardy in seinem Essay *A Mathematician's Apology* die Worte

> »The proof is by *reductio ad absurdum*, and *reductio ad absurdum*, which Euclid loved so much, is one of a mathematician's finest weapons. It is a far finer gambit than any chess gambit: a chess player may offer the sacrifice of a pawn or even a piece, but a mathematician offers *the game*.«

Zur Schaffung von Klarheit muss man zunächst zwei inhaltlich verschiedene Arten des Widerspruchsbeweises unterscheiden. Präzisieren lässt sich diese Unterscheidung anhand der Regeln

$$\frac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}, \qquad \frac{\Gamma, \neg A \vdash \bot}{\Gamma \vdash A}.$$

Die linke Regel stellt die stets verfügbare Negationseinführung dar, die man auch als *Widerlegung durch Widerspruch* bezeichnen kann. In der rechten Regel, der klassischen *Reductio ad absurdum*, gelangt man zunächst per Negationseinführung von $\Gamma, \neg A \vdash \bot$ zu $\Gamma \vdash \neg\neg A$, und daraufhin zu $\Gamma \vdash A$. Die Beseitigung der Doppelnegation ist allerdings lediglich in der klassischen Logik verfügbar, in der intuitionistischen gilt sie dagegen als unzulässig.

Manche stellen die Regeln in einer Form dar, in der die Kontradiktion nicht explizit auftaucht. Wir erhalten sie als zulässige Regeln, indem der Einführung der Kontradiktion direkt ihre Beseitigung angeschlossen wird. Es findet sich

$$\frac{\Gamma, A \vdash \neg B \qquad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash \neg A}, \qquad \frac{\Gamma, \neg A \vdash \neg B \qquad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A}.$$

Wie zuvor ist die linke Form allgemein verfügbar, die rechte dagegen nur bei Vorhandensein der Beseitigung der Doppelnegation.

1.7.2. Klassische Kontraposition

Eine weitere Regel ist die umgekehrte Kontraposition

$$\frac{\Gamma \vdash \neg A \Rightarrow \neg B}{\Gamma \vdash B \Rightarrow A}.$$

Sie verlangt ebenfalls die klassische Logik. Mit ihr lässt sich nämlich die klassische Reductio ad absurdum herleiten:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Gamma, \neg A \vdash \bot}{\Gamma, \neg A, \top \vdash \bot}\,\text{Abschwächung}}{\Gamma, \neg A \vdash \neg\top}\,\text{Neg-Einf.}}{\Gamma \vdash \neg A \Rightarrow \neg\top}\,\text{Subj-Einf.}}{\Gamma \vdash \top \Rightarrow A}\,\text{bedenklich} \qquad \dfrac{}{\top \vdash \top}\,\text{Axiom}}{\Gamma, \top \vdash A}\,\text{Modus ponens}}{\Gamma \vdash A}\,\text{Kürzung der Tautologie}$$

Weil alle anderen Schlüsse unbedenklich sind, kann die umgekehrte Kontraposition als der bedenkliche Schritt identifiziert werden. In der klassischen Logik ist sie allerdings zulässig. Die Herleitung unter Verwendung der Beseitigung der Doppelnegation sei dem Leser überlassen. Als kleiner Tipp sei aber gegeben, dass die Einführung der Doppelnegation unter allen Umständen zulässig ist, wie man sich unschwer überzeugt.

Wir schreiben die Abkürzungen DNE für die Beseitigung der Doppelnegation, LEM für den Satz vom ausgeschlossenen Dritten und EFQ für ex falso quodlibet. Eine Alternative zu DNE bietet LEM zuzüglich EFQ. Manche Beweise verkürzen sich damit, andere verlängern sich. Tatsächlich lässt sich DNE aus LEM zuzüglich EFQ herleiten. Umgekehrt lässt sich sowohl LEM als auch EFQ aus DNE herleiten. Es tut sich die Frage auf, ob sich EFQ aus LEM herleiten lässt. Die Antwort darauf lautet nein. Eine ausführliche Untersuchung findet man in [6].

1.7.3. Notwendige und hinreichende Bedingungen

Manchmal trifft man auf die Ausdrucksweise, eine Bedingung B sei für eine Aussage A notwendig. Sie sagt aus, dass $\neg B$ zu $\neg A$ führt. Falls die Bedingung verletzt ist, kann die Aussage unmöglich gelten. Es liegt demnach die Implikation $\neg B \Rightarrow \neg A$ vor. Sie wird im Sinne der klassischen Logik verstanden. Man hat also

$$(A \Rightarrow B) \;\Leftrightarrow\; (B \text{ ist notwendig für } A).$$

Die Ausdrucksweise, eine Bedingung B sei für A hinreichend, sagt aus, dass B die Aussage A bereits nach sich zieht. Es liegt demnach die Implikation $B \Rightarrow A$ vor. Man hat also

$$(B \Rightarrow A) \;\Leftrightarrow\; (B \text{ ist hinreichend für } A).$$

Ist eine Bedingung B sowohl notwendig als auch hinreichend für A, liegt eine Äquivalenz vor. Man hat also

$$(A \Leftrightarrow B) \;\Leftrightarrow\; (B \text{ ist notwendig und hinreichend für } A).$$

Tabelle 1.5.: Verneinung von Aussagen

$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$	$\neg(\forall x\colon A(x)) \Leftrightarrow (\exists x\colon \neg A(x))$	$\neg\bot \Leftrightarrow \top$
$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$	$\neg(\exists x\colon A(x)) \Leftrightarrow (\forall x\colon \neg A(x))$	$\neg\top \Leftrightarrow \bot$
$\neg(A \Rightarrow B) \Leftrightarrow A \wedge \neg B$	$\neg(A \Leftrightarrow B) \Leftrightarrow A \oplus B$	$\neg\neg A \Leftrightarrow A$

1.7.4. Verneinung von Aussagen

Die Verneinung von All- und Existenzaussagen besitzt Umformungen, die analog zu den de morganschen Gesetzen vonstatten gehen. In Tabelle 1.5 ist die jeweilige Gesetzmäßigkeit hinter das jeweilige de morgansche Gesetz gestellt. Die Beweise sind in Tabelle 1.6 ausgeführt.

In Worten gilt eine Aussageform $A(x)$ genau dann nicht für jedes x, wenn sich mindestens ein x findet, für das $A(x)$ nicht gilt. Und es existiert genau dann kein $A(x)$ erfüllendes x, wenn $A(x)$ für jedes x nicht gilt.

In Verbindung mit der Ersetzungsregel ermöglichen die in Tabelle 1.5 aufgeführten Gesetze die äquivalente Umformung verneinter Aussagen, die solange fortgeführt werden kann, bis die Verneinung durch alle Junktoren und Quantoren der Aussage gedrungen ist. Mit der Entfaltung von Def. 3.1, die später auf S. 73 eingeführt wird, findet sich zum Beispiel die Umformung

$$\begin{aligned}
\neg(\forall x \in M\colon A(x)) &\iff \neg(\forall x\colon x \in M \Rightarrow A(x)) \\
&\iff (\exists x\colon \neg(x \in M \Rightarrow A(x))) \\
&\iff (\exists x\colon x \in M \wedge \neg A(x)) \\
&\iff (\exists x \in M\colon A(x)).
\end{aligned}$$

Es stellt sich somit heraus, dass die Verneinung der beschränkten Quantifizierung ebenfalls analog zu den de morganschen Gesetzen umgeformt wird. Es gilt

$$\begin{aligned}
&\neg(\forall x \in M\colon A(x)) \Leftrightarrow \exists x \in M\colon \neg A(x), \\
&\neg(\exists x \in M\colon A(x)) \Leftrightarrow \forall x \in M\colon \neg A(x).
\end{aligned}$$

Tabelle 1.6.: Zur Verneinung von All- und Existenzaussagen

$$\dfrac{\dfrac{\dfrac{\dfrac{}{1 \equiv \exists x\colon A(x)} \quad \dfrac{\dfrac{\dfrac{}{2 \equiv \forall x\colon \neg A(x)}}{2 \vdash \neg A(u)} \quad \dfrac{}{3 \equiv A(u)}}{2, 3 \vdash \bot}}{1, 2 \vdash \bot}}{2 \vdash \neg\exists x\colon A(x)}}{\vdash (\forall x\colon \neg A(x)) \Rightarrow \neg(\exists x\colon A(x))} \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{1 \equiv \neg\exists x\colon A(x)} \quad \dfrac{\dfrac{}{2 \equiv A(x)}}{2 \vdash \exists x\colon A(x)}}{1, 2 \vdash \bot}}{1 \vdash \neg A(x)}}{1 \vdash \forall x\colon \neg A(x)}}{\vdash \neg(\exists x\colon A(x)) \Rightarrow (\forall x\colon \neg A(x))}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{1 \equiv \exists x\colon \neg A(x)} \quad \dfrac{\dfrac{\dfrac{}{2 \equiv \neg A(u)} \quad \dfrac{\dfrac{}{3 \equiv \forall x\colon A(x)}}{3 \vdash A(u)}}{2, 3 \vdash \bot}}{1, 3 \vdash \bot}}{1 \vdash \neg\forall x\colon A(x)}}{\vdash (\exists x\colon \neg A(x)) \Rightarrow \neg(\forall x\colon A(x))} \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{1 \equiv \neg\forall x\colon A(x)} \quad \dfrac{\dfrac{\dfrac{\dfrac{}{2 \equiv \neg\exists x\colon \neg A(x)}}{2 \vdash \forall x\colon \neg\neg A(x)}}{2 \vdash \neg\neg A(x)}}{2 \vdash A(x)}}{2 \vdash \forall x\colon A(x)}}{1, 2 \vdash \bot}}{1 \vdash \neg\neg\exists x\colon \neg A(x)}}{1 \vdash \exists x\colon \neg A(x)}}{\vdash \neg(\forall x\colon A(x)) \Rightarrow (\exists x\colon \neg A(x))}$$

2. Semantik

2.1. Die klassische Semantik der Aussagenlogik

2.1.1. Die Erfüllungsrelation

Bislang trat die Logik in der Form eines formalen Systems in Erscheinung. Gegenstand eines solchen Systems sind im Allgemeinen *Wörter* einer formalen Sprache; im natürlichen Schließen sind das die Sequenzen. Einige Wörter, die *Axiome*, werden als gegeben vorausgesetzt. Unter Anwendung von *Ableitungsregeln*, auch *Inferenzregeln* genannt, das sind die Schlussregeln, leitet man aus bereits abgeleiteten Wörtern weitere Wörter der Sprache ab. In diesem Sinne handelt es sich um ein rein syntaktisches System.

Zum tieferen Verständnis muss man sich im Fortgang damit beschäftigen, welche inhaltliche Bedeutung den logischen Aussagen beigemessen wird. Der hierfür wesentliche Schritt besteht in der Definition einer passenden *Semantik*.

Gegenstand der Semantik der Logik ist der Wahrheitsgehalt von Aussagen. Man hat gefunden, dass es sich mit der Frage nach dem Wesen der Wahrheit schwierig verhält. Wir wollen daher an dieser Stelle gar nicht erst versuchen, sie zu ergründen. Stattdessen tritt Wahrheit für uns zunächst lediglich im leicht fassbaren Rahmen der zweiwertigen booleschen Algebra auf.

In der klassischen Semantik der Aussagenlogik herrscht das *Bivalenzprinzip*, das besagt, dass jede Aussage entweder *wahr* oder *falsch* sein muss, also einen von zwei Wahrheitswerten haben muss. Eine Aussage kann nicht *ein wenig wahr* oder *halbwegs wahr* sein, noch kann sie eine von mehreren unterschiedlichen gleichwertigen Wahrheiten haben. Wir schreiben kurz 0 für falsch und 1 für wahr. Enthält eine Formel logische Variablen, kommt ihr ein Wahrheitswert zu, sobald alle Variablen durch eine Interpretation mit einem Wahrheitswert belegt wurden.

Die Art und Weise, wie einer Formel ein Wahrheitswert zukommt, präzisiert die Erfüllungsrelation. Sie wird als Rekursion über den Formelaufbau definiert. Der Wahrheitswert einer Formel ist hierbei einzig und allein durch die Wahrheitswerte ihrer Teilformeln bestimmt.

Tabelle 2.1.: Wahrheitstafel der Junktoren

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
0	0	1	0	0	1	1
1	0	0	0	1	0	0
0	1	1	0	1	1	0
1	1	0	1	1	1	1

Definition 2.1 (Erfüllungsrelation).
Eine *Interpretation* I ist eine Funktion, die jede atomare logische Variable P mit einem Wahrheitswert $I(P) \in \{0, 1\}$ belegt. Man definiert $I \vDash A$, sprich »I erfüllt A«, rekursiv als

$$
\begin{aligned}
(I \vDash \bot) &:\Leftrightarrow 0, & (I \vDash A \wedge B) &:\Leftrightarrow ((I \vDash A) \wedge (I \vDash B)),\\
(I \vDash \top) &:\Leftrightarrow 1, & (I \vDash A \vee B) &:\Leftrightarrow ((I \vDash A) \vee (I \vDash B)),\\
(I \vDash P) &:\Leftrightarrow I(P), & (I \vDash A \Rightarrow B) &:\Leftrightarrow ((I \vDash A) \Rightarrow (I \vDash B)),\\
(I \vDash \neg A) &:\Leftrightarrow \neg(I \vDash A), & (I \vDash A \Leftrightarrow B) &:\Leftrightarrow ((I \vDash A) \Leftrightarrow (I \vDash B)).
\end{aligned}
$$

Die rechte Seite der jeweiligen Festsetzung ist metalogisch zu verstehen und per Wahrheitstafel definiert, siehe Tabelle 2.1. Die Schreibweise $I \nvDash A$ ist gleichbedeutend mit $\neg(I \vDash A)$. Eine Interpretation wird auch als *Modell* bezeichnet. Man nennt sie *Modell* einer Formel, falls sie die Formel erfüllt. Andernfalls spricht man von einem *Kontramodell* oder *Gegenmodell* der Formel.

Definition 2.2. Für einen Kontext $\Gamma = \{A_1, \ldots, A_n\}$ setzt man

$$(I \vDash \Gamma) :\Leftrightarrow (I \vDash A_1) \wedge \ldots \wedge (I \vDash A_n).$$

2.1.2. Gültigkeit einer Formel

Eine wichtige Rolle spielen *allgemeingültige* Formeln, die man in der Aussagenlogik auch als *Tautologien* bezeichnet. Sie sind immer wahr, unabhängig davon, mit welchem Wahrheitswert ihre logischen Variablen belegt werden.

Als allgemeinere Begrifflichkeit wollen wir auf einen Kontext Γ bezogen gültige Formeln A betrachten. Die Idee hierbei ist, dass wenn die Formeln des Kontextes als wahr angenommen werden, die Formel A ebenfalls wahr sein muss. Trifft dies auf A zu, schreibt man $\Gamma \vDash A$, gelesen »im Kontext Γ ist A gültig«, oder auch »Γ zieht A nach sich«. Die Bezeichnung *logische Folgerung* oder *logische Konsequenz* ist ebenfalls verbreitet.

Tabelle 2.2.: Wahrheitstafel der Tautologie zur Kontraposition

a	b	$\neg b$	$\neg a$	$a \Rightarrow b$	$\neg b \Rightarrow \neg a$	$(a \Rightarrow b) \Leftrightarrow (\neg b \Rightarrow \neg a)$
0	0	1	1	1	1	1
1	0	1	0	0	0	1
0	1	0	1	1	1	1
1	1	0	0	1	1	1

Definition 2.3 (Gültige Formel).
Eine Formel A heißt *gültig* im Kontext Γ, wenn jede Interpretation, die sämtliche Formeln von Γ erfüllt, auch A erfüllt. Metalogisch

$$(\Gamma \vDash A) \;:\Leftrightarrow\; \forall I\colon (I \vDash \Gamma) \Rightarrow (I \vDash A).$$

Eine im leeren Kontext gültige aussagenlogische Formel A nennt man wie gesagt Tautologie. Statt $\emptyset \vDash A$ schreibt man auch kurz $\vDash A$. Wie bei Sequenzen schreibt man auch $\Gamma, A, B \vDash C$ statt $\Gamma \cup \{A, B\} \vDash C$.

2.1.3. Wahrheitstafeln

Obgleich der Variablenvorrat unendlich groß sein darf, enthält eine Formel von den Variablen nur endlich viele. Insofern sind für eine Formel in einem Kontext auch nur endlich viele Interpretationen relevant. Sind insgesamt n Variablen vorhanden, sind es 2^n Interpretationen.

Eine Interpretation I mit der Auswertung $I \vDash A$ ist nichts anderes als eine Zeile der Wahrheitstafel der Formel A. Eine Formel ist genau dann tautologisch, wenn in der Ergebnisspalte in jeder Zeile eine 1 steht.

Ein Beispiel. Die Wahrheitstafel 2.2 bestätigt

$$\vDash (a \Rightarrow b) \Leftrightarrow (\neg b \Rightarrow \neg a).$$

Die Tafel führt zusätzlich die Teilformeln auf, was bei längeren Formeln recht mühselig erscheinen mag. Eine geschickte Methode zur Reduzierung des Schreibaufwands erspart die Teilformeln, und setzt ihre Wahrheitswerte dafür schlicht unter die Junktoren, denn Ziffern benötigen nicht viel Platz.

Die Prüfung einer logischen Folgerung per Wahrheitstafel wird ermöglicht durch die metalogische Beziehung

$$(A_1, \ldots, A_n \vDash A) \;\Leftrightarrow\; (\vDash A_1 \wedge \ldots \wedge A_n \Rightarrow A).$$

Nämlich findet sich die äquivalente Umformung

$$\begin{aligned}
(A_1, \ldots, A_n \vDash A) &\iff_{(1)} (\forall I\colon (I \vDash A_1) \wedge \ldots \wedge (I \vDash A_n) \Rightarrow (I \vDash A)) \\
&\iff_{(2)} (\forall I\colon (I \vDash A_1 \wedge \ldots \wedge A_n) \Rightarrow (I \vDash A)) \\
&\iff_{(3)} (\forall I\colon I \vDash A_1 \wedge \ldots \wedge A_n \Rightarrow A) \\
&\iff_{(4)} (\vDash A_1 \wedge \ldots \wedge A_n \Rightarrow A).
\end{aligned}$$

Hierbei gilt (1), (4) gemäß Def. 2.3, 2.2 und (2), (3) gemäß Def. 2.1.

2.1.4. Korrektheit des natürlichen Schließens

Jede Schlussregel und jedes Axiom besitzt eine semantische Entsprechung. Die Beweise dafür werden nun auf metalogischer Ebene mittels natürlichem Schließen selbst erbracht, was wie eine Art Zirkelschluss erscheinen mag.

Zu den Grundsequenzen und zur Abschwächungsregel findet sich:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{I \vDash \Gamma \cup \{A\}}\,1}{(I \vDash \Gamma) \wedge (I \vDash A)}\,\text{Def. 2.2}}{I \vDash A}}{\dfrac{(I \vDash \Gamma \cup \{A\}) \Rightarrow (I \vDash A)}{\Gamma \cup \{A\} \vDash A}\,\text{Def. 2.3}}\,{\sim}1 \qquad \dfrac{\dfrac{\dfrac{\Gamma \vDash A}{(I \vDash \Gamma) \Rightarrow (I \vDash A)}\,\text{Def. 2.3} \quad \dfrac{\dfrac{}{I \vDash \Gamma \cup \Gamma'}\,1}{I \vDash \Gamma}\,\text{Def. 2.2}}{I \vDash A}}{\dfrac{(I \vDash \Gamma \cup \Gamma') \Rightarrow (I \vDash A)}{\Gamma \cup \Gamma' \vDash A}\,\text{Def. 2.3}}\,{\sim}1$$

Die Prüfung der restlichen Entsprechungen sei dem Leser überlassen.

Satz 2.1 (Korrektheit des natürlichen Schließens).
Ist die Sequenz $\Gamma \vdash A$ ableitbar, so muss auch $\Gamma \vDash A$ gelten.

Beweis. Strukturelle Induktion über die Konstruktion von Beweisbäumen. Induktionsanfänge sind die semantischen Entsprechungen der Grundsequenzen. Induktionsschritte sind die semantischen Entsprechungen der Schlussregeln. Die Beweise der Entsprechungen wurden bereits diskutiert. □

Tabelle 2.3.: Einbettung in die gewöhnliche Algebra

Modern	$\neg a$	$a \wedge b$	$a \vee b$	$a \Rightarrow b$
Boole	$1 - a$	ab	$a + b(1 - a)$	$1 - a + ab$

2.1.5. Logische Äquivalenz

Definition 2.4 (Äquivalente Formeln).
Die Äquivalenz zweier Formeln A, B ist definiert als

$$(A \equiv B) \;:\Leftrightarrow\; (\vDash A \Leftrightarrow B).$$

Eine Äquivalenz besteht genau dann, wenn jede der beiden Formeln eine logische Folgerung der anderen ist. Das heißt, es besteht die metalogische Beziehung

$$(\vDash A \Leftrightarrow B) \;\Leftrightarrow\; (A \vDash B) \wedge (B \vDash A).$$

Mit den semantischen Entsprechungen der Schlussregeln findet sich nämlich:

$$\dfrac{\dfrac{\vDash A \Leftrightarrow B}{\vDash A \Rightarrow B} \quad \dfrac{}{A \vDash A}}{A \vDash B} \qquad \dfrac{\dfrac{A \vDash B}{\vDash A \Rightarrow B} \quad \dfrac{B \vDash A}{\vDash B \Rightarrow A}}{\vDash A \Leftrightarrow B}$$

Satz 2.2. Es ist $A \equiv B$ eine Äquivalenzrelation. Das heißt, es gilt

$$A \equiv A, \qquad \text{(Reflexivität)}$$
$$(A \equiv B) \Rightarrow (B \equiv A), \qquad \text{(Symmetrie)}$$
$$(A \equiv B) \wedge (B \equiv C) \Rightarrow (A \equiv C). \qquad \text{(Transitivität)}$$

Der Beweis sei dem Leser als kleine Übung überlassen.

Der Satz 2.2 vermittelt, dass Formeln mit Äquivalenzen so umgeformt werden dürfen, wie Terme mit Termumformungen. Es finden sich im Fortgang eine Reihe von grundlegenden Äquivalenzen, die Regeln der *booleschen Algebra*. Sie wurde erstmals in der Mitte des 19. Jahrhunders vom britischen Mathematiker George Boole in seiner Abfassung *The Mathematical Analysis of Logic* und seinem späteren Buch *An Investigation of The Laws of Thought* beschrieben. Boole beschreibt allerdings, anders als heute üblich, eine Einbettung der logischen Operationen in die gewöhnliche Algebra, siehe Tabelle 2.3.

Logische Äquivalenz im absoluten Sinne ist nicht unter allen Umständen der Weisheit letzter Schluss. In der Schaltalgebra tun sich Problemstellungen auf, wo

Tabelle 2.4.: Die Regeln der booleschen Algebra.

Konjunktion	Disjunktion	Bezeichnung
$A \wedge 0 \equiv 0$	$A \vee 1 \equiv 1$	Extremalgesetze
$A \wedge \neg A \equiv 0$	$A \vee \neg A \equiv 1$	Komplementärgesetze
$A \wedge A \equiv A$	$A \vee A \equiv A$	Idempotenzgesetze
$A \wedge 1 \equiv A$	$A \vee 0 \equiv A$	Neutralitätsgesetze
$A \wedge B \equiv B \wedge A$	$A \vee B \equiv B \vee A$	Kommutativgesetze
$A \wedge (B \wedge C) \equiv (A \wedge B) \wedge C$	$A \vee (B \vee C) \equiv (A \vee B) \vee C$	Assoziativgesetze
$\neg(A \wedge B) \equiv \neg A \vee \neg B$	$\neg(A \vee B) \equiv \neg A \wedge \neg B$	De morgansche Gesetze
$A \wedge (A \vee B) \equiv A$	$A \vee (A \wedge B) \equiv A$	Absorptionsgesetze

der Wahrheitswert einer Formel A in einzelnen Zeilen der Wahrheitstafel keine Rolle spielt, man spricht von *Don't-Care-Zellen*. Es sei X eine Formel, die genau in diesen Zeilen wahr ist, in allen anderen falsch. Man möchte A nun beispielsweise zu B vereinfachen. Unter normalen Umständen sollte diese Vereinfachung die Äquivalenz $A \equiv B$ einhalten. Bei Vorhandensein der irrelevanten Zellen genügt jedoch die weniger strenge Forderung

$$\neg X \vDash A \Leftrightarrow B.$$

Wir haben es hier mit einer relativen Äquivalenz zu tun. Auch bei ihr handelt es sich um eine Äquivalenzrelation, wobei X beliebig ist, aber fest sein muss.

2.1.6. Die Einsetzungsregel

Satz 2.3 (Einsetzungsregel).
Ist die Formel A allgemeingültig, so führt die simultane Ersetzung einiger atomarer Variablen durch Formeln bei ihr zu einer weiteren allgemeingültigen Formel. Metalogisch

$$(\vDash A) \Rightarrow (\vDash A[P_1 := B_1, \ldots, P_n := B_n]).$$

Beweis. Die Bestimmung von $I \vDash A[\ldots]$ läuft in derselben Weise ab wie die von $I \vDash A$, außer dass $I(P_k)$ durch $I \vDash B_k$ zu ersetzen ist. Sofern A allgemeingültig ist, gilt $I \vDash A$ unabhängig davon, ob $I(P_k)$ wahr oder falsch ist. Ergo muss $I \vDash A[\ldots]$ unabhängig davon gelten, ob $I \vDash B_k$ wahr oder falsch ist. □

Ich mag daran erinnern, dass bei der Substitution *jedes* Vorkommen der Variable durch dieselbe Formel zu ersetzen ist.

Zum Beispiel erhält man zu der simultanen Ersetzung $a := A$ und $b := B$ aus

$$\vDash (a \Rightarrow b) \Leftrightarrow (\neg b \Rightarrow \neg a) \quad \text{das Schema} \quad \vDash (A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A).$$

Mit dem Vollständigkeitssatz infolge das Theoremschema

$$\vdash (A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A).$$

Hiermit gewinnt man kurzum die Regeln

$$\frac{\Gamma \vdash A \Rightarrow B}{\Gamma \vdash \neg B \Rightarrow \neg A}, \qquad \frac{\Gamma \vdash \neg B \Rightarrow \neg A}{\Gamma \vdash A \Rightarrow B}.$$

Die Vorgehensweise ermöglicht summa summarum die Auffindung von zulässigen Schlussregeln durch geistloses Ausfüllen von Wahrheitstafeln, was allerdings auf die klassische Aussagenlogik beschränkt bleibt.

2.1.7. Wahrheitsfunktionen

Mit der Bindung ihrer atomaren Variablen gehen aus den Formeln der Aussagenlogik Wahrheitsfunktionen hervor. Die Disjunktion zweier Aussagen wird zum Beispiel vermittelt durch die Funktion

$$f\colon \{0,1\} \times \{0,1\} \to \{0,1\}, \quad f(a,b) := (a \vee b).$$

Jede Wahrheitsfunktion in n Variablen ist durch ihre Wahrheitstafel charakterisiert, die aus 2^n Zeilen besteht, da der Definitionsbereich so viele unterschiedliche Tupel enthält. Ein Tupel von Wahrheitswerten wird auch als Bitfolge betrachtet.

Zwei Formeln A, B sind genau dann äquivalent, wenn sie durch dieselben Interpretationen erfüllt werden, sich also in der Wahrheitstafel gleich verhalten. Man überzeugt sich davon unschwer mit der metalogischen Umformung

$$\begin{aligned}(A \equiv B) &\iff (\vDash A \Leftrightarrow B) \iff (\forall I\colon I \vDash A \Leftrightarrow B)\\ &\iff (\forall I\colon (I \vDash A) \Leftrightarrow (I \vDash B)).\end{aligned}$$

Demnach charakterisieren äquivalente Formeln dieselbe Wahrheitsfunktion. Man kann dies auch so betrachten, dass die Wahrheitsfunktion die Äquivalenzklasse all ihrer Formeln repräsentiert. Unter den Formeln gibt es nun einen besonderen Vertreter, die *disjunktive Normalform*, kurz DNF, die eigentlich nichts anderes als eine

direkte Kodierung der Wahrheitstafel ist. Insofern ließt sich an der Wahrheitstafel unmittelbar die DNF der Formel ab.

Die Normalform einer Formel lässt sich unter Umständen vereinfachen. Bei der DNF der Disjunktion ist bereits klar, welche Formel Resultat der Vereinfachung sein müsste. Es findet sich

$$\begin{aligned}&(a \wedge \neg b) \vee (\neg a \wedge b) \vee (a \wedge b) \equiv (a \wedge \neg b) \vee ((\neg a \vee a) \wedge b)\\ &\equiv (a \wedge \neg b) \vee (1 \wedge b) \equiv (a \wedge \neg b) \vee b \equiv (a \vee b) \wedge (\neg b \vee b)\\ &\equiv (a \vee b) \wedge 1 \equiv a \vee b.\end{aligned}$$

In der Schaltalgebra ist die Vereinfachung der DNF von großer Wichtigkeit, da sie den maßgeblichen Schritt zur Ermittelung von Schaltungen mit einer minimalen Zahl von Gattern darstellt. Aus dieser Anforderung heraus wurden systematische Verfahren zur Vereinfachung entwickelt. Für wenige Variablen stellt das sogenannte Karnaugh-Veitch-Diagramm ein geschicktes Hilfsmittel dar.

Mit dem Verfahren nach Quine und McCluskey lassen sich Formeln mit beliebig vielen Variablen mit einem Computer automatisch vereinfachen. Der Algorithmus ist allerdings von exponentieller Laufzeit, dessen Ausführung also von einer kombinatorischen Explosion überschattet. Eine wesentliche Verbesserung darf man auch nicht erwarten, da die Problemstellung der Vereinfachung als NP-vollständig befunden wurde. Für die Formeln mit sehr vielen Variablen liegt das Wissen über die beste Vereinfachung somit im Schleier der Dunkelheit.

2.2. Die klassische Semantik der Logik erster Stufe

2.2.1. Strukturen

Wie in der Aussagenlogik soll den Formeln eine Bedeutung zukommen. Eine Formel der Prädikatenlogik ist allerdings komplizierter als eine Formel der Aussagenlogik. Statt nur atomare Variablen mit Wahrheitswerten zu belegen, sind nun unterschiedliche Arten von Symbolen vorhanden, die mit einer Bedeutung versehen werden müssen.

Es gibt erstens die logischen Verknüpfungen, denen eine für die Logik spezifische feste Funktion zukommt. Ebenso ist die Gleichheit zweier Terme fest definiert. Zweitens können in einer Formel Symbole für Funktionen und Relationen vorkommen, darunter fallen auch Konstanten. Sie werden je nach gewählter Struktur anders mit Funktionen und Relationen belegt. Drittens darf eine Formel freie Variablen enthalten. Sie werden durch eine extra definierte Belegungsfunktion mit

Werten belegt. Bei der Interpretation einer quantifizierten Teilformel wird die Belegungsfunktion modifiziert.

Je nach Festlegung des logisch-mathematischen Systems und dessen inhaltlicher Deutung haben die mathematischen Terme Werte in einer bestimmten Menge U, die wir *Universum*, *Domäne*, *Träger*, *Grundbereich* oder *Diskursuniversum* nennen. Die Elemente von U nennen wir *Individuen* oder *Objekte*. Ich möchte noch einmal daran erinnern, dass der beschriebene Kalkül der Prädikatenlogik die Menge U als nichtleer fordert.

Eine *Struktur* M sei das Universum U zusammen mit einer Interpretationsfunktion, die jedem Funktionssymbol f von n Argumenten eine Funktion $f^M\colon U^n \to U$ und jedem Relationssymbol R von n Argumenten eine Relation $R^M\colon U^n \to \{0,1\}$ zuordnet. Konstanten deuten wir als Funktionen von null Argumenten.

Man kann eine Struktur auf unterschiedliche Art kodieren, wobei unter Umständen die Hilfsbegriffe *Signatur* und *Ariätsfunktion* gebraucht werden. Die Signatur beschreibt hierbei die Menge der Funktionssymbole und die Menge Relationssymbole. Die Aritätsfunktion ordnet jedem Funktionssymbol und jedem Relationssymbol seine Stelligkeit zu, das ist die Anzahl der Argumente.

Beispiel 1. Ein Kalkül für Halbordnungen. Zusätzlich zu den Regeln und Axiomen der Prädikatenlogik sollen als mathematische Axiome demnach noch die Axiome für Halbordnungen gelten. Die Signatur enthalte dazu lediglich das Relationssymbol $\leq$. Eine mögliche Struktur M ist die Menge $U := \mathcal{P}(G)$ zu irgendeiner Grundmenge G, zusammen mit der Setzung $\leq^M \;:=\; \subseteq$, so dass je zwei Mengen $A, B \in U$ ein Wahrheitswert $A \subseteq B$ zugeordnet wird. Man notiert so eine Struktur üblicherweise einfach schludrig als $(U, \leq^M)$, konkret $(\mathcal{P}(G), \subseteq)$.

Beispiel 2. Ein Kalkül für Gruppen. Zusätzlich zu den Regeln und Axiomen der Prädikatenlogik sollen als mathematische Axiome demnach noch die Axiome für Gruppen gelten. Die Signatur enthalte dazu ein Konstantensymbol für das neutrale Element, die Verknüpfung von Elementen als zweistelliges Funktionssymbol und die Invertierungsfunktion als einstelliges Funktionssymbol. Eine Struktur notiert man dann als $(G, \cdot, e, \mathrm{inv})$. Eine mögliche Struktur ist $(\mathbb{R}, +, 0, \mathrm{inv})$ mit $\mathrm{inv}(x) := -x$, die Gruppe der reellen Zahlen bezüglich Addition. Eine andere mögliche Struktur ist $(S(X), \circ, \mathrm{id}, \mathrm{inv})$ mit $\mathrm{inv}(f) := f^{-1}$, die Gruppe der bijektiven Selbstabbildungen auf der Menge X bezüglich der Verkettung, auch symmetrische Gruppe genannt. Es ist hier id die identische Abbildung und f^{-1} die Umkehrabbildung von f.

Beispiel 3. Ein Kalkül für die Mengenlehre. Der unerfahrene Leser mag diesbezüglich zunächst das Kapitel über Mengenlehre studieren und später hierher zurückkommen. Das Diskursuniversum soll das Mengenuniversum sein. Nun verhält

es sich allerdings so, dass die Zusammenfassung aller denkbaren Mengen eine echte Klasse bildet, die keine Menge mehr sein kann. Man könnte die Definition des Diskursuniversums so modifizieren, dass es eine beliebige Klasse sein darf, was aber dazu führen würde, dass dieses nicht mehr ohne Ausnahme als Element eines Mengensystems auftreten darf. Wir gehen dieser Komplikation durch die Betrachtung des Grothendieckuniversums $U = V_\kappa$ für ZFC oder $U = V_{\kappa+1}$ für die Morse-Kelly-Mengenlehre aus dem Weg, wobei mit den V_α die kumulative Hierarchie und mit κ eine als Ordinalzahl betrachtete unerreichbare Kardinalzahl gemeint ist. Die Einzelheiten erläutert Shulman [32]. Einziges Relationssymbol ist nun $\in$ für die Elementrelation, und Funktionssymbole gibt es gar keine. Das mag verwirrend erscheinen, wo die Mengenlehre doch von Relationen und Funktionen durchdrungen ist. Diese stellen bei näherer Betrachtung aber Objekte des Universums dar.

Aber ist die leere Menge nicht ein Konstantensymbol, die Inklusion $\subseteq$ nicht ein weiteres Relationssymbol und Operatoren wie $\cap$ und $\cup$ Funktionssymbole? An sich stimmt das. Das formale System der Mengenlehre kann um diese Symbole erweitert werden, die minimalistische Fassung des Systems kommt jedoch allein mit $\in$ aus. Alle weiteren Symbole werden durch ihre Definition ersetzt, obgleich dadurch unter Umständen exorbitant lange Formeln entstehen. Zum Beispiel wird die Formel $A \cap B \subseteq A \cup B$ expandiert zu

$$\forall x\colon x \in A \wedge x \in B \Rightarrow x \in A \vee x \in B.$$

2.2.2. Interpretationen

Zur Interpretation eines Terms t mit freien Variablen betrachtet man außerdem noch eine *Belegung* β, die jeder Variablen x einen Wert $\beta(x) \in U$ zuordnet. Geläufig ist ebenfalls die Bezeichnung *Zuweisung*.

Man nennt nun $I = (M, \beta)$ eine *Interpretation* der Formel A. Wir definieren $I(t)$ für einen Term t rekursiv über den Termaufbau als

$$\begin{aligned} I(x) &:= \beta(x), \\ I(f(t_1, \ldots, t_n)) &:= f^M(I(t_1), \ldots, I(t_n)), \end{aligned}$$

wobei x eine Variable und f ein n-stelliges Funktionssymbol ist. Die Belegung mit einer modifiziert belegten Variable wird diesbezüglich definiert als

$$\beta[y{:=}u](x) := \begin{cases} u, & \text{wenn } x = y, \\ \beta(x), & \text{sonst.} \end{cases}$$

Für $I = (M, \beta)$ sei $I[y{:=}u] := (M, \beta[y{:=}u])$.

Um Missverständnissen aus dem Weg zu gehen, sollte ich erwähnen, dass die Begriffe *Variable* und *Atom* in der Prädikatenlogik anders verstanden werden als in der Aussagenlogik. Die atomaren Variablen der Aussagenlogik entsprechen nullstelligen Relationssymbolen, denen bei einer Interpretation Relationen von null Argumenten zukommt, das sind schlicht die konstanten Wahrheitswerte. Dagegen meint Variable in der Prädikatenlogik üblicherweise eine Individuenvariable. Eine atomare Formel ist in der Prädikatenlogik nicht nur ein nullstelliges Relationssymbol, sondern jede Applikation eines Relationssymbols und jede Gleichung.

Die Erfüllungsrelation definiert man analog zur Aussagenlogik, also

$$(I \vDash A \wedge B) \;:\Leftrightarrow\; (I \vDash A) \wedge (I \vDash B)$$

und so weiter. Zusätzlich setzt man

$$\begin{aligned}
&(I \vDash t_1 = t_2) :\Leftrightarrow I(t_1) = I(t_2),\\
&(I \vDash R(t_1, \ldots, t_n)) :\Leftrightarrow R^M(I(t_1), \ldots, I(t_n)),\\
&(I \vDash \forall x\colon A) :\Leftrightarrow \forall u \in U\colon I[x{:=}u] \vDash A,\\
&(I \vDash \exists x\colon A) :\Leftrightarrow \exists u \in U\colon I[x{:=}u] \vDash A.
\end{aligned}$$

Man definiert hiermit die semantische Folgerung analog zur Aussagenlogik als

$$(\Gamma \vDash A) \;:\Leftrightarrow\; \forall I\colon (I \vDash \Gamma) \Rightarrow (I \vDash A).$$

Wir nennen A *allgemeingültig*, falls $\emptyset \vDash A$, was wir wieder durch $\vDash A$ abkürzen wollen.

Sofern die Formel A geschlossen ist, das heißt, keine freien Variablen besitzt, spielt die Belegung der Variablen noch keine Rolle, sie entsteht eigentlich erst bei der rekursiven Bestimmung von $I \vDash A$ mit dem Einstieg in eine quantifizierte Formel. Bei einer einer geschlossenen Formel A darf man insofern $M \vDash A$ statt $M, \beta \vDash A$ schreiben. Gilt $M \vDash A$, nennt man M ein *Modell für* A.

Ich will erklären, warum die Interpretation der Quantoren so durchgeführt wurde wie beschrieben. Intuitiv erfüllt I die Formel $\forall x\colon A$ genau dann, wenn I die Formel $A[x := u]$ für jedes $u \in U$ erfüllt. Allerdings ist u kein Term der logischen Sprache, womit $A[x := u]$ keine Formel mehr wäre. Wir könnten nicht einmal für jedes u eine Konstante zur Sprache hinzufügen, weil U überabzählbar sein darf, während eine gewöhnliche Sprache nur abzählbar viele Symbole in Form von Zeichenketten umfassen kann. Die gewählte Ausweg aus der Problematik besteht darin, dass die Belegung $x := u$ erst einmal in der Interpretation gespeichert wird, statt sie direkt als Substitution auf die Formel anzuwenden.

2.2.3. Korrektheit des natürlichen Schließens

Wir müssen nun wieder aufzeigen, dass der Kalkül des natürlichen Schließens bezüglich der definierten Semantik korrekt ist. Dies geschieht wieder per struktureller Induktion über den Aufbau des Beweises. Für die bereits in der Aussagenlogik vorhandenen Regeln können wir die dort gemachte Argumentation unverändert übernehmen. Zu bestätigen verbleibt die Korrektheit der Regeln zur Einführung und Beseitigung der Quantoren, und ggf. noch die Korrektheit der Regeln zum Umgang mit Gleichheiten.

Satz 2.4. Es ist $I[x := I(t)] \vDash A$ äquivalent zu $I \vDash A[x := t]$.

Beweis. Man unternimmt hierzu eine strukturelle Induktion über den Formelaufbau von A. Die Induktionsvoraussetzungen sind jeweils unschwer zu verarbeiten. Im Wesentlichen verbleibt nun

$$I[x := I(t)](v) = I(v[x := t])$$

für jede Individuenvariable v zu zeigen. Man nimmt dazu die Fallunterscheidung zwischen $x = v$ und $x \neq v$ vor. Im ersten Fall gilt

$$I[x := I(t)](v) = I(t) = I(x[x := t]) = I(v[x := t]),$$

und im zweiten Fall gilt

$$I[x := I(t)](v) = I(v) = I(v[x := t]). \square$$

Satz 2.5. Aus $\Gamma \vDash \forall x\colon A$ folgt $\Gamma \vDash A[x := t]$.

Beweis. Laut der Prämisse muss zu jedem I mit $I \vDash \Gamma$ auch $I[x := u] \vDash A$ für jedes $u \in U$ gelten. Weitere Prämisse ist ein I mit $I \vDash \Gamma$. Die Allaussage wird mit diesem I und mit $u := I(t)$ spezialisiert. Zu zeigen verbleibt

$$(I[x := I(t)] \vDash A) \Rightarrow (I \vDash A[x := t]).$$

Diese gilt gemäß Satz 2.4. $\square$

Satz 2.6. Aus $\Gamma \vDash A[x := t]$ folgt $\Gamma \vDash \exists x\colon A$.

Beweis. Laut der Prämisse muss zu jedem I mit $I \vDash \Gamma$ auch $I \vDash A[x := t]$ gelten. Weitere Prämisse ist ein I mit $I \vDash \Gamma$. Die Allaussage wird mit diesem I spezialisiert. Zu zeigen verbleibt

$$(I \vDash A[x := t]) \Rightarrow \exists u \in U\colon (I[x := u] \vDash A).$$

Gewählt wird $u := I(t)$. Nun greift wieder Satz 2.4. $\square$

Satz 2.7. Aus $\Gamma \vDash A$ und $x \notin \mathrm{FV}(\Gamma)$ folgt $\Gamma \vDash \forall x\colon A$.

Beweis. Laut der Prämisse muss zu jedem I mit $I \vDash \Gamma$ auch $I \vDash A$ gelten. Weitere Prämisse ist ein I mit $I \vDash \Gamma$. Wegen $x \notin \mathrm{FV}(\Gamma)$ ist $I \vDash \Gamma$ äquivalent zu $I[x := u] \vDash \Gamma$. Demnach lässt sich die Allaussage mit $I[x := u]$ spezialisieren, womit man $I[x := u] \vDash A$ erhält. Ergo gilt

$$\forall u \in U\colon (I[x := u] \vDash A), \text{ also } I \vDash \forall x\colon A. \qquad \square$$

Satz 2.8. Aus $\Gamma \vDash \exists x\colon A$ und $\Gamma \cup \{A\} \vDash B$ mit $x \notin \mathrm{FV}(\Gamma \cup \{B\})$ folgt $\Gamma \vDash B$.

Beweis. Laut der ersten Prämisse existiert zu jedem I mit $I \vDash \Gamma$ ein $u \in U$ mit $I[x := u] \vDash A$. Laut der zweiten Prämisse muss zu jedem I mit $I \vDash \Gamma$ und $I \vDash A$ auch $I \vDash B$ gelten. Wegen $x \notin \mathrm{FV}(\Gamma \cup \{B\})$ ist $I \vDash \Gamma$ zu $I[x := u] \vDash \Gamma$ sowie $I \vDash B$ zu $I[x := u] \vDash B$ äquivalent. Dritte Prämisse ist ein I mit $I \vDash \Gamma$. Die erste Allaussage wird mit I spezialisiert, die zweite mit $I[x := u]$. Mit den vorliegenden Aussagen gelingt der Schluss auf $I[x := u] \vDash B$. Ergo gilt $I \vDash B$, wie gewünscht. $\square$

2.3. Semantik der Modallogik

2.3.1. Die relationale Semantik der Modallogik

Mit der Entwicklung der Modallogiken ging die Frage einher, wie modallogische Formeln semantisch interpretiert werden können. Besonders wichtig sind die nach Saul Kripke benannten *Kripke-Semantiken*, auch *relationale Semantiken* genannt. Die Interpretation der atomaren Aussagen wird hierbei durch eine Welt parametrisiert. Eine Aussage kann also in einer bestimmten Welt wahr sein, während sie in einer anderen falsch ist. Als Logik der Metasprache soll dabei die klassische Prädikatenlogik dienen.

Eine *Kripke-Struktur* sei ein Tripel $M = (W, R, V)$. Hierbei ist W eine Menge von Welten, R eine zweistellige Relation auf W, und V eine Wertung, die jeder atomaren Variable a parametrisiert durch die Welt $w \in W$ einen Wahrheitswert $V(w, a) \in \{0, 1\}$ zuordnet. Man nennt R die *Zugänglichkeitsrelation*, wobei $R(w, w')$ gedeutet wird als »w' ist von w aus zugänglich«. Die Relation muss nicht unbedingt symmetrisch sein, womit man unter Umständen nicht mehr von w' aus in die ursprüngliche Welt w zurückgelangt.

Das Paar $I = (M, w)$, bestehend aus einer Struktur und einer Welt, betrachten wir als Interpretation. Man legt die Erfüllung einer atomaren Variable a bezüglich einer Interpretation fest als

$$(M, w \vDash a) :\Leftrightarrow V(w, a).$$

Die Erfüllung der Junktoren wird analog zur klassischen Semantik definiert, also

$$(I \vDash A \wedge B) :\Leftrightarrow (I \vDash A) \wedge (I \vDash B)$$

und so weiter, wobei die rechte Seite gemäß den herkömmlichen Wahrheitstafeln ausgewertet wird. Für die beiden modalen Junktoren legt man nun fest

$$(M, w \vDash \Box A) :\Leftrightarrow \forall w' \in W : R(w, w') \Rightarrow (M, w' \vDash A),$$
$$(M, w \vDash \Diamond A) :\Leftrightarrow \exists w' \in W : R(w, w') \wedge (M, w' \vDash A).$$

In Worten ist die Aussage $\Box A$ genau dann wahr in der Welt w, wenn A in jeder von w aus zugänglichen möglichen Welt wahr ist. Es ist $\Diamond A$ genau dann wahr in der Welt w, wenn A in mindestens einer von w aus zugänglichen möglichen Welt wahr ist.

Die semantische Folgerung wird wieder in der herkömmlichen Form

$$(\Gamma \vDash A) \;:\Leftrightarrow\; \forall I : (I \vDash \Gamma) \Rightarrow (I \vDash A)$$

definiert. Ausgeschrieben lautet sie demnach

$$(\Gamma \vDash A) \;:\Leftrightarrow\; \forall M : \forall w \in W(M) : (M, w \vDash \Gamma) \Rightarrow (M, w \vDash A).$$

Man spricht gelegentlich auch von der *lokalen* semantischen Folgerung, um sie von einer weiteren, der selten diskutierten *globalen* semantischen Folgerung abgrenzen zu können. Zur Unterscheidung will ich die globale Folgerung als $\vDash^g$ notieren. Man definiert sie als

$$(\Gamma \vDash^g A) \;:\Leftrightarrow\; \forall M : (M \vDash \Gamma) \Rightarrow (M \vDash A),$$

wobei $M \vDash A$ für $\forall w \in W(M) : (M, w \vDash A)$ stehen soll.

Ist Γ die leere Menge, koinzidieren die lokale und die globale Folgerung. Das heißt, $\emptyset \vDash A$ gilt genau dann, wenn $\emptyset \vDash^g A$. Man schreibt wieder $\vDash A$ als Kurzform von $\emptyset \vDash A$ und sagt dazu, die Formel A sei *allgemeingültig*.

Bei der tieferen Untersuchung, welche Beziehungen zwischen Kalkül und Semantik bestehen, zerlegt man eine Kripke-Struktur $M = (W, R, V)$ in die zwei Teile $F = (W, R)$ und V. Man nennt F den *Rahmen*.

Gezeigt werden muss nun die Korrektheit des natürlichen Schließens für die Modallogik. Dies geschieht wieder per struktureller Induktion über den Formelaufbau. Bei den herkömmlichen Schlussregeln verläuft der Beweis identisch wie in der klassischen Semantik der Aussagenlogik. Zu bestätigen verbleibt insofern lediglich noch die Korrektheit der Nezessisierungsregel und die Allgemeingültigkeit des Axiomenschemas K.

Satz 2.9. Aus $\vDash A$ folgt $\vDash \Box A$.

Beweis. Gemäß der Prämisse gilt $M, w' \vDash A$ für jedes w'. Demnach gilt erst recht $R(w, w') \Rightarrow (M, w' \vDash A)$. Ergo gilt $\vDash \Box A$. Quod erat demonstrandum. □

Satz 2.10. Das Schema $\Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B)$ ist allgemeingültig.

Beweis. Seien M, w fest, aber beliebig. Zu zeigen ist

$$M, w \vDash \Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B),$$

was gemäß der Definition der Erfüllung äquivalent umgeformt wird zu

$$(M, w \vDash \Box(A \Rightarrow B)) \Rightarrow (M, w \vDash \Box A) \Rightarrow (M, w \vDash \Box B).$$

Zu zeigen ist daher $M, w' \vDash B$ unter Annahme von $R(w, w')$ und

$$\forall w' \colon R(w, w') \Rightarrow (M, w' \vDash A \Rightarrow B),$$
$$\forall w' \colon R(w, w') \Rightarrow (M, w' \vDash A).$$

Die beiden Allaussagen werden mit w' und $R(w, w')$ spezialisiert. Laut der ersten Aussage $M, w' \vDash A \Rightarrow B$ folgt $M, w' \vDash B$ aus $M, w' \vDash A$. Laut zweiten gilt $M, w' \vDash A$. Ergo gilt $M, w' \vDash B$. Quod erat demonstrandum. □

2.3.2. Die Standardübersetzung

Die relationale Semantik der diskutierten Modallogiken deutet bereits an, dass Modallogik etwas mit der Logik erster Stufe zu tun haben müsste. Es scheint, dass in gewisser Weise $\Box A$ der Allquantifizierung, und $\Diamond A$ der Existenzquantifizierung entspricht. Die im Folgenden beschriebene *Standardübersetzung* modallogischer Formeln in die Logik erster Stufe präzisiert diesen Gedankengang. Mithin stellen sich die diskutierten Modallogiken als Spezialfälle der Logik erster Stufe heraus. Auf diese Weise erhält man eine Möglichkeit, modallogische Theoreme mit Beweisassistenten zu formulieren, die lediglich die herkömmliche Logik verstehen.

Die Standardübersetzung ist rekursiv definiert als

$$\mathrm{ST}_x(a) := P_a(x), \qquad \mathrm{ST}_x(A \Rightarrow B) := \mathrm{ST}_x(A) \Rightarrow \mathrm{ST}_x(B),$$
$$\mathrm{ST}_x(\neg A) := \neg \mathrm{ST}_x(A), \qquad \mathrm{ST}_x(A \Leftrightarrow B) := \mathrm{ST}_x(A) \Leftrightarrow \mathrm{ST}_x(b),$$
$$\mathrm{ST}_x(A \wedge B) := \mathrm{ST}_x(A) \wedge \mathrm{ST}_x(B), \qquad \mathrm{ST}_x(\Box A) := \forall y \colon R(x, y) \Rightarrow \mathrm{ST}_y(A),$$
$$\mathrm{ST}_x(A \vee B) := \mathrm{ST}_x(A) \vee \mathrm{ST}_x(B), \qquad \mathrm{ST}_x(\Diamond A) := \forall y \colon R(x, y) \wedge \mathrm{ST}_y(A).$$

Die Formeln $\bot, \top$ bleiben fix, also $\mathrm{ST}_x(\bot) := \bot$ und $\mathrm{ST}_x(\top) := \top$.

Hierbei ist R ein zweistelliges Prädikatsymbol, und P_a zu jedem a ein einstelliges Prädikatsymbol. Eine Interpretation (M, w) mit Kripke-Struktur $M = (W, R, V)$ und Welt $w \in W$ lässt sich als Interpretation $(M, \beta[x := w])$ für die übersetzten Formeln deuten, dergestalt dass das Symbol P_a mit $w \mapsto V(w, a)$ interpretiert wird, und das Symbol R mit der Zugänglichkeitsrelation R. Demzufolge nimmt W die Rolle des Diskursuniversums ein. Die Belegung β spielt keine Rolle, da x die einzige freie Variable ist. Man gelangt nun zum folgenden Resultat.

Satz 2.11. Es gilt $M, w \vDash A$ genau dann, wenn $M, \beta[x := w] \vDash \mathrm{ST}_x(A)$.

Beweis. Per struktureller Induktion über den Formelaufbau von A. Es sei $\beta\frac{w}{x}$ Kurzschreibweise für $\beta[x := w]$. Den Induktionsanfang liefert die Umformung

$$(M, w \vDash a) \iff V(w, a) \iff (M, \beta\tfrac{w}{x} \vDash P_a(x)).$$

Zur Formel $A \wedge B$ liegt die Äquivalenz von $M, w \vDash C$ und $M, \beta\frac{w}{x} \vDash \mathrm{ST}_x(C)$ jeweils für $C = A$ und $C = B$ als Induktionsvoraussetzung vor. Es findet sich hiermit die äquivalente Umformung

$$\begin{aligned}(M, w \vDash A \wedge B) &\iff (M, w \vDash A) \wedge (M, w \vDash B)\\ &\iff (M, \beta\tfrac{w}{x} \vDash \mathrm{ST}_x(A)) \wedge (M, \beta\tfrac{w}{x} \vDash \mathrm{ST}_x(B))\\ &\iff M, \beta\tfrac{w}{x} \vDash \mathrm{ST}_x(A) \wedge \mathrm{ST}_x(B)\\ &\iff M, \beta\tfrac{w}{x} \vDash \mathrm{ST}_x(A \wedge B).\end{aligned}$$

Bei den restlichen herkömmlichen Junktoren verläuft die Argumentation analog. Zur Formel $\Box A$ liegt die Äquivalenz von $M, w' \vDash A$ und $M, \beta\frac{w'}{y} \vDash \mathrm{ST}_y(A)$ als Induktionsvoraussetzung vor. Es findet sich hiermit die äquivalente Umformung

$$\begin{aligned}(M, w \vDash \Box A) &\iff \forall w' \colon R(w, w') \Rightarrow (M, w' \vDash A)\\ &\iff \forall w' \colon R(w, w') \Rightarrow (M, \beta\tfrac{w'}{y} \vDash \mathrm{ST}_y(A))\\ &\iff \forall w' \colon (M, \beta\tfrac{w}{x}\tfrac{w'}{y} \vDash R(x, y)) \Rightarrow (M, \beta\tfrac{w}{x}\tfrac{w'}{y} \vDash \mathrm{ST}_y(A))\\ &\iff (M, \beta\tfrac{w}{x} \vDash \forall y \colon R(x, y) \Rightarrow \mathrm{ST}_y(A))\\ &\iff (M, \beta\tfrac{w}{x} \vDash \mathrm{ST}_x(\Box A)).\end{aligned}$$

Bei $\Diamond A$ verläuft die Argumentation analog. □

Mithin gilt in unmittelbarer Konsequenz der

Satz 2.12. Es gilt $\Gamma \models A$ genau dann, wenn $\mathrm{ST}_x(\Gamma) \models \mathrm{ST}_x(A)$.

Vermittels dieser Äquivalenz erhält man sogleich den

Satz 2.13. Aus $\vdash \mathrm{ST}_x(A)$ folgt $\models A$.

Kraft der Standardübersetzung in die Logik erster Stufe ergibt sich insofern wie gewünscht ein alternativer korrekter Kalkül für die diskutierten Modallogiken. Eine tiefergehende Untersuchung zur Standardübersetzung findet man in [21].

3. Mengenlehre

3.1. Grundbegriffe

3.1.1. Der Mengenbegriff

Eine *Menge* darf man sich wie einen Beutel vorstellen, der einzelne Objekte enthält. Die Objekte heißen *Elemente* der Menge. Jedoch gilt es hierbei zu beachten, dass es sich mit einer Menge nicht gänzlich wie mit einem Beutel verhält, in dem dasselbe Objekt mehrmals zu finden sein kann.

> »*Unter einer Menge verstehen wir jede Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.*«
> — Georg Cantor, 1895 (redigiert aus [27])

Obwohl blumig anmutend, fassen diese Worte das Konzept recht gut auf den Punkt. Wichtig ist hier das Wort *wohlunterschieden*, das uns zu verstehen gibt, dass ein Element nicht mehrmals in einer Menge enthalten sein kann. *Eine Menge ist genau dadurch festgelegt, welche Elemente sie enthält. Sie enthält Elemente weder mehrmals, noch in einer bestimmten Reihenfolge.*

Es gibt die *leere Menge*, notiert als $\emptyset$ oder $\{\}$. Man darf sie sich wie einen leeren Beutel vorstellen. Dagegen enthält die Menge $\{\emptyset\}$ genau ein Element. Es ist ein Beutel, der den leeren Beutel enthält.

Wir schreiben kurz $x \in A$ für »x ist ein Element von A«, auch »x gehört zu A« oder »x liegt in A«. Es steht $x \notin A$ für $\neg x \in A$, gelesen $\neg(x \in A)$.

Für eine endliche Menge $\{x_1, \ldots, x_n\}$ definiert man

$$x \in \{x_1, \ldots, x_n\} \;:\Leftrightarrow\; x = x_1 \vee \ldots \vee x = x_n.$$

3.1.2. Gleichheit von Mengen

Es wurde gesagt, eine Menge sei dadurch charakterisiert, welche Elemente sie enthält. Um diesen Sachverhalt näher zu erfassen, muss geklärt werden, wie es sich mit der Gleichheit von Mengen verhält. Intuitiv sieht man zwei Mengen als gleich

an, wenn sie exakt dieselben Elemente enthalten. Allerdings stellt sich die Frage, wie dies als logische Aussage wiedergegeben wird, weil man die Elemente der beiden Mengen nicht einfach auflisten und paarweise vergleichen kann, da die Auflistungen eine unterschiedliche Reihenfolge der Elemente aufweisen mögen. Die Antwort darauf lautet, man verlangt, dass jedes Element der einen Menge auch in der anderen zu finden sein wird und umgekehrt.

Es werden die Axiome der Logik mit Gleichheit gefordert, laut denen die Gleichheit reflexiv ist und eine Ersetzungsregel besitzt. Ihnen hinzugefügt wird das

Axiom 3.1 (Extensionalität).
Für je zwei Mengen A, B soll gelten

$$(\forall x\colon x \in A \Leftrightarrow x \in B) \Rightarrow A = B.$$

Die Umkehrung des Axioms braucht nicht unbedingt gefordert werden, da sie bereits aus den beiden allgemeinen Axiomen der Gleichheit hervorgeht. Alternativ kann man die Gleichheit per Äquivalenz

$$A = B \;:\Leftrightarrow\; (\forall x\colon x \in A \Leftrightarrow x \in B)$$

definieren. Der Leser wird mühelos bestätigen, dass die Gleichheit dahingehend die Reflexivität, mehr noch, die Axiome einer Äquivalenzrelation erfüllt. Allerdings kann ein Teil des Schemas zur Ersetzungsregel nun anscheinend nicht mehr hergeleitet werden. Dieses sollte zumindest für atomare Formeln bestehen, also für jedes Argument der Relation $\in$. Was bei näherer Betrachtung fehlt, ist die Aussage

$$A = B \Rightarrow (\forall M\colon A \in M \Rightarrow B \in M).$$

Dieser Sachverhalt wird dahingehend zusätzlich als Axiom gefordert.

Es ist zulässig, ein Element bei der aufzählenden Angabe einer Menge mehrmals aufzuführen. Dies ändert allerdings nichts daran, dass ein Element stets nur einmal in einer Menge enthalten ist. Zum Beispiel gilt $\{\emptyset, \emptyset\} = \{\emptyset\}$. Mit der Definition der aufzählenden Angabe und dem Idempotenzgesetz der Aussagenlogik findet sich nämlich die äquivalente Umformung

$$x \in \{\emptyset, \emptyset\} \iff x = \emptyset \vee x = \emptyset \iff x = \emptyset \iff x \in \{\emptyset\}.$$

3.1.3. Beschränkte Quantifizierung

In der Mathematik erstreckt sich die Quantifizierung meist nicht über das gesamte Diskursuniversum, sondern bleibt auf eine bestimmte Menge beschränkt. Eine extra Notation macht dies ergonomisch, wobei eine Erweiterung der logischen Sprache hierfür nicht nötig ist. Die beschränkte Quantifizierung wird logisch auf eine unbeschränkte zurückgeführt.

Definition 3.1 (Beschränkte Quantifizierung).
Für jede Menge M und jede Aussageform $A(x)$ setzt man

$$(\forall x \in M\colon A(x)) \;:\Leftrightarrow\; (\forall x\colon x \in M \Rightarrow A(x)),$$
$$(\exists x \in M\colon A(x)) \;:\Leftrightarrow\; (\exists x\colon x \in M \wedge A(x)).$$

Die Aussage $\forall x \in \emptyset\colon A(x)$ ist allgemeingültig, man spricht von der *leeren Wahrheit*, engl. *vacuous truth*. Via ex falso quodlibet erhält man nämlich:

$$\dfrac{\dfrac{\dfrac{}{\vdash \neg x \in \emptyset} \quad \dfrac{}{x \in \emptyset \vdash x \in \emptyset}}{x \in \emptyset \vdash \bot}}{\dfrac{x \in \emptyset \vdash A(x)}{\vdash \forall x\colon x \in \emptyset \Rightarrow A(x)}}\,\text{EFQ}$$

Viele Regeln zur beschränkten Quantifizierung sind analog zu den Regeln der unbeschränkten. Beispielsweise gilt

$$(\forall x \in M\colon A(x) \wedge B(x)) \;\Leftrightarrow\; (\forall x \in M\colon A(x)) \wedge (\forall x \in M\colon B(x)).$$

Man muss allerdings Vorsicht walten lassen. Nicht bei jeder Äquivalenz liegt eine direkte Analogie vor. Zwar besteht für eine Formel A, in der x nicht frei vorkommt, die Äquivalenz

$$(\exists x\colon A) \;\Leftrightarrow\; A.$$

Die Analogie ist jedoch von der ein klein wenig intrikateren Form

$$(\exists x \in M\colon A) \;\Leftrightarrow\; M \neq \emptyset \wedge A.$$

Diese Beziehung erklärt sich durch die Umformung

$$(\exists x \in M\colon A) \Leftrightarrow (\exists x\colon x \in M \wedge A) \Leftrightarrow (\exists x\colon x \in M) \wedge A \Leftrightarrow M \neq \emptyset \wedge A.$$

Die letzte Umformung gilt, weil $M \neq \emptyset$ gleichbedeutend mit $\exists x\colon x \in M$ ist.

3.1.4. Teilmengen

Gehört jedes Element einer Menge A auch zu einer Menge B, nennt man A eine *Teilmenge* von B. Man sagt auch, B umfasse A, oder B sei eine Obermenge von A. Eine *echte Teilmenge* sei A dann, wenn zusätzlich $A \neq B$ gilt. Die Menge der geraden Zahlen ist eine echte Teilmenge der ganzen Zahlen. Jede Menge ist eine Teilmenge von sich selbst, jedoch keine echte.

Die Menge der Quadrate ist eine Teilmenge der Vierecke, genauer eine Teilmenge der Rechtecke und auch eine Teilmenge der Rhomben. Weder ist die Menge der Rechtecke eine Teilmenge der Rhomben, noch ist die Menge der Rhomben eine Teilmenge der Rechtecke. Allerdings ist sowohl die Menge der Rechtecke als auch die der Rhomben eine Teilmenge der Parallelogramme.

Definition 3.2 (Inklusion).
Man definiert $A \subseteq B$, gelesen »A ist eine Teilmenge von B«, als

$$A \subseteq B \;:\Leftrightarrow\; (\forall x\colon x \in A \Rightarrow x \in B).$$

Unschwer bestätigt sich die Äquivalenz

$$A = B \;\Leftrightarrow\; A \subseteq B \wedge B \subseteq A.$$

Diese Gegebenheit dient häufig als Hilfestellung beim Beweis der Gleichheit zweier Mengen. Man zeigt $x \in A \Rightarrow x \in B$ und $x \in B \Rightarrow x \in A$ für festes, aber beliebiges x. Und damit ist bereits klar, dass $A = B$ sein muss.

Die Symbolik $A \subset B$ soll den Umstand ausdrücken, dass A eine *echte* Teilmenge von B ist, also sowohl $A \subseteq B$ als auch $A \neq B$ gilt. Außerhalb dieses Buches sollte man dafür aber besser $A \subsetneq B$ notieren, da $A \subset B$ leider auch in der Bedeutung von $A \subseteq B$ geläufig ist. Ich bin davon nicht so begeistert, weil die Harmonie mit der Notation für Halbordnungen dadurch verloren geht. Man überzeugt sich nämlich unschwer davon, dass die Inklusion die Axiome einer Halbordnung erfüllt. Für die Mengenlehre ist dieser Umstand aber von großer Bedeutung.

3.1.5. Komprehension

Es wäre an der Zeit, die Wege der Mengenbildung zu erkunden, auf denen eine Vielzahl unterschiedlicher Mengen erreicht werden kann. Zunächst steht die Komprehension als allgemeines Mittel zur Verfügung. Aus ihr gliedern sich im Fortgang verschiedene Mengenoperationen aus.

Die Abgrenzung einer Menge durch Aufzählung ihrer Elemente kommt lediglich für hinreichend kleine endliche Mengen infrage. Um sich von dieser Einengung zu befreien, beschreibt man Mengen allgemeiner durch Eigenschaften. Zu einer Eigenschaft denkt man sich hierbei die Menge der Elemente, die diese Eigenschaft erfüllen. Zur Eigenschaft *ist eine gerade ganze Zahl*,

$$A(x) \;:\Leftrightarrow\; x \text{ ist eine gerade ganze Zahl} \;\Leftrightarrow\; \exists k \in \mathbb{Z}\colon x = 2k,$$

erhält man zum Beispiel die Menge der geraden Zahlen. Diesbezüglich wird ihre Ausformung symbolisiert durch die Notation

$$\{x \mid A(x)\} = \{\ldots, -4, -2, 0, 2, 4, \ldots\}.$$

Es wird $\{x \mid A(x)\}$ gelesen als »die Klasse der x, für die $A(x)$ gilt« oder »die Menge der x, für die $A(x)$ gilt«.

Definition 3.3 (Komprehension).
Zu einer Eigenschaft $A(x)$ definiert man die Klasse $\{x \mid A(x)\}$ gemäß

$$u \in \{x \mid A(x)\} :\Leftrightarrow A(u).$$

Man muss mit der Komprehension ein wenig vorsichtig umgehen, denn nicht jede Klasse ist eine Menge. Die russellsche Klasse $R := \{x \mid x \notin x\}$ ist das klassische Beispiel. Logisch gilt laut dieser Festlegung die Allaussage

$$\forall x\colon x \in R \Leftrightarrow x \notin x.$$

Angenommen, R wäre eine Menge. Insofern dürfte man die Aussage $R \in M$ bezüglich einer Menge M formulieren, also speziell $R \in R$. Spezialisierung der Allaussage mit $x := R$ führt nun kurzerhand zu

$$R \in R \Leftrightarrow R \notin R.$$

Für jede Formel A gilt allerdings das Theorem

$$\vdash (A \Leftrightarrow \neg A) \Rightarrow \bot.$$

Insgesamt ergibt sich so ein Beweis der Kontradiktion. Irgendetwas kann also nicht gut sein. Diese von Bertrand Russell im Jahre 1901 entdeckte Verwicklung, die seit jeher den Namen *russellsche Antinomie* trägt, brachte Gottlob Freges logizistisches Programm in unangenehme Schwierigkeiten. Frege versuchte, eine Reduktion der Mathematik auf die Logik zu unternehmen, wurde daraufhin aber von Russell brieflich in Kenntnis gesetzt, dass sein Werk *Grundgesetze der Arithmetik* in wesentlicher Weise von der Antinomie unterhöhlt wird. Russell führte das Programm fort, und veröffentlichte schließlich die *Principia Mathematica*, die der Antinomie mithilfe einer Typentheorie aus dem Weg geht. [22]

Russell formulierte seine Antinomie anschaulich in Gestalt des Barbier-Paradoxons. *Ein Barbier rasiere all jene, und nur jene, die sich nicht selbst rasieren. Rasiert sich der Barbier selbst?* Mit der Zeit verbreiteten sich viele Abwandlungen. Das Wesen der Antinomie liegt in der problematischen Selbstbezüglichkeit, aus der auch das

Lügner-Paradoxon hervorgeht. In dürren Worten geht es so: *Pinocchio sagt, »Meine Nase wächst jetzt.«*

Man begegnet der Problematik, indem man R als eine echte Klasse ansieht. Für sie darf die Aussage $R \in M$ nicht formulierbar, oder zumindest nicht ableitbar sein. Ich will sogleich näher auf Klassen eingehen, zuvor aber eine weniger allgemeine Form der Komprehension, die *Aussonderung*, aufführen. Sie verhält sich gutartig, dergestalt dass sie nicht mehr ermöglicht, als das Ausfiltern von Elementen aus einer gegebenen Menge.

Definition 3.4 (Aussonderung).
Zu einer Menge M und einer Eigenschaft $A(x)$ definiert man

$$u \in \{x \in M \mid A(x)\} \;:\Leftrightarrow\; u \in M \wedge A(u).$$

Bekräftigt wird die Gutartigkeit durch das

Axiom 3.2 (Schema der Aussonderung).
Ist M eine Menge, so auch $\{x \in M \mid A(x)\}$ für jede Eigenschaft $A(x)$.

Aussonderung schränkt das Argumentieren mehr ein als nötig. Einen durchdachten Weg aus der Antinomie, der uns ungehindertes Bilden von Komprehensionen zurückbringt, bietet das Konzept der Klassen. Generell soll jede denkbare Ansammlung, also auch jede Menge eine *Klasse* sein.

Alternativ bestünde auch die Möglichkeit, Mengen und Klassen als zwei unterschiedliche *Sorten* aufzufassen, worunter man eine Art Typsystem versteht, das Mengen und Klassen zur Unterscheidung mit einem unterschiedlichen Typ behaftet. Die Elementbeziehung $x \in A$ darf diesbezüglich nur geformt werden, wenn x von der Sorte Menge ist, womit die russellsche Antinomie ebenfalls unterbunden würde. Diesen Weg wollen wir allerdings zugunsten der folgenden, ein wenig eleganteren Ausführungen nicht beschreiten.

Wir legen ein Universum $\mathcal{U}$ fest, das alle Mengen als Elemente enthält. Dieses, häufig auch mit V abgekürzte, ist selbst eine Klasse, die Klasse aller Mengen, auch genannt die *universelle Klasse* oder *Allklasse*. Eine Klasse sei genau dann eine *Menge*, wenn sie ein Element der universellen Klasse ist, andernfalls spricht man von einer *echten Klasse*. Man darf sich eine echte Klasse demnach als eine Ansammlung vorstellen, die in irgendeiner Weise, gleich welcher Art das vonstatten gehen mag, zu reichhaltig ist, um als Menge gelten zu können.

Statt die universelle Klasse als Grundobjekt zu betrachten, wird sie logisch mit der Forderung

$$u \in \mathcal{U} \;:\Leftrightarrow\; \exists C\colon u \in C$$

auf die Elementbeziehung zurückgeführt. Das heißt, eine Klasse u sei genau dann eine Menge, wenn sie Element irgendeiner Klasse ist. Es wird nun eine gutartige Komprehension festgelegt.

Definition 3.5 (Klassenkomprehension).
Zu jeder Eigenschaft $A(x)$ ist die Klasse $\{x \mid A(x)\}$ definiert gemäß

$$u \in \{x \mid A(x)\} :\Leftrightarrow u \in \mathcal{U} \wedge A(u).$$

Jede Klasse steht demnach nicht für eine beliebige Aussageform, sondern ist als eine Aussonderung aus der universellen Klasse interpretierbar. Die Komprehension wird hierdurch dahingehend entschärft, dass das Argument der russellschen Antinomie nicht mehr zu einer Kontradiktion führt. Dennoch bleibt die Argumentation wichtig, denn sie zeigt eben, dass die russellsche Klasse eine echte Klasse ist. Die Teilmengenbeziehung übernehmen wir unverändert, sprechen aber nun von der Teilklassenbeziehung. Daraufhin folgt sofort, dass jede Klasse eine Teilklasse der universellen Klasse sein muss. Es erscheint weiterhin vernünftig, jede Teilmenge einer Menge auch als Menge anzusehen. Wäre die universelle Klasse eine Menge, so müsste die russellsche Klasse als Teilklasse ebenso eine Menge sein, was aber nicht stimmt, wie wir gerade zuvor festgestellt haben. Ergo muss die universelle Klasse ebenfalls eine echte sein.

Vermittels der Komprehension ist im Weiteren die leere Menge definierbar. Darüber hinaus können wir die universelle Klasse gleichfalls, insofern sie zuvor gemäß der genannten Existenzaussage erklärt war, via Komprehension darstellen. Wir legen die beiden Klassen, die einander Gegenpole sind, fest als

$$\begin{aligned} \emptyset &:= \{x \mid \bot\} = \{x \mid x \neq x\}, \\ \mathcal{U} &:= \{x \mid \top\} = \{x \mid x = x\}. \end{aligned}$$

Die später erläuterten Begriffe Vereinigung, Schnitt, Differenz, Komplement und das kartesische Produkt sind eigentlich allgemein für Klassen definiert. Die Gleichheit von Mengen und die Teilmengenbeziehung wird wie gesagt zur Gleichheit von Klassen und Teilklassenbeziehung erweitert. Mithin sind Relationen und Abbildungen zwischen Klassen definierbar. Als Faustregel darf man sich merken, dass jeder Begriff zunächst für Klassen definiert wird, es sich aber, sofern diese als Elemente einer Klasse auftreten, um Mengen handeln muss.

Im Fortgang tut sich irgendwann die Frage auf, welche Gesetzmäßigkeiten eigentlich wesentlich sind, um die allgemeine Mengenlehre ausarbeiten zu können. Mehr noch tut sich damit einhergehend die Frage auf, ob diese Gesetzmäßigkeiten wirklich unbedenklich sind. Zur Beantwortung der ersten Frage stellt man ein System von Axiomen auf, wir sprechen dann von einer *axiomatischen Mengenlehre*.

Davon wurde mehrere geschaffen, die zueinander in nichttrivialer Weise in Beziehung stehen. Mit der zweiten Frage verhält es sich schwierig.

Die beiden Ansätze Aussonderung und Klassenkomprehension spiegeln sich in unterschiedlichen axiomatischen Mengenlehren wider. Das System ZFC, für *Zermelo-Fraenkel with Choice*, gewährt lediglich die Aussonderung. Man wird sich dabei fragen, wie größere Mengen überhaupt zu bilden sind, wenn doch aus Mengen nur Teilmengen ausgesondert werden. Dafür stehen zwei weitere Axiome zur Verfügung, das Axiom der Potenzmenge und das Axiom der Ersetzung.

Befindet man Klassenkomprehension als zulässig, führt dies zum System MK, für *Morse-Kelley*. Gleichwohl bleiben das Axiom der Potenzmenge und das Axiom der Ersetzung weiterhin verfügbar, um Mengen von echten Klassen abgrenzen zu können. Wem dieses System zu allgemein ist, dem steht als strengere Fassung noch das System NBG, für *Neumann-Bernays-Gödel*, zur Verfügung. Man konnte zeigen, dass jede in NBG beweisbare Aussage, die nur Mengen betrifft, ebenfalls in ZFC beweisbar sein muss.

Der Leser mag sich fragen, warum axiomatische Mengenlehre nicht direkt im Grundstudium an Hochschulen unterrichtet wird, schließlich sind Axiomensysteme doch sonst der Standard. Ist sie unpragmatisch? Dem ist nicht Fall. Zwar ist sie ein wenig umständlicher auszuarbeiten als die zur Abgrenzung als naiv bezeichnete Mengenlehre, es hält sich meines Erachtens aber sehr in Grenzen.

Um sich präziser auszudrücken, kann man sich bei der Notation dazu zwingen, jede Variable, die in einer Formel frei vorkommen darf, explizit aufzuführen. Dahingehend soll $A(x_1, \ldots, x_n)$ für eine Formel stehen, in der ausschließlich die Variablen $x_1, \ldots, x_n$ frei vorkommen dürfen. Man mag dies gern als $A(\vec{x})$, auch in der Variante $A(\mathbf{x})$ oder $A(\underline{x})$, abkürzen. Diesbezüglich notiert man

$$\{x \in M \mid A(x, \vec{y})\} = \{x \in M \mid A(x, y_1, \ldots, y_n)\}$$

für die Menge der $x \in M$, die $A(x, \vec{y})$ erfüllen, wobei die freien Variablen $y_1, \ldots, y_n$ als Parameter bezeichnet werden, die natürlich selbst Mengen darstellen, insofern in der reinen Mengenlehre allein Mengen als Objekte auftreten.

3.1.6. Mengenoperationen

Definition 3.6 (Potenzmenge).
Die *Potenzmenge* einer Menge M ist definiert als

$$\mathcal{P}(M) := \{A \mid A \subseteq M\}.$$

Zum Beispiel ist $\mathcal{P}(\emptyset) = \{\emptyset\}$ und $\mathcal{P}(\{0\}) = \{\emptyset, \{0\}\}$. Des Weiteren

$$\mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\},$$
$$\mathcal{P}(\{0,1,2\}) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \{0,1,2\}\}.$$

Die Teilmengenbeziehung darf als eine Art Ordnung zwischen Mengen betrachtet werden, jedoch nicht als eine Totalordnung, das heißt, bei einigen Mengen A, B gilt weder $A \subseteq B$ noch $B \subseteq A$. So ist wie gesagt weder die Menge der Rechtecke eine Teilmenge der Rhomben, noch umgekehrt.

Die Potenzmenge einer Grundmenge G bildet mit der Teilmengenbeziehung eine halbgeordnete Menge, engl. *poset* für *partially ordered set*. Das heißt, alle Mengen $A, B, C \in \mathcal{P}(G)$ erfüllen die drei Axiome

$$A \subseteq A, \qquad \text{(Reflexivität)}$$
$$A \subseteq B \wedge B \subseteq A \Rightarrow A = B, \qquad \text{(Antisymmetrie)}$$
$$A \subseteq B \wedge B \subseteq C \Rightarrow A \subseteq C. \qquad \text{(Transitivität)}$$

Weil die Potenzmenge keine Aussonderung ist, bedarf es eines extra Axioms, um sie als Menge begreifen zu können.

Axiom 3.3 (Potenzmenge).
Ist M eine Menge, so auch $\mathcal{P}(M)$.

Zusammen mit dem Axiom der Vereinigung bietet es einen Weg, um größere Menge aus kleineren zu erzeugen. Später wird sich ihnen diesbezüglich noch das Axiom der Ersetzung hinzugesellen.

Definition 3.7 (Schnitt, Vereinigung, Differenz).
Zu zwei Mengen A, B definiert man

$$A \cap B := \{x \mid x \in A \wedge x \in B\}, \qquad \text{(\textit{Schnittmenge})}$$
$$A \cup B := \{x \mid x \in A \vee x \in B\}, \qquad \text{(\textit{Vereinigungsmenge})}$$
$$A \setminus B := \{x \mid x \in A \wedge x \notin B\}. \qquad \text{(\textit{Differenzmenge})}$$

Sind A, B Teilmengen einer Grundmenge G, so sind auch $A \cap B$, $A \cup B$ und $A \setminus B$ Teilmengen der Grundmenge. Der Beweis zu $A \cap B \subseteq G$ ist:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{x \in A \cap B \vdash x \in A \cap B}}{x \in A \cap B \vdash x \in A \wedge x \in B}}{x \in A \cap B \vdash x \in A} \qquad \dfrac{\dfrac{\dfrac{}{A \subseteq G \vdash A \subseteq G}}{A \subseteq G \vdash \forall x\colon x \in A \Rightarrow x \in G}}{A \subseteq G \vdash x \in A \Rightarrow x \in G}}{\dfrac{A \subseteq G, x \in A \cap B \vdash x \in G}{\dfrac{A \subseteq G \vdash \forall x\colon x \in A \cap B \Rightarrow x \in G}{A \subseteq G \vdash A \cap B \subseteq G}}}$$

In so pedantischer Ausführlichkeit findet man Beweise in Büchern nicht vor. Erstens wird man stillschweigend zulässige Schlussregeln zur Verkürzung aufstellen. So nimmt der Baum die konzise Form

$$\dfrac{\dfrac{\dfrac{}{x \in A \cap B \vdash x \in A \cap B}}{x \in A \cap B \vdash x \in A} \qquad \dfrac{}{A \subseteq G \vdash A \subseteq G}}{\dfrac{A \subseteq G, x \in A \cap B \vdash x \in G}{A \subseteq G \vdash A \cap B \subseteq G}} \qquad \dfrac{\dfrac{\dfrac{}{x \in A \cap B}\,{\scriptstyle 1}}{x \in A} \qquad A \subseteq G}{\dfrac{x \in G}{A \cap B \subseteq G}\,{\scriptstyle \sim 1}}$$

an. Zweitens formuliert der Mathematiker den Beweis meist in Worten: Um $A \cap B \subseteq G$ zu zeigen, muss $x \in G$ aus $x \in A \cap B$ abgeleitet werden. Mit $x \in A \cap B$ gilt erst recht $x \in A$. Wegen $A \subseteq G$ ist somit $x \in G$, was zu zeigen war. □

Definition 3.8 (Komplement).
Bezüglich einer Grundmenge G heißt $A^{\mathrm{c}} := G \setminus A$ *Komplementärmenge*.

Es zeigt sich elementar, dass die Potenzmenge einer Grundmenge G mit den Operationen $A \cap B$, $A \cup B$ die Axiome einer booleschen Algebra erfüllt, wobei $\emptyset$ das Nullelement und G selbst das Einselement ist. Es gelten somit analoge Regeln wie in der klassischen Aussagenlogik. Wie die Notation suggeriert, entspricht der Schnitt der Konjunktion, die Vereinigung der Disjunktion und das Komplement der Negation.

Wie in jedem Verband ist in einer booleschen Algebra $(M, \wedge, \vee)$ für $a, b \in M$ eine Halbordnung $a \leq b$ definiert, indem $a \leq b$ und $a \wedge b = a$ als äquivalent angesehen werden. Bei der Mengenalgebra entpuppt sie sich als nichts anderes als die Inklusion. Das heißt, es gilt

$$A \subseteq B \;\Leftrightarrow\; A \cap B = A \;\Leftrightarrow\; A \cup B = B.$$

Die Beziehung $A \setminus B = A \cap B^{\mathrm{c}}$ ist eine recht dienliche. Vermöge ihr können die Operationen mit der Differenzmenge ebenfalls mit der booleschen Algebra diskutiert werden. So ermöglicht sie die Rechnung

$$\begin{aligned} A \setminus (B \cap C) &= A \cap (B \cap C)^{\mathrm{c}} = A \cap (B^{\mathrm{c}} \cup C^{\mathrm{c}}) \\ &= (A \cap B^{\mathrm{c}}) \cup (A \cap C^{\mathrm{c}}) = (A \setminus B) \cup (B \setminus C). \end{aligned}$$

Tabelle 3.1.: Logische Entsprechungen der Mengenoperationen

0	1	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$	$A \oplus B$
$\emptyset$	G	A^c	$A \cap B$	$A \cup B$	$A^c \cup B$	$(A \triangle B)^c$	$A \triangle B$

Dieses flippende Distributivgesetz ist also dadurch zu erklären, dass das Distributivgesetz des Schnittes an das de morgansche Gesetz angefügt wird. Mit dem Idempotenzgesetz $A = A \cap A$, zuzüglich dem Kommutativ- und Assoziativgesetz des Schnittes findet sich weiterhin

$$\begin{aligned} A \setminus (B \cup C) &= A \cap (B \cup C)^c = A \cap B^c \cap C^c = A \cap A \cap B^c \cap C^c \\ &= A \cap B^c \cap A \cap C^c = (A \setminus B) \cap (B \setminus C). \end{aligned}$$

Die alternative Bestätigung dieser Formeln mittels natürlichem Schließen oder logischer Äquivalenzumformung bietet eine leichte Übung. Die Bestätigung von

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

verläuft analog.

In gewisser Hinsicht spiegeln Mengen logische Aussagen wider. Zu jeder logischen Verknüpfung gehört eine Mengenoperation, siehe Tabelle 3.1. Die Subjunktion $A \Rightarrow B$ wird dabei via $\neg A \vee B$ übersetzt, was die klassische Logik verlangt. Man kann nun darüber hinaus einen Zusammenhang zwischen der Inklusion und den Sequenzen herstellen. Und zwar verhält sich die Beziehung

$$A_1 \cap \ldots \cap A_m \subseteq B_1 \cup \ldots \cup B_n.$$

analog zur allgemeinen Sequenz

$$A_1, \ldots, A_m \vdash B_1, \ldots, B_n.$$

Mehr noch, jede Schlussregel des natürlichen Schließens besitzt eine direkte Entsprechung. So verhält sich die Regel

$$\frac{C \vdash A \wedge B}{C \vdash A} \quad \text{analog zu} \quad C \subseteq A \cap B \Rightarrow C \subseteq A.$$

Aber Vorsicht, der Sequenz $\{C_1\} \cup \{C_2\} \vdash A$ entspricht $C_1 \cap C_2 \subseteq A$.

Diese Sichtweise ermöglicht es, das natürliche Schließen durch Eulerdiagramme zu veranschaulichen. Umgekehrt hilft die Betrachtung als Sequenz beim Urteilen, ob eine Menge Teil einer anderen ist.

Für mehrere Mengen definiert man

$$\bigcap_{i=1}^{n} A_i := A_1 \cap A_2 \cap \ldots \cap A_n, \qquad \bigcup_{i=1}^{n} A_i := A_1 \cup A_2 \cup \ldots \cup A_n.$$

Pedantiker mögen die Schreibweise mit den Auslassungspunkten als ungenau empfinden. Zufriedenstellend ist für sie die Erklärung, dass es sich um die rekursive Festlegung

$$\bigcap_{i=1}^{1} A_i := A_1, \qquad \bigcap_{i=1}^{n} A_i := A_n \cap \bigcap_{i=1}^{n-1} A_i$$

handelt. Regeln wie $B \cup \bigcap_{i=1}^{n} A_i = \bigcap_{i=1}^{n} (B \cup A_i)$ kann man nun per Induktion über n beweisen. Es geht fast trivial vonstatten, sobald man das Prinzip verstanden hat. Im Wesentlichen weiten sich hier die Regeln der booleschen Algebra von den zweistelligen auf die mehrstelligen Operationen aus.

Man kann Vereinigung und Schnitt aber noch wesentlich allgemeiner fassen.

Definition 3.9 (Allgemeine Vereinigung).
Sei $\mathcal{A}$ eine Menge von Mengen. Die *Vereinigung* der $A \in \mathcal{A}$ ist

$$\bigcup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A := \{x \mid \exists A \in \mathcal{A} \colon x \in A\}.$$

Für $\mathcal{A} = \emptyset$ ist $\bigcup \mathcal{A} = \emptyset$. Die Disjunktion findet ihre Entsprechung genau in der Vereinigung von zwei Mengen. Dazu passend findet der Existenzquantor seine Entsprechung genau in der Vereinigung beliebig vieler Mengen. Aus diesem Grund weiten sich die Regeln der booleschen Algebra auf die allgemeine Vereinigung aus. Zum Beispiel lautet das Distributivgesetz für Mengen

$$B \cap \bigcup_{A \in \mathcal{A}} A = \bigcup_{A \in \mathcal{A}} (B \cap A).$$

Entfaltung der Definition führt nämlich zur logischen Äquivalenz

$$x \in B \wedge (\exists A \in \mathcal{A} \colon x \in A) \;\Leftrightarrow\; (\exists A \in \mathcal{A} \colon x \in B \wedge x \in A).$$

Ihr Beweis gelingt mühelos.

Definition 3.10 (Allgemeiner Schnitt).
Sei $\mathcal{A}$ eine nichtleere Menge von Mengen. Der *Schnitt* der $A \in \mathcal{A}$ ist

$$\bigcap \mathcal{A} = \bigcap_{A \in \mathcal{A}} A := \{x \mid \forall A \in \mathcal{A} \colon x \in A\}.$$

Im Gegensatz zur Vereinigung wurde der Schnitt $\bigcap \mathcal{A}$ für $\mathcal{A} = \emptyset$ undefiniert gelassen. Hier gibt es zwei Möglichkeiten. Zum einen könnte man die Bedingung $\mathcal{A} \neq \emptyset$ einfach fallen lassen, womit sich im allgemeinen Mengenuniversum beim leeren Schnitt die Allklasse $\{x \mid \top\}$ ergibt, die jedoch keine Menge ist.

Aus diesen Grund gibt es noch die alternative Definition

$$\bigcap \mathcal{A} := \{x \in G \mid \forall A \in \mathcal{A} \colon x \in A\}.$$

Hierzu ist eine Grundmenge G festzulegen, so dass $\mathcal{A} \subseteq \mathcal{P}(G)$ gilt, oder man setzt $G := \bigcup \mathcal{A}$, wobei sich da die Frage nach der Nützlichkeit stellt.

Eine Menge von Mengen nennt man ein *Mengensystem*. Hiervon zu unterscheiden ist der Begriff der *indizierten Mengenfamilie*, kurz *Familie*. Sie stellt eine Verallgemeinerung einer Folge von Mengen dar. In ihr darf dieselbe Menge mehrmals vorkommen. Man kann Schnitt und Vereinigung auch für Familien definieren, was aber eigentlich keine wesentliche Verallgemeinerung zu den obigen Festlegungen darstellt, wie ich im Folgenden diskutieren möchte.

Eine *Familie* (A_i) von Mengen A_i mit $i \in I$ ist eine Abbildung $A \colon I \to Z$, wobei Z eine Zielmenge ist, welche die A_i als Elemente enthält. Die Menge I wird in diesem Zusammenhang auch *Indexmenge* genannt. Man definiert

$$\bigcup_{i \in I} A_i := \bigcup A(I) = \bigcup \{X \mid \exists i \in I \colon X = A_i\} = \{x \mid \exists i \in I \colon x \in A_i\},$$

wobei mit $A(I)$ das Bild von I unter A gemeint ist. Man bekommt

$$\begin{aligned}\bigcup_{i \in I} A_i &= \{x \mid \exists X \colon X \in \{X \mid \exists i \in I \colon X = A_i\} \wedge x \in X\} \\ &= \{x \mid \exists X \colon (\exists i \in I \colon X = A_i) \wedge x \in X\} \\ &= \{x \mid \exists X \colon \exists i \in I \colon X = A_i \wedge x \in X\} = \{x \mid \exists i \in I \colon x \in A_i\}.\end{aligned}$$

Für $I \neq \emptyset$ definiert man entsprechend

$$\bigcap_{i \in I} A_i := \bigcap A(I) = \{x \mid \forall i \in I \colon x \in A_i\}.$$

Die Operation über eine Familie $(A_i)_{i \in I}$ kann also auf die jeweilige Operation über das System $A(I)$ zurückgeführt werden.

Später nützlich ist der

Satz 3.4. Es gilt $\bigcup_{i\in I\cup J} A_i = (\bigcup_{i\in I} A_i) \cup (\bigcup_{i\in J} A_j)$.

Beweis. Es findet sich die äquivalente Umformung

$$\begin{aligned} x \in \bigcup_{i\in I\cup J} A_i &\iff (\exists i\colon i \in I \cup J \wedge x \in A_i) \\ &\iff (\exists i\colon (i \in I \wedge x \in A_i) \vee (i \in J \wedge x \in A_i)) \\ &\iff (\exists i\colon i \in I \wedge x \in A_i) \vee (\exists i\colon i \in J \wedge x \in A_i) \\ &\iff x \in \bigcup_{i\in I} A_i \vee x \in \bigcup_{i\in J} A_i \\ &\iff x \in (\bigcup_{i\in I} A_i) \cup (\bigcup_{i\in J} A_i). \square \end{aligned}$$

Gelegentlich hat man es mit einer *disjunkten Vereinigung* zu tun. Sie ist bedeutsam in der Theorie der Kardinalzahlen und in der Informatik bei algebraischen Datentypen sowie deren Bezug zur Beweistheorie. Genauer gesagt hantiert man in der Informatik diesbezüglich nicht mit Mengen, sondern mit Typen, die sich in gewissen Zügen analog zu Mengen verhalten. Die disjunkte Vereinigung zweier Mengen kennzeichnet jedes Element vor der Vereinigung mit einem Tag, das die Information liefert, aus welcher der Mengen es entstammt. Man setzt

$$A \sqcup B := \{(0, a) \mid a \in A\} \cup \{(1, b) \mid b \in B\}.$$

Die Zahlen $0, 1$ sind hier die *Tags* oder *Diskriminatoren*. Anstelle der Zahlen könnten genauso gut zwei beliebige unterschiedliche Elemente als Tags verwendet werden. Beispielsweise ginge auch

$$A \sqcup B = \{(\text{Grün}, a) \mid a \in A\} \cup \{(\text{Blau}, b) \mid b \in B\}.$$

In der Informatik verwendet man auch gern left, right als Tags.

Mit jeder disjunkten Vereinigung ist eine Fallunterscheidung verbunden. Liegt ein $x \in A \sqcup B$ vor, so muss entweder $x = (0, a)$ für ein $a \in A$ oder $x = (1, b)$ für ein $b \in B$ sein. Bei einer gewöhnlichen Vereinigung besteht dagegen kein ausschließendes Oder.

Wir fassen zwei Objekte x, y zu einem *geordneten Paar* (x, y) zusammen. Die beiden Objekte müssen nicht unbedingt etwas miteinander zu tun haben, sie dürfen völlig verschiedener Art sein. Im Unterschied zu einer Menge spielt bei Paaren die Reihenfolge eine Rolle, auch darf dasselbe Objekt zweimal vorkommen. Im Paar $t = (x, y)$ ist genau die Information über x, y enthalten. Das heißt, es lassen sich x, y aus dem Paar extrahieren. Man schreibt dafür $t_1 = x$ und $t_2 = y$. Die Schreibweise t_i heißt *Indizierung*, es ist darin i der *Index*. Näher betrachtet vermittelt je Index i eine Projektionsabbildung π_i die Indizierung, also $t_i = \pi_i(t)$.

Zwei Paare seien definitionsgemäß genau dann gleich, wenn sie komponentenweise gleich sind,

$$(x, y) = (x', y') \;:\Leftrightarrow\; x = x' \wedge y = y'.$$

Die genannten Eigenschaften charakterisieren den Begriff Paar im Wesentlichen, mehr müssen wir nicht wissen. Man hat sich trotzdem auch mal überlegt, wie Paare in der reinen Mengenlehre dargestellt werden können, wo alle Objekte Mengen sein sollen. Nach Kuratowski sind Paare kodierbar als

$$(x, y) := \{\{x\}, \{x, y\}\}.$$

Die beiden Projektionen sind diesbezüglich darstellbar als

$$\pi_1(t) = \bigcup \bigcap t,$$
$$\pi_2(t) = (\bigcap \bigcup t) \cup ((\bigcup \bigcup t) \setminus (\bigcup \bigcap t)).$$

Unter dem allgemeineren Begriff *Tupel* fassen wir eine beliebige endliche Zahl von Objekten in einer bestimmten Reihenfolge zusammen. Tupel aus drei Objekten heißen *Tripel*, die aus vier heißen *Quadrupel*. Analog zu den Paaren ist ihre Gleichheit definiert als

$$(x_1, \ldots, x_n) = (x_1', \ldots, x_n') \;:\Leftrightarrow\; x_1 = x_1' \wedge \ldots \wedge x_n = x_n'.$$

Man kann Tupel folgendermaßen als Schachtelung von Paaren darstellen, woraufhin aber zwei unterschiedliche Arten von Paaren vorhanden sind. Zur Unterscheidung will ich die gewöhnlichen Paare als $\langle x, y\rangle$ notieren. Man legt das leere Tupel fest als $() := \emptyset$, die weiteren rekursiv durch

$$(x_1, \ldots, x_n, x_{n+1}) := \langle (x_1, \ldots, x_n), x_{n+1}\rangle.$$

Für ein Tripel ergibt sich beispielsweise

$$(x, y, z) = \langle (x, y), z\rangle = \langle \langle (x), y\rangle, z\rangle = \langle \langle \langle \emptyset, x\rangle, y\rangle, z\rangle.$$

Zu zwei Mengen X, Y kann man nun die Menge aller Paare betrachten, deren erste Komponente X entstammt, und deren zweite Komponente Y entstammt. Sie heißt *Produktmenge* oder *kartesisches Produkt* der Mengen X, Y.

Definition 3.11 (Produktmenge).
Das *kartesische Produkt* zweier Mengen X, Y ist die Menge

$$X \times Y := \{(x, y) \mid x \in X \wedge y \in Y\}.$$

Genau genommen handelt es sich hierbei um eine Bildmenge, was bedeutet, dass sich im rechten Term Existenzquantoren verstecken. Ausführlich ausgeschrieben lautet der Term

$$\begin{aligned} X \times Y &= \{t \mid \exists x \in X \colon \exists y \in Y \colon t = (x, y)\} \\ &= \{t \mid \exists x \colon \exists y \colon x \in X \wedge y \in Y \wedge t = (x, y)\}. \end{aligned}$$

Wie jede Bildmenge ist das Produkt als Vereinigung darstellbar. Es ist

$$X \times Y = \bigcup_{x \in X} \bigcup_{y \in Y} \{(x, y)\}.$$

Die beiden kurzen Identitäten $X \times \emptyset = \emptyset$ und $\emptyset \times Y = \emptyset$ gehen unmittelbar aus der Definition hervor. Es verhält sich analog wie mit der Multiplikation einer Zahl mit null. Eine sinnvolle Sichtweise, wie die Theorie der Kardinalzahlen lehrt.

Satz 3.5. Ist $A \subseteq X$ und $B \subseteq Y$, dann ist $A \times B \subseteq X \times Y$.

Beweis. Es liege t in $A \times B$. Laut Definition existieren mithin $a \in A$ und $b \in B$, so dass $t = (a, b)$. Wegen $A \subseteq X$ ist aber auch $a \in X$ und wegen $B \subseteq Y$ ist auch $b \in Y$. Daher existieren $a \in X$ und $b \in Y$, so dass $t = (a, b)$. Gemäß Definition heißt das, $t \in X \times Y$. Gemäß Definition ist $A \times B$ somit eine Teilmenge von $X \times Y$. □

Satz 3.6. Es gilt $X \times (A \cup B) = (X \times A) \cup (X \times B)$.

Beweis. Es ginge zu Fuß. Aber Satz 3.4 ermöglicht die kurze Termumformung

$$\begin{aligned} X \times (A \cup B) &= \bigcup_{x \in X} \bigcup_{y \in A \cup B} \{(x, y)\} = \bigcup_{x \in X} \left(\left(\bigcup_{y \in A} \{(x, y)\} \right) \cup \left(\bigcup_{y \in B} \{(x, y)\} \right) \right) \\ &= \left(\bigcup_{x \in X} \bigcup_{y \in A} \{(x, y)\} \right) \cup \left(\bigcup_{x \in X} \bigcup_{y \in B} \{(x, y)\} \right) = (X \times A) \cup (X \times B). \end{aligned}$$ □

Da sich die Existenzquantifizierung nicht in allgemeiner Weise distributiv bezüglich der Konjuktion verhält, mag es intrikat erscheinen, dass der folgende Satz dennoch gilt.

Satz 3.7. Es gilt $X \times (A \cap B) = (X \times A) \cap (X \times B)$.

Beweis. Wir wollen sodann auch einmal den Beweis betrachten, um zu verstehen, was da vonstatten geht.

Es sei $t \in X \times (A \cap B)$. Dann existieren $x \in X$ und $y \in A \cap B$, so dass $t = (x, y)$. Ergo ist sowohl $y \in A$ als auch $y \in B$. Und somit sowohl $t \in X \times A$ als auch $t \in X \times B$. Ergo ist $t \in (X \times A) \cap (X \times B)$. Diese Implikation schafft der Existenzquantor auch in allgemeiner Weise. Betrachten wir nun die kritische umgekehrte Richtung.

Es sei $t \in (X \times A) \cap (X \times B)$. Dann ist sowohl $t \in X \times A$ als auch $t \in X \times B$. Ergo existiert ein $x \in X$ und $y \in A$ mit $t = (x, y)$. Weiterhin ein $x' \in X$ und $y' \in B$ mit $t = (x', y')$. Hier liegt der Hase im Pfeffer: Wegen $(x, y) = t = (x', y')$ muss $x = x'$ und $y = y'$ sein. Wir wissen nun, $y \in A \cap B$, ergo $t \in X \times (A \cap B)$. □

Ob eine Aussageform in zwei Variablen über das Durchlaufen aller Paare Allquantifizierung erfährt, oder über das Durchlaufen der Variablen selbst, macht keinen Unterschied. Dieses Schönfinkeln der auf Produktmengen beschränkten Quantifizierung mag am Beispiel endlicher Mengen recht einsichtig sein. Die allgemeine Bestätigung vermittelt der

Satz 3.8. Es ist $\forall t \in X \times Y \colon A(t)$ äquivalent zu $\forall x \in X \colon \forall y \in X \colon A(x, y)$.

Beweis. Zur Implikation von links nach rechts. Aus festen, aber beliebigen $x \in X$ und $y \in Y$ muss $A(x, y)$ abgeleitet werden. Mit dem aus ihnen geformten Paar $t := (x, y)$ bezeugen sie $t \in X \times Y$. Hiermit erhält man $A(t)$, also $A(x, y)$ aus der Voraussetzung.

Zur Implikation von rechts nach links. Aus festem, aber beliebigem $t \in X \times Y$ ist $A(t)$ abzuleiten. Mit $t \in X \times Y$ existieren $x \in X$ und $y \in Y$ mit $t = (x, y)$. Hiermit erhält man $A(x, y)$ aus der Voraussetzung. Wegen der Gleichheit von t und (x, y) stimmt $A(x, y)$ mit $A(t)$ überein. □

Ein analoger Sachverhalt besteht bei der Existenzquantifizierung.

Satz 3.9. Es ist $\exists t \in X \times Y \colon A(t)$ äquivalent zu $\exists x \in X \colon \exists y \in Y \colon A(x, y)$.

Beweis. Zur Implikation von links nach rechts. Laut der Voraussetzung ist irgendein $t \in X \times Y$ mit $A(t)$ vorhanden, infolge $x \in X$ und $y \in X$ mit $t = (x, y)$, womit $A(x, y)$ gilt. Ergo existieren $x \in X$ und $y \in Y$ mit $A(x, y)$.

Zur Implikation von rechts nach links. Laut Voraussetzung sind $x \in X$ und $y \in Y$ mit $A(x, y)$ vorhanden. Mit dem aus ihnen geformten Paar $t := (x, y)$ bezeugen sie $t \in X \times Y$. Laut Festlegung ist $A(x, y)$ dasselbe wie $A(t)$, womit t die Existenz eines $t \in X \times Y$ mit $A(t)$ bezeugt. □

Mithin bestehen die korrespondierenden Termumformungen

$$\bigcap_{t\in I\times J} A_t = \bigcap_{i\in I}\bigcap_{j\in J} A_{ij}, \qquad \bigcup_{t\in I\times J} A_t = \bigcup_{i\in I}\bigcup_{j\in J} A_{ij}.$$

Sie bestätigen sich mühelos via

$$\begin{aligned} x \in \textstyle\bigcap_{t\in I\times J} A_t &\iff \forall t \in I\times J\colon x\in A_t \iff \forall i\in I\colon \forall j\in J\colon x\in A_{ij} \\ &\iff \forall i\in I\colon x\in \textstyle\bigcap_{j\in J} A_{ij} \iff x\in \bigcap_{i\in I}\bigcap_{j\in J} A_{ij}. \end{aligned}$$

Hier ist A_{ij} eine Kurzschreibweise für $A_{(i,j)}$ bzw. $A_{i,j}$.

3.2. Abbildungen

3.2.1. Der Abbildungsbegriff

Unter einer *Abbildung*, auch *Funktion* genannt, verstehen wir ganz allgemein eine Zuordnung von Elementen einer Definitionsmenge zu Elementen einer Zielmenge, bei der zu *jedem* Element der Definitionsmenge *genau ein* Element der Zielmenge gehört. Es hat allerdings ein wenig gedauert, bis diese Vorstellung als die Förderliche erkannt wurde.

In der historischen Entwicklung ging ihr die speziellere Vorstellung voran, dass eine Größe, etwa eine physikalische Größe, in einer rechnerischen Abhängigkeit von einer anderen Größe steht. So steht die Periodendauer der Schwingung eines Pendels in einer bestimmten rechnerischen Abhängigkeit von der Pendellänge.

Eine recht pragmatische Vorstellung von einer Funktion vermittelt das Modell der Black Box, das soll eine Rechenmaschine sein, deren innere Mechaniken bzw. Elektroniken unbekannt bleiben. Speist man ein Argument x in die Black Box ein, spuckt sie daraufhin einen Wert $f(x)$ aus. Speist man abermals dasselbe Argument ein, spuckt sie abermals denselben Wert aus.

Der Begriff der Abbildung ist für die Mathematik zentral.

Definition 3.12 (Abbildung).
Eine *Abbildung* $f\colon X \to Y$ ist eine Relation $f = (X, Y, G)$ mit $G \subseteq X \times Y$, die die beiden Eigenschaften

$$\forall x \in X\colon \exists y \in Y\colon (x, y) \in G,$$
$$\forall x \in X\colon \forall y, y' \in Y\colon (x, y) \in G \wedge (x, y') \in G \Rightarrow y = y'$$

erfüllt. Man nennt G den *Graph*, X die *Definitions-* und Y die *Zielmenge*.

Genau genommen möchte man die Abbildung f eigentlich nicht als mit dem Tripel (X, Y, G) identisch sehen, sondern als dasjenige Objekt, dessen innere Information aus diesem Tripel besteht.

Zuweilen wird der Graph von f auch einfach als f bezeichnet. Sollte dies verwirrend sein, schreibt man ausführlicher G_f oder $\mathrm{Graph}(f)$.

Mit $\forall x \in X\colon \exists! y \in Y\colon (x, y) \in G_f$ ist eine Relation f kurzum eine Funktion, was laut Satz 1.3 äquivalent zu den beiden Eigenschaften ist.

Statt $(x, y) \in G_f$ schreibt man üblicherweise $y = f(x)$. Es wird $f(x)$ gelesen als »f von x« oder »das Bild von x unter f«. Es wird $f\colon X \to Y$ gelesen als »f ist eine Abbildung von X nach Y«. Besitzt ein $y \in Y$ ein $x \in X$ mit $y = f(x)$, nennt man x ein Urbildelement zu y.

Alle denkbaren Abbildungen $X \to Y$ fasst man wiederum zu einer Menge zusammen, die Y^X oder $\mathrm{Abb}(X, Y)$ notiert wird. Zu jeder Menge M sei $|M|$ die Anzahl ihrer Elemente. Es seien X, Y endlich. Zu jedem Argument $x \in X$ gibt es nun $|Y|$ Möglichkeiten, einen Funktionswert festzulegen. Hat X zwei Elemente, gibt es für das erste Elemente $|Y|$ Möglichkeiten, und für das zweite auch nochmals $|Y|$. Da man beide unabhängig wählen kann, multiplizieren sich die Möglichkeiten zu $|Y|^2$. Allgemein betrachtet findet sich $|Y^X| = |Y|^{|X|}$, was eine gewisse Rechtfertigung für die gewählte Notation liefert.

Es stellt sich die Frage, ob die Formel denn auch im Trivialfall $X = \emptyset$ stimmt. Insofern man $0^0 := 1$ definiert, so dass allgemein $|Y|^0 = 1$ gilt, lautet die Antwort ja, weil $Y^\emptyset = \{\emptyset\}$ ist. Überzeugung leistet die folgende kurze Überlegung. Soll $f\colon \emptyset \to Y$ sein, muss f als Relation eine Teilmenge von $\emptyset \times Y = \emptyset$ sein. Die einzige Teilmenge der leeren Menge ist die leere Menge selbst. Sie ordnet wie gefordert jedem Element der leeren Definitionsmenge genau einen Wert zu. Das mag eigenartig wirken, ist aber, wir erinnern uns an das Prinzip der leeren Wahrheit, der richtige Schluss.

Anders verhält es sich mit $Y = \emptyset$ bei $X \neq \emptyset$. Hier gilt $0^{|X|} = 0$. Auch in diesem Fall stimmt die Formel, weil $\emptyset^X = \emptyset$ ist. Wie zuvor ist die einzige mögliche Relation die leere. Sie müsste allerdings jedem $x \in X$ ein $y \in \emptyset$ zuordnen, was widersprüchlich ist. Ergo ist die Menge der Abbildungen $X \to \emptyset$ leer, sofern X nichtleer ist.

3.2.2. Bild, Urbild

Wird ein jedes Element einer Teilmenge der Definitionsmenge in eine Abbildung geschickt, formen die Werte eine neue Menge, die *Bildmenge* der Teilmenge unter der Abbildung.

Definition 3.13 (Bild).
Die *Bildmenge* einer Menge $A \subseteq X$ unter der Abbildung $f\colon X \to Y$ ist

$$f(A) := \{f(x) \mid x \in A\} = \{y \mid \exists x \in A\colon y = f(x)\}.$$

Insofern $y = f(x)$ als äquivalent zu $y \in \{f(x)\}$ befunden wird, ergibt sich auch die Darstellung $f(A) = \bigcup_{x \in A} \{f(x)\}$. Insbesondere in Programmiersprachen treten endliche Mengen als Objekte auf. Die Berechnung verläuft mithin gemäß der rekursiven Festlegung

$$f(\emptyset) := \emptyset, \qquad f(\{x_1, \ldots, x_n\}) := f(\{x_1, \ldots, x_{n-1}\}) \cup \{f(x_n)\},$$

für die auch die Bezeichnung $\mathrm{map}(f, A)$ gebräuchlich ist.

Die analytische Geometrie sieht Figuren als Punktmengen. Abbildungen transformieren die Punktmengen, woraus neue Figuren hervorgehen. Ein einfaches Beispiel bietet die Abbildung

$$f\colon \mathbb{R} \to \mathbb{R}^2, \quad f(t) := \begin{pmatrix} p_x + v_x t \\ p_y + v_y t \end{pmatrix}.$$

In der Hinsicht, dass die reellen Zahlen $\mathbb{R}$ als eine Zahlengerade aufgefasst werden, formt ihr Bild $f(\mathbb{R})$ eine Gerade der euklidischen Ebene. Die passende Wahl der Parameter (p_x, p_y) als Basispunkt und (v_x, v_y) als Geschwindigkeitsvektor gestattet es, jede beliebige Gerade der Ebene zu beschreiben. Man nennt in der üblichen Terminologie aber auch t den Parameter und f eine Parametergerade, wobei das Bild $f(\mathbb{R})$ auch Spur genannt wird. Diese Sprechweisen entspringen der Sichtweise, dass $f(t)$ ein durch die Zeit t parametrisierter Punkt ist, der mit dem Lauf der Zeit eine Spur zieht. Der Geschwindigkeitsvektor charakterisiert die Bewegung des Punktes, für die Spur ist dagegen nur dessen Richtung von Bedeutung.

Das *Urbild* eines Wertes gibt Auskunft, welche Elemente der Definitionsmenge unter der Abbildung zu dem Wert führen. Allgemeiner gibt das Urbild einer Menge von Werten Auskunft, welche Elemente der Definitionsmenge in die Menge führen.

Definition 3.14 (Urbild).
Das *Urbild* einer Menge B unter $f\colon X \to Y$ ist

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\} = \{x \in X \mid \exists y \in B\colon y = f(x)\}.$$

Dem Wesen des Abbildungsbegriffs entspringend, zeichnet sich die Urbildoperation durch Verträglichkeit mit den Mengenoperationen aus, was bei der Bildoperation nur zum Teil stimmt.

Satz 3.10. Für jede Abbildung f und beliebige Mengen A, B, A_i gilt

$$\begin{aligned} f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B), & f^{-1}\Big(\bigcap_{i\in I} A_i\Big) &= \bigcap_{i\in I} f^{-1}(A_i), \\ f^{-1}(A \cup B) &= f^{-1}(A) \cup f^{-1}(B), & f^{-1}\Big(\bigcup_{i\in I} A_i\Big) &= \bigcup_{i\in I} f^{-1}(A_i). \end{aligned}$$

Beweis. Den Beweis verschafft die äquivalente Umformung

$$\begin{aligned} x \in f^{-1}\Big(\bigcap_{i\in I} A_i\Big) &\iff f(x) \in \bigcap_{i\in I} A_i \iff (\forall i \in I\colon f(x) \in A_i) \\ &\iff (\forall i \in I\colon x \in f^{-1}(A_i)) \iff x \in \bigcap_{i\in I} f^{-1}(A_i). \end{aligned}$$

Bei der Vereinigung verläuft die Umformung analog. □

Der Satz lehrt die wichtige Folgerung, dass die Urbilder zweier disjunkter Mengen ebenfalls disjunkt sind. Allgemein überführt die Urbildoperation eine disjunkte

Zerlegung einer Menge in disjunkte Urbilder. Ist die Zielmenge am feinsten zerlegt, besteht sie also aus einelementigen Mengen, nennt man die Urbilder dieser Mengen auch Fasern. Man beachte, dass ein Urbild oder eine Faser leer sein kann.

Ist die Zielmenge bezüglich einer Ordnungsrelation angeordnet, nennt man die Fasern auch *Niveaumengen*. Diese Begrifflichkeit betrifft vor allem die Funktionen $f\colon X \to \mathbb{R}$ mit $X \subseteq \mathbb{R}^n$. Eine Faser $f^{-1}(\{c\})$, auch als $f(x) = c$ beschrieben, nennt man auch eine *implizite Funktion*. Sie sind in der analytischen Geometrie und der mehrdimensionalen Analysis von großer Bedeutung.

Ich möchte noch einmal auf das Bild eingehen. Zu welchem Ausmaß die Bildoperation mit den Mengenoperationen Schnitt und Vereinigung verträglich ist, geben die folgenden Sätze aufschluss.

Satz 3.11. Für jedes $f\colon X \to Y$ und $A, B, A_i \subseteq X$ gilt

$$\begin{aligned} f(A \cap B) &\subseteq f(A) \cap f(B), & f\Big(\bigcap_{i\in I} A_i\Big) &\subseteq \bigcap_{i\in I} f(A_i),\\ f(A \cup B) &= f(A) \cup f(B), & f\Big(\bigcup_{i\in I} A_i\Big) &= \bigcup_{i\in I} f(A_i). \end{aligned}$$

Beweis. Es gelte $y \in f(\bigcap_{i\in I} A_i)$. Somit ist ein $x \in \bigcap_{i\in I} A_i$ mit $f(x) = y$ vorhanden. Laut Definition des Schnittes gilt $x \in A_i$ zu jedem i, womit zu jedem i ein $x \in A_i$ mit $f(x) = y$ vorliegt. Ergo gilt $y \in f(A_i)$ für jedes i, und somit $y \in \bigcap_{i\in I} f(A_i)$.

Es gelte $y \in f(\bigcup_{i\in I} A_i)$. Somit ist ein $x \in \bigcup_{i\in I} A_i$ mit $f(x) = y$ vorhanden. Laut Definition der Vereinigung liegt x in mindestens einer der A_i, womit zu mindestens einem i ein $x \in A_i$ mit $f(x) = y$ vorliegt. Ergo gilt $y \in f(A_i)$ für mindestens ein i, und somit $y \in \bigcup_{i\in I} f(A_i)$.

Umgekehrt gelte $y \in \bigcup_{i\in I} f(A_i)$. Somit gilt $y \in f(A_i)$ für mindestens ein i. Zu diesem i ist also ein $x \in A_i$ mit $y = f(x)$ vorhanden. Ergo liegt ein $x \in \bigcup_{i\in I} A_i$ mit $y = f(x)$ vor, womit $y \in f(\bigcup_{i\in I} A_i)$ ist. □

Als kurze Folgerung der vorangegangenen Sätze ergibt sich, dass sich sowohl die Bild- als auch die Urbildoperation bezüglich der Inklusion monoton steigend verhalten. Das heißt, es gilt der

Satz 3.12. Für jedes $f\colon X \to Y$ und $A, B \subseteq X$ bzw. $A', B' \subseteq Y$ gilt

$$\begin{aligned} A \subseteq B &\Rightarrow f(A) \subseteq f(B),\\ A' \subseteq B' &\Rightarrow f^{-1}(A') \subseteq f^{-1}(B'). \end{aligned}$$

Beweis. Es ist $A \subseteq B$ äquivalent zu $B = A \cup B$. Somit gilt

$$f(B) = f(A \cup B) = f(A) \cup f(B),$$

womit $f(A) \subseteq f(B)$ ist. Beim Urbild verläuft die Argumentation analog. □

Ein Beispiel mit $f(A \cap B) \neq f(A) \cap f(B)$ findet sich leicht. Sei dazu $X := \{a, b\}$ eine Menge, die zwei verschiedene Elemente enthält und $Y := \{y\}$ eine Einermenge. Die Abbildung $f : X \to Y$ ist dadurch bereits eindeutig festgelegt. Zu $A := \{a\}$ und $B := \{b\}$ gilt nun

$$f(A \cap B) = f(\emptyset) = \emptyset \neq f(A) \cap f(B) = \{y\} \cap \{y\} = \{y\}.$$

Die einfachst mögliche nicht-injektive Abbildung f habe ich dabei gezielt aus dem folgenden Grund als Gegenbeispiel gewählt.

Satz 3.13. Es gilt $f(A) \cap f(B) \subseteq f(A \cap B)$ genau dann für alle $A, B \subseteq X$, wenn $f : X \to Y$ injektiv ist.

Beweis. Es gelte $f(a) = f(b)$. Mit den Setzungen $A := \{a\}$ und $B := \{b\}$ ist

$$f(A) \cap f(B) = \{f(a)\} \cap \{f(b)\} = \{f(a)\} \cap \{f(a)\} = \{f(a)\} \neq \emptyset$$

Laut Prämisse ist $f(A \cap B)$ eine Obermenge von $f(A) \cap f(B)$, also $f(A \cap B) \neq \emptyset$, womit $A \cap B \neq \emptyset$, womit zwingend $a = b$ sein muss.

Umgekehrt sei f injektiv. Es gelte $y \in f(A) \cap f(B)$, also $y \in f(A)$ und $y \in f(B)$. Somit existiert ein $a \in A$ mit $f(a) = y$ und ein $b \in B$ mit $f(b) = y$. Da f injektiv ist, erhält man nun $a = b$ aus $y = f(a) = f(b)$. Ergo gilt $a \in A \cap B$, womit es die Aussage $y \in f(A \cap B)$ bezeugt. □

3.2.3. Komposition

Definition 3.15 (Komposition).
Sei $f : X \to Y$ und $g : Y \to Z$. Ihre *Verkettung*, gelesen »g nach f«, ist

$$(g \circ f) : X \to Z, \quad (g \circ f)(x) := g(f(x)).$$

Oft liegt die Situation $f : X \to Y$ und $g : Y' \to Z$ mit $Y \subseteq Y'$ vor. Das ist aber nicht weiter schlimm. Es darf dann $g \circ f := g|_Y \circ f$ gesetzt werden, wobei mit $g|_Y$ die Einschränkung des Definitionsbereichs von g auf Y gemeint ist.

Definition 3.16 (Einschränkung).
Sei $f : X \to Y$ und $A \subseteq X$. Die *Einschränkung* von f auf A ist

$$f|_A : A \to Y, \quad f|_A(x) := f(x).$$

Die Abbildung $\mathrm{id}_X : X \to X$ mit $\mathrm{id}_X(x) := x$ heißt *identische Abbildung*. Sie verhält sich bei der Komposition neutral, das heißt, bezüglich $f : X \to Y$ gilt $f \circ \mathrm{id}_X = f$ und $\mathrm{id}_Y \circ f = f$.

Mit der Bildmenge unter einer Komposition verhält es sich ganz analog wie mit dem Funktionswert eines einzelnen Elements. Sie lässt sich als die Hintereinanderschaltung der jeweiligen Mengenabbildungen darstellen, denn es gilt der

Satz 3.14. Es gilt $(g \circ f)(A) = g(f(A))$ für jede Menge $A \subseteq \mathrm{dom}(f)$.

Beweis. Mit der Entfaltung von Def. 3.13, 3.3, und der Anwendung von Axiom 3.1 nimmt die Aussage die Form

$$(\exists x \in A : z = (g \circ f)(x)) \;\Leftrightarrow\; (\exists y \in f(A) : z = g(y))$$

an. Angenommen, die linke Seite gilt. Dann liegt ein $x \in A$ mit $z = g(f(x))$ vor. Sei nun $y := f(x)$, dann gilt $y \in f(A)$ und $z = g(y)$, womit die Existenzaussage der rechten Seite erfüllt wird.

Angenommen, die rechte Seite gilt. Dann liegt ein $y \in f(A)$ mit $z = g(y)$ vor. Infolge existiert laut Def. 3.13 ein $x \in A$ mit $y = f(x)$. Nun gilt $z = g(f(x))$, womit x ebenfalls ein Zeuge für die linke Existenzaussage ist. □

Beim Urbild unter einer Komposition dreht sich die Reihenfolge um.

Satz 3.15. Es gilt $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$ für jede Menge $B \subseteq \mathrm{cod}(g)$.

Beweis. Die äquivalente Umformung

$$\begin{aligned} x \in (g \circ f)^{-1}(B) &\iff (g \circ f)(x) \in B \iff g(f(x)) \in B \\ &\iff f(x) \in g^{-1}(B) \iff x \in f^{-1}(g^{-1}(B)). \end{aligned}$$ □

Man kann diesen Beweis aber auch so führen, dass die Abbildungen in allgemeiner Weise als Relationen betrachtet werden. Die Gleichheit $y = f(x)$ ist bei einer Relation nicht erklärt, man hat nur $(x, y) \in f$. Unbeschadet dessen bleibt die Umformung durchführbar via

$$\begin{aligned} x \in (g \circ f)^{-1}(B) &\iff (\exists z \in B : (x, z) \in g \circ f) \\ &\iff (\exists z \in B : \exists y : (y, z) \in g \wedge (x, y) \in f) \\ &\iff (\exists y : (\exists z \in B : (y, z) \in g) \wedge (x, y) \in f) \\ &\iff (\exists y : y \in g^{-1}(B) \wedge (x, y) \in f) \\ &\iff x \in f^{-1}(g^{-1}(B)). \end{aligned}$$

3.2.4. Injektionen, Surjektionen, Bijektionen

Definition 3.17 (Injektion).
Eine Abbildung $f\colon X \to Y$ heißt injektiv, wenn

$$\forall x, x' \in X\colon f(x) = f(x') \Rightarrow x = x'.$$

Erinnern wir uns an den Abschnitt *Logik mit Gleichheit*, fällt auf, dass die Definition der Injektivität als Umkehrung der Ersetzungsregel betrachtbar ist. Das heißt, für eine auf den Werten der Terme t, t' definierte Injektion f gilt die Äquivalenz

$$t = t' \;\Leftrightarrow\; f(t) = f(t').$$

Die Injektionen vermitteln somit genau die *Äquivalenzumformungen* von Gleichungen. Wie sich aus Axiom 3.1 in Verbindung mit Def. 3.4 ergibt, erfährt die Lösungsmenge einer Bestimmungsgleichung durch sie keine Veränderung. Man notiert

$$L := \{x \in G \mid t = t'\} = \{x \in G \mid f(t) = f(t')\}$$

für die Lösungsmenge L der Gleichung $t = t'$ in der Variable x, die die Werte der Grundmenge G durchläuft. Aufgrund dessen schaffen sie ein wesentliches Werkzeug zum Lösen von Bestimmungsgleichungen.

Es muss X bei $f\colon X \to Y$ nicht notwendigerweise die Grundmenge sein. Wichtig ist allein, dass die Terme t, t' ausschließlich Werte in X annehmen können. Beispielsweise ist das Quadrieren auf den reellen Zahlen zwar nicht injektiv, bei Einschränkung der Definitionsmenge auf die nichtnegativen reellen Zahlen allerdings schon. So initiiert es die äquivalente Umformung

$$|x| = |x-2| \Leftrightarrow x^2 = (x-2)^2 = x^2 - 4x + 4 \Leftrightarrow 0 = -4x + 4 \Leftrightarrow x = 1.$$

Außerdem darf die Variable der Bestimmungsgleichung in einer Äquivalenzumformung als Parameter auftauchen. Die injektive Funktion

$$f_a\colon \mathbb{R} \to \mathbb{R}, \quad f_a(\xi) := \xi + a$$

beschreibt zum Beispiel den Sachverhalt, dass eine Zahl a zu beiden Seiten einer Gleichung addiert werden darf,

$$t = t' \;\Leftrightarrow\; t + a = t' + a.$$

Hier darf insbesondere auch $a := x$ als Parameter eingesetzt werden. Bei

$$f_a\colon \mathbb{R} \to \mathbb{R}, \quad f_a(\xi) := a\xi$$

gilt es allerdings $a \neq 0$ zu berücksichtigen. Das heißt, die Setzung $a := x$ liefert nur dann eine Äquivalenzumformung, wenn die Grundmenge, die x durchläuft, eine Teilmenge von $\mathbb{R} \setminus \{0\}$ ist. Andernfalls müsste man diesen Umstand durch eine Fallunterscheidung künstlich herstellen.

Erwähnenswert ist weiterhin, dass die gemachten Begriffe bereits Gleichungen in mehreren Variablen und Gleichungssysteme umfassen. Eine Gleichung in zwei Variablen liegt vor, wenn die Grundmenge aus Paaren besteht. Ein System von zwei Gleichungen liegt vor, wenn die die Terme t, t' jeweils ein Paar zum Wert haben. Diese formale Beschreibung müsste man allerdings als ein wenig oberflächlich befinden, kämen nicht weitere Erörterungen hinzu. Tiefersinnig wäre es an sich ohne Pointe, tauchte nur eine der Variablen in der Gleichung auf. Gleichermaßen mag ein System in zwei Variablen erst reizvoll sein, wenn die Gleichungen in nichttrivialer Weise miteinander verwoben sind.

Die Injektionen sind genau die Abbildungen, die rückgängig gemacht werden können. Rückgängig machen einer Abbildung $f\colon X \to Y$ ist so zu verstehen, dass eine Umkehrabbildung g zur Verfügung stehen soll, so dass $g(f(x)) = x$ gilt. Gilt dies für jedes $x \in X$, nennt man g eine *Linksinverse* zu f.

Satz 3.16. Eine Abbildung g mit $g \circ f = \mathrm{id}$ heißt Linksinverse von f. Eine Abbildung $f\colon X \to Y$ mit nichtleerem X ist genau dann injektiv, wenn sie mindestens eine Linksinverse besitzt.

Beweis. Sei g eine Linksinverse von f. Seien x, x' fest, aber beliebig, und sei $f(x) = f(x')$. Mithin gilt $g(f(x)) = g(f(x'))$. Weil g eine Linksinverse ist, hat man aber $g(f(x)) = x$ und $g(f(x')) = x'$, womit sich wie gewünscht $x = x'$ ergibt.

Sei f injektiv. Mit $X \neq \emptyset$ liegt irgendein $a \in X$ vor. Sofern $y \in f(X)$ ist, liegt außerdem ein x mit $y = f(x)$ vor. Es wird g festgelegt per Fallunterscheidung

$$g(y) := \begin{cases} x, & \text{wenn } y \in f(X), \\ a, & \text{wenn } y \notin f(X). \end{cases}$$

Sie ist verhält sich so, dass $g(f(x)) = x$ für jedes $x \in X$ gilt, denn mit $x \in X$ ist $f(x) \in f(X)$. Somit existiert mit g eine Linksinverse. □

Die Festlegung von g kann allerdings im Allgemeinen keine konstruktive im Sinne von berechenbare sein. Hierzu muss man lediglich wissen, dass das Halteproblem unentscheidbar ist. Gemeint ist, dass es kein Verfahren gibt, das zu jeder Turingmaschine und jeder Eingabe entscheidet, ob die Maschine für die Eingabe irgendwann hält. Es wird nun $f\colon \mathbb{N} \to \mathbb{N}$ definiert als $f(n) := 0$, sofern die Turingmaschine in genau n Schritten hält, sonst $f(n) := n + 1$. Die Berechnung von f unternimmt die

Listing 3.1: Berechnung einer Linksinverse

```
def inv(f, X):
    a = next(iter(X))
    def g(y):
        for x in X:
            if f(x) == y: return x
        return a
    return g
```

Maschine insofern selbst, womit es sich bei f um eine berechenbare Funktion handelt. Außerdem ist f injektiv, wie man unschwer erkennen kann. Zur Festlegung von $g(0)$ muss aber in Erfahrung gebracht werden, ob und in wie vielen Schritten die Maschine hält. Da die gegebene Turingmaschine nicht vorab bekannt sein soll, müsste das Halteproblem gelöst werden. Ergo ist g nicht berechenbar. [34]

Die aufgezeigte Problematik besteht indes nur hinsichtlich des Unendlichen. Für eine endliche Definitionsmenge stellt sich die Linksinverse sehr wohl als berechenbar dar. Das Listing 3.1 zeigt ein konkretes Programm, das die Linksinverse einer jeden Funktion f bestimmt, die auf einer endlichen Menge X definiert ist. Wäre f zudem surjektiv, dürfte X sogar unendlich sein, ohne dass sich das Programm jemals in einer endlosen Schleife verfängt.

Definition 3.18 (Surjektion).
Eine Abbildung $f\colon X \to Y$ heißt *surjektiv*, wenn $f(X) = Y$ ist.

Weil $f(X) \subseteq Y$ allgemeingültig ist, genügt es generell, $Y \subseteq f(X)$ zu zeigen.

Definition 3.19 (Bijektion).
Eine Abbildung heißt *bijektiv*, wenn sie sowohl injektiv als auch surjektiv ist.

Satz 3.17. Jede Bijektion f besitzt eine eindeutig bestimmte Abbildung, welche sowohl ihre einzige Linksinverse als auch ihre einzige Rechtsinverse ist. Man nennt sie die Umkehrabbildung f^{-1}.

Beweis. Sei $f\colon X \to Y$ bijektiv. Es existiert somit mindestens eine Linksinverse g und mindestens eine Rechtsinverse h. Weil die Verkettung das Assoziativgesetz erfüllt, darf man rechnen

$$g = g \circ \mathrm{id}_Y = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_X \circ h = h.$$

Sei nun g' eine weitere Linksinverse und h' eine weitere Rechtsinverse. Wiederholt man die obige Rechnung abermals, erhält man $g' = h$ und $g' = h'$. Ergo gilt $g = h = g' = h'$. □

3.2.5. Allgemeines Mengenprodukt

Nachdem nun der Begriff der Abbildung bereits erklärt wurde, kann eine Begriffsverallgemeinerung des kartesischen Produktes erörtert werden, die Tupel mit Folgen und Funktionen in Beziehung setzt. Eine Folge darf man einerseits als Funktion betrachten, deren Definitionsbereich die natürlichen Zahlen sind. Andererseits ist eine Folge indizierbar wie ein Tupel. Das bringt uns auf die Idee, Tupel wie Folgen als Funktionen aufzufassen.

Die Ziffern der Zahl 2520 sind in Little-Endian-Konvention das Tupel

$$t = (0, 2, 5, 2) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N} \times \mathbb{N}.$$

Die Indexmenge $I := \{0, 1, 2, 3\}$ enthält die Stellenwerte der Ziffern. Es verhält sich nun so, dass die Funktion

$$f\colon I \to \mathbb{N}, \quad f(0) := 0,\ f(1) := 2,\ f(2) := 5,\ f(3) := 2$$

genau die Information enthält, die das Tupel charakterisiert. Es ist $f(i) = t_i$ die Indizierung des Tupels. So gesehen ist jedes Tupel durch eine Funktion kodiert. Führt man diese Überlegung fort, gelangt man schließlich zum allgemeinen kartesischen Produkt

$$\prod_{i\in I} X_i := \{f\colon I \to \bigcup_{i\in I} X_i \mid \forall i \in I\colon f(i) \in X_i\}.$$

Der Formelsalat sagt im Wesentlichen nur, dass eine Abbildung $f \in \prod_{i\in I} X_i$ einen in X_i liegenden Funktionswert $f(i)$ besitzt.

Diese allgemeinen Produkte von Mengen treten in der Mathematik nur sehr sporadisch in Erscheinung. Eine Rolle spielen sie in der Theorie der Kardinalzahlen. In der abhängigen Typentheorie besitzen sie allerdings ein direktes, allgegenwärtiges Analogon, den *Typ abhängiger Funktionen*.

Bei einer Gleichsetzung wie $\prod_{i\in\{1,2\}} X_i \cong X_1 \times X_2$ ist zu beachten, dass es sich eigentlich nicht um eine Gleichung handelt, denn Abbildungen sind ja nicht dasselbe wie Paare. Allerdings gehört zu jedem Paar wie gesagt in kanonischer Weise genau eine gleichartige Abbildung. Es besteht eine Isomorphie, was durch die Tilde über dem Gleichheitszeichen angedeutet wird. Damit verbunden liegt ein Isomorphismus vor, mit dem sich ein Objekt der einen Seite in ein Objekt der anderen

Seite überführen lässt. Das ist so ähnlich wie die Übersetzung eines Textes von der einen in die andere Sprache. Die Gestalt ändert sich, aber der Inhalt bleibt gleich.

Intuitiv sollte $\prod_{i\in I} X_i$ nichtleer sein, sofern jedes X_i nichtleer ist. Dennoch lässt sich dieser Sachverhalt nicht ohne Weiteres ableiten. Es handelt sich um ein Axiom der Mengenlehre, das sogenannte *Auswahlaxiom*. Es besagt, dass zu jeder Familie (X_i) nichtleerer Mengen eine Auswahlfunktion f existiert, die zu jeder Menge X_i ein Element $f(i) \in X_i$ auswählt. Eine gründliche Untersuchung des Auswahlaxioms zeigt auf, dass es sich mit den Unendlichkeiten schwierig verhält. Unter bestimmten Umständen ist das Axiom nicht vonnöten. Außerdem, so lässt sich zeigen, impliziert es den Satz vom ausgeschlossenen Dritten.

3.2.6. Definite Kennzeichnung

Sofern zu einer vorgelegten Aussageform $A(x)$ nicht nur ein diese erfüllendes Objekt existiert, sondern das Objekt auch eindeutig bestimmt ist, würden wir gerne über dieses Objekt reden können. Diesbezüglich stehe die Symbolik $\iota x\colon A(x)$ für »dasjenige x, das $A(x)$ erfüllt«. Sie entstammt der *Principia Mathematica* und wird *definite Kennzeichung* oder *definite Beschreibung* genannt. Mit den Mitteln der Mengenlehre ist dieses Objekt darstellbar als die Schnittmenge

$$(\iota x\colon A(x)) := \bigcap\{x \mid A(x)\}.$$

Nämlich ist $\{x \mid A(x)\}$ im Falle der eindeutigen Existenz eine Einermenge. Für jede Einermenge $\{a\}$ gilt

$$\bigcap\{a\} = \bigcup\{a\} = a.$$

Hierfür ist allerdings eine reine Mengenlehre erforderlich, in der keine sogenannten Urelemente auftauchen, die selbst keine Mengen sind. Durchliefe die Variable x Urelemente, obgleich es sich in Wirklichkeit nur um ein einziges handelt, wäre die Schnittmenge nicht definiert.

Näher betrachtet sind eindeutige Existenz und Reduzierung auf eine Einermenge bedeutungsgleich, das soll heißen, es besteht die Äquivalenz

$$(\exists!x\colon A(x)) \;\Leftrightarrow\; \exists x\colon \{y \mid A(y)\} = \{x\}.$$

Löst man die Gleichheit darin per Extensionalität auf, findet sich die kurze Charakterisierung

$$(\exists!x\colon A(x)) \;\Leftrightarrow\; \exists x\colon \forall y\colon (A(y) \Leftrightarrow x = y).$$

Sie zeigt sich bereits in der reinen Logik mit Gleichheit als beweisbar, ohne dafür Mengenlehre zur Verfügung stellen zu müssen.

Liegt eine Abbildung $f\colon X \to Y$ vor, unterstellt man die Bedeutung des Terms $f(x)$ üblicherweise wortlos als offenkundig. In der formalen Mengenlehre muss man diese Applikation allerdings erst einmal definieren. Wir legen fest,

$$f(x) := (\iota y\colon (x,y) \in f) = \bigcap\{y \mid (x,y) \in f\}.$$

Unter dieser Festlegung stellt sich heraus, dass die Äquivalenz

$$y = f(x) \;\Leftrightarrow\; (x,y) \in f$$

die linke Seite nicht erklärt, sondern vielmehr für $x \in X$ beweisbar ist. Stünde die Festlegung nicht zur Verfügung, müsste man Aussagen, in denen $f(x)$ vorkommt, umständlich durch Existenzaussagen kodieren. Beispielsweise stünde $f(x) \in B$ für die Existenz eines $y \in B$ mit $(x,y) \in f$. Kommen mehrere Applikationen vor, führt dies im Allgemeinen zu geschachtelten Existenzaussagen. Höchst umständlich wäre es, würde man in einer längeren Rechnung jede Applikation einschließlich der arithmetischen Operationen durch eine Existenzaussage kodieren.

3.3. Relationen

3.3.1. Relationen im Allgemeinen

Beim Erlangen von tieferliegenden Einsichten in Probleme und Zusammenhänge spielt das Aufspüren und Klären unterschiedlicher Beziehungen eine wesentliche Rolle. Zwei Objekte können auf unterschiedliche Art und Weise in Beziehung stehen. In der Mathematik beschäftigt man sich mit Beziehungen, die zu den zwei Objekten eine Aussage trifft, der ein Wahrheitsgehalt beigemessen wird. Zwei Zahlen x, y sind gleich. Eine Zahl x ist kleiner als eine Zahl y. Der Betrag der Differenz zweier Zahlen x, y ist kleiner als eine bstimmte Konstante. Der Abstand der Punkte $x = (x_1, x_2)$ und $y = (y_1, y_2)$ ist kleiner als eine bestimmte Konstante. Zwei Stühle x, y eines Stuhlkreises sind benachbart. Die Schüler x, y gehen in dieselbe Klasse. Es gibt einen Weg, der vom Ort x zum Ort y führt.

Die eindrückliche Vielfalt möglicher Beziehungen macht es ja unerlässlich, eine allgemeine Auffassung von ihnen zu erhalten. Eine zweistellige Relation R schafft ein Beziehungsgefüge zwischen zwei Mengen X, Y. Wir sagen, ein Objekt $x \in X$ steht bezüglich R zu einem Objekt $y \in Y$ in Beziehung, wenn die Aussage $R(x,y)$ eine wahre ist. Ein Beispiel wäre $R(x,y) := (x < y)$, wobei $X = Y$ die Menge der ganzen Zahlen sei.

Definition 3.20 (Zweistellige Relation).
Es seien X, Y zwei Mengen. Jede Teilmenge $R \subseteq X \times Y$ heißt Relation zwischen X und Y. Zur Relation soll die Kenntnis von X, Y dazugehören. Insofern ist das durch das Tripel (X, Y, R) kodierte Objekt die Relation und R ihr Graph.

Eine Relation ist auch interpretierbar als wahrheitswertige Funktion

$$1_R \colon X \times Y \to \{0, 1\}, \quad 1_R(x, y) := [(x, y) \in R].$$

Sie ist die Indikatorfunktion der Teilmenge R bezüglich der Grundmenge $X \times Y$. Anstelle $(x, y) \in R$ oder $1_R(x, y) = 1$ schreiben wir schlicht $R(x, y)$ für die Aussage, dass x und y bezüglich R in Relation stehen.

Eine Relation mit $X \neq Y$ heißt *heterogen*, eine mit $X = Y$ heißt *homogen*. Wir werden uns fast ausschließlich mit homogenen Relationen beschäftigen. Man darf allerdings sagen, dass auch die heterogenen in der Mathematik weit verbreitet sind, denn jede Abbildung ist betrachtbar als Relation mit speziellen Eigenschaften. Genauer ist eine Abbildung eine linkstotale und rechtseindeutige Relation. Linkstotal heißt, dass jedes $x \in X$ zu mindestens einem $y \in Y$ in Beziehung steht. Rechtseindeutig heißt, dass ein $y \in Y$ höchstens zu einem $x \in X$ in Beziehung steht. Es handelt sich also lediglich um eine Sprechweise für die bereits bekannten definierenden Eigenschaften.

Relationen lassen sich in Pfeildiagrammen darstellen. Man zeichnet hierbei einem Pfeil von einem $x \in X$ zu einem $y \in Y$, falls x, y in Relation stehen.

3.3.2. Äquivalenzrelationen

Manchmal interessiert man sich nicht für die gänzliche Gleichheit zweier Objekte. Die *Äquivalenzrelationen* verallgemeinern den Gleichheitsbegriff dahingehend, dass zwei Objekte schon dann als *gleichartig* angesehen werden, wenn sie in in einer bestimmten Eigenschaft übereinstimmen. Um welche Eigenschaft es sich dabei handelt, bestimmt die Relation.

Definition 3.21 (Äquivalenzrelation).
Seien A eine Menge und seien $x, y, z \in A$. Sei $R(x, y) := (x \sim y)$ eine Relation. Man nennt R *Äquivalenzrelation*, wenn gilt:

$$\begin{aligned} &x \sim x, && (\textit{Reflexivität}) \\ &x \sim y \Rightarrow y \sim x, && (\textit{Symmetrie}) \\ &x \sim y \wedge y \sim z \Rightarrow x \sim z. && (\textit{Transitivität}) \end{aligned}$$

Definition 3.22 (Äquivalenzklasse).
Sei M eine Menge und $x \sim y$ eine Äquivalenzrelation für $x, y \in M$. Die Menge

$$[a] := \{x \in M \mid x \sim a\}$$

nennt man die *Äquivalenzklasse* zum *Repräsentanten* $a \in M$.

Satz 3.18 (Äquivalenzrelation induziert Zerlegung).
Eine Menge wird durch eine Äquivalenzrelation in disjunkte Äquivalenzklassen zerlegt, lat. partitioniert.

Beweis. Sei M die Menge und $x \sim y$ die Äquivalenzrelation. Zu zeigen ist, dass kein Element von M in mehr als einer Äquivalenzklasse vorkommt. Seien $a, b, c \in M$, sei $c \in [a]$ und $c \in [b]$. Aufgrund von $c \sim a$ sowie $c \sim b$ und der Transitivität gilt

$$x \in [a] \iff x \sim a \iff x \sim c \iff x \sim b \iff x \in [b].$$

Man hat also

$$(\forall x \in M \colon x \in [a] \Leftrightarrow x \in [b]), \iff [a] = [b].$$

Wenn also $[a] \neq [b]$ ist, kann nicht gleichzeitig $c \in [a]$ und $c \in [b]$ sein. □

Satz 3.19 (Zerlegung induziert Äquivalenzrelation).
Sei M eine Menge. Die Familie (A_k) von Mengen $A_k \subseteq M$ bilde eine Zerlegung von M, d. h. dass die Vereinigung aller A_k die Menge M überdeckt und dass paarweise $A_i \cap A_j = \emptyset$ für $i \neq j$ ist. Dann ist

$$x \sim y :\Leftrightarrow \exists k \colon x \in A_k \wedge y \in A_k$$

eine Äquivalenzrelation auf M.

Beweis. Da die A_k die Menge M überdecken, muss es für ein beliebiges $x \in M$ mindestens eine Menge A_k geben, so dass $x \in A_k$. Daher gilt $x \sim x$.

Die Symmetrie ergibt sich trivial.

Zur Transitivität. Voraussetzung ist $x \sim y$ und $y \sim z$. Es gibt also ein i mit $x \in A_i$ und $y \in A_i$. Außerdem gibt es ein j mit $y \in A_j$ und $z \in A_j$. Somit gilt

$$\exists i \colon \exists j \colon x \in A_i \wedge y \in A_i \wedge y \in A_j \wedge z \in A_j.$$

Wegen

$$A_i \cap A_j = \emptyset \Leftrightarrow \forall y \colon (y \in A_i \wedge y \in A_j \Leftrightarrow 0)$$

für $i \neq j$ kann $y \in A_i \wedge y \in A_j$ aber nur erfüllt sein, wenn $i = j$ ist. Daher ergibt sich

$$\exists i \colon x \in A_i \wedge z \in A_i,$$

das heißt $x \sim z$. □

Definition 3.23 (Quotientenmenge).
Für eine gegebene Äquivalenzrelation wird die aus allen Äquivalenzklassen bestehende Menge

$$M/\sim := \{[x] \mid x \in M\}$$

als *Quotientenmenge* oder *Faktormenge* bezeichnet.

Definition 3.24 (Quotientenabbildung).
Für eine gegebene Äquivalenzrelation ist die Projektion

$$\pi \colon M \to M/\sim, \quad \pi(x) := [x]$$

surjektiv und wird *Quotientenabbildung* genannt.

Definition 3.25 (Repräsentantensystem).
Für eine gegebene Äquivalenzrelation auf M nennt man eine Teilemenge $A \subseteq M$ ein *vollständiges Repräsentantensystem*, wenn die Einschränkung $\pi|_A$ bijektiv ist, wobei mit π die Quotientenabbildung gemeint ist.

Repräsentantensysteme ermöglichen die einfache Handhabung von Äquivalenzklassen. Möchte man wissen, ob ein Element x in der Äquivalenklasse $[a]$ enthalten ist, dann braucht man bloß zu überprüfen, ob $x \sim a$ ist. Außerdem besitzt die Quotientenabbildung nun eine Darstellung $p \colon M \to A$, dergestalt dass $\pi = \pi|_A \circ p$. Warum sollte das von Bedeutung sein? Nun, Äquivalenzklassen fallen oft unendlich groß aus. In der Kombinatorik treten zwar auch endliche Äquivalenzklassen auf, diese werden trotzdem schnell unzugänglich groß. Die Äquivalenzklassen und die Quotientenabbildung muss man also als abstrakte mathematische Objekte betrachten. Abstrakte mathematische Objekte müssen wir erst über eine Darstellung zugänglich machen, und genau dies ermöglicht ein Repräsentantensystem.

Satz 3.20 (Charakterisierung von Äquivalenzklassen).
Sei auf der Menge M eine Äquivalenzrelation gegeben. Eine Teilmenge $A \subseteq M$

ist genau dann eine Äquivalenzklasse, wenn

1. $A \neq \emptyset$,
2. $x, y \in A \Rightarrow x \sim y$,
3. $x \in A \wedge y \in M \wedge x \sim y \Rightarrow y \in A$.

Beweis. Angenommen, A ist eine Äquivalenzklasse. Dann gibt es definitionsgemäß ein a mit $A = [a]$. Daher ist mindestens $a \in A$ und somit $A \neq \emptyset$. Mit $x, y \in A$ ergibt sich $A = [x] = [y]$. Aufgrund von

$$x \sim y \;\Leftrightarrow\; [a] = [b]$$

muss somit $x \sim y$ sein. Sei nun $x \in A$ und $y \in M$ mit $x \sim y$. Es folgt $A = [x] = [y]$. Daher muss $y \in A$ sein.

Umgekehrt angenommen, die drei Eigenschaften sind erfüllt. Zu zeigen ist, dass es ein a gibt mit $A = [a]$. Da A gemäß 1. nichtleer ist, enthält es mindestens ein Element, dieses nennen wir a. Für jedes weitere Element $x \in A$ ergibt sich $x \sim a$, da sonst 2. verletzt sein würde. Schließlich muss man noch wissen, ob $x \in A$, wenn $x \sim a$ und $x \in M$ ist. Dies ist aber mit 3. gesichert. Es gibt also tatsächlich ein a mit $A = \{x \in M \mid x \sim a\}$. □

Eine große Fülle von Äquivalenzrelationen lässt sich auf die folgende einfache Art konstruieren. Hat man eine beliebige Abbildung $f\colon X \to Y$, dann sind die Urbilder $f^{-1}(\{y_1\})$ und $f^{-1}(\{y_2\})$ disjunkt, sofern $y_1 \neq y_2$, denn

$$f^{-1}(\{y_1\}) \cap f^{-1}(\{y_2\}) = f^{-1}(\{y_1\} \cap \{y_2\}) = f^{-1}(\emptyset) = \emptyset.$$

Demnach ist gemäß

$$Z = X/{\sim} = \{f^{-1}(\{y\}) \mid y \in f(X)\}$$

eine Zerlegung des Definitionsbereichs X gegeben und somit auch eine Äquivalenzrelation. Für $x_1, x_2 \in X$ gilt

$$x_1 \sim x_2 \;\Leftrightarrow\; f(x_1) = f(x_2).$$

Ist f zudem surjektiv, dann gehört zu jedem Element von Y genau eine Äquivalenzklasse. Demnach definiert f dann eine verallgemeinerte Quotientenabbildung, da die Elemente von Y die Äquivalenzklassen charakterisieren. Die Bijektion $\varphi\colon Z \to Y$ hat dabei die Eigenschaft $f = \varphi \circ \pi$. Sofern Y für uns zugänglich ist, resultiert hieraus auch eine verallgemeinerte Darstellung der Quotientenabbildung, denn

$$f = \varphi \circ \pi = \varphi \circ (\pi|_A \circ p) = (\varphi \circ \pi|_A) \circ p.$$

Nun ist $\varphi \circ \pi|_A$ auch bijektiv, weil φ und $\pi|_A$ es sind. Somit charakterisiert Y ein vollständiges Repräsentantensystem.

Was bisher erläutert wurde, mag recht abstrakt erscheinen. Wir haben aber eigentlich ein recht intuitives Verständnis für diese Begrifflichkeiten. Ein Bilderbuchbeispiel für eine Quotientenmenge bieten die Klassen einer Schule. Zwei Schüler seien genau dann äquivalent, wenn sie in dieselbe Klasse gehen. Die Äquivalenzklasse eines Schülers ist dann schlicht seine Schulklasse. Die Menge der Schüler der Schule wird in disjunkte Schulklassen zerlegt. Die Menge dieser Schulklassen bildet die Quotientenmenge. Ein vollständiges Repräsentantensystem entsteht zum Beispiel durch die Wahl eines Klassensprechers in jeder Klasse.

Ein weiteres typisches Beispiel für eine Äquivalenzrelation ist die Kongruenz modulo m, die elementar in der Zahlentheorie und Gruppentheorie vorkommt. Die Äquivalenzklassen sind hier die Restklassen. Die Reste bilden ein kanonisches vollständiges Repräsentantensystem. Das Bilden des Restes zu einer Zahl ist eine Darstellung der Quotientenabbildung.

3.3.3. Operationen auf Äquivalenzklassen

Äquivalenzklassen werden später wichtig sein für die Formulierung von Konstruktionen. Bei diesen Konstruktionen ist eine Abbildung zwischen Quotientenmengen erforderlich. Weil die Äquivalenzklassen dabei über Repräsentanten dargestellt sind, liegt es nahe, auch die Abbildung über Repräsentanten zu definieren. Dies wirft die Frage nach der *Wohldefiniertheit* auf. Darunter versteht man, dass die Abbildung auch tatsächlich unabhängig von den gewählten Repräsentanten ist. Was das genau bedeutet, wird im folgenden Abschnitt erklärt.

Gegeben seien zwei Quotientenmengen $M/\sim$ und $M'/\sim'$. Eine vorhandene Abbildung $f\colon M \to M'$ induziert dann eventuell gemäß

$$f\colon M/\sim \longrightarrow M'/\sim', \quad f([a]) \coloneqq [f(a)]$$

eine Abbildung zwischen den Quotientenmengen. Kommt es dabei nicht zu einem Widerspruch, liegt also eine Abbildung vor, spricht man von *Wohldefiniertheit*. Hierfür darf der Funktionswert nicht vom gewählten Repräsentant abhängen, d. h. die Bedingung

$$\forall x \in [a]\colon f(x) \in [f(a)]$$

muss erfüllt sein. Anders formuliert,

$$x \sim a \;\Rightarrow\; f(x) \sim' f(a).$$

Für mehrstellige Abbildungen ist das Vorgehen analog. Eine Abbildung $f\colon M^2 \to M'$ induziert

$$f\colon (M/\!\sim)^2 \to M'/\!\sim', \quad f([a],[b]) := [f(a,b)],$$

sofern

$$x \sim a \wedge y \sim b \;\Rightarrow\; f(x,y) \sim' f(a,b).$$

Bei den Konstruktionen kommen in der Regel zweistellige Abbildungen (mit $M = M'$) vor, weil die Verknüpfungen von Elementen der algebraischen Strukturen zweistellig sind. Diese Verknüpfungen werden im nächsten Abschnitt besprochen.

3.3.4. Kongruenzrelationen

Definition 3.26 (Kongruenzrelation).
Gegeben sei eine Menge M, auf der eine zweistellige Verknüpfung $*\colon M^2 \to M$ definiert ist. Eine Äquivalenzrelation auf M nennt man *Kongruenzrelation*, wenn die induzierte Verknüpfung $[a] * [b] := [a * b]$ wohldefiniert ist.

Bei einer Kongruenzrelation sagt man »a ist kongruent zu b« anstelle von »a ist äquivalent zu b« und schreibt $a \equiv b$ anstelle von $a \sim b$. Eigentlich kann man den Begriff für eine beliebige Stelligkeit definieren. Es besteht jedoch zunächst nur Bedarf an zweistelligen Verknüpfungen.

Im Folgenden schreiben wir für die Verknüpfung kurz ab anstelle $a*b$. Das spart ein wenig Schreibaufwand und ist so üblich, solange keine Verwechslungsgefahr mit einer bereits auf andere Art definierten Multiplikation besteht.

Satz 3.21. Sei M eine Struktur aus der Liste Magma, Monoid, Halbgruppe, Gruppe, kommutatives Monoid, kommutative Gruppe. Sei $\equiv$ eine Kongruenzrelation auf M und φ die zugehörige Quotientenabbildung. Dann bildet die Quotientenmenge $M/\!\equiv$ bezüglich der induzierten Verknüpfung $\varphi(a)\varphi(b) := \varphi(ab)$ ebenfalls eine Struktur derselben Art und φ ist ein Homomorphismus.

Beweis. Im Folgenden seien a', b', c' beliebige Elemente der Quotientenmenge. Weil φ surjektiv ist, gibt es immer $a, b, c \in M$ mit $a' = \varphi(a)$, $b' = \varphi(b)$ und $c' = \varphi(c)$.

Die Verknüpfung auf M sei abgeschlossen. Dann ist

$$a'b' = \varphi(a)\varphi(b) = \varphi(ab) \in M/\!\equiv .$$

Somit ist die Quotientenmenge bezüglich der induzierten Verknüpfung abgeschlossen.

Die Verknüpfung auf M erfülle das Assoziativgesetz. Dann gilt

$$(a'b')c' = \varphi(ab)\varphi(c) = \varphi(abc) = \varphi(a)\varphi(bc) = a'(b'c').$$

Die induzierte Verknüpfung erfüllt somit ebenfalls das Assoziativgesetz.

Die Verknüpfung auf M habe ein neutrales Element e. Dann gilt

$$\varphi(a) = \varphi(ea) = \varphi(e)\varphi(a), \quad \varphi(a) = \varphi(ae) = \varphi(a)\varphi(e).$$

Demzufolge besitzt $M/\equiv$ mit $e' := \varphi(e)$ ebenfalls ein neutrales Element.

Zur Verknüpfung auf M gebe es zu jedem Element ein inverses. Dann gilt

$$\varphi(e) = \varphi(aa^{-1}) = \varphi(a)\varphi(a^{-1}), \quad \varphi(e) = \varphi(a^{-1}a) = \varphi(a^{-1})\varphi(a).$$

Demzufolge gibt es auf der Quotientenstruktur mit $\varphi(a)^{-1} := \varphi(a^{-1})$ ebenfalls zu jedem Element ein inverses.

Die Verknüpfung auf M sei kommutativ. Dann gilt

$$a'b' = \varphi(a)\varphi(b) = \varphi(ab) = \varphi(ba) = \varphi(b)\varphi(a) = b'a'.$$

Somit ist die Verknüpfung auf der Quotientenstruktur $M/\equiv$ ebenfalls kommutativ. □

Satz 3.22. Satz 3.21 gilt auch für Ringe, unitäre Ringe, kommutative Ringe und kommutative unitäre Ringe, sofern die Relation eine Kongruenzrelation sowohl bezüglich der additiven als auch der multiplikativen Verknüpfung ist.

Beweis. Sei $(R, +, \cdot)$ der Ring und $\equiv$ die Kongruenzrelation. Gemäß Satz 3.21 ist $(R/\equiv, +)$ eine kommutative Gruppe und $(R/\equiv, \cdot)$ eine Halbgruppe. Bei einem unitären Ring ist $(R/\equiv, \cdot)$ ein Monoid, und bei einem kommutativen unitären Ring ein kommutatives Monoid.

Es verbleiben noch die Distributivgesetze zu prüfen. Sei φ die Quotientenabbildung. Man rechnet

$$\begin{aligned} a'(b' + c') &= \varphi(a)(\varphi(b) + \varphi(c)) = \varphi(a)(\varphi(b + c)) = \varphi(a(b + c)) = \varphi(ab + ac) \\ &= \varphi(ab) + \varphi(ac) = \varphi(a)\varphi(b) + \varphi(a)\varphi(c) = a'b' + a'c'. \end{aligned}$$

Die Rechnung zum Rechtsdistributivgesetz ist analog.

Damit ist der Satz gezeigt, und ferner ist gezeigt dass φ ein Ringhomomorphismus ist. Und für einen unitären Ring ist φ Eins-erhaltend, wie bereits aus Satz 3.21 hervorgeht. □

3.3.5. Ordnungsrelationen

Bei vielen Untersuchungen genügt es nicht, Elemente einer Menge nur vergleichen zu können. Man möchte zusätzlich in Erfahrung bringen können, ob Elemente in einer bestimmten Weise geordnet sind, in einer bestimmten Weise in Reihenfolge stehen. Hierfür definiert man *Ordnungsrelationen*, von denen es verschiedene Arten gibt, je nachdem, welche Axiome sie erfüllen. Auf einer Menge können auch mehrere unterschiedliche Ordungsrelationen derselben Art definiert werden.

Eine in vielen Bereichen der Mathematik hiesige Art von Relation ist die *Halbordnung*, auch *Partialordnung* genannt. Eine wichtige Unterart der Halbordnungen stellen die *Totalordnungen* dar. So ist die Totalordnung der reellen Zahlen in der Analysis von zentraler Bedeutung.

Definition 3.27 (Halbordnung).
Eine auf einer Menge M definierte Relation $\leq$ heißt *Halbordnung*, wenn

$$\begin{aligned}
&\forall x \in M\colon x \leq x, && (\textit{Reflexivität})\\
&\forall x, y \in M\colon x \leq y \wedge y \leq x \Rightarrow x = y, && (\textit{Antisymmetrie})\\
&\forall x, y, z \in M\colon x \leq y \wedge y \leq z \Rightarrow x \leq z. && (\textit{Transitivität})
\end{aligned}$$

Ist $\leq$ eine Halbordnung auf M, nennt man M eine halbgeordnete Menge und kodiert dies als Struktur $(M, \leq)$, um klarzustellen, bezüglich welcher Relation die Menge halbgeordnet ist.

Die zu einer Halbordnung zugehörige strenge Halbordnung wird definiert durch

$$x < y \;:\Leftrightarrow\; x \leq y \wedge x \neq y.$$

Eine Relation R auf M heißt *irreflexiv*, wenn

$$\neg\exists x \in M\colon R(x, x),$$

in Worten, wenn kein Element zu sich selbst in Relation steht. Man bestätigt mühelos, dass die strenge Halbordnung irreflexiv und transitiv ist.

Definition 3.28 (Totalordnung).
Eine Halbordnung $\leq$ auf M heißt *Totalordnung*, wenn

$$\forall x, y \in M\colon x \leq y \vee y \leq x. \qquad (\textit{Totalität})$$

Das geläufigste Beispiel für eine Totalordnung ist die herkömmliche Ordnung der reellen Zahlen.

Wie bei Halbordnungen ist die zu einer Totalordnung zugehörige strenge Totalordnung erklärt durch

$$x < y \;:\Leftrightarrow\; x \leq y \wedge x \neq y.$$

Satz 3.23. Die strenge Totalordnung ist *trichotom*, das heißt,

$$\forall x, y \in M\colon (x < y) \oplus (x = y) \oplus (y < x).$$

Beweis. Zunächst machen wir uns klar,

$$A \oplus B \oplus C \;\Leftrightarrow\; (A \wedge \neg B \wedge \neg C) \vee (\neg A \wedge B \wedge \neg C) \vee (\neg A \wedge \neg B \wedge C).$$

In dem Fall, wo eine der drei Aussagen gezeigt wird, sind also außerdem die beiden anderen Aussagen zu widerlegen.

Es wird eine Fallunterscheidung bezüglich der Gleichheit unternommen. Im Fall $x = y$ führt laut Definition der strengen Totalordnung sowohl $x < y$ als auch $y < x$ zu einem Widerspruch. Es sei nun $x \neq y$ der Fall. Laut Totalität liegt $x \leq y \vee y \leq x$ vor. Es sei $x \leq y$ der Unterfall. Sofort folgt $x < y$ gemäß Definition der strengen Totalordnung. Angenommen, es wäre $y < x$, dann wäre neben $x \leq y$ laut Definition der strengen Totalordnung mithin auch $y \leq x$, womit laut Antisymmetrie $x = y$ wäre, was aber im Widerspruch zu $x \neq y$ steht. Die Argumentation im Unterfall $y \leq x$ verläuft analog. □

Satz 3.24. Ist $\leq$ eine Totalordnung auf M, dann gilt

$$\forall x, y \in M\colon \neg(x \leq y) \Leftrightarrow y < x.$$

Beweis. Von links nach rechts. Es gelte $\neg(x \leq y)$. Laut Totalität muss dann $y \leq x$ sein. Angenommen, es wäre $x = y$, dann wäre insbesondere $x \leq y$, was widersprüchlich zur Voraussetzung ist. Also muss $x \neq y$ sein, und somit $y < x$.

Von rechts nach links. Es gelte $y < x$, also $y \leq x$ und $x \neq y$. Angenommen, es wäre $x \leq y$, dann wäre $x = y$ gemäß Antisymmetrie, was aber im Widerspruch zu $x \neq y$ steht. □

Sei $(M, \leq)$ halbgeordnet und $A \subseteq M$. Wir legen die Kurzschreibweisen

$$A \leq s \;:\Leftrightarrow\; \forall x \in A\colon x \leq s,$$
$$s \leq A \;:\Leftrightarrow\; \forall x \in A\colon s \leq x$$

fest. Ein s mit $A \leq s$ heißt *obere Schranke* von A. Existiert mindestens ein s mit dieser Eigenschaft, nennt man A *nach oben beschränkt*. Ein s mit $s \leq A$ heißt *untere Schranke* von A. Existiert mindestens ein s mit dieser Eigenschaft, nennt man A *nach unten beschränkt*.

Definition 3.29 (Maximum, Minimum).
Sei $(M, \leq)$ halbgeordnet und $A \subseteq M$. Man legt fest

$$y = \max(A) :\Leftrightarrow y \in A \wedge A \leq y,$$
$$y = \min(A) :\Leftrightarrow y \in A \wedge y \leq A.$$

Statt vom Maximum spricht man auch vom größten Element, anstelle vom Minimum vom kleinsten Element. Die vier Sprechweisen deuten hin auf den

Satz 3.25. Eine Teilmenge einer halbgeordneten Menge besitzt höchstens ein Maximum und höchstens ein Minimum.

Beweis. Sei $y = \max(A)$ und $y' = \max(A)$. Laut Definition des Maximums sind dann $y, y' \in A$, wobei $\forall x \in A\colon x \leq y$ und $\forall x \in A\colon x \leq y'$. Die erste Allaussage wird mit $x := y'$ spezialisiert, die zweite mit $x := y$. Somit gilt $y' \leq y$ und $y \leq y'$, ergo $y = y'$ gemäß der Antisymmetrie. Zum Minimum analog. □

Definition 3.30 (Supremum, Infimum).
Sei $(M, \leq)$ halbgeordnet und $A \subseteq M$. Man legt fest

$$\sup(A) := \min\{s \in M \mid A \leq s\}, \qquad (\textit{kleinste obere Schranke})$$
$$\inf(A) := \max\{s \in M \mid s \leq A\}. \qquad (\textit{größte untere Schranke})$$

Entfaltung von min und max führt zu

$$y = \sup(A) \;\Leftrightarrow\; A \leq y \wedge \forall s \in M\colon A \leq s \Rightarrow y \leq s,$$
$$y = \inf(A) \;\Leftrightarrow\; y \leq A \wedge \forall s \in M\colon s \leq A \Rightarrow s \leq y.$$

Definition 3.31 (Maximales Element, Minimales Element).
Sei $(M, \leq)$ halbgeordnet und $A \subseteq M$. Man legt fest

$$y \text{ ist ein } \textit{maximales Element} \text{ von } A :\Leftrightarrow y \in A \wedge \neg\exists x \in A\colon y < x,$$
$$y \text{ ist ein } \textit{minimales Element} \text{ von } A :\Leftrightarrow y \in A \wedge \neg\exists x \in A\colon x < y.$$

Sie lassen sich am Hasse-Diagramm veranschaulichen. Ein minimales Element ist eines, in das keine Kanten hineinlaufen. Ein maximales Element ist eines, aus dem keine Kanten herauslaufen. Weiterhin gilt der

Satz 3.26. Existiert ein Maximum, so ist dieses das einzige maximale Element. Existiert ein Minimum, so ist dieses das einzige minimale Element.

Beweis. Sei y das Maximum von A. Wir zeigen, dass y ein maximales Element ist, wofür $\neg\exists x \in A\colon y < x$ zu bestätigen ist. Das heißt, wir müssen uns von

$$(\forall x \in A\colon x \leq y), (\exists x \in A\colon y < x) \vdash \bot$$

überzeugen. Also angenommen, es läge ein $x \in A$ mit $y < x$ vor, womit $y \leq x$ und $x \neq y$ wäre. Die Allaussage wird mit diesem x spezialisiert. Aus $x \leq y$ und $y \leq x$ folgt $x = y$ per Antisymmetrie, was im Widerspruch zu $x \neq y$ steht.

Weiterhin gilt es noch zu zeigen, dass kein weiteres maximales Element existieren kann. Sei y' ein weiteres maximales Element. Weil y das Maximum ist, muss $y' \leq y$ gelten. Wegen $y \neq y'$ gilt außerdem $y' < y$. Weil y' maximal sein soll, dürfte kein $x \in A$ mit $y' < x$ zu finden sein, doch wir haben so eines mit $x := y$, was absurd ist. Zum Minimum verläuft der Beweis analog. □

Es drängt sich einigen Lesern nun womöglich auf, leichtfertig die Umkehrungen für allgemeingültig zu erachten. Ich meine die Aussage »Existiert ein einziges maximales Element, muss es sich bei diesem um das Maximum handeln.« Bei unendlichen Mengen braucht das aber nicht unbedingt richtig sein. Nämlich kann neben dem maximalen Element außerdem noch ein unendlicher Aufstieg vorliegen, womit kein Maximum existiert. Analog kann neben dem minimalen Element außerdem noch ein unendlicher Abstieg vorliegen. Im Englischen sind die Abkürzungen ACC und DCC geläufig, die für *Ascending Chain Condition* und *Descending Chain Condition* stehen. Die ACC ist bei Abhandensein unendlicher aufsteigender Ketten

$$x_0 < x_1 < x_2 < \ldots$$

erfüllt. Analog ist die DCC bei Abhandensein unendlicher absteigender Ketten

$$x_0 > x_1 > x_2 > \ldots$$

erfüllt. Unter der ACC bzw. DCC ist die jeweilige Umkehrung tatsächlich richtig.

Ein wichtiges Beispiel für eine Halbordnung ist die Teilmengenbeziehung. Man folgert mühelos, dass das Maximum einer Menge, sofern existent, zugleich ihr Supremum sein muss. Analog ist das Minimum einer Menge, sofern existent, zugleich ihr Infimum.

Satz 3.27. Sei G eine Grundmenge und $\mathcal{A} \subseteq \mathcal{P}(G)$ mit $\mathcal{A} \neq \emptyset$, das heißt, eine nichtleere Menge von Teilmengen von G. Auf $(\mathcal{P}(G), \subseteq)$ gilt

$$\bigcap \mathcal{A} = \inf(\mathcal{A}), \qquad \bigcup \mathcal{A} = \sup(\mathcal{A}).$$

Beweis. Zum Infimum. Entfaltung bringt die Aussage in die Form

$$(\forall A \in \mathcal{A} \colon \bigcap \mathcal{A} \subseteq A) \wedge (\forall S \in \mathcal{A} \colon (\forall A \in \mathcal{A} \colon S \subseteq A) \Rightarrow S \subseteq \bigcap \mathcal{A}).$$

Zur ersten Aussage der Konjunktion gelangt man durch Bestätigung von

$$A \in \mathcal{A}, x \in \bigcap \mathcal{A} \vdash x \in A,$$

was aber gerade aus der definierenden Eigenschaft von $\bigcap \mathcal{A}$ hervorgeht. Zur zweiten Aussage der Konjunktion gelangt man durch Bestätigung von

$$S \in \mathcal{A}, (\forall A \in \mathcal{A} \colon S \subseteq A), x \in S \vdash \forall A \in \mathcal{A} \colon x \in A.$$

Also angenommen, $A \in \mathcal{A}$, dann erhält man $S \subseteq A$ vermittels der vorhandenen Allaussage. Mit $x \in S$ folgt somit die gewünschte Aussage $x \in A$.

Zum Supremum. Entfaltung bringt die Aussage in die Form

$$(\forall A \in \mathcal{A} \colon A \subseteq \bigcup \mathcal{A}) \wedge (\forall S \in \mathcal{A} \colon (\forall A \in \mathcal{A} \colon A \subseteq S) \Rightarrow \bigcup \mathcal{A} \subseteq S).$$

Zur ersten Aussage der Konjunktion gelangt man durch Bestätigung von

$$A \in \mathcal{A}, x \in A \vdash \exists A' \in \mathcal{A} \colon x \in A'.$$

Man wähle dazu schlicht die Menge $A' := A$, bei der die geforderte Eigenschaft bereits vorausgesetzt ist. Zur zweiten Aussage der Konjunktion gelangt man durch Bestätigung von

$$S \in M, (\forall A \in \mathcal{A} \colon A \subseteq S), (\exists A' \in \mathcal{A} \colon x \in A') \vdash x \in S.$$

Es liegt ein $A' \in \mathcal{A}$ mit $x \in A'$ vor. Die Allaussage wird mit $A := A'$ spezialisiert, womit man $A' \subseteq S$ erhält. Mit $x \in A'$ folgt somit die gewünschte Aussage $x \in S$. □

3.3.6. Monotone Abbildungen

Definition 3.32 (Monotone Abbildung).
Es seien $(M, \leq), (M', \leq')$ halbgeordnet. Eine Abbildung $f\colon M \to M'$ heißt

monoton steigend,	wenn $\forall x, y \in M\colon x \leq y \Rightarrow f(x) \leq' f(y)$,
monoton fallend,	wenn $\forall x, y \in M\colon x \leq y \Rightarrow f(y) \leq' f(x)$,
streng monoton steigend,	wenn $\forall x, y \in M\colon x < y \Rightarrow f(x) <' f(y)$,
streng monoton fallend,	wenn $\forall x, y \in M\colon x < y \Rightarrow f(y) <' f(x)$.

Beispiele. Die Folge $a\colon \mathbb{N} \to \mathbb{R}$ mit $a_n := n$ ist streng monoton steigend. Monotonie von Folgen spielt eine wichtige Rolle in der Konvergenztheorie.

Die reelle Funktion $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ mit $f(x) := x^2$ ist streng monoton steigend. Monotonie reeller Funktionen spielt eine wichtige Rolle in der Analysis.

Es sei G eine endliche Grundmenge. Die Abbildung

$$f\colon M \to M', \quad f(A) := |A|, \quad M := \mathcal{P}(G), \quad M' := \{0, \ldots, |M|\}$$

ist monoton steigend bezüglich $(M, \subseteq)$ und $(M', \leq)$.

Satz 3.28. Es sei $(M, \leq)$ eine totalgeordnete und $(M', \leq')$ eine halbgeordnete Menge. Ist $f\colon M \to M'$ streng monoton, so ist f auch injektiv.

Beweis. Gezeigt wird für streng monoton steigendes f die Kontraposition

$$x \neq y \Rightarrow f(x) \neq f(y).$$

Es gelte $x \neq y$. Gemäß Trichotomie ist dann entweder $x < y$ oder $y < x$. Im Fall $x < y$ ist $f(x) <' f(y)$, und somit $f(x) \neq f(y)$. Im Fall $y < x$ ist $f(y) <' f(x)$, und somit ebenfalls $f(x) \neq f(y)$. □

Satz 3.29. Seien $(M, \leq), (M', \leq')$ totalgeordnete Mengen. Ist $f\colon M \to M'$ streng monoton steigend (fallend) und surjektiv, so ist f bijektiv und f^{-1} ebenfalls streng monoton steigend (fallend).

Beweis. Laut Satz 3.28 ist f bijektiv. Wir betrachten die Kontraposition,

$$(x <' y \Rightarrow f^{-1}(x) < f^{-1}(y)) \iff (f^{-1}(y) \leq f^{-1}(x) \Rightarrow y \leq' x).$$

Es gelte $f^{-1}(y) \leq f^{-1}(x)$. Weil f streng monoton steigend ist, gilt $f(f^{-1}(y)) \leq' f(f^{-1}(x))$, ergo $y \leq' x$. Für fallende Abbildungen verläuft der Beweis analog. □

Beispiel. Wir wollen die reelle Funktion $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ mit $f(x) := x^2$ als injektiv wissen. Hierfür wird

$$x \geq 0, x < y \vdash x^2 < y^2$$

gezeigt. Aus $0 \leq x$ und $x < y$ erhält man zunächst $y > 0$. Multipliziert man $x < y$ nun beidseitig mit y, gelangt man zu $xy < y^2$. Im Fall $x = 0$ ist $x^2 = 0$ und $xy = 0$, ergo $x^2 < y^2$. Im Fall $x > 0$ multipliziert man $x < y$ beidseitig mit x und erhält $x^2 < xy$. Per Transitivität findet sich ebenfalls $x^2 < y^2$.

Die Antwort auf die Frage, ob f surjektiv ist, vermittelt der Zwischenwertsatz. Hierfür ist zunächst der Begriff der stetigen Funktion zu erklären. So wie es scheint, sind streng genommen bereits Mittel der Analysis erforderlich, um einfache algebraische Problemstellungen lösen zu können. Es geht ja am Ende zumal darum, die beidseitige Anwendung von f als Äquivalenzumformung einer Gleichung oder einer Ungleichung zu erkennen.

3.4. Kardinalzahlen

3.4.1. Gleichmächtigkeit

Definition 3.33 (Gleichmächtigkeit).
Zwei Mengen seien *gleichmächtig*, wenn zwischen diesen mindestens eine Bijektion existiert.

Die Aussage »A ist gleichmächtig zu B« notiert man kurz $|A| = |B|$.

Fast trivial ist, dass es sich bei der Gleichmächtigkeit um eine Äquivalenzrelation handelt. Seien dazu A, B, C beliebig. Eine Bijektion $A \to A$ liefert die identische Abbildung, womit Reflexivität besteht. Liegt eine Bijektion $A \to B$ vor, so auch ihre bijektive Umkehrabbildung $B \to A$, womit Symmetrie besteht. Schließlich ist die Verkettung einer Bijektion $A \to B$ mit einer Bijektion $B \to C$ ebenfalls bijektiv, womit Transitivität besteht.

Die kontraintuitive Eigenart des Unendlichen wird gerne am *hilbertschen Hotel* veranschaulicht. Es handelt sich um ein Hotel mit unendlich vielen Zimmern, die mit den positiven ganzen Zahlen nummeriert sind. Anders als ein endliches, kann es auch dann weitere Gäste aufnehmen, wenn es bereits voll belegt ist.

Die Zimmer im hilbertschen Hotel seien mit 1, 2, 3 usw. nummeriert und gänzlich jeweils mit einem Gast belegt. Nun kommt ein neuer Gast mit der Nummer 0 hinzu. Um auch ihn im Hotel unterbringen zu können, wechselt jeder Gast in das nächste Zimmer. Formal handelt es sich um die Zuordnung

$$f\colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 1}, \quad f(n) := n + 1.$$

Hierbei kommt es, begründet durch die Endlosigkeit des Hotels, nie zu einer Komplikation. Es gibt kein letztes Zimmer, dessen Gast nun mangels freiem Zimmer im Flur stünde. Die Funktion f ist ersichtlich bijektiv, denn *jedes* Zimmer des Hotels wird belegt, und zwar von *nicht mehr als einem* Gast. Das bedeutet aber, dass die nichtnegativen ganzen Zahlen, obwohl sie ein weiteres Element enthalten, zu den positiven ganzen Zahlen gleichmächtig sind.

Allgemeiner kann auf diese Weise jede endliche Zahl weiterer Gäste im Hotel untergebracht werden. *Die Mächtigkeit einer abgezählten unendlichen Menge erhöht sich nicht durch die Hinzunahme einer endlichen Zahl weiterer Elemente.*

3.4.2. Kardinalzahlarithmetik

Satz 3.30. Die Addition $|A| + |B| := |A \cup B|$ ist für $A \cap B = \emptyset$ wohldefiniert.

Beweis. Zu zeigen ist, dass das Ergebnis nicht von der Wahl der Repräsentanten A, B abhängt. Das heißt, es ist zu zeigen, dass $|A \cup B| = |A' \cup B'|$ aus $|A| = |A'|$

und $|B| = |B'|$ mit $A' \cap B' = \emptyset$ folgt. Nach Voraussetzung existieren Bijektionen $f_1 \colon A \to A'$ und $f_2 \colon B \to B'$. Wie bei der Summe, dem algebraischen Datentyp zur disjunkten Vereinigung, konstruiert man eine Fallunterscheidung

$$f \colon A \cup B \to A' \cup B', \quad f(x) := \begin{cases} f_1(x) & \text{für } x \in A, \\ f_2(x) & \text{für } x \in B. \end{cases}$$

Die Abbildung f ist injektiv. Sei dazu y ein beliebiges Bild. Weil A', B' disjunkt sind, ist entweder $y \in A'$, womit

$$f(x_1) = f(x_2) = y \in A' \implies y = f_1(x_1) = f_1(x_2) \implies x_1 = x_2,$$

oder $y \in B'$, womit

$$f(x_1) = f(x_2) = y \in B' \implies y = f_2(x_1) = f_2(x_2) \implies x_1 = x_2.$$

Zusammengefasst folgt $x_1 = x_2$ aus $f(x_1) = f(x_2)$.

Auch ist f surjektiv. Sei dazu y ein beliebiges Element der Zielmenge. Weil A', B' disjunkt sind, ist entweder $y \in A'$, womit ein x mit $y = f_1(x)$ vorliegt, oder $y \in B'$, womit ein x mit $y = f_2(x)$ vorliegt. Zusammengefasst existiert zu jedem $y \in A' \cup B'$ ein x mit $y = f(x)$. □

Satz 3.31. Die Multiplikation $|A| \cdot |B| := |A \times B|$ ist wohldefiniert.

Beweis. Zu zeigen ist, dass $|A \times B| = |A' \times B'|$ aus $|A| = |A'|$ und $|B| = |B'|$ folgt. Nach Voraussetzung existieren Bijektionen $f_1 \colon A \to A'$ und $f_2 \colon B \to B'$. Man konstruiert mit ihnen die Bijektion

$$f \colon A \times B \to A' \times B', \quad f(a, b) := (f_1(a), f_2(b)).$$

Sie ist injektiv, denn für alle Paare $(a_1, b_1), (a_2, b_2) \in A \times B$ gilt

$$\begin{aligned} f(a_1, b_1) = f(a_2, b_2) &\iff (f_1(a_1), f_2(b_1)) = (f_1(a_2), f_2(b_2)) \\ &\iff f_1(a_1) = f_1(a_2) \wedge f_2(b_1) = f_2(b_2) \\ &\iff a_1 = a_2 \wedge b_1 = b_2 \iff (a_1, b_1) = (a_2, b_2). \end{aligned}$$

Sie ist auch surjektiv. Sei dazu $(a', b') \in A' \times B'$ beliebig. Wegen $a' \in A'$ existiert ein a mit $a' = f_1(a)$. Wegen $b' \in B'$ existiert ein b mit $b' = f_2(b)$. Laut Konstruktion liegt mit (a, b) somit ein Zeuge für $(a', b') = f(a, b)$ vor. □

Satz 3.32. Die Potenzierung $|B|^{|A|} := |B^A|$ ist wohldefiniert.

Beweis. Zu zeigen gilt, dass $\mathrm{Abb}(A,B) = \mathrm{Abb}(A',B')$ aus $|A| = |A'|$ und $|B| = |B'|$ folgt, wobei wir $\mathrm{Abb}(A,B)$ statt B^A schreiben. Nach Voraussetzung existieren Bijektionen $f_1 \colon A \to A'$ und $f_2 \colon B \to B'$. Man konstruiert mit ihnen die Bijektion

$$F \colon \mathrm{Abb}(A,B) \to \mathrm{Abb}(A',B'), \quad F(f) := f_2 \circ f \circ f_1^{-1}.$$

Sie ist injektiv, weil

$$F(f) = F(g) \iff f_2 \circ f \circ f_1^{-1} = f_2 \circ g \circ f_1^{-1} \iff f = g.$$

Die letzte Umformung gilt, weil Bijektionen ja linkskürzbar und rechtskürzbar sind.

Sie ist auch surjektiv. Dazu ist zu jedem f' ein Zeuge f für $f' = F(f)$ gesucht. Man darf $f := f_2^{-1} \circ f' \circ f_1$ setzen, denn

$$f_2^{-1} \circ f' \circ f_1 = f \iff f' \circ f_1 = f_2 \circ f \iff f' = f_2 \circ f \circ f_1^{-1} = F(f). \qquad \square$$

3.4.3. Der Satz von Cantor

Satz 3.33. Es sei X eine beliebige Menge. Ihre Potenzmenge $\mathcal{P}(X)$ ist zur Menge $\mathrm{Abb}(X,\{0,1\})$ der binärwertigen Abbildungen gleichmächtig.

Beweis. Jede Menge $A \subseteq X$ wird kodiert durch ihre Indikatorfunktion

$$1_A \colon X \to \{0,1\}, \quad 1_A(x) := [x \in A] = \begin{cases} 1, & \text{wenn } x \in A, \\ 0, & \text{wenn } x \notin A. \end{cases}$$

Sie vermitteln die kanonische Bijektion

$$\varphi \colon \mathcal{P}(X) \to \mathrm{Abb}(X,\{0,1\}), \quad \varphi(A) := 1_A.$$

Zur Injektivität. Def. 3.17 verlangt die Bestätigung von

$$\varphi(A) = \varphi(B) \Rightarrow A = B, \text{ das heißt } 1_A = 1_B \Rightarrow A = B.$$

Die linke Seite enfaltet sich damit, dass zwei Abbildungen genau dann gleich sind, wenn sie in jedem ihrer Funktionswerte übereinstimmen. Die rechte Seite gewinnt man per Extensionalität, das ist Axiom 3.1. Dies führt zu

$$(\forall x \colon 1_A(x) = 1_B(x)) \Rightarrow (\forall x \colon x \in A \Leftrightarrow x \in B).$$

Die Aussage stimmt, denn man darf umformen

$$1_A(x) = 1_B(x) \iff [x \in A] = [x \in B] \iff (x \in A \Leftrightarrow x \in B).$$

Zur Surjektivität. Es ist hierfür zu prüfen, dass $\mathrm{Abb}(X, \{0,1\})$ eine Teilmenge der Bildmenge $\varphi(\mathcal{P}(X))$ ist. Entfaltung von Def. 3.2 und Def. 3.13 führt zu

$$\forall f\colon \Big(f \in \mathrm{Abb}(X, \{0,1\}) \Rightarrow \exists A \in \mathcal{P}(X)\colon f = \varphi(A)\Big).$$

Dem Existenzquantor genügt die »Einsfaser«

$$A := f^{-1}(\{1\}) = \{x \in X \mid f(x) \in \{1\}\} = \{x \in X \mid f(x) = 1\}.$$

Es gilt $f = 1_A$, denn

$$1_A(x) = [x \in A] = [x \in \{x \mid f(x) = 1\}] = [f(x) = 1] = f(x). \square$$

Eine Teilmenge wird durch ihre Indikatorfunktion charakterisiert. Stellen wir uns beispielsweise einen gedrückten Akkord als Teilmenge einer Klaviatur vor. Die Indikatorfunktion ist hier eine endliche Folge, entspricht also einem Tupel, welches genau in den gedrückten Tasten den Wert 1 besitzt. Das Tupel ist zwar nicht der Akkord selbst, kodiert aber genau die Information des Akkords.

Satz 3.34 (Satz von Cantor).
Jede Menge ist weniger mächtig als ihre Potenzmenge.

Beweis. Sei X eine Menge. Zu zeigen ist $|X| < |\mathcal{P}(X)|$. Mit $x \mapsto \{x\}$ liegt unschwer Einsichtig eine Injektion $X \to \mathcal{P}(X)$ vor, denn aus $\{x\} = \{x'\}$ folgt $x = x'$.

Die Widerlegung der Existenz einer Surjektion $f\colon X \to \mathcal{P}(X)$ klärt uns nun darüber auf, dass erst recht keine Bijektion bestehen kann. Man erreicht dies nach Cantor durch das Diagonalargument zweiter Art. Hierfür definiert man die Diagonalmenge

$$D := \{x \in X \mid x \notin f(x)\}.$$

Als Aussonderung aus X ist D eine Teilmenge von X, also ein Element von $\mathcal{P}(X)$. Weil f als surjektiv angenommen wird, muss $D \in f(X)$ sein. Das heißt, es muss ein $x \in X$ mit $D = f(x)$ vorliegen. Weil es sich dabei um eine Gleichheit zwischen Mengen handelt, ist insbesondere $x \in D$ äquivalent zu $x \in f(x)$. Laut der Definition von D ist $x \in D$ andererseits, weil $x \in X$ vorliegt, äquivalent zu $x \notin f(x)$. Summa summarum besteht die widersprüchliche Äquivalenz

$$x \in f(x) \Leftrightarrow x \notin f(x). \square$$

Hierneben gibt es auch noch die folgende speziellere Formulierung des Sachverhaltes. Es ist $\mathbb{N}$ weniger mächtig als $\{0,1\}^{\mathbb{N}}$. Angenommen, es wäre $f\colon \mathbb{N} \to \{0,1\}^{\mathbb{N}}$ eine Abzählung. Sie ordnet jeder natürlichen Zahl n eine Binärfolge

$$(f(n)_0, f(n)_1, f(n)_2, \ldots)$$

zu. Nun kann man allerdings eine weitere Binärfolge konstruieren, die sich von allen anderen unterscheidet, im Widerspruch zur Annahme, f würde alle Binärfolgen abzählen. Dies geschieht wieder per Cantors Diagonalargument zweiter Art. Man definiert dafür die Folge

$$d\colon \mathbb{N} \to \{0,1\}, \quad d_n := 1 - f(n)_n.$$

Listet man die Folgen tabellarisch auf, je Zeile eine Folge, handelt es sich bei (d_n) um die negierte Diagonale, womit sie sich von jeder Folge der Abzählung mindestens in dieser unterscheiden muss.

Zu klären verbleibt, wie die allgemeine Form des Beweises mit der speziellen zusammenhängt. Hierfür erinnern wir uns, dass zu jeder Teilmenge von $\mathbb{N}$ in natürlicher Weise genau eine Indikatorfunktion gehört, womit eine Bijektion zwischen $\mathcal{P}(\mathbb{N})$ und $\{0,1\}^{\mathbb{N}}$ hergestellt wird. Diese Indikatorfunktionen sind nichts anderes als die Binärfolgen. Demnach konstituiert

$$D := \{n \in \mathbb{N} \mid f(n)_n = 0\}$$

die analoge Beschreibung der Diagonalmenge. Wäre f surjektiv, gäbe es ein n mit $1_D = f(n)$, womit insbesondere $1_D(n) = f(n)_n$ wäre. Demnach ist $f(n)_n = 1$ äquivalent zu $1_D(n) = 1$, was $n \in D$ heißt, also $f(n)_n = 0$. Die widersprüchliche Äquivalenz gestaltet sich nun also als

$$f(n)_n = 1 \Leftrightarrow f(n)_n = 0.$$

Die letztendliche Klärung liefert die Rechnung

$$1_D(n) = [f(n)_n = 0] = [\neg f(n)_n = 1] = 1 - [f(n)_n = 1] = 1 - f(n)_n = d_n.$$

Kurzum ist $1_D = d$. Die Folge der negierten diagonalen Bits entpuppt sich als Indikatorfunktion der Diagonalmenge.

Man kann sich so eine Binärfolge als einen Pfad eines unendlich großen Binärbaums vorstellen. Von jedem Knoten gehen zwei Zweige aus, wobei das jeweilige Bit der Binärfolge kodiert, welcher der Zweige zu beschreiten ist. Das Abzählbare stellt man sich vor wie eine unendliche Reihe von Straßenlaternen, die man alle

auf einmal betrachtet. Die Potenzmenge des Abzählbaren dagegen wie die Blätter des Baums, von denen jedes in den unendlich vielen Schritten des Durchschreitens seiner kodierenden Binärfolge erreicht würde.

Ähnlich wie das hilbertsche Hotel besitzt so ein Baum die kontraintuitive Eigenschaft, dass jeder seiner Teilbäume genau so viele Blätter kodiert wie der gesamte Baum.

Die bisher gemachten Feststellungen sind schließlich als Folgerungen eines allgemeinen Satzes darstellbar, dessen Formulierung zunächst auf Mengen beschränkt bleiben soll. Für die allgemeine Fassung müssten wir zunächst die Begrifflichkeiten der Kategorientheorie einführen, was an dieser Stelle aber den Rahmen sprengen würde.

Satz 3.35 (Fixpunktsatz von Lawvere für Mengen).
Existiert eine Surjektion $X \to \mathrm{Abb}(X, Y)$, muss jede Abbildung $Y \to Y$ mindestens einen Fixpunkt besitzen.

Beweis. Es sei $\varphi\colon Y \to Y$ beliebig. Laut Voraussetzung existiert eine Surjektion $f\colon X \to \mathrm{Abb}(X, Y)$. Man setzt

$$d\colon X \to Y, \quad d(x) := \varphi(f(x)(x)).$$

Weil f surjektiv ist, existiert ein $a \in X$ mit $f(a) = d$. Es ist mit $y_a := f(a)(a)$ ein Fixpunkt gefunden, denn

$$\varphi(y_a) = \varphi(f(a)(a)) = d(a) = f(a)(a) = y_a. \qquad \square$$

Für $|Y| \geq 2$ kann man immer eine Abbildung ohne Fixpunkt finden. Speziell für $X := \mathbb{N}$ und $Y := \{0, 1\}$ findet sich $\varphi(y) := 1 - y$, womit d zu der zuvor diskutierten negierten Diagonale wird.

Ein weiteres Beispiel. Es sei $X = V_\alpha$ eine Menge der Von-Neumann-Hierarchie und $Y := \{0, 1\}$. Die Abbildung

$$f\colon X \to \mathrm{Abb}(X, Y), \quad f(x)(u) := (u \in x)$$

ordnet jeder Menge x ihr Prädikat $P = f(x)$ zu, so dass sie umgekehrt als

$$x = \{u \in X \mid P(u)\}$$

beschrieben ist, da f injektiv ist, denn es gilt $(x \in X \Rightarrow x \subseteq X)$ und daraufhin

$$f(x) = f(x') \iff (\forall u\colon (u \in x) = (u \in x')) \iff x = x'$$

gemäß dem Prinzip der Extensionalität, das die Gleichheit von Mengen definiert. Angenommen, f wäre surjektiv, das heißt, jedes Prädikat beschriebe eine Menge. Dann hätte jede boolesche Funktion $\varphi\colon Y \to Y$ einen Fixpunkt. Aber die Verneinung $\varphi(y) := \neg y$ besitzt keinen Fixpunkt. Das heißt, es existiert keine Menge $a \in X$, die die Gleichung

$$f(a)(a) = \varphi(f(a)(a)), \iff (a \in a) = (a \notin a),$$

erfüllt. Wir haben hier eine gewisse Abart der russellschen Antinomie in leichter Umformulierung. Die Diagonalmenge $D = \{x \in X \mid d(x)\}$ entspricht hierbei der russellschen Klasse. Demnach muss $D \notin X$ gelten.

3.5. Induktive Mengen

3.5.1. Modellierung der natürlichen Zahlen

Wir wollen uns nun eingehender mit dem Wesen der Mengen beschäftigen, über die die Induktion verläuft. Der Leser mag diesen Abschnitt beim ersten Lesen überspringen.

Als Grundmenge werde die Menge der reellen Zahlen gewählt. Die Basis sei die Einermenge $B = \{0\}$ und die einzige Produktionsregel sei die Nachfolgerabbildung $S(n) := n + 1$. Die Menge der induktiven Mengen definieren wir als

$$\mathcal{A} := \{A \subseteq \mathbb{R} \mid B \subseteq A \wedge \forall n \in A\colon S(n) \in A\}.$$

In Worten enthält sie alle Mengen, die die Basis als Untermenge haben und bezüglicher der Nachfolgerabbildung abgeschlossen sind.

Wir definieren die natürlichen Zahlen als die Schnittmenge

$$\mathbb{N} := \bigcap \mathcal{A}.$$

Aufgrund von Satz 3.27 ist $\mathbb{N}$ das Infimum der induktiven Mengen. Wir zeigen nun, dass $\mathbb{N}$ selbst eine der induktiven Mengen ist, womit sie das Minimum, die kleinste induktive Menge sein muss. Das bedeutet, jede andere induktive Menge muss eine Obermenge von $\mathbb{N}$ sein.

Satz 3.36. Es ist $\mathbb{N} \in \mathcal{A}$.

Beweis. Der Beweis ist abermals rein technischer Natur. Die Aussage $\mathbb{N} \in \mathcal{A}$ wird zunächst bezüglich der Definition von $\mathcal{A}$ entfaltet, das bringt sie in die Form

$$B \subseteq \mathbb{N} \wedge \forall n \in \mathbb{N}\colon S(n) \in \mathbb{N}.$$

Erste Aussage der Konjunktion. Entfaltung von $B \subseteq \mathbb{N}$ führt zu

$$\forall n\colon n \in B \Rightarrow \forall A \in \mathcal{A}\colon n \in A.$$

Mit $A \in \mathcal{A}$ ist $B \subseteq A$, somit ist mit $n \in B$ auch $n \in A$.

Entfaltung der zweiten Aussage der Konjunktion führt zu

$$\forall n\colon (\forall A \in \mathcal{A}\colon n \in A) \Rightarrow (\forall A \in \mathcal{A}\colon S(n) \in A),$$

die via

$$(\forall A \in \mathcal{A}\colon n \in A), A \in \mathcal{A} \vdash S(n) \in A.$$

bewiesen wird. Mit $A \in \mathcal{A}$ erhält man zum einen $n \in A$ und laut Definition von $\mathcal{A}$ zum anderen $\forall n \in A\colon S(n) \in A$. Mit $n \in A$ folgt also $S(n) \in A$. □

Aus der Beschreibung als Schnitt geht in der Tat hervor, dass das Prinzip der Induktion der Menge der natürlichen Zahlen innewohnt. Wir arbeiten es folgenderweise heraus. Es sei $A \subseteq \mathbb{N}$. Es sei außerdem gezeigt worden, dass A induktiv ist. Allgemein folgt aber $\bigcap \mathcal{A} \subseteq A$ aus $A \in \bigcap \mathcal{A}$. Weil ja $\mathbb{N}$ der Schnitt aller induktiven Mengen ist, muss also $\mathbb{N} \subseteq A$ gelten, und damit $A = \mathbb{N}$. Zusammengefasst gilt

$$0 \in A \wedge (\forall n \in A\colon S(n) \in A) \Rightarrow A = \mathbb{N},$$

was aber nichts anderes als das Schema der Induktion ist. Das heißt, wir haben das Schema der Induktion aus der Definition von $\mathbb{N}$ abgeleitet.

Bislang wurden die natürlichen Zahlen mithilfe der reellen Zahlen beschrieben, die ihrerseits ein Gebilde komplizierter Art sind. Es ist aber auch machbar, die induktiven Mengen in kurzen Zügen mit den elementaren Mitteln der Mengenlehre zu modellieren. Wir definieren hierbei die Null als die leere Menge und die Nachfolgerabbildung als

$$S(n) := n \cup \{n\}.$$

Infolgedessen gilt

$$0 = \emptyset, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}, \quad \text{usw.}$$

Nun definieren wir $\mathbb{N}$ wieder als den Schnitt der induktiven Mengen, mit dem Unterschied, dass wir als Grundgesamtheit nicht die Potenzmenge der reellen Zahlen, sondern die Klasse aller Mengen wählen. Man spricht von der Modellierung der natürlichen Zahlen nach John von Neumann, der $S(\alpha) = \alpha \cup \{\alpha\}$ sogar allgemeiner als den Nachfolger einer Ordinalzahl α betrachtete.

Einige Logiker sehen die aufgezeigte Herangehensweise, $\mathbb{N}$ als kleinste induktive Menge zu definieren, als problematisch an. Dies bedingt nämlich, dass die gesuchte Menge $\mathbb{N}$ bereits im Mengensystem der induktiven Mengen enthalten ist. Wir haben es hier mit einer gewissen Art von Selbstbezüglichkeit zu tun, man spricht von einer *imprädikativen* Definition.

3.5.2. Der dedekindsche Rekursionssatz

Satz 3.37 (Rekursionssatz von Dedekind).
Zu einer X Menge seien $x_0 \in X$ und $\varphi\colon X \to X$ festgelegt. Es bezeichne $S(n) = n + 1$ die Nachfolgerabbildung. Die Funktionalgleichung

$$\forall n \in \mathbb{N}\colon f(S(n)) = \varphi(f(n))$$

besitzt zum Anfangswert $f(0) = x_0$ eine eindeutige Lösung $f\colon \mathbb{N} \to X$.

Beweis. Die Existenz bereitet gewisse Schwierigkeiten. Kümmern wir uns zunächst lieber um die Eindeutigkeit. Es seien f, g zwei Lösungen. Wir müssen zu $f = g$ gelangen, das heißt, $f(n) = g(n)$ für alle n. Induktion über n. Im Anfang ist $f(0) = x_0$ und $g(0) = x_0$, ergo $f(0) = g(0)$. Im Induktionsschritt ist $f(S(n)) = g(S(n))$ aus der Induktionsvoraussetzung $f(n) = g(n)$ zu folgern. Vermöge der Funktionalgleichung findet sich

$$f(S(n)) = \varphi(f(n)) = \varphi(g(n)) = g(S(n)).$$

Das war's schon. Zur Existenz betrachten wir das Mengensystem

$$\mathcal{A} := \{A \subseteq \mathbb{N} \times X \mid (0, x_0) \in A \wedge \forall (n, x) \in A\colon (S(n), \varphi(x)) \in A\}.$$

Es ist $\mathcal{A}$ nichtleer, weil $\mathbb{N} \times X \in \mathcal{A}$ ist. Sei $G := \bigcap \mathcal{A}$. Man überzeugt sich von $G \in \mathcal{A}$, womit G die kleinste Menge des Systems sein muss. Wir bestätigen nun, dass G der Graph einer Abbildung $f\colon \mathbb{N} \to X$ sein muss. Dazu wird

$$\forall n \in \mathbb{N}\colon \exists! x \in X\colon (n, x) \in G$$

vermittels Induktion über n gezeigt. Im Anfang ist $(0, x_0) \in G$, weil das System nur aus solchen Mengen besteht. Angenommen, es gäbe noch ein weiteres $x_0' \neq x_0$ mit $(0, x_0') \in G$. Dann wäre $G \setminus \{(0, x_0')\} \in \mathcal{A}$ im Widerspruch dazu, dass G die kleinste Menge ist. Nämlich gilt

$$\forall (m, x) \in G \setminus \{(0, x_0')\}\colon (S(m), \varphi(x)) \in G \setminus \{(0, x_0')\}$$

erfüllt, denn es ist $(S(m), \varphi(x)) \neq (0, x_0')$ für jedes m wegen $0 \notin S(\mathbb{N})$.

Zum Schritt. Induktionsvoraussetzung ist die Existenz genau eines x mit $(n, x) \in G$. Wegen $G \in \mathcal{A}$ gilt somit auch $(S(n), \varphi(x)) \in G$. Angenommen, es gäbe ein weiteres $x' \neq \varphi(x)$ mit $(S(n), x') \in G$. Man setze $A := G \setminus \{(S(n), x')\}$. Es wäre nun $A \in \mathcal{A}$ im Widerspruch dazu, dass G die kleinste Menge ist. Nämlich gilt

$$(0, x_0) \in A \wedge \forall (m, u) \in A \colon (S(m), \varphi(u)) \in A.$$

Die erste Aussage muss stimmen, weil $(0, x_0)$ nicht $(S(n), x')$ sein kann, da $S(n) \neq 0$ für jedes n. Zur zweiten Aussage ist im Wesentlichen $(S(m), \varphi(u)) \neq (S(n), x')$ zu zeigen. Angenommen, $(S(m), \varphi(u)) = (S(n), x')$, das heißt, $S(m) = S(n)$ und $\varphi(u) = x'$. Weil S injektiv ist, müsste dann $m = n$ sein. Demnach müsste (n, x) mit (m, u) übereinstimmen, weil ja nach Voraussetzung x eindeutig bestimmt ist. Mit $x = u$ folgt daraufhin $\varphi(x) = \varphi(u)$. Aber aus $x' \neq \varphi(x)$ und $\varphi(u) = x'$ folgt $\varphi(x) \neq \varphi(u)$, was widersprüchlich ist.

Es darf G demzufolge als Graph einer Abbildung f gedeutet werden. Die Aussage $G \in \mathcal{A}$ bedeutet diesbezüglich nichts anderes als

$$f(0) = x_0 \wedge \forall n \in \mathbb{N} \colon f(S(n)) = \varphi(f(n)),$$

womit eine Lösung existiert. □

Der letzte Teil des Beweises ist ein wenig haarig. Da findet ein Widerspruchsbeweis innerhalb eines Widerspurchsbeweises statt, und ich habe ein paar kleine Schritte ausgelassen, um die Darstellung nicht zu sehr mit Kleinkram aufzublähen.

Der Beweis soll zurückgehen auf [36], wobei die moderne Fassung in [37], [38] zu finden ist. Dedekind selbst bewies den Satz in [35]. Allerdings setzt er dabei die zuvor definierte Ordnung der natürlichen Zahlen voraus. Beschränkt man sich auf die Peano-Axiome allein, würde der Beweis zirkulär, sollte die Ordnung über die rekursiv festgelegte Addition definiert worden sein. Diese Problematik wird in [39] näher diskutiert.

3.5.3. Fixpunkteigenschaft kleinster induktiver Mengen

Zu einer Menge $A \subseteq G$ und Nachfolgerabbildung $S \colon G \to G$ sei

$$F(A) := \{0\} \cup S(A) = \{0\} \cup \{m \in G \mid \exists n \in A \colon m = S(n)\}.$$

Diesbezüglich finden sich einige Mengen A, die die Gleichung $F(A) = A$ erfüllen, also Fixpunkte der Abbildung F sind. Zum Beispiel ist dies $A := \mathbb{N}$, aber auch $A := \mathbb{Z}$,

wobei die reellen Zahlen $G := \mathbb{R}$ als Grundmenge dienen sollen und $S(x) := x + 1$ zu jeder reellen Zahl x gesetzt wird.

Eine Menge A mit der Eigenschaft $F(A) \subseteq A$ bezeichnet man als *Präfixpunkt*. Tatsächlich ist eine Menge genau dann ein Präfixpunkt, wenn sie induktiv ist. Nämlich findet sich die äquivalente Umformung

$$\begin{aligned}
& m \in F(A) \Rightarrow m \in A \\
\iff & (m = 0 \vee \exists n \in A\colon m = S(n)) \Rightarrow m \in A \\
\iff & \neg(m = 0 \vee \exists n \in A\colon m = S(n)) \vee m \in A \\
\iff & m \neq 0 \wedge (\forall n \in A\colon m \neq S(n)) \vee m \in A \\
\iff & (m \neq 0 \vee m \in A) \wedge (\forall n \in A\colon m \neq S(n) \vee m \in A) \\
\iff & (m = 0 \Rightarrow m \in A) \wedge (\forall n \in A\colon m = S(n) \Rightarrow m \in A) \\
\iff & 0 \in A \wedge \forall n \in A\colon S(n) \in A.
\end{aligned}$$

Demzufolge stimmt das System der Präfixpunkte mit dem System der induktiven Mengen überein. Mithin ist $\mathbb{N}$ der kleinste Präfixpunkt

$$\mathbb{N} = \bigcap \{A \subseteq G \mid F(A) \subseteq A\}.$$

Ebenso ist $\mathbb{N}$ der kleinste Fixpunkt, wie folgendermaßen ersichtlich wird. Hat man zwei beliebige Mengensysteme $\mathcal{A}, \mathcal{B}$ mit $\mathcal{A} \subseteq \mathcal{B}$, so gilt $\bigcap \mathcal{B} \subseteq \bigcap \mathcal{A}$. Wir erinnern uns außerdem daran, dass die Schnittmenge eine untere Schranke ist, also $\bigcap \mathcal{A} \subseteq A$ für jedes $A \in \mathcal{A}$ gilt. Speziell gilt dies für

$$\mathcal{A} := \{A \mid F(A) = A\}, \quad \mathcal{B} := \{A \mid F(A) \subseteq A\}.$$

Mit $\mathbb{N} \in \mathcal{A}$ erhält man somit

$$\mathbb{N} = \bigcap \mathcal{B} \subseteq \bigcap \mathcal{A} \subseteq \mathbb{N}.$$

Ergo muss $\mathbb{N} = \bigcap \mathcal{A}$ sein, der kleinste Fixpunkt von F.

Liegt nun in allgemeiner Weise eine induktive Struktur mit Basismenge B und einer Menge Φ von Nachfolgerabbildungen vor, ist der Fortsetzungsschritt definiert gemäß

$$F(A) := B \cup \bigcup \{f(A^n) \mid (n, f) \in \Phi\}.$$

Auch in dieser allgemeinen Situation lässt sich analog nachrechnen, dass das System der Präfixpunkte mit dem System der induktiven Mengen übereinstimmt. Und es ergibt sich wieder, dass der kleinste Fixpunkt A mit dem kleinsten Präfixpunkt

$X := \bigcap \mathcal{B}$ übereinstimmt. Es verbleibt dafür lediglich $X \in \mathcal{A}$, also $F(X) = X$ zu zeigen, da dann wieder das Argument $\bigcap \mathcal{A} \subseteq X$ greift. Da X Präfixpunkt ist, verbleibt nur noch $X \subseteq F(X)$ zu zeigen. Man rechnet zunächst nach, dass F eine bezüglich der Inklusion monoton steigende Abbildung ist, das heißt, für beliebige Mengen A_1, A_2 gilt

$$A_1 \subseteq A_2 \Rightarrow F(A_1) \subseteq F(A_2).$$

Mithin folgt $F(F(X)) \subseteq F(X)$ aus $F(X) \subseteq X$. Das heißt aber, $F(X) \in \mathcal{B}$, womit $\bigcap \mathcal{B} \subseteq F(X)$, gelten muss, weil $\bigcap \mathcal{B}$ eine untere Schranke von $\mathcal{B}$ ist. Mit Blick auf die Definition von X wird $X \subseteq F(X)$ nun schließlich ersichtlich.

Führt man Überlegungen dieser Art ganz abstrakt für Halbordnungen bzw. Verbände durch, gelangt man zum Fixpunktsatz von Knaster und Tarski. Dieser besagt, in einem vollständigen Verband bildet die Menge der Fixpunkte einer monoton steigenden Abbildung ebenfalls einen vollständigen Verband. Mithin existiert ein kleinster Fixpunkt.

Man kann sich den natürlichen Zahlen auch von unten nähern. Es gilt

$$F(\emptyset) = \{0\}, \quad F(F(\emptyset)) = \{0, 1\}, \quad F(F(F(\emptyset))) = \{0, 1, 2\}, \quad \text{usw.}$$

Man überlegt sich nun den Grenzwert

$$\mathbb{N} = \bigcup \{F^n(\emptyset) \mid n \in \mathbb{N}\}, \quad F^0(A) := A, \quad F^{n+1}(A) := F(F^n(A)).$$

Allerdings ist diese Beziehung wohl nicht zur Definition von $\mathbb{N}$ nutzbar. Zum einen wäre sie zirkulär, da $\mathbb{N}$ innerhalb der Definition vorkäme. Zum anderen ist die Iterierte F^n rekursiv definiert, wofür wir eigentlich einen Rekursionssatz benötigen, der $\mathbb{N}$ ebenfalls als definiert voraussetzt. Wie dem auch sei, die Entfaltung der Gleichung führt jedenfalls zu

$$m \in \mathbb{N} \Leftrightarrow \exists A\colon m \in A \wedge \exists n \in \mathbb{N}\colon A = F^n(\emptyset).$$

Wegen $m \in F^{m+1}(\emptyset)$ kann man $A := F^{m+1}(\emptyset)$ und $n := m + 1$ setzen, um die rechte Seite aus der linken zu folgern. Und wegen $F^n(\emptyset) \subseteq \mathbb{N}$ zu jedem $n \in \mathbb{N}$ folgert man die linke Seite aus der rechten.

3.5.4. Hüllenoperatoren

Die kleinste induktive Menge wurde als Schnitt des Systems der induktiven Mengen erhalten. Man darf sie als eine Hülle der fundierenden Menge betrachten, denn

sie schließt die fundierende Menge insofern ab, als vermittels der Produktionsregeln keine weiteren Elemente mehr produziert werden können, die außerhalb der Hülle lägen. Der deduktive Abschluss einer Formelmenge stellt ebenfalls eine Hülle dar, wie bereits diskutiert wurde. Weitere Beispiele für Hüllen sind die lineare Hülle einer Teilmenge eines Vektorraums, die abgeschlossene Hülle einer Teilmenge eines topologischen Raums, die konvexe Hülle einer Teilmenge eines euklidischen Vektorraums und die kleensche Hülle einer Menge von Zeichenketten. In der Maßtheorie ist die kleinste σ-Algebra, die eine Teilmenge der Grundmenge umfasst, eine Hülle der Teilmenge.

Zur Schaffung von Ordnung wollen wir einige allgemeine Prinzipien zu Hüllen ausarbeiten.

Definition 3.34 (Hüllenoperator).
Zu einer Grundmenge U wird eine Abbildung $C\colon \mathcal{P}(U) \to \mathcal{P}(U)$ als Hüllenoperator, engl. *closure operator* bezeichnet, sofern gilt

$$\begin{aligned}
&\forall X \subseteq U\colon X \subseteq C(X), && \text{(Extensivität)}\\
&\forall X, Y \subseteq U\colon X \subseteq Y \Rightarrow C(X) \subseteq C(Y), && \text{(Monotonie)}\\
&\forall X \subseteq U\colon C(C(X)) \subseteq C(X). && \text{(Abgeschlossenheit)}
\end{aligned}$$

In der Konsequenz geht $C(X) \subseteq C(C(X))$ zu jedem $X \subseteq U$ aus der Anfügung der Monotonie an die Extensivität hervor. Somit erweitert sich die Abgeschlossenheit postwendend zur Idempotenz

$$\forall X \subseteq U\colon C(C(X)) = C(X).$$

Ein Mengensystem $\mathcal{A} \subseteq \mathcal{P}(U)$ wird *Hüllensystem* genannt, sofern

$$U \in \mathcal{A} \wedge (\forall \mathcal{T} \subseteq \mathcal{A}\colon \mathcal{T} \neq \emptyset \Rightarrow \bigcap \mathcal{T} \in \mathcal{A}).$$

Satz 3.38. Ist $\mathcal{A}$ ein Hüllensystem, so ist

$$C(X) := \bigcap\{A \in \mathcal{A} \mid X \subseteq A\}$$

ein Hüllenoperator.

Beweis. Die Extensivität bestätigt sich unschwer, indem $x \in C(X)$ aus $x \in X$ mit x fest, aber beliebig abgeleitet wird. Die Monotonie zeigt sich analog. Zu bemerken ist hierbei, dass die Extensivität und Monotonie für jedes beliebige Mengensystem gilt, da die definierende Eigenschaft des Hüllensystems bisher nicht benötigt wurde. Kommen wir zur interessanteren Abgeschlossenheit. Man überzeugt sich

zunächst von $C(X) \in \mathcal{A}$ zu jedem $X \subseteq U$. Dazu wird die definierende Eigenschaft des Hüllensystems mit $\mathcal{T} := \{A \in \mathcal{A} \mid X \subseteq A\}$ spezialisiert, die als Aussonderung eine Teilmenge von $\mathcal{A}$ ist. Es ist $\mathcal{T}$ außerdem nichtleer, da $U \in \mathcal{T}$, wie aus $U \in \mathcal{A}$ und der Prämisse $X \subseteq U$ hervorgeht.

Mit dieser Vorbereitung gelingt der folgende Gedankengang. Die Entfaltung des ersten C in der zu beweisenden Abgeschlossenheit bringt diese in die Form

$$\bigcap\{A \in \mathcal{A} \mid C(X) \subseteq A\} \subseteq C(X).$$

Nun ist die Schnittmenge aber eine untere Schranke bezüglich der Inklusion. Das heißt, es verbleibt bloß noch

$$C(X) \in \{A \in \mathcal{A} \mid C(X) \subseteq A\}$$

zu bestätigen. Es wurde $C(X) \in \mathcal{A}$ zuvor gezeigt, und $C(X) \subseteq C(X)$ gilt gemäß Reflexivität der Inklusion. □

Satz 3.39. Das System der Induktiven Mengen ist ein Hüllensystem.

Beweis. Dabei dürfen die per se vorhandene fundierende Menge B und das Argument des Hüllenoperators als unterschiedlich angesehen werden, das macht die Formulierung ein wenig allgemeiner, als wenn wir diese beiden koinzidieren lassen. Als erstes wird also $B \subseteq \bigcap \mathcal{T}$ gezeigt. Aus $\mathcal{T} \subseteq \mathcal{A}$ folgt $\bigcap \mathcal{A} \subseteq \bigcap \mathcal{T}$, womit bereits $B \subseteq \bigcap \mathcal{A}$ genügt. Es verbleibt demnach die Bestätigung der Aussage

$$x \in B \vdash \forall A \in \mathcal{A} \colon x \in A,$$

die aber ersichtlich ist, denn $B \subseteq A$ ist schlicht in der Aussage $A \in \mathcal{A}$ enthalten.

Außerdem muss man sich noch von $f(x) \in \bigcap \mathcal{T}$ überzeugen, wobei $x \in (\bigcap \mathcal{T})^n$ bezüglich $(n, f) \in \Phi$ mit $f \colon U^n \to U$ vorausgesetzt wird. Demnach muss $f(x) \in A$ zu einem beliebigen $A \in \mathcal{T}$ bestätigt werden. Aus $\mathcal{T} \subseteq \mathcal{A}$ folgt $A \in \mathcal{A}$, womit

$$\forall (n, f) \in \Phi \colon \forall x \in A^n \colon f(x) \in A.$$

Diese Allaussage wird mit dem vorliegenden (n, f) und x spezialisiert. Es verbleibt demnach $x \in A^n$ zu bestätigen, was wir aus $(\bigcap \mathcal{T})^n \subseteq A^n$ folgern, die wiederum aus $\bigcap \mathcal{T} \subseteq A$ folgt. Man macht sich an dieser Stelle wieder zunutze, dass der Schnitt eine untere Schranke bezüglich der Inklusion ist. Dafür muss $A \in \mathcal{T}$ sein, was aber vorausgesetzt werden durfte. □

3.5.5. Induktive Mengen allgemein

Allgemein sei G eine Grundmenge, die die gesuchte Menge enthält. Es sei $B \subseteq G$ die Menge von Grundelementen. Es bezeichne Φ die Menge der Funktionen, die die Produktionsregeln kodieren, wobei $(n, f) \in \Phi$ eine Funktion $f\colon G^n \to G$ bedeuten soll. Die Menge der induktiven Mengen ist dann

$$\mathcal{A} := \{A \subseteq G \mid B \subseteq A \wedge \forall (n, f) \in \Phi\colon \forall x \in A^n\colon f(x) \in A\}.$$

Die Setzung $G := \mathbb{R}$, $B := \{0\}$ und $F := \{S\}$ mit $S(x) := x + 1$ liefert wie gesagt die natürlichen Zahlen via $\bigcap \mathcal{A} = \mathbb{N}$.

Es zeigt sich wieder, dass $\bigcap \mathcal{A}$ die kleinste induktive Menge ist. Der Beweis von $\bigcap \mathcal{A} \in \mathcal{A}$ läuft hierbei völlig analog zu dem von Satz 3.36 ab. Aufgrund dessen will ich darauf verzichten, diesen technischen Klimbim nochmals auszuführen.

3.5.6. Frei erzeugte induktive Mengen

Wie bei den natürlichen Zahlen besteht ein wesentlicher Schritt darin, Funktionen rekursiv auf der kleinsten induktiven Menge zu definieren. Hierbei kann es allerdings zu Komplikationen kommen. Zum einen wäre die Rekursion nicht wohldefiniert, wenn die Funktion der jeweiligen Produktionsregel keine injektive ist. Zum anderen könnten sich zwei der Rekurrenzen des Systems widersprechen, so dass die Lösungsmenge des Systems leer bleiben muss.

Ich will diese Umstände näher betrachten. Sei hierzu f die gesuchte Funktion und f_i die Funktion der jeweiligen Produktionsregel. Das System bestehe aus Rekurrenzen der Form

$$f(f_i(x)) = \varphi_i(x, f(x)).$$

Eine Implementierung im Computer führt die Berechnung demnach aus als

$$f(y) := \varphi_i(x, f(x)), \quad x := f_i^{-1}(y),$$

wofür f_i aber injektiv sein muss. Angenommen, es gälte zu zwei x_1, x_2 die Gleichheit $f_i(x_1) = f_j(x_2)$ zu $i \neq j$. Dies zöge nach sich

$$\varphi_i(x_1, f(x_1)) = f(f_i(x_1)) = f(f_j(x_2)) = \varphi_j(x_2, f(x_2)).$$

Wären die linke und rechte Seite verschieden, hätte das System keine Lösung. Um diese Komplikationen ausschließen zu können, wird gefordert, dass die f_i paarweise disjunkte Bildmengen haben.

Zuletzt muss man folgendes beachten. Es seien Rekursionsanfänge zu den $x_k \in B$ definiert als $f(x_k) := y_k$. Angenommen, es findet sich aber auch ein x mit $x_k = f_i(x)$, dann müsste gelten

$$y_k = f(x_k) = f(f_i(x)) = \varphi_i(x, f(x)).$$

Wäre die linke und rechte Seite verschieden, hätte das System keine Lösung. Um diese Komplikationen ausschließen zu können, wird gefordert, dass zu jedem f_i die Bildmenge disjunkt zu B ist.

Definition 3.35 (Frei erzeugte induktive Menge).
Es sei G die Grundmenge und $A \subseteq G$ die kleinste induktive Menge zu $B \subseteq A$. Man nennt A *frei erzeugt* aus B, wenn für alle $(n, f), (m, g) \in \Phi$ gilt,

1. die Einschränkung von f auf A^n ist injektiv,
2. die Bildmengen $f(A^n)$ und $g(A^m)$ sind disjunkt, sofern $f \neq g$,
3. die Bildmenge $f(A^n)$ ist disjunkt zu B.

3.5.7. Wohlfundierte Induktion

Aus der gewöhnlichen Induktion hatten wir bereits die starke Induktion hergeleitet. Wie sich herausstellt, kann diese abstrakter gefasst werden, was zum Begriff der *wohlfundierten Induktion* führt. Mit ihr erhält man ein sehr allgemeines Induktionsprinzip, das neben der gewöhnlichen und der strukturellen die transfinite Induktion umfasst.

Eine als Ordnungsrelation betrachtete Relation wollen wir *wohlfundiert* nennen, wenn jede ihrer nichtleeren Teilmengen ein minimales Element enthält. Den Begriff des minimalen Elements entnehmen wir aus Def. 3.31 auf S. 110, obgleich die Relation nicht unbedingt eine Halbordnung sein muss.

Definition 3.36 (Wohlfundierte Relation).
Eine auf M definierte Relation $<$ heißt *wohlfundiert*, wenn

$$\forall A \subseteq M\colon A \neq \emptyset \Rightarrow \exists y \in A\colon \neg\exists x \in A\colon x < y.$$

Die Symbole $<, \prec$ werden gegenüber $\leq, \preceq$ üblicherweise für Relationen mit dem Attribut *streng* verwendet. Das sind solche, die die Diagonale ausschließen, man nennt sie *irreflexiv*. Dass die Symbolik hier gerechtfertigt ist, zeigt der

Satz 3.40. Jede wohlfundierte Relation ist irreflexiv.

Beweis. Sei $(M, <)$ wohlfundiert. Zu zeigen ist $\forall u \in M\colon \neg(u < u)$. Als Hilfsmittel dafür dient die allgemeine Äquivalenz $E(u) \Leftrightarrow \exists x\colon x = u \wedge E(x)$, bei der $E(x)$ eine beliebige Aussageform sein darf. Sei u fest, aber beliebig. Die Bedingung aus Def. 3.36 wird nun mit der Einermenge $A := \{u\}$ spezialisiert. Daraufhin gelingt die Umformung

$$\begin{aligned}&\exists y \in \{u\}\colon \neg\exists x \in \{u\}\colon x < y \iff \exists y\colon y = u \wedge \neg\exists x\colon x = u \wedge x < y\\ &\iff \neg\exists x\colon x = u \wedge x < u \iff \neg(u < u).\ \square\end{aligned}$$

Die Idee bei einer wohlfundierten Relation soll sein, dass man beim Herabsteigen nie in einen infiniten Regress gerät. Unabhängig davon, welcher Weg und welche Knotenpunkte gewählt werden, wird man immer zu einem minimalen Element gelangen. Diese minimalen Elemente werden als die Basis, als das Fundament für die Induktion und die Rekursion dienen.

Satz 3.41. Ist $(M, <)$ wohlfundiert, existiert keine unendliche streng monoton absteigende Kette, das heißt, es gilt die *Descending Chain Condition*

$$\neg\exists f \in \mathrm{Abb}(\mathbb{N}, M)\colon \forall n \in \mathbb{N}\colon f(n+1) < f(n).$$

Beweis. Angenommen, es läge eine unendlich absteigende Funktion f vor. Da das Bild $f(\mathbb{N})$ eine nichtleere Teilmenge von M und M wohlfundiert ist, muss ein minimales Element $y \in f(\mathbb{N})$ existieren, wozu ein n mit $y = f(n)$ gehört. Nun wäre aber $f(n+1) < y$, was im Widerspruch zur Minimalität von y steht. $\square$

Satz 3.42. Die Descending Chain Condition ist genau dann erfüllt, wenn jede monoton absteigende Kette ab irgendeiner Stelle stabilisiert, das heißt,

$$\forall f \in \mathrm{Abb}(\mathbb{N}, M)\colon (\forall n\colon f(n+1) \leq f(n)) \Rightarrow \exists m\colon \forall n \geq m\colon f(m) = f(n).$$

Beweis. Es sei $f\colon \mathbb{N} \to M$ eine beliebige Folge mit $f(n+1) \leq f(n)$ für jedes n. Angenommen, f enthielte unendlich viele Stellen n_i mit $f(n_i+1) \neq f(n_i)$, also $f(n_i+1) < f(n_i)$. Dann ließe sich die Teilfolge $g\colon \mathbb{N} \to M$ mit $g(i) := n_i$ bilden, die der Descending Chain Condition widerspricht. Ergo kann es von den Stellen n_i nur endlich viele geben, womit die Folge f ab einer Stelle stabilisieren muss.

Zur Umkehrung. Angenommen, es läge f mit $f(n+1) < f(n)$ für jedes n vor. Für dieses f gilt erst recht $f(n+1) \leq f(n)$ für jedes n. Laut Prämisse muss f somit ab einem m stabilisieren, womit $f(m+1) = f(m)$, was im Widerspruch zur strengen Monotonie von f steht. $\square$

Satz 3.43 (Wohlfundierte Induktion).
Ist $(M, <)$ wohlfundiert, gilt für jede Teilmenge $A \subseteq M$ das Prinzip

$$(\forall x \in M\colon (\forall y \in M\colon y < x \Rightarrow y \in A) \Rightarrow x \in A) \Rightarrow A = M.$$

Beweis. Dieser wird ganz analog zum dem von Satz 1.6 auf S. 37 geführt. Es ist genau dann $A = M$, wenn das Komplement $B := A^{\mathrm{c}} = M \setminus A$ leer ist. Man unternimmt die äquivalente Umformung

$$\begin{aligned}
&B \neq \emptyset \Rightarrow (\exists x \in B\colon \neg\exists y \in B\colon y < x)\\
\iff &(\forall x \in B\colon \exists y \in B\colon y < x) \Rightarrow B = \emptyset\\
\iff &(\forall x \in M\colon x \notin A \Rightarrow \exists y \in B\colon y < x) \Rightarrow A = M\\
\iff &(\forall x \in M\colon (\forall y \in B\colon \neg y < x) \Rightarrow x \in A) \Rightarrow A = M.
\end{aligned}$$

Und schließlich

$$\begin{aligned}
(\forall y \in B\colon \neg y < x) &\iff (\forall y \in M\colon y \notin A \Rightarrow \neg y < x)\\
&\iff (\forall y \in M\colon y < x \Rightarrow y \in A).\ \square
\end{aligned}$$

Da ausschließlich äquivalente Umformungen vorgenommen wurden, stellt sich das Induktionsprinzip sogar als äquivalent zur Wohlfundiertheit heraus.

Zu einer abstrakten Fassung des Induktionsprinzips findet man folgendermaßen. Das Urbild

$$R^{-1}(y) := \{x \mid R(x, y)\} = \{x \mid x < y\}$$

nennt man auch das *Anfangsstück* bis y bezüglich R. Wir setzen

$$F(A) := \{x \mid R^{-1}(x) \subseteq A\}.$$

Nun mag das Induktionsprinzip formuliert werden als

$$\forall A \subseteq M\colon F(A) \subseteq A \Rightarrow A = M.$$

Eine Menge wollen wir diesbezüglich wieder *induktiv* nennen, wenn sie die Prämisse $F(A) \subseteq A$ erfüllt. Das Induktionsprinzip besagt damit, dass jede induktive Menge mit M übereinstimmen muss, oder anders ausgedrückt, dass auf $\mathcal{P}(M)$ keine anderen Präfixpunkte von F als M selbst vorhanden sind.

Die gewöhnliche Induktion. Es ist $(\mathbb{N}, \prec)$ mit $n \prec m :\Leftrightarrow m = S(n)$ wohlfundiert, wobei mit $S(n)$ wieder der Nachfolger von n gemeint ist. Als Induktionsprinzip erhält man hierbei die gewöhnliche Induktion. Dies wird ersichtlich, wenn $\mathbb{N}$ in $\mathbb{N} = \{0\} \cup \mathbb{N}_{\geq 1}$ zerlegt und daraufhin die allgemeine Äquivalenz

$$(\forall x \in A \cup B\colon E(x)) \Leftrightarrow (\forall x \in A\colon E(x)) \wedge (\forall x \in B\colon E(x))$$

herangezogen wird. Außerdem ergibt sich in diesem Fall $R^{-1}(m) = S^{-1}(\{m\})$ als Anfangsstück. Die Prämisse des Induktionsprinzips nimmt demnach die Form

$$F(A) \subseteq A \Leftrightarrow (\forall m \in \{0\} \colon S^{-1}(\{m\}) \subseteq A \Rightarrow m \in A) \wedge (\forall m \in \mathbb{N}_{\geq 1} \colon S^{-1}(\{m\}) \subseteq A \Rightarrow m \in A)$$

an. Die linke Seite der Konjuktion vereinfacht sich zu $0 \in A$, da $S^{-1}(\{0\}) = \emptyset$ gilt, insofern 0 keinen Vorgänger besitzt. Auf der rechte Seite nutzt man den Umstand aus, dass $m \in \mathbb{N}_{\geq 1}$ mit $\exists n \in \mathbb{N} \colon m = S(n)$ gleichbedeutend ist. Weil S injektiv ist, ergibt sich nun $S^{-1}(S(\{n\})) = \{n\}$, womit $S^{-1}(\{m\}) \subseteq A$ äquivalent zu $n \in A$ ist. Ergo findet sich $n \in A \Rightarrow S(n) \in A$, wobei dies für ein ganz allgemeines n gelten muss, weil m allgemein gehalten war. Wem diese Umformungen zu unpräzise dargelegt sind, der mag sie gerne zur Übung nochmals formal nachrechnen.

Die starke Induktion. Es ist $(\mathbb{N}, <)$ wohlfundiert, wobei mit $<$ die gewöhnliche Ordnung auf $\mathbb{N}$ gemeint ist. Als Induktionsprinzip erhält man unmittelbar die starke Induktion.

4. Elemente der Algebra

4.1. Gruppentheorie

4.1.1. Elementare Gesetzmäßigkeiten

Die Gruppentheorie klärt uns tiefer über das Wesen von Symmetrien auf. Ich will Gruppen zunächst axiomatisch einführen, damit wir die elementaren Begriffe später bereits parat haben. Was unter einer Symmetrie zu verstehen sei, und in welchem Bezug sie zu Gruppen stehen, möchte ich daraufhin im Fortgang erörtern. Die Begriffe sollen eigentlich mit Blick auf die Idee der Symmetrie motiviert werden. Damit die längere Diskussion die Ausarbeitung der abstrakten Regeln nicht so sehr fragmentiert, möchte ich diese Ausarbeitung allerdings vorziehen.

Definition 4.1 (Gruppe).
Sei G eine Menge und $*\colon G \times G \to \Omega$ eine Verknüpfung. Die Menge G bildet bezüglich der Verknüpfung eine Gruppe $(G, *)$, wenn die folgenden Axiome erfüllt sind:

(E) Es darf $\Omega = G$ sein, d. h., die Verknüpfung führt nicht aus G heraus.

(A) Das Assoziativgesetz $a * (b * c) = (a * b) * c$ gilt für alle $a, b, c \in G$.

(N) Es gibt ein neutrales Element e, so dass $g * e = e * g = g$ für jedes $g \in G$ gilt.

(I) Zu jedem $g \in G$ gibt es ein Element $h \in G$ mit $g * h = h * g = e$, wobei e ein neutrales Element ist. Dieses h wird inverses Element zu g genannt.

Anstelle von $g * h$ schreibt man auch kurz gh. Für das inverse Element zu g schreibt man g^{-1}. Es gibt auch Gruppen, bei denen die Verknüpfung als Addition geschrieben wird, da schreibt man $g + h$ anstelle von gh und ng anstelle von g^n für $n \in \mathbb{Z}$. Für das inverse Element $-g$ anstelle von g^{-1}.

Definition 4.2 (Abelsche Gruppe).
Zwei Elemente a, b kommutieren, wenn $a * b = b * a$ ist. Eine Gruppe G heißt abelsch oder kommutativ, wenn alle Elemente der Gruppe kommutieren, d. h.

wenn das Kommutativgesetz $\forall a, b \in G\colon a * b = b * a$ erfüllt ist.

Bei den allermeisten Verknüpfungen, die als Addition geschrieben werden, ist das Kommutativgesetz erfüllt.

Satz 4.1. Das neutrale Element einer Gruppe ist eindeutig bestimmt, d. h. es kann keine zwei unterschiedlichen neutralen Elemente geben.

Beweis. Seien e und e' zwei neutrale Elemente. Zu zeigen ist, dass dann schon $e' = e$ gilt. Nach Voraussetzung gilt $ae = a$ und $e'b = b$ für alle a, b. Setzt man $a := e'$ und $b := e$ ein, dann ergibt sich $e' = e'e = e$. □

Satz 4.2. In jeder Gruppe gilt die Linkskürzbarkeit und die Rechtskürzbarkeit. Damit ist gemeint, sowohl aus $ga = gb$ als auch aus $ag = bg$ folgt $a = b$.

Beweis. Man multipliziert die Gleichung $ga = gb$ beidseitig mit g^{-1}. Anschließende Anwendung des Assoziativgesetzes, gefolgt von $g^{-1}g = e$ zuzüglich $ea = e$ und $eb = e$ stellt die Folgerung

$$ga = gb \implies g^{-1}ga = g^{-1}gb \implies ea = eb \implies a = b$$

her. Bei $ag = bg$ verläuft der Beweis analog. □

Kürzbarkeit bedeutet, eine Multiplikation auf beiden Seiten rückgängig machen zu können. Das Rückgänig-machen-können ist wiederum die charakteristische Eigenschaft einer injektiven Abbildung. Unter diesem Aspekt gesehen bedeutet die Kürzbarkeit, dass die Links- und Rechts-Translation

$$\begin{aligned} l_g &\colon G \to G, \quad l_g(x) := gx, \\ r_g &\colon G \to G, \quad r_g(x) := xg \end{aligned}$$

injektiv sind. Wie man leicht nachrechnet, sind sie sogar bijektiv. Für die Umkehrabbildungen gilt $(l_g)^{-1} = l_{g^{-1}}$ und $(r_g)^{-1} = r_{g^{-1}}$.

Satz 4.3. Zu jedem Element ist das inverse Element eindeutig bestimmt, das heißt, es kann keine zwei unterschiedlichen inversen Elemente geben.

Beweis. Seien h und h' invers zu g. Dann gilt $gh = e$ und $gh' = e$. Daher ist $gh = gh'$. Gemäß Linkskürzbarkeit folgt daraus $h = h'$. □

Definition 4.3 (Untergruppe).
Sei $(G, *)$ eine Gruppe und $U \subseteq G$. Man nennt U Untergruppe von G, kurz $U \leq G$, wenn $(U, *)$ die Gruppenaxiome bezüglich derselben Verknüpfung $*$ erfüllt.

Satz 4.4 (Untergruppenkriterium).
Sei G eine Gruppe. Eine nichtleere Teilmenge $H \subseteq G$ ist eine Untergruppe von G, wenn mit $a, b \in H$ auch $ab \in H$ und mit $a \in H$ auch $a^{-1} \in H$ ist.

Beweis. Da H nichtleer ist, gibt es mindestens ein Element $a \in H$. Nach Voraussetzung ist dann auch $a^{-1} \in H$, und daher auch das neutrale Element $e = aa^{-1} \in H$.

Das Assoziativgesetz gilt in H, weil es in G gilt. Die Abgeschlossenheit und die Existenz der inversen Elemente stehen direkt in der Voraussetzung. Damit sind alle Axiome überprüft. □

Definition 4.4 (Homomorphismus zwischen Gruppen).
Seien $(G, *)$ und $(G', *')$ zwei Gruppen. Eine Abbildung $\varphi\colon G \to G'$ wird Homomorphismus genannt, wenn die Gleichung $\varphi(a * b) = \varphi(a) *' \varphi(b)$ für alle $a, b \in G$ erfüllt ist.

Satz 4.5. Sei $\varphi\colon G \to G'$ ein Homomorphismus. Sind $e \in G$ und $e' \in G'$ die neutralen Elemente, dann gilt $e' = \varphi(e)$. Außerdem ist $\varphi(g)^{-1} = \varphi(g^{-1})$ für jedes $g \in G$.

Beweis. Es gilt $e'\varphi(e) = \varphi(e) = \varphi(ee) = \varphi(e)\varphi(e)$. Kürzen ergibt $e' = \varphi(e)$. Daraus folgt

$$e' = \varphi(e) = \varphi(g^{-1}g) = \varphi(g^{-1})\,\varphi(g).$$

Damit bekommt man

$$\varphi(g)^{-1} = e'\varphi(g)^{-1} = \varphi(g^{-1})\,\varphi(g)\,\varphi(g)^{-1} = \varphi(g^{-1})$$ □

Satz 4.6. Sei $\varphi\colon G \to G'$ ein Homomorphismus. Die Bildmenge $\varphi(G)$ ist eine Untergruppe von G'.

Beweis. Zu prüfen sind die Voraussetzungen des Untergruppenkriteriums. Wegen $\varphi(a)\varphi(b) = \varphi(ab) \in \varphi(G)$ und $\varphi(a)^{-1} = \varphi(a^{-1}) \in \varphi(G)$ sind diese erfüllt. □

Für injektive, surjektive, bijektive Homomorphismen gibt es eigene Bezeichnungen. Die injektiven nennt man Monomorphismen, die surjektiven Epimorphismen und die bijektiven Isomorphismen.

Gibt es zwischen zwei Gruppen G, G' einen Isomorphismus, dann nennt man die beiden Gruppen isomorph zueinander, man schreibt dafür $G \simeq G'$. Zwei Gruppen die isomorph zueinander sind, sind im Wesentlichen gleich. Isomorphie ist eine Äquivalenzrelation.

Monomorphismen charakterisieren die Einbettung einer Gruppe in eine andere Gruppe. Man kann Einbettungen als Verallgemeinerung der Untergruppenbeziehung sehen. Hat man nämlich einen Monomorphismus $\varphi\colon H \to G$, dann erhält man bei Einschränkung der Zielmenge auf die Bildmenge einen Isomorphismus, d. h. es gilt $H \simeq \varphi(H)$. Die Gruppen H und $\varphi(H)$ sind also im Wesentlichen gleich. Andererseits ist $\varphi(H) \leq G$ gemäß Satz 4.6.

4.1.2. Gruppenaktionen

Definition 4.5 (Linksaktion).
Eine Abbildung $\varphi\colon G \times X \to X$ heißt Gruppenlinksaktion, kurz Linksaktion, wenn für das neutrale Element $e \in G$ und alle $g, h \in G$ gilt

$$\varphi(e, x) = x, \qquad \varphi(gh, x) = \varphi(g, \varphi(h, x)).$$

Anstelle von $\varphi(g, x)$ schreibt man für gewöhnlich einfach gx, bzw. $g + x$ bei einer additiv geschriebenen Verknüpfung.

Definition 4.6 (Rechtsaktion).
Eine Abbildung $\varphi\colon X \times G \to X$ heißt Gruppenrechtsaktion, kurz Rechtsaktion, wenn für das neutrale Element $e \in G$ und alle $g, h \in G$ gilt

$$\varphi(x, e) = x, \qquad \varphi(x, gh) = \varphi(\varphi(x, g), h).$$

Bei diesen Axiomen ist für X eine beliebige Menge zugelassen. Es kann auch $X = G$ sein. Beispiele dafür haben wir bereits kennengelernt, nämlich ist die Linkstranslation (4.1.1) eine Linksaktion und die Rechtstranslation (4.1.1) eine Rechtsaktion.

Korollar 4.7. Jede Aktion $\varphi\colon G \times X \to X$ ist ein Homomorphismus $\varphi\colon G \to S(X)$ mit $\varphi(g)(x) := \varphi(g, x)$. Hierbei ist $S(X)$ die Menge der Bijektionen $X \to X$, diese bildet bezüglich Verkettung eine Gruppe.

Beweis. Für jedes x gilt

$$\varphi(gh)(x) = \varphi(gh, x) = \varphi(g, \varphi(h, x)) = \varphi(g)(\varphi(h)(x)) = (\varphi(g) \circ \varphi(h))(x).$$

Folglich ist $\varphi(gh) = \varphi(g) \circ \varphi(h)$. Außerdem ist $\varphi(g)$ bijektiv mit $\varphi(g)^{-1} = \varphi(g^{-1})$, denn

$$\varphi(g^{-1}) \circ \varphi(g) = \varphi(g^{-1}g) = \varphi(e) = \mathrm{id},$$
$$\varphi(g) \circ \varphi(g^{-1}) = \varphi(gg^{-1}) = \varphi(e) = \mathrm{id}\,. \qquad \square$$

4.1.3. Symmetrie

Nach längerer Beschäftigung mit der Gruppentheorie wird man sich irgendwann fragen, was Gruppen eigentlich sind. Wie sich herausstellt sind Gruppen eng mit dem Begriff Symmetrie verbunden. Um das erklären zu können, müssen wir erst einmal herausarbeiten, was man unter Symmetrie versteht.

In der Geometrie ist eine Symmetrie eines Objektes eine Deckabbildung, das ist eine Abbildung durch die dem Objekt keine Veränderung widerfährt, in dem Sinn dass sich das alte und das neue Objekt genau überdecken. Zwar darf dabei jedem Punkt des Objektes ein Punkt an anderem Ort zugeordnet werden, jedoch verändert sich das Objekt insgesamt nicht.

Sei also $M \subseteq \mathbb{R}^2$ ein geometrisches Objekt, dargestellt als Teilmenge der Koordinatenebene. Eine Symmetrie ist dann eine Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(M) = M$. Liegen zwei solche Abbildungen f, g vor, dann ist

$$(g \circ f)(M) = g(f(M)) = g(M) = M,$$

also ist $g \circ f$ auch eine Symmetrie. Drehungen und Spiegelungen lassen sich auch punktweise rückgängig machen, sind also bijektiv. Dies wollen wir für alle Symmetrien fordern. Klar ist außerdem, dass die identische Abbildung id eine Deckabbildung ist, und die Verkettung von Abbildungen das Assoziativgesetz erfüllt. Die Symmetrien eines Objektes bilden demnach eine Gruppe, die *Symmetriegruppe* dieses Objektes.

Die Symmetriegruppen sind Untergruppen einer allgemeinen Gruppe, der *symmetrischen Gruppe*. Die symmetrische Gruppe ist die Menge

$$S(X) := \{f\colon X \to X \mid f \text{ ist bijektiv}\},$$

in Worten: die Mengen der bijektiven Selbstabbildungen. Eine Abbildung $f\colon X \to Y$ heißt Selbstabbildung, wenn $X = Y$ gilt. In unserem Fall ist $X = \mathbb{R}^2$.

Die Menge $S(X)$ bildet bezüglich Verkettung eine Gruppe, das ist ganz klar, weil die Verkettung das Assoziativgesetz erfüllt und $S(X)$ genau so definiert ist, dass es zu jedem Element $f \in S(X)$ auch ein Inverses bezüglich Verkettung gibt, das ist f^{-1}, die Umkehrabbildung zu f.

Sei U eine Untergruppe von $S(X)$ und $\varphi\colon U \times X \to X$ mit $\varphi(f, x) := f(x)$. Bei φ handelt es sich um eine Gruppenaktion, denn $\varphi(\mathrm{id}, x) = \mathrm{id}(x) = x$ und

$$\varphi(g \circ f, x) = (g \circ f)(x) = g(f(x)) = \varphi(g, \varphi(f, x)).$$

Für eine endliche Menge X bezeichnet man die Untergruppen von $S(X)$ als Permutationsgruppen. Man kann ohne Beschränkung der Allgemeinheit $X := \{1, \ldots, n\}$ und $S_n := S(X)$ setzen, das heißt eigentlich bloß, dass jedem Element von X eine Nummer gegeben wird.

4.2. Ringtheorie

4.2.1. Elementare Gesetzmäßigkeiten

Es gibt in der Mathematik Objekte wie Restklassen, Matrizen und Polynome, für die wie bei den ganzen Zahlen eine Addition und eine Multiplikation definiert ist. Die Addition und Multiplikation von zwei Matrizen ergibt z. B. wieder eine Matrix. In jedem Fall genügen die Addition und Multiplikation einem bestimmten Muster, den Ring-Axiomen. Das legt nahe, aus den Axiomen allgemeine Rechenregeln und Gesetzmäßigkeiten abzuleiten, die somit in allen Ringen gültig sind.

Wir erhalten dadurch als neues Werkzeug ein verallgemeinertes Rechnen. Das ist für uns ganz besonders wichtig, da eine enorme Anzahl von mathematischen Strukturen die Struktur eines Rings enthält. Z. B. ist jeder Körper auch ein Ring. Die rationalen, reellen und komplexen Zahlen bilden jeweils einen Körper. Allein schon dieser Umstand, dass die wichtigsten grundlegenden Zahlenbereiche einen Körper bilden, macht es sinnvoll, Ringe und Körper näher zu studieren.

Ringe sind außerdem bedeutsam als Grundlage für die Konzepterweiterungen Modul und assoziative Algebra. Diesen beiden Begriffen ist auf bestimmte Art geometrische Information eingeimpft, sie sind von großer Tragweite in der linearen Algebra. Z. B. ist jeder Vektorraum, und damit insbesondere jeder euklidische Vektorraum ein Modul. Beispiele für assoziative Algebren sind die Tensoralgebra, die äußere Algebra und die Clifford-Algebra.

Überraschend treten auch in der Analysis solche geometrisch motivierten Konzepte auf. So wurde die Analysis zur Funktionalanalysis weiterentwickelt, die auch mit Vektorräumen arbeitet. Als assoziative Algebren kommen hier die Banachalgebren hinzu.

Neben kontinuierlichen Strukturen sind für die Algebra auch diskrete Strukturen wie Restklassenringe typisch. Die Restklassenringe bilden eine Grundlage für die Zahlentheorie.

Schließlich sind Ringe auch tief in der abstrakten Algebra verwurzelt. Es scheint so, als ergäbe sich dort eine nur schwer überschaubare Fülle von Strukturen. Das mag richtig sein, allerdings bringen die mit der axiomatischen Methode gewonnenen allgemeinen Gesetzmäßigkeiten eine gewisse Ordnung.

Definition 4.7 (Ring).
Eine Struktur $(R, +, \cdot)$ heißt Ring, wenn

1. $(R, +)$ eine kommutative Gruppe ist,
2. $(R, \cdot)$ eine Halbgruppe ist,

3. die Distributivgesetze $a(b+c) = ab + ac$ und $(a+b)c = ac + bc$ für alle $a, b, c \in R$ erfüllt sind.

Es gibt hier einen Unterschied zwischen Linksdistributivgesetz und Rechtsdistributivgesetz, weil die Multiplikation nicht kommutativ sein braucht.

Definition 4.8 (Unitärer Ring).
Ein Ring $(R, +, \cdot)$ heißt unitär oder Ring mit Eins, wenn $(R, \cdot)$ ein Monoid ist.

D. h. ein unitärer Ring ist ein Ring R, in dem es ein Einselement e gibt, so dass $e \cdot a = a \cdot e = a$ für alle $a \in R$. Man kann $e = 1$ schreiben, muss aber beachten, dass damit ein abstraktes Element gemeint ist. Unter Umständen verbietet sich das auch aufgrund von Zweideutigkeit. Z. B. ist im Matrizenring das Einselement die Einheitsmatrix. Diese schreibt man E oder I und nicht 1, um sie von der dort ebenfalls vorkommenden Skalarmultiplikation mit der Zahl 1 unterscheiden zu können.

Korollar 4.8. Sei R ein Ring und $0 \in R$ das Nullelement, dann gilt $0 \cdot a = 0$ und $a \cdot 0 = 0$ für jedes $a \in R$.

Beweis. Man rechnet

$$0a = 0a + 0 = 0a + 0a - 0a = (0+0)a - 0a = 0a - 0a = 0.$$

Für $a \cdot 0$ ist die Rechnung analog. □

Korollar 4.9. Sei R ein Ring und $a, b \in R$, dann gilt $(-a)b = -(ab) = a(-b)$.

Beweis. Man rechnet

$$\begin{aligned}(-a)b &= (-a)b + 0 = (-a)b + ab - (ab) = ((-a) + a)b - (ab) \\ &= 0b - (ab) = 0 - (ab) = -(ab).\end{aligned}$$

Für $a(-b)$ ist die Rechnung analog. □

Korollar 4.10. Sei R ein Ring und $a, b \in R$, dann gilt $(-a)(-b) = ab$.

Beweis. Mit dem letzten Korollar und $-(-x) = x$ rechnet man

$$(-a)(-b) = -((-a)b) = -(-(ab)) = ab. \quad \square$$

Definition 4.9 (Einheitengruppe).
Ist R ein Ring mit Eins e, dann ist die Menge der Einheiten definiert als

$$R^* := \{a \in R \mid \text{es gibt ein } b \in R \text{ mit } ab = ba = e\}.$$

Weil $(R, \cdot)$ ein Monoid ist, muss $(R^*, \cdot)$ eine Gruppe sein, denn die Forderung dass jedes Element mutiplikativ invertierbar ist, ist das letzte Axiom einer multiplikativ geschriebenen Gruppe.

Die Gruppe $\mathbb{Z}^* = \{-1, 1\}$ ist trivial. Ein recht interessantes Beispiel für eine Einheitengruppe ist die allgemeine lineare Gruppe, das ist die Gruppe der invertierbaren quadratischen Matrizen. In der linearen Algebra weiß man, eine quadratische Matrix ist genau dann invertierbar, wenn ihre Determinante nicht verschwindet, das heißt, es gilt

$$(K^{n\times n})^* = \mathrm{GL}(n, K) := \{A \in K^{n\times n} \mid \det(A) \neq 0\}.$$

Hierbei ist K ein beliebiger Körper, z. B. $K = \mathbb{R}$ oder $K = \mathbb{C}$. Es ist ja so, dass der Matrizenraum $K^{m\times n}$ kanonisch isomorph zum Vektorraum $\mathrm{Hom}(K^n, K^m)$ ist, welcher aus allen linearen Abbildungen $K^n \to K^m$ besteht. Um es in einfachen Worten auszudrücken: Multiplikation mit einer Matrix ist eine lineare Abbildung, und jede lineare Abbildung zwischen Koordinatenräumen lässt sich eindeutig als Matrix darstellen. Für $m = n$ handelt es sich um Endomorphismen. Sind diese bijektiv, spricht man von Automorphismen. Demnach ist $\mathrm{GL}(n, K)$ kanonisch isomorph zur Automorphismengruppe $\mathrm{Aut}(K^{n\times n})$. Diese Gruppe besteht aus allen Symmetrien, welche die Vektorraumstruktur respektieren. Darin enthalten sind Untergruppen von Symmetrien wie Spiegelungen und Drehungen.

5. Zahlenbereiche

5.1. Die natürlichen Zahlen

5.1.1. Modelle der natürlichen Zahlen

Zahlen spielen in den Mathematik eine maßgebliche Rolle, und dies nicht nur bei der quantitativen Erfassung von Größen, sondern auch bei der logischen Klärung der Gegenstände. Zum Beispiel setzen wesentliche Teile der Analysis, der linearen Algebra und der Stochastik den Begriff der reellen Zahlen voraus. Und dies sind die Kerngebiete der Mathematik, die das Fundament für die Natur- und Ingenieurwissenschaften bilden.

Den auf den Zahlenbereichen definierten arithmetischen Operationen wohnen bestimmte Rechenregeln inne. Anfangs mag man diese Regeln im Rahmen einer Theoriefindung klären und daraufhin zu Grundregeln gelangen, die man axiomatisch voraussetzt. Stattdessen will ich die Zahlenbereiche aber sogleich vermittels Mengenlehre modellieren und aufzeigen, dass die Modelle die Grundregeln erfüllen. Die Zahlenbereiche werden dabei schrittweise aufeinander aufgebaut.

Wir starten mit den natürlichen Zahlen.

Definition 5.1 (Peano-Axiome).
Eine Menge $\mathbb{N}$ ist als Struktur $(\mathbb{N}, 0, s)$ mit $0 \in \mathbb{N}$ und $s\colon \mathbb{N} \to \mathbb{N}$ ein Modell der *natürlichen Zahlen*, wenn gilt

(P1) s ist injektiv,

(P2) $\forall n \in \mathbb{N}\colon s(n) \neq 0$,

(P3) $A(0) \wedge (\forall n \in \mathbb{N}\colon A(n) \Rightarrow A(s(n))) \Rightarrow (\forall n \in \mathbb{N}\colon A(n))$.

Zusätzlich muss, wie bereits diskutiert, der Rekursionssatz vorausgesetzt werden, um die arithmetischen Operationen definieren zu können. Das Axiom (P3), das das Prinzip der Induktion beschreibt, habe ich als Schema gefasst, um in der Logik erster Stufe verbleiben zu können. Es vermittelt zu jeder Aussageform $A(n)$ ein Axiom. Die ursprüngliche Fassung der Peano-Axiome substituiert A allerdings gegen eine Prädikatvariable und setzt somit die Logik zweiter Stufe voraus.

Definition 5.2 (Addition natürlicher Zahlen).
Die *Addition* zweier natürlicher Zahlen ist rekursiv definiert als

$$a + 0 := a, \qquad a + s(b) := s(a + b).$$

Definition 5.3 (Multiplikation natürlicher Zahlen).
Die *Multiplikation* zweier natürlicher Zahlen ist rekursiv definiert als

$$a \cdot 0 := 0, \qquad a \cdot s(b) := a \cdot b + a.$$

Diese rekursiven Bildungsvorschriften lassen sich direkt so im Computer implementieren. Gleichwohl ist das Prozedere außerordentlich ineffizient und somit praktisch nur für sehr kleine Zahlen brauchbar.

Definition 5.4 (Ordnung der natürlichen Zahlen).
Die *Ordnung* der natürlichen Zahlen ist definiert als

$$a \leq b \;:\Leftrightarrow\; \exists n \in \mathbb{N} \colon a + n = b.$$

Werden die natürlichen Zahlen als formales System dargestellt, in dem gerechnet werden kann, entfällt der Rekursionssatz vorläufig aus der Betrachtung. Dafür werden Addition und Multiplikation axiomatisch erklärt. Man nennt dieses System die *Peano-Arithmetik*, kurz PA. Beschränkt man sich dabei auf die intuitionistische Logik, spricht man von der *Heyting-Arithmetik*, kurz HA.

Definition 5.5 (Peano-Arithmetik).
Die Struktur $(\mathbb{N}, 0, s, +, \cdot, \leq)$ sei ein Modell der *Peano-Arithmetik*, wenn $(\mathbb{N}, 0, s)$ die Peano-Axiome erfüllt und zusätzlich für alle $a, b \in \mathbb{N}$ gilt

$$a + 0 = a, \qquad a + s(b) = s(a + b), \qquad a \leq b \Leftrightarrow \exists n \in \mathbb{N} \colon a + n = b.$$
$$a \cdot 0 = 0, \qquad a \cdot s(b) = a \cdot b + a.$$

Für die natürlichen Zahlen finden sich diverse Modelle, darunter jenes bereits diskutierte, das sie als endliche Ordinalzahlen darstellt. Dahingehend tut sich die Frage auf, ob jedes dieser Modelle gleichberechtigt ist.

Definition 5.6 (Isomorphismus zwischen Peano-Strukturen).
Eine Bijektion $f \colon \mathbb{N} \to \mathbb{N}'$ ist ein Isomorphismus zwischen zwei Peano-Strukturen $(\mathbb{N}, 0, s)$ und $(\mathbb{N}', 0', s')$, wenn $f(0) = 0'$ und $f(s(n)) = s'(f(n))$ gilt.

Man rechnet unschwer nach, dass die Umkehrabbildung eines Isomorphismus' ebenfalls ein Isomorphismus ist.

Satz 5.1 (Isomorphiesatz von Dedekind).
Je zwei Modelle der natürlichen Zahlen sind Isomorph.

Beweis. Es seien $(\mathbb{N}, 0, s)$, $(\mathbb{N}', 0', s')$ zwei Modelle der natürlichen Zahlen. Vermittels des Rekursionssatzes definiert man

$$\begin{aligned} f\colon \mathbb{N} \to \mathbb{N}', &\quad f(0) := 0', &\quad f(s(n)) := s'(f(n)), \\ g\colon \mathbb{N}' \to \mathbb{N}, &\quad g(0') := 0, &\quad g(s'(n)) := s(g(n)), \\ h\colon \mathbb{N} \to \mathbb{N}, &\quad h(0) := 0, &\quad h(s(n)) := s(h(n)). \end{aligned}$$

Man rechnet unschwer nach, dass $g \circ f = h$ und $h = \mathrm{id}$ gilt. Analog erhält man $f \circ g = \mathrm{id}$, womit f, g bijektiv sind. Und per Definition sind f, g Isomorphismen zwischen Peano-Strukturen. □

Man macht sich außerdem klar, dass die rekursive Definition des Isomorphismus in Def. 5.6 enthalten ist. Das heißt, der Isomorphismus ist zudem noch eindeutig bestimmt. Insofern es also nur einen, und somit einen ausgezeichneten Isomorphismus gibt, kann man sagen, dass je zwei Modelle der natürlichen Zahlen kanonisch isomorph sind.

Satz 5.2. Mit $1 := s(0)$ gilt $s(a) = a + 1$.

Beweis. Es findet sich $a + 1 = a + s(0) = s(a + 0) = s(a)$. □

Satz 5.3 (Assoziativgesetz der Addition).
Für alle $a, b, c \in \mathbb{N}$ gilt $(a + b) + c = a + (b + c)$.

Beweis. Induktion über c. Im Anfang $c = 0$ gilt

$$(a + b) + c = (a + b) + 0 = a + b = a + (b + 0) = a + (b + c).$$

Mit der Induktionsvoraussetzung $(a + b) + c = a + (b + c)$ findet sich

$$(a+b)+s(c) = s((a+b)+c) \overset{\text{IV}}{=} s(a+(b+c)) = a+s(b+c) = a+(b+s(c)). \qquad \square$$

Satz 5.4 (Neutrales Element der Addition).
Für alle $a \in \mathbb{N}$ gilt $0 + a = a + 0 = a$.

Beweis. Per Definition gilt $a+0 = a$. Zu $0+a$ per Induktion über a. Im Anfang $a = 0$ ist $0 + a = 0$ per Definition. Zum Schritt. Induktionsvoraussetzung sei $0 + a = a$. Man findet

$$0 + s(a) = s(0 + a) \overset{\text{IV}}{=} s(a). \qquad \square$$

Satz 5.5 (Kommutativgesetz der Addition).
Für alle $a, b \in \mathbb{N}$ gilt $a + b = b + a$.

Beweis. Zunächst $a + 1 = 1 + a$ per Induktion über a. Im Anfang $a = 0$ gilt die Aussage gemäß Satz 5.4. Zum Schritt. Man findet

$$1 + s(a) = s(1 + a) \overset{\text{IV}}{=} s(a + 1) = a + s(1).$$

Nun $a + b = b + a$ per Induktion über b. Der Fall $b = 0$ gilt gemäß Satz 5.4. Anfang sei $b = 1$. Dieser wurde zuvor gezeigt. Zum Schritt findet sich

$$\begin{aligned} s(b) + a &= (b + 1) + a = b + (1 + a) = b + (a + 1) = b + s(a) \\ &= s(b + a) \overset{\text{IV}}{=} s(a + b) = a + s(b). \end{aligned}$$ □

Satz 5.6 (Distributivgesetz der Multiplikation).
Für alle $a, b, c \in \mathbb{N}$ gilt $(a + b)c = ac + bc$.

Beweis. Induktion über c. Im Anfang $c = 0$ resultieren beide Seiten der Gleichung im Wert 0. Zum Schritt findet sich

$$(a+b)s(c) = (a+b)c+(a+b) \overset{\text{IV}}{=} ac+bc+a+b = ac+a+bc+b = as(c)+bs(c).$$ □

Satz 5.7. Für alle $a \in \mathbb{N}$ gilt $0a = 0$.

Beweis. Induktion über a. Im Anfang $a = 0$ ist per Definition $0a = 0$. Der Schritt ist

$$0s(a) = 0a + 0 \overset{\text{IV}}{=} 0 + 0 = 0.$$

Satz 5.8 (Neutrales Element der Multiplikation).
Für alle $a \in \mathbb{N}$ gilt $a \cdot 1 = a$ und $1 \cdot a = a$.

Beweis. Die Formel $a \cdot 1 = a$ folgt unmittelbar aus der Definition und bereits bekannten Regeln.

Die Formel $1 \cdot a = a$ per Induktion über a. Im Anfang $a = 0$ folgt die Regel unmittelbar aus der Definition. Der Schritt ist

$$1s(a) = 1a + 1 \overset{\text{IV}}{=} a + 1 = s(a).$$ □

Satz 5.9 (Kommutativgesetz der Multiplikation).
Für alle $a, b \in \mathbb{N}$ gilt $ab = ba$.

Beweis. Induktion über b. Im Anfang $b = 0$ gilt die Regel gemäß Lemma 5.7. Der Schritt ist

$$as(b) = ab + a \overset{\text{IV}}{=} ba + a = ba + 1a = (b+1)a = s(b)a. \qquad \square$$

Satz 5.10. Aus $a \leq b$ folgt $a + c \leq b + c$.

Beweis. Mit der Prämisse liegt ein n mit $a + n = b$ vor. Somit folgt $a + c + n = b + c$, womit n auch ein Zeuge für $a + c \leq b + c$ ist. $\square$

5.2. Die ganzen Zahlen

5.2.1. Konstruktion

Definition 5.7 (Ganze Zahlen).
Auf $\mathbb{N}_0 \times \mathbb{N}_0$ wird die folgende Äquivalenzrelation definiert:

$$(x, y) \sim (x', y') \;:\Leftrightarrow\; x + y' = x' + y.$$

Die Quotientenmenge $\mathbb{Z} := (\mathbb{N}_0 \times \mathbb{N}_0)/\!\sim$ nennt man die *ganzen Zahlen*.

Satz 5.11 (Ring der ganzen Zahlen).
Die Operationen

$$[(x, y)] + [(x', y')] := [(x + x', y + y')],$$
$$[(x, y)] \cdot [(x', y')] := [(xx' + yy', xy' + x'y)]$$

sind auf $\mathbb{Z}$ wohldefiniert und $(\mathbb{Z}, +, \cdot)$ bildet einen kommutativen unitären Ring.

Beweis. Wohldefiniert heißt, dass das Ergebnis der Operationen nicht von den gewählten Repräsentanten der Argumente abhängig ist. Sei dazu $(x, y) \sim (a, b)$ und $(x', y') \sim (a', b')$. Zu zeigen ist nun

$$(x + x', y + y') \sim (a + a', b + b'),$$

was laut Definition zu

$$(x + x') + (b + b') = (a + a') + (y + y').$$

äquivalent ist. Gemäß Voraussetzung ist $x + b = a + y$ und $x' + b' = a' + y'$. Man bekommt damit auf der linken Seite

$$x + x' + b + b' = x + b + x' + b' = a + y + a' + y',$$

was wiederum mit der rechten Seite übereinstimmt.

Mit der Multiplikation verhält es sich etwas komplizierter. Zu Vereinfachung wird zunächst gezeigt:

$$\begin{aligned}
&[(x, y)] \cdot [(x', y')] = [(a, b)] \cdot [(x', y')], \\
&\iff (xx' + yy', xy' + yx') \sim (ax' + by', ay' + bx') \\
&\iff xx' + yy' + ay' + bx' = ax' + by' + xy' + yx' \\
&\iff (x + b)x' + (a + y)y' = (a + y)x' + (x + b)y'.
\end{aligned}$$

Diese Gleichung ist gemäß Voraussetzung $(x, y) \sim (a, b)$ bzw. $x + b = a + y$ erfüllt.

Analog bestätigt man

$$[(a, b)] \cdot [(x', y')] = [(a, b)] \cdot [(a', b')].$$

Gemäß Transitivität ergibt sich somit

$$[(x, y)] \cdot [(x', y')] = [(a, b)] \cdot [(a', b')].$$

Es ist nun zu bestätigen, dass $(\mathbb{Z}, +)$ eine kommutative Gruppe ist. Das Assoziativgesetz:

$$\begin{aligned}
&([(x, y)] + [(x', y')]) + [(x'', y'')] = [(x + x', y + y')] + [(x'', y'')] \\
&= [(x + x' + x'', y + y' + y'')] = [(x, y)] + [(x' + x'', y' + y'')] \\
&= [(x, y)] + ([(x', y')] + [(x'', y'')]).
\end{aligned}$$

Das neutrale Element ist $[(0, 0)]$:

$$[(x, y)] + [(0, 0)] = [(x + 0, y + 0)] = [(x, y)].$$

Das inverse Element zu $[(x, y)]$ ist $[(y, x)]$, denn es gilt

$$\begin{aligned}
&[(x, y)] + [(y, x)] = [(x + y, y + x)] = [(0, 0)] \\
&\iff (x + y, y + x) \sim (0, 0) \iff x + y + 0 = y + x + 0.
\end{aligned}$$

Das Kommutativgesetz:

$$[(x, y)] + [(x', y')] = [(x + x', y + y')] = [(x' + x, y' + y)] = [(x', y')] + [(x, y)].$$

Es ist nun zu bestätigen, dass $(\mathbb{Z}, \cdot)$ ein kommutatives Monoid bildet. Das Assoziativgesetz:

$$\begin{aligned}
&([(x,y)] \cdot [(x',y')]) \cdot [(x'',y'')] = [(xx'+yy', xy'+x'y)] \cdot [(x'',y'')] \\
&= [(xx'x''+x''yy'+xy'y''+x'yy'',\ xx'y''+yy'y''+xx''y'+x'x''y)] \\
&= [(x,y)] \cdot [(x'x''+y'y'', x'y''+x''y')] = [(x,y)] \cdot ([(x',y')] \cdot [(x'',y'')]).
\end{aligned}$$

Das Kommutativgesetz:

$$\begin{aligned}
&[(x,y)] \cdot [(x',y')] = [(xx'+yy',\ xy'+yx')] \\
&= [(x'x+y'y,\ x'y+xy')] = [(x',y')] \cdot [(x,y)].
\end{aligned}$$

Das neutrale Element ist $[(1,0)]$, denn es gilt

$$[(x,y)] \cdot [(1,0)] = [(x \cdot 1 + y \cdot 0,\ 1 \cdot y + x \cdot 0)] = [(x,y)].$$

Schließlich ist noch das Distributivgesetz zu bestätigen. Man findet

$$\begin{aligned}
&[(a,b)] \cdot ([(x,y)] + [(x',y')]) = [(a,b)] \cdot [(x+x', y+y')] \\
&= [(ax+ax'+by+by',\ ay+ay'+bx+bx')] \\
&= [(ax+by, ay+bx)] + [(ax'+by', ay'+bx')] \\
&= [(a,b)] \cdot [(x,y)] + [(a,b)] \cdot [(x',y')].
\end{aligned}$$

Somit sind alle Axiome bestätigt. □

Definition 5.8 (Monoidhomomorphismus).
Seien $(M,+)$ und $(M',+')$ zwei Monoide. Eine Abbildung $\varphi\colon M \to M'$ heißt Monoidhomomorphismus, wenn für alle $a, b \in M$ gilt

$$\varphi(a+b) = \varphi(a) + \varphi(b)$$

und $\varphi(0) = 0'$ ist.

Einen injektiven Homomorphismus nennt man Monomorphismus. Ein Monomorphismus charakterisiert eine Einbettung einer Struktur als Unterstruktur einer anderen.

Satz 5.12 (Einbettung der natürlichen Zahlen in die ganzen).
Die Abbildung $\varphi\colon \mathbb{N}_0 \to \mathbb{Z}$ mit $\varphi(n) := [(n,0)]$ ist ein Monoidmonomorphismus.

Beweis. Es ergibt sich

$$\varphi(a+b) = [(a+b,0)] = [(a,0)] + [(b,0)] = \varphi(a) + \varphi(b).$$

Außerdem ist $\varphi(0) = [(0,0)]$, und $[(0,0)]$ ist das neutrale Element von $(\mathbb{Z},+)$.

Schließlich ist noch die Injektivität zu prüfen:

$$[(a,0)] = \varphi(a) = \varphi(b) = [(b,0)] \iff (a,0) \sim (b,0)$$
$$\iff a+0 = b+0 \iff a = b. \square$$

Anstelle von $\varphi(n) = [(n,0)]$ darf man daher einfach schreiben $n = [(n,0)]$. Außerdem definiert man $a - b := a + (-b)$. Daraus ergibt sich nun

$$[(x,y)] = [(x,0)] + [(0,y)] = [(x,0)] - [(y,0)] = x - y.$$

Die umständliche Schreibweise $[(x,y)]$ wird ab jetzt nicht mehr benötigt.

Definition 5.9 (Totalordnung der ganzen Zahlen).
Auf $\mathbb{Z}$ wird die folgende strenge Totalordnung definiert:

$$[(x,y)] < [(x',y')] :\Leftrightarrow x + y' < x' + y.$$

Satz 5.13 (Einbettung der Totalordnung).
Die Abbildung φ aus Satz 5.12 genügt der Forderung

$$a < b \Rightarrow \varphi(a) < \varphi(b).$$

Beweis. Nach den Definitionen ist

$$\varphi(a) < \varphi(b) \iff [(a,0)] < [(b,0)] \iff a + 0 < 0 + b \iff a < b. \square$$

5.3. Die rationalen Zahlen

5.3.1. Konstruktion

Definition 5.10 (Rationale Zahlen).
Auf $\mathbb{Z} \times \mathbb{N}_1$ wird die folgende Äquivalenzrelation definiert:

$$(x,y) \sim (x',y') :\Leftrightarrow xy' = x'y.$$

Die Quotientenmenge $\mathbb{Q} := (\mathbb{Z} \times \mathbb{N}_1)/\sim$ nennt man die *rationalen Zahlen*.

Für die Äquivalenzklasse $[(x,y)]$ schreibt man $\frac{x}{y}$.

Satz 5.14 (Körper der rationalen Zahlen).
Die Operationen

$$\frac{x}{y} + \frac{x'}{y'} := \frac{xy' + x'y}{yy'}, \qquad \frac{x}{y} \cdot \frac{x'}{y'} := \frac{xx'}{yy'}$$

sind auf $\mathbb{Q}$ wohldefiniert und $(\mathbb{Q}, +, \cdot)$ bildet einen Körper.

Beweis. Wohldefiniert bedeutet, dass das Ergebnis der Operationen nicht von den gewählten Repräsentanten der Argumente abhängig ist. Sei dazu $(a, b) \sim (x, y)$ und $(a', b') \sim (x', y')$. Zu zeigen ist nun

$$\begin{aligned}
&(ab' + a'b, bb') \sim (xy' + x'y, yy'),\\
&\iff (ab' + a'b)(yy') = (xy' + x'y)(bb')\\
&\iff ab'yy' + a'byy' = xy'bb' + x'ybb'.
\end{aligned}$$

Substituiert man $ay = xb$ und $a'y' = x'b'$ auf der linken Seite der Gleichung, ergibt sich die rechte Seite. Zu zeigen ist weiterhin

$$(aa', bb') \sim (xx', yy') \iff aa'yy' = xx'bb'.$$

Wieder wird linke Seite der Gleichung über $ay = xb$ und $a'y' = x'b'$ in die rechte Seite überführt. Die Wohldefiniertheit der Operationen ist damit bestätigt.

Es bleibt zu prüfen, dass $(\mathbb{Q}, +, \cdot)$ allen Körperaxiomen genügt. Das neutrale Element der Addition ist $0/1$, denn es gilt

$$\frac{x}{y} + \frac{0}{1} = \frac{x \cdot 1 + 0 \cdot y}{y \cdot 1} = \frac{x}{y}.$$

Das neutrale Element der Multiplikation ist $1/1$, denn es gilt

$$\frac{x}{y} \cdot \frac{1}{1} = \frac{x \cdot 1}{y \cdot 1} = \frac{x}{y}.$$

Die Assoziativität der Addition ergibt sich ohne größere Umstände:

$$\begin{aligned}
\left(\frac{x}{y} + \frac{x'}{y'}\right) + \frac{x''}{y''} &= \frac{xy' + x'y}{yy'} + \frac{x''}{y''} = \frac{xy'y'' + x'yy'' + x''yy'}{yy'y''},\\
\frac{x}{y} + \left(\frac{x'}{y'} + \frac{x''}{y''}\right) &= \frac{x}{y} + \frac{x'y'' + x''y'}{y'y''} = \frac{xy'y'' + x'yy'' + x''yy'}{yy'y''}.
\end{aligned}$$

Die Assoziativität der Multiplikation ist etwas einfacher:

$$\left(\frac{x}{y} \cdot \frac{x'}{y'}\right) \cdot \frac{x''}{y''} = \frac{xx'}{yy'} \cdot \frac{x''}{y''} = \frac{xx'x''}{yy'y''} = \frac{x}{y} \cdot \frac{x'x''}{y'y''} = \frac{x}{y} \cdot \left(\frac{x'}{y'} \cdot \frac{x''}{y''}\right).$$

Das Kommutativgesetz der Addition:

$$\frac{x}{y} + \frac{x'}{y'} = \frac{xy' + x'y}{yy'} = \frac{x'y + xy'}{y'y} = \frac{x'}{y'} + \frac{x}{y}.$$

Das Kommutativgesetz der Multiplikation:

$$\frac{x}{y'} \cdot \frac{x'}{y'} = \frac{xx'}{yy'} = \frac{x'x}{y'y} = \frac{x'}{y'} \cdot \frac{x}{y}.$$

Das additiv inverse Element zu x/y ist $(-x)/y$, denn es gilt

$$\frac{x}{y} + \frac{-x}{y} = \frac{xy + (-x)y}{y^2} = \frac{0}{y^2} = \frac{0}{1}.$$

Das multiplikativ inverse Element zu x/y mit $x \neq 0$ ist y/x, denn es gilt

$$\frac{x}{y} \cdot \frac{y}{x} = \frac{xy}{xy} = \frac{1}{1}.$$

Schließlich findet bestätigt man noch das Distributivgesetz:

$$\frac{a}{b} \cdot \left(\frac{x}{y} + \frac{x'}{y'}\right) = \frac{a}{b} \cdot \frac{xy' + x'y}{yy'} = \frac{axy' + ax'y}{byy'},$$
$$\frac{ax}{by} + \frac{ax'}{by'} = \frac{axby' + ax'by}{byby'} = \frac{b}{b} \cdot \frac{axy' + ax'y}{byy'}.$$

Hierbei beachtet man, dass $b/b = 1/1$ das multiplikativ neutrale Element ist. □

Definition 5.11 (Ringhomomorphismus).
Seien $(R, +, *)$ und $(R', +', *')$ zwei Ringe. Die Abbildung $\varphi\colon R \to R'$ heißt *Ringhomomorphismus*, wenn für alle $a, b \in R$ gilt:

$$\varphi(a + b) = \varphi(a) +' \varphi(b), \qquad \varphi(a * b) = \varphi(a) *' \varphi(b).$$

Besitzt R ein Einselement 1 und R' ein Einselement $1'$, dann nennt man φ *Einserhaltend*, wenn $\varphi(1) = 1'$ ist.

Einen injektiver Homomorphismus wird Monomorphismus genannt. Ein Monomorphismus charakterisiert eine Einbettung einer Unterstruktur in eine andere Struktur.

Satz 5.15 (Einbettung der ganzen Zahlen in die rationalen).
Sei $\varphi\colon \mathbb{Z} \to \mathbb{Q}$ mit $\varphi(z) := z/1$. Die Abbildung φ ist ein Eins-erhaltender Ringmonomorphismus.

Beweis. Die Erhaltung des Einselements ergibt sich trivial. Ferner findet man

$$\varphi(a+b) = \frac{a+b}{1} = \frac{a\cdot 1 + b\cdot 1}{1\cdot 1} = \frac{a}{1} + \frac{b}{1} = \varphi(a) + \varphi(b)$$

und

$$\varphi(ab) = \frac{ab}{1} = \frac{ab}{1\cdot 1} = \frac{a}{1}\cdot\frac{b}{1} = \varphi(a)\cdot\varphi(b). \qquad \square$$

Gemäß der Einbettung können wir die ganze Zahl z ab jetzt mit der rationalen Zahl $z/1$ identifizieren. Das heißt, man schreibt einfach $z = z/1$ statt $\varphi(z) = z/1$.

Definition 5.12 (Division rationaler Zahlen).
Wie in jedem Körper ist die Division für $a, b \in \mathbb{Q}$ definiert als $a/b := ab^{-1}$.

Die Division ist also gerade die Multiplikation des Kehrwertes des Nenners:

$$\frac{x}{y} / \frac{x'}{y'} = \frac{x}{y}\cdot\frac{y'}{x'}.$$

Die Division muss natürlich mit der Notation für rationale Zahlen kompatibel sein, sonst dürfte man nicht die gleiche Schreibweise verwenden. Zur Unterscheidung schreiben wir Division für einen Augenblick mit Doppelstrich als $a//b$. Man findet

$$\frac{x}{y} = \frac{x}{1}\cdot\frac{1}{y} = \frac{x}{1} // \frac{y}{1} = x//y.$$

Tatsächlich führt beides zum gleichen Ergebnis.

Da die rationalen Zahlen einen Körper bilden, gilt $a/a = aa^{-1} = 1$ für jede rationale Zahl a.

Satz 5.16 (Addition, Subtraktion, Multiplikation von Brüchen).
Seien a, b, c, d rationale Zahlen mit $b \neq 0$ und $d \neq 0$. Es gilt

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}, \qquad \frac{a}{b} - \frac{c}{d} = \frac{ad-bc}{bd}, \qquad \frac{a}{b}\cdot\frac{c}{d} = \frac{ac}{bd}.$$

Der Beweis wird dem Leser überlassen.

6. Ein kategorieller Blick auf die Logik

6.1. Grundbegriffe

6.1.1. Kategorien

Logisches Schließen weist enge Bezüge zu Typentheorien und zu Konstruktionen in der Mengenlehre auf. Im Folgenden wird eine Abstraktion erklärt, die bei der tieferen Klärung dieser Bezüge Hilfe leistet.

Definition 6.1 (Kategorie). Eine *Kategorie* ist ein Tripel $\mathbf{C} = (\mathrm{Ob}, \mathrm{Hom}, \circ)$, sofern die folgenden beiden Axiome erfüllt sind:

1. Für $f\colon A \to B$, $g\colon B \to C$, $h\colon C \to D$ gilt das Assoziativgesetz $h \circ (g \circ f) = (h \circ g) \circ f$.
2. Für jedes Objekt X existiert die Identität $\mathrm{id}_X\colon X \to X$, so dass $f \circ \mathrm{id}_A = \mathrm{id}_B \circ f = f$ für alle Objekte A, B und $f\colon A \to B$.

Die Elemente der Klasse Ob nennt man *Objekte*. Die Elemente der Klasse Hom nennt man *Morphismen*. Die Schreibweise $f\colon X \to Y$ ist gleichbedeutend mit $f \in \mathrm{Hom}(X, Y)$, wobei $X, Y \in \mathrm{Ob}$. Mit $\mathrm{Hom}(X, Y)$ ist die Teilklasse von Hom gemeint, die alle Morphismen von X nach Y enthält. Man schreibt $\mathrm{dom}(f) = X$ und $\mathrm{cod}(f) = Y$. Die Verknüpfung $g \circ f$, gelesen »g nach f«, ist definiert für $\mathrm{cod}(f) = \mathrm{dom}(g)$. Man nennt sie die *Verkettung* von g und f.

Nun gut, man macht hier zunächst zwei Beobachtungen. Erstens erinnern die Axiome an die Regeln für die Verkettung von Abbildungen. In der Tat bilden die Mengen mit den Abbildungen als Morphismen eine Kategorie.

Zweitens erinnern die Axiome an die Monoid-Axiome, haben aber den Unterschied, dass die Morphismen, die verkettet werden sollen, kompatibel sein müssen. Wie sich unschwer herausstellt, bilden die Monoide genau die Kategorien, die ein einziges Objekt besitzen. Die Morphismen werden dabei als die Elemente des Objektes X gedeutet, die Verkettung als deren Verknüpfung. Zwei Morphismen f, g müssen hierbei zwingend kompatibel sein, denn $\mathrm{cod}(f) = X$ und $\mathrm{dom}(g) = X$. Hiermit liegt ein erstes Beispiel vor, in dem die Morphismen nicht notwendigerweise Abbildungen darstellen.

Satz 6.1 (Kategorie der Mengen).
Sei Ω das Mengenuniversum und für $A, B \in \Omega$ sei $\mathrm{Hom}(A, B) := \mathrm{Abb}(A, B)$. Sei $g \circ f$ die Verkettung von Abbildungen. Dann bildet **Set** $:= (\Omega, \mathrm{Hom}, \circ)$ eine Kategorie.

Beweis. Trivial. □

Satz 6.2 (Kategorie der Gruppen).
Sei Ω die Klasse aller Gruppen und für $G, H \in \Omega$ sei $\mathrm{Hom}(G, H)$ die Klasse der Homomorphismen von G nach H. Sei $g \circ f$ die Verkettung von Homomorphismen. Dann bildet **Group** $:= (\Omega, \mathrm{Hom}, \circ)$ eine Kategorie.

Beweis. Homomorphismen sind Abbildungen, die Axiome daher wie bei der Kategorie der Mengen erfüllt. Die Verkettung zweier Homomorphismen ist ja auch ein Homomorphismus. □

Entsprechend bilden Ringe mit Ringhomomorphismen, Körper mit Körperhomomorphismen, Vektorräume mit Vektorraumhomomorphismen usw. Kategorien. Des Weiteren bilden die endlichen Mengen, Gruppen, Ringe jeweils eine Kategorie.

6.1.2. Funktoren

Nun ist es so, dass Gruppen auch Mengen und Homomorphismen auch Abbildungen sind. Die Kategorie der Gruppen ist gewissermaßen in der Kategorie der Mengen enthalten. Um das zu präzisieren, benötigen wir den Begriff des Vergissfunktors.

Definition 6.2 (Kovarianter Funktor).
Sind $\mathbf{C}, \mathbf{D}$ Kategorien, dann nennt man $F\colon \mathbf{C} \to \mathbf{D}$ einen *kovarianten Funktor*, wenn jedem Objekt X von $\mathbf{C}$ ein Objekt $F(X)$ von $\mathbf{D}$ zugeordnet wird und jedem Morphismus $f \in \mathrm{Hom}_{\mathbf{C}}(X, Y)$ ein ein Morphismus $F(f) \in \mathrm{Hom}_{\mathbf{D}}(F(X), F(Y))$ zugeordnet wird, so dass die folgenden beiden Verträglichkeitsaxiome erfüllt sind:

$$F(g \circ f) = F(g) \circ F(f),$$
$$F(\mathrm{id}_X) = \mathrm{id}_{F(X)}\,.$$

Definition 6.3 (Kontravarianter Funktor).
Wie beim kovarianten Funktor, mit dem Unterschied $F(g \circ f) = F(f) \circ F(g)$.

Bemerkung. Die Notation ist überladen. Nämlich ist die Zuordnung $F\colon \mathrm{Ob}(\mathbf{C}) \to \mathrm{Ob}(\mathbf{D})$ zu unterscheiden von

$$\tilde{F}\colon \mathrm{Hom}_{\mathbf{C}}(X, Y) \to \mathrm{Hom}_{\mathbf{D}}(F(X), F(Y)).$$

Das Paar $(F, \tilde{F})$ kodiert dann eigentlich den Funktor $\mathbf{C} \to \mathbf{D}$.

Satz 6.3 (Vergissfunktor).
Sei $F\colon$ **Group** $\to$ **Set** mit $F((G, *, e)) := G$, und jedem Gruppenhomomorphismus

$$\varphi\colon (G, *, e) \to (G', *', e')$$

sei die Abbildung $F(\varphi)\colon G \to G'$ mit $F(\varphi)(x) := \varphi(x)$ zugeordnet. Bei F handelt es sich um einen kovarianten Funktor.

Beweis. Es gilt $F(\mathrm{id})(x) = \mathrm{id}(x)$, und daher $F(\mathrm{id}) = \mathrm{id}$. Außerdem gilt

$$F(\varphi_2 \circ \varphi_1)(x) = (\varphi_2 \circ \varphi_1)(x) = \varphi_2(\varphi_1(x)) = F(\varphi_2)(F(\varphi_1)(x)) = (F(\varphi_2) \circ F(\varphi_1))(x),$$

und daher $F(\varphi_2 \circ \varphi_1) = F(\varphi_2) \circ F(\varphi_1)$. □

Satz 6.4. Sei $P(X) = 2^X$ die Potenzmenge von X. Dann ist wie folgt ein kovarianter Funktor gegeben:

$$P\colon \mathbf{Set} \to \mathbf{Set}, \quad P(X) := 2^X, \quad P(f)(M) := f(M),$$

wobei f eine beliebige Abbildung und $f(M)$ die Bildmenge von M unter f ist.

Beweis. Nach Satz 3.14 gilt

$$P(g \circ f)(M) = (g \circ f)(M) = g(f(M)) = P(g)(P(f)(M)) = (P(g) \circ P(f))(M).$$

Daher ist $P(g \circ f) = P(g) \circ P(f)$. Außerdem ist

$$P(\mathrm{id}_X)(M) = \mathrm{id}_X(M) = M = \mathrm{id}_{P(X)}(M)$$

und daher $P(\mathrm{id}_X) = \mathrm{id}_{P(X)}$. □

Zum Funktor P kommt noch ein weiterer Aspekt hinzu. Für eine Abbildung f kann man ganz pedantisch das Bild $f(x)$ von der Bildmenge $f(\{x\})$ unterscheiden. Aufgrund der Gleichung $f(\{x\}) = \{f(x)\}$ verschwimmt diese Unterscheidung aber gewissermaßen. Die Abbildungen f und $P(f)$ verhalten sich also gewissermaßen gleich. Man kann sagen, dass f auf ganz natürliche Art und Weise die Abbildung $P(f)$ zugeordnet ist. Definiert man

$$\eta(X)\colon X \to 2^X, \quad \eta(X)(x) := \{x\},$$

dann kommutiert das folgende Diagramm:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle \eta(X)}\downarrow & & \downarrow{\scriptstyle \eta(Y)} \\ 2^X & \xrightarrow[P(f)]{} & 2^Y \end{array}$$

D. h. es gilt $\eta(Y) \circ f = P(f) \circ \eta(X)$. Die Zuordnung η ist eine sogenannte natürliche Transformation.

Definition 6.4 (Natürliche Transformation).
Seien **C**, **D** Kategorien und $F, G\colon \mathbf{C} \to \mathbf{D}$ Funktoren. Dann schreibt man $\eta\colon F \to G$ und nennt η *natürliche Transformation*, wenn die folgenden beiden Axiome erfüllt sind:

1. Jedes Objekt X von **C** bekommt einen Morphismus $\eta(X)\colon F(X) \to G(X)$.
2. Für jeden Morphismus $f\colon X \to Y$ gilt $\eta(Y) \circ F(f) = G(f) \circ \eta(X)$.

Die zweite Bedingung lässt sich übersichtlich als kommutierendes Diagramm darstellen:

$$\begin{array}{ccc} F(X) & \xrightarrow{F(f)} & F(Y) \\ {\scriptstyle \eta(X)}\downarrow & & \downarrow{\scriptstyle \eta(Y)} \\ G(X) & \xrightarrow[G(f)]{} & G(Y) \end{array}$$

Ein weiteres Beispiel ergibt sich bezüglich Äquivalenzrelationen in Erinnerung an die auf S. 105 erklärten wohldefinierten Abbildungen. Eine Abbildung $f\colon M \to M'$ heiße *induzierend* bezüglich $(M, \sim)$, $(M, \sim')$, wenn für alle $a, b \in M$ gilt

$$a \sim b \Rightarrow f(a) \sim' f(b).$$

Satz 6.5. Die Paare $(M, \sim)$, bestehend aus Menge und Äquivalenzrelation, bilden mit den induzierenden Abbildungen als Morphismen bezüglich Verkettung eine Kategorie.

Beweis. Die identische Abbildung ist offensichtlich induzierend. Hat man neben $f\colon M \to M'$ eine weitere induzierende Abbildung $g\colon M' \to M''$, dann folgt $g(y) \sim''$

$g(b)$ aus $y \sim' b$. Aus $x \sim a$ folgt mit $y := f(x)$ und $b := f(a)$ somit $g(f(x)) \sim'' g(f(x))$. Daher ist auch $g \circ f$ induzierend. □

Genau dann wenn f induzierend ist, existiert eine induzierte Abbildung

$$I(f)\colon M/{\sim} \longrightarrow M'/{\sim'}, \text{ so dass } I(f) \circ \pi = \pi' \circ f,$$

wobei π, π' jeweils die kanonische Projektion ist.

Satz 6.6. Bei der Induktion I handelt es sich um einen kovarianten Funktor.

Beweis. Man betrachte das folgende kommutierende Diagramm:

$$\begin{array}{ccccc} M & \xrightarrow{f} & M' & \xrightarrow{g} & M'' \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi'} & & \downarrow{\scriptstyle \pi''} \\ M/{\sim} & \xrightarrow[I(f)]{} & M'/{\sim'} & \xrightarrow[I(g)]{} & M''/{\sim''} \end{array}$$

Die Induktion I besitzt die Eigenschaften

$$\begin{aligned} &I(f) \circ \pi = \pi' \circ f, \\ &I(g) \circ \pi' = \pi'' \circ g, \\ &I(g \circ f) \circ \pi = \pi'' \circ (g \circ f). \end{aligned}$$

Damit kann man nun rechnen

$$I(g \circ f) \circ \pi = \pi'' \circ g \circ f = I(g) \circ \pi' \circ f = I(g) \circ I(f) \circ \pi. \tag{6.1}$$

Infolge gilt $I(g \circ f) = I(g) \circ I(f)$, da die kanonische Projektion π eine Surjektion ist. Aus der Forderung $I(\mathrm{id}) \circ \pi = \pi \circ \mathrm{id} = \pi$ ergibt sich $I(\mathrm{id}) = \mathrm{id}$, da π surjektiv ist. □

Die Abbildung $\eta((M, \sim)) := \pi$, die jeder Menge mit Äquivalenzrelation ihre kanonische Projektion zuordnet, ist eine natürliche Transformation.

Funktoren haben einen allgemeine Eigenschaft.

Satz 6.7. Wird ein Funktor auf einen Isomorphismus angewendet, ist das Resultat wieder ein Isomorphismus.

Beweis. Dieser ist direkt aus den Definitionen zu erhalten. Wir betrachten nur einen kovarianten Funktor F. Der Beweis für einen kontravarianten Funktor verläuft analog.

Sei $f\colon X \to Y$ ein beliebiger Isomorphismus im Definitionsbereich des Funktors. Laut Definition existert eine Inverse, das heißt, ein $g\colon Y \to X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$. Gemäß der definierenden Eigenschaft eines Funktors darf man rechnen

$$\mathrm{id}_{F(X)} = F(\mathrm{id}_X) = F(g \circ f) = F(g) \circ F(f),$$
$$\mathrm{id}_{F(Y)} = F(\mathrm{id}_Y) = F(f \circ g) = F(f) \circ F(g).$$

Somit ist $F(f)$ ein Isomorphismus mit Inverse $F(g)$. □

Wird beispielsweise der Vergissfunktor von Gruppen zu Mengen auf einen Gruppenisomorphismus angewendet, ist das Resultat zwingend eine Bijektion, da die Bijektionen die Isomorphismen in der Kategorie der Mengen sind.

6.1.3. Anfangs- und Endobjekte

Definition 6.5 (Anfangsobjekt, Endobjekt, Nullobjekt).
Es sei $\mathbf{C}$ eine Kategorie. Ein $A \in \mathrm{Ob}(\mathbf{C})$ heißt *Anfangsobjekt*, wenn es zu jedem Objekt $X \in \mathrm{Ob}(\mathbf{C})$ genau einen Morphismus $A \to X$ gibt. Ein $E \in \mathrm{Ob}(\mathbf{C})$ heißt *Endobjekt*, wenn es zu jedem Objekt $X \in \mathrm{Ob}(\mathbf{C})$ genau einen Morphismus $X \to E$ gibt. Ein Objekt heißt *Nullobjekt*, wenn es sowohl Anfangsobjekt als auch Endobjekt ist.

Mengen. Wir untersuchen zunächst die Kategorie der Mengen. Anfangsobjekt bedeutet hier eine Menge A, bei der es zu jeder Menge X genau eine Abbildung $A \to X$ gibt. Betrachten wir zunächst endliche Mengen, dann ergibt sich aufgrund von Satz 7.10 auf S. 180 ja die Bedingung $|X|^{|A|} = 1$. Das geht nur, wenn $|A| = 0$ ist, und das bedeutet $A = \emptyset$. Die einzige Abbildung in $\mathrm{Abb}(\emptyset, X)$ ist die leere Abbildung, und dies bleibt auch dann richtig, wenn X gänzlich beliebig ist. Somit haben wir die leere Menge als einziges Anfangsobjekt identifiziert.

Endobjekt bedeutet eine Menge E, so dass es zu jeder Menge X genau eine Abbildung $X \to E$ gibt. Wieder beschränken wir uns zunächst auf endliche Mengen und nutzen Satz 7.10. Wir erhalten die Bedingung $|E|^{|X|} = 1$. Das geht nur, wenn $|E| = 1$ ist. Jede Menge mit einem Element ist also Endobjekt, denn allgemein gibt es dann nur eine einzige Abbildung, nämlich die konstante Abbildung. Dies bleibt auch dann richtig, wenn X gänzlich beliebig ist.

Ein Nullobjekt existiert offenbar nicht.

Benutzen wir doch $1 := \{\emptyset\}$ als kanonisches Endobjekt. Interessant ist, dass man zu einer Menge X jedes Element $x \in X$ mit der Abbildung $x\colon 1 \to X$ identifizieren kann, für die $x(\emptyset) = x$ gilt. Zu einer Abbildung $f\colon X \to Y$ können wir die Zuordnung $f(x) = y$ bzw. $(x, y) \in f$ nun in der Form $f \circ x = y$ beschreiben.

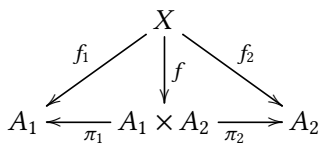
6.1.4. Produkt und Koprodukt

Ein wichtiger Begriff der Theorie ist das Produkt von Objekten. Weil es sich dabei um eine Verallgemeinerung des kartesischen Produktes von Mengen handelt, möchte ich die Zusammenhänge zunächst am vertrauten Schauplatz der Mengen betrachten.

Zu zwei Mengen A_1, A_2 können wir das Produkt $A_1 \times A_2$ bilden. Man definiert die Projektionen auf die Komponenten als

$$\pi_1 \colon A_1 \times A_2 \to A_1, \quad \pi_1((x,y)) \coloneqq x,$$
$$\pi_2 \colon A_1 \times A_2 \to A_2, \quad \pi_2((x,y)) \coloneqq y.$$

Nun betrachten wir Abbildungen $f_1 \colon X \to A_1$ und $f_2 \colon X \to A_2$. Zunächst sei $X \coloneqq \{\emptyset\}$. Die jeweilige Abbildung pickt dann ein Element aus der jeweiligen Menge heraus, das sind $a_1 \coloneqq f_1(\emptyset)$ und $a_2 \coloneqq f_2(\emptyset)$. Nun ist eine Abbildung f gesucht, sodass das Diagramm



kommutiert. Die Bedingungen an f sind also $\pi_i \circ f = f_i$ für $i \in \{1,2\}$. Damit ist aber eindeutig festgelegt, dass $f(\emptyset) = (a_1, a_2)$ sein muss, denn ein Tupel ist durch die Komponenten festgelegt, und die sind $\pi_i(f(\emptyset)) = f_i(\emptyset) = a_i$ für $i \in \{1,2\}$. Somit ist f eindeutig bestimmt.

Die Betrachtung kann man genauso für eine allgemeine Menge X führen, weil die Argumentation dann jeweils für jedes Element von X gilt. Wieder ist f eindeutig bestimmt.

Gelegentlich wird f als $f = f_1 \times f_2$ notiert.

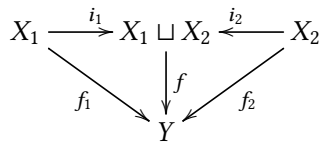
Definition 6.6 (Produkt).
Sei $\mathbf{C}$ eine Kategorie und seien Y_1, Y_2 Objekte von $\mathbf{C}$. Ein Objekt Y von $\mathbf{C}$ mit Projektionen $\pi_1 \colon Y \to Y_1$ und $\pi_2 \colon Y \to Y_2$ heißt *Produkt*, wenn zu jedem Objekt X von $\mathbf{C}$ und allen Morphismen $f_1 \colon X \to Y_1$ und $f_2 \colon X \to Y_2$ genau ein Morphismus $f \colon X \to Y$ existiert, so dass $f_1 = \pi_1 \circ f$ und $f_2 = \pi_2 \circ f$.

Das kartesische Produkt $Y \coloneqq Y_1 \times Y_2$ ist ein Produkt in der Kategorie der Mengen. Wir schreiben $f(x) = y$ mit $y = (y_1, y_2)$. Nun ist $\pi_1(y) = y_1$ und $\pi_2(y) = y_2$, laut Forderung soll also $y_1 = f_1(x)$ und $y_2 = f_2(x)$ sein. Dadurch ist f mit $f(x) \coloneqq (f_1(x), f_2(x))$ eindeutig festgelegt.

Zu zwei Mengen können wir weiterhin die disjunkte Vereinigung $X_1 \sqcup X_2$ bilden. Wir rekapitulieren, dass zu ihr die beiden kanonischen Injektionen

$$i_1 \colon X_1 \to X_1 \sqcup X_2, \quad i_1(x) := (1, x),$$
$$i_2 \colon X_2 \to X_1 \sqcup X_2, \quad i_2(x) := (2, x)$$

gehören. Man stellt sich nun die Frage, was das Wesensmerkmal der disjunkten Vereinigung ist. Das ist doch, dass zu jedem ihrer Elemente die Information vorliegt, ob es aus der linken oder der rechten Menge entstammt. Das heißt, es muss eine Abbildung geben, die auf den Tag projiziert. Betrachten wir dazu die Abbildungen $f_1 \colon X_1 \to Y$ und $f_2 \colon X_2 \to Y$ mit $Y := \{1, 2\}$ und $f_k(x) := k$. Mit der Abbildung f_k gelangt man von X_k also direkt zum Tag k. Nun ist eine Abbildung f gesucht, so dass das Diagramm



kommutiert. Das heißt, es soll $f \circ i_k = f_k$ für $k \in \{1, 2\}$ sein. Das macht $f((1, x)) = 1$ und $f((2, x)) = 2$. Dadurch ist f eindeutig bestimmt. Es ist die gesuchte Projektion auf den Tag.

Definition 6.7 (Koprodukt).
Sei $\mathbf{C}$ eine Kategorie und seien X_1, X_2 Objekte von $\mathbf{C}$. Ein Objekt X von $\mathbf{C}$ mit Morphismen $i_1 \colon X_1 \to X$ und $i_2 \colon X_2 \to X$ heißt *Koprodukt*, wenn zu jedem Objekt Y von $\mathbf{C}$ und allen Morphismen $f_1 \colon X_1 \to Y$ und $f_2 \colon X_2 \to Y$ genau ein Morphismus $f \colon X \to Y$ existiert, so dass $f_1 = f \circ i_1$ und $f_2 = f \circ i_2$.

Die disjunkte Vereinigung $X := X_1 \sqcup X_2$ mit den Injektionen $i_1(x) := (1, x)$ und $i_2(x) := (2, x)$ ist ein Koprodukt in der Kategorie der Mengen. Es gilt schon mal

$$f(x) = \textbf{match } x \begin{cases} (1, x) \mapsto y_1, \\ (2, x) \mapsto y_2. \end{cases}$$

Laut Forderung soll außerdem $y_1 = f_1(x)$ und $y_2 = f_2(x)$ sein, wodurch f eindeutig festgelegt ist.

6.1.5. Exponentialobjekte

Die Notation B^A stehe für die Menge der Abbildungen $A \to B$. Es soll nun die Applikation einer Abbildung auf ein Argument als eigenständige Operation

$$\varepsilon \colon B^A \times A \to B, \quad \varepsilon(f, a) := f(a)$$

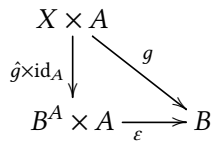
gedacht werden. Einer zweistelligen Abbildung $g\colon X \times A \to B$ lässt sich die Abbildung

$$\hat{g}\colon X \to B^A, \quad \hat{g}(x)(a) := g(x,a).$$

zuordnen. Diesen Vorgang nennen wir Currying. Man findet nun

$$g(x,a) = \hat{g}(x)(a) = \varepsilon(\hat{g}(x),a) = (\varepsilon \circ (\hat{g} \times \mathrm{id}_A))(x,a).$$

Die Gleichung $g = \varepsilon \circ (\hat{g} \times \mathrm{id}_A)$ ist also für jede Abbildung g erfüllt, was bedeutet, dass das Diagramm



kommutiert.

Definition 6.8 (Exponentialobjekt).
Sei **C** eine Kagegorie, in der das Produkt je zweier Objekte existiert. Zu zwei Objekten A, B von **C** heißt ein Objekt B^A von **C** zusammen mit einem Morphismus $\varepsilon\colon B^A \times A \to B$ *Exponentialobjekt*, wenn es zu jedem Objekt X von **C** und Morphismus $g\colon X \times A \to B$ genau einen Morphismus $\hat{g}\colon X \to B^A$ gibt, so dass $\varepsilon \circ (\hat{g} \times \mathrm{id}_A) = g$ gilt.

Satz 6.8. Es besteht die Isomorphie $\mathrm{Hom}_{\mathbf{C}}(X \times A, B) \cong \mathrm{Hom}_{\mathbf{C}}(X, B^A)$.

Beweis. Es sei $\lambda(g) := \hat{g}$ bezüglich Def. 6.8. Zu zeigen ist, dass es sich bei λ um einen Isomorphismus handelt. Für jedes $h\colon X \to B^A$ sei dazu

$$\lambda'(h) := \varepsilon \circ (h \times \mathrm{id}_A)\colon X \times A \to B.$$

Zu zeigen ist, dass λ' der inverse Morphismus zu λ ist. Def. 6.8 sichert nun direkt zu, dass $\lambda'(\lambda(g)) = g$ gelten muss. Zu bestätigen verbleibt $\lambda(\lambda'(h)) = h$. Die Allaussage in Def. 6.8 wird hierzu spezialisiert mit $g := \varepsilon \circ (h \times \mathrm{id}_A)$. Nun wissen wir aber nicht nur, dass $\hat{g}$ existieren muss, wir können es mit der Setzung $\hat{g} := h$ angeben, denn dieses erfüllt die Gleichung

$$\varepsilon \circ (\hat{g} \times \mathrm{id}_A) = \varepsilon \circ (h \times \mathrm{id}_A).$$

Infolge gilt

$$\lambda(\lambda'(h)) = \lambda(\varepsilon \circ (h \times \mathrm{id}_A)) = \lambda(\varepsilon \circ (\hat{g} \times \mathrm{id}_A)) = \lambda(g) = \hat{g} = h. \qquad \square$$

6.2. Beweise als Terme

6.2.1. Kartesisch abgeschlossene Kategorien

Definition 6.9 (Kartesisch abgeschlossene Kategorie).
Eine Kategorie $\mathbf{C}$ heißt *kartesisch abgeschlossen*, wenn

1. sie ein Terminalobjekt enthält,
2. je zwei Objekte A, B von $\mathbf{C}$ ein Produkt $A \times B$ in $\mathbf{C}$ besitzen,
3. je zwei Objekte A, B von $\mathbf{C}$ ein Exponential B^A in $\mathbf{C}$ besitzen.

Dass die Kategorie der Mengen kartesisch abgeschlossen ist, wurde bereits während der Diskussion der drei Begrifflichkeiten nachgerechnet. Die Kategorie der endlichen Mengen ist ebenfalls kartesisch abgeschlossen. Sind A, B endlich, ist ja $A \times B$ und B^A ebenfalls eine endliche Menge.

6.2.2. Die BHK-Interpretation

Im Bestreben, die Struktur der Schlussregeln zu ergründen, tritt mit der Zeit das trübe Muster zutage, dass der Aufbau von Beweisen gleichermaßen abläuft, wie der von Elementen bestimmter geläufiger Mengenkonstruktionen. Ein Beweis der Konjunktion $A \wedge B$ ist zum Beispiel erbracht, wenn ein Beweis a der Aussage A und ein Beweis b der Aussage B gefunden wurde. Dies gilt natürlich nur unter der Voraussetzung, dass der Beweis der Konjunktion durch ihre Einführungsregel zustande gekommen ist. Das heißt, wenn überhaupt, kann es sich nur um die Normalform der Beweise handeln. Wird nun die Evidenz für $A \wedge B$ vorgelegt, muss aus dieser sowohl a als auch b extrahiert werden können. Der Beweis von $A \wedge B$ ist demnach das Paar (a, b). Man erkennt nun, dass sich die Einführungsregel

$$\frac{\vdash A \qquad \vdash B}{\vdash A \wedge B} \quad \text{analog zu} \quad \frac{a \in [A] \qquad b \in [B]}{(a, b) \in [A] \times [B]}$$

verhält. Hierbei bezeichne $[A]$ die Menge der Beweise der Aussage A, wobei die Gleichheit $[A \wedge B] = [A] \times [B]$ gefordert wird.

Bei der Subjunktion tritt nun die Analogie zwischen

$$\frac{\vdash A \Rightarrow B \qquad \vdash A}{\vdash B} \quad \text{und} \quad \frac{f \in \mathrm{Abb}([A], [B]) \qquad a \in [A]}{f(a) \in [B]}$$

in Erscheinung, wobei $[A \Rightarrow B] = \mathrm{Abb}([A], [B])$ gefordert wird.

Ein Beweis der Disjunktion $A \vee B$ gilt bereits als erbracht, wenn ein Beweis zumindest einer der beiden Aussagen A, B gefunden wurde. Wird nun die Evidenz von $A \vee B$ zur Prüfung vorgelegt, muss man erfahren dürfen, zu welcher der beiden

Aussagen der Beweis gefunden wurde. Diese Unterscheidung schafft die disjunkte Vereinigung $[A \vee B] = [A] \sqcup [B]$. Bei der Disjunktion besteht die Analogie also zwischen

$$\frac{\vdash A}{\vdash A \vee B}, \quad \frac{\vdash B}{\vdash A \vee B} \quad \text{und} \quad \frac{a \in [A]}{(1, a) \in [A] \sqcup [B]}, \quad \frac{b \in [B]}{(2, b) \in [A] \sqcup [B]}.$$

Solange dies die einzigen Regeln zur Einführung der Disjunktion bleiben, wird eine Aussage auch nur dann als wahr angesehen werden dürfen, wenn ein Beweis zumindest einer der beiden Aussagen A, B vorgelegt wurde. Infolge stellt sich der Satz vom ausgeschlossenen Dritten unzulässig heraus, die Logik somit als intuitionistisch.

Die gerade geschilderten Überlegungen manifestieren sich in der *BHK-Interpretation der intuitionistischen Logik*, benannt nach Luitzen Brouwer, Arend Heyting und Andrei Kolmogorow.

Ein gern herangezogenes Beispiel, das erfahrbar macht, wie es sich mit der intuitionistischen Logik verhält, bietet die anschließende Überlegung zu Rechtssystemen. Betrachten wir den folgenden Fall. Aus dem Kühlschrank wurde ein Kuchen entwendet. Außerdem liegen belastbare Beweismittel vor, die belegen, dass Alice oder Bob dafür verantwortlich ist. Allerdings fanden sich weder Indizien, die eine der beiden Personen belasten, noch fand sich ein Alibi, das eine der beiden entlasten würde. Somit bleibt für beide die Unschuldsvermutung erhalten. Würde man nun von den beiden je einen halben Kuchen als Schadensersatz fordern, käme dies einer Kollektivstrafe gleich. Wird das Konzept von Kollektivstrafen verworfen, führt dies insofern dazu, dass dem Rechtssystem die intuitionistische Logik innewohnen muss.

Zur Berücksichtigung von Hypothesen wird ein Urteil der Form $H \vdash A$ als $\vdash H \Rightarrow A$ gedacht. Zufolge dieser Überlegung nimmt die Regel

$$\frac{H \vdash A \qquad H \vdash B}{H \vdash A \wedge B} \quad \text{die Form} \quad \frac{a \colon [H] \to [A] \qquad b \colon [H] \to [B]}{(x \mapsto (a(x), b(x))) \colon [H] \to [A] \times [B]}$$

an. Mit der Einbeziehung von Hypothesen wird schließlich die Subjunktionseinführung fassbar. Zu

$$\frac{H, A \vdash B}{H \vdash A \Rightarrow B} \quad \text{findet sich} \quad \frac{f \colon [H] \times [A] \to [B]}{(x \mapsto a \mapsto f(x, a)) \colon [H] \to [A] \to [B]}$$

Mit der Notation $x \mapsto a \mapsto f(x, a)$ ist die Funktion g gemeint, für die gilt

$$g(x)(a) = f(x, a).$$

Der Funktionswert von g an der Stelle x soll also wiederum eine Funktion sein, dessen Funktionswert an der Stelle a der Wert $f(x,a)$ ist. Der Übergang von f zu g wird *Schönfinkeln*, im Englischen *currying*, genannt, benannt nach Moses Schönfinkel und Haskell Brooks Curry. Das logische Gegenstück lässt sich auch als das Theoremschema

$$(H \wedge A \Rightarrow B) \Rightarrow (H \Rightarrow (A \Rightarrow B)),$$

darstellen, das Schema der *Exportation*.

6.2.3. Sequenzen als Morphismenklassen

Bislang waren Abbildungen zwischen Mengen Gegenstand der Betrachtung. Uns hält aber nichts davon ab, diese als Morphismen zwischen Objekten zu sehen. Dieser Abstraktionsschritt soll nun als nächstes vorgenommen werden.

Mit der Idee, das Urteil $A \vdash B$ als Vorhandensein eines Morphismus $A \longrightarrow B$ zu interpretieren, gelangt man zur kategoriellen Semantik. Dabei werden nur solche Sequenzen betrachtet, die eine einzige Vorderformel besitzen, was aber keine wesentliche Einschränkung darstellt, da die Vorderformeln immer zu einer Konjunktion zusammengefasst werden können. Ist die Ansammlung der Vorderformeln leer, erhält man als leere Konjunktion die tautologische Formel. Das heißt, ein Urteil der Form $\vdash A$ wird als $\top \vdash A$ betrachtet.

Jede Aussage wird als ein Objekt interpretiert, und Urteile sind Morphismen zwischen den Aussagen. Die Konjunktion $A \wedge B$ wird interpretiert als Produkt $A \times B$, die Disjunktion $A \vee B$ als Koprodukt $A \sqcup B$. Die kontradiktorische Formel $\bot$ wird als Anfangsobjekt 0 interpretiert, die tautologische Formel $\top$ als ein Terminalobjekt 1. Bei der Subjunktion $A \Rightarrow B$ muss man nun ein wenig achtsam sein. Ihr entspricht nicht etwa die Morphismenklasse $A \longrightarrow B$, sondern wie bei allen Formeln ein Objekt, das Exponentialobjekt B^A. Die Negation $\neg A$ wird als $A \Rightarrow \bot$ betrachtet, also durch 0^A interpretiert.

Der Überlegung nach darf das Urteil $\vdash A$ genau dann gefällt werden, wenn ein Morphismus $1 \longrightarrow A$ existiert. Dieser Morphismus wählt ein Element aus A aus, womit A bewohnt sein muss. Allgemein stellen wir Aussagen gemäß Tabelle 6.1 als Objekte dar.

> **Satz 6.9.** Bildet die Übersetzung aus Tabelle 6.1 in eine bikartesisch abgeschlossene Kategorie ab, so gilt für die intuitionistische Logik die Folgerung
>
> $$(\Gamma \vdash A) \;\Rightarrow\; (\exists f \colon [\Gamma] \longrightarrow [A]).$$

Tabelle 6.1.: Übersetzung von Aussagen in Objekte

$[A \wedge B] := [A] \times [B]$	$[\bot] := 0$
$[A \vee B] := [A] \sqcup [B]$	$[\top] := 1$
$[A \Rightarrow B] := [B]^{[A]}$	$[A \Leftrightarrow B] := [B]^{[A]} \times [A]^{[B]}$
$[\neg A] := 0^{[A]}$	$[\{A_1, \ldots, A_n\}] := [A_1] \times \ldots \times [A_n]$

Beweis. Es wird eine strukturelle Induktion über den Aufbau von Beweisen durchgeführt. Zum Induktionsanfang. Die Regel

$$\frac{}{A \vdash A} \quad \text{fordert} \quad \frac{}{\exists f \colon [A] \to [A]}.$$

Dies zeigt sich mit der Setzung $f := \mathrm{id}$. Die Abschwächungsregel

$$\frac{\Gamma \vdash B}{\Gamma, A \vdash B} \quad \text{fordert} \quad \frac{\exists f \colon [\Gamma] \to [B]}{\exists g \colon [\Gamma] \times [A] \to [B]}.$$

Mit der Prämisse, die Induktionsvoraussetzung ist, muss man zur Konklusion gelangen, was kurzum mit der Setzung $g := f \circ \pi_1$ erreicht wird, wobei π_1 gemäß der universellen Eigenschaft des Produktes zur Verfügung steht.

Die Konjunktionseinführung

$$\frac{\Gamma \vdash A \qquad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad \text{fordert} \quad \frac{\exists f_1 \colon [\Gamma] \to [A] \qquad \exists f_2 \colon [\Gamma] \to [B]}{\exists f \colon [\Gamma] \to [A] \times [B]}.$$

Man legt $f := (f_1, f_2)$ vor, was gemäß der universellen Eigenschaft des Produktes gebildet werden darf. Die Beseitigungsregel

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \text{fordert} \quad \frac{\exists f \colon [\Gamma] \to [A] \times [B]}{\exists g \colon [\Gamma] \to [A]}.$$

Man setzt $g := \pi_1 \circ f$, wobei π_1 gemäß der universellen Eigenschaft des Produktes zur Verfügung steht.

Die Subjunktionseinführung

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \to B} \quad \text{fordert} \quad \frac{\exists f \colon [\Gamma] \times [A] \to [B]}{\exists g \colon [\Gamma] \to [B]^{[A]}}.$$

Man hat $g := \lambda f$, wobei λ der angegebene Isomorphismus zu Satz 6.8 ist. Die Beseitigungsregel

$$\frac{\Gamma \vdash A \Rightarrow B \qquad \Gamma \vdash A}{\Gamma \vdash B} \quad \text{fordert} \quad \frac{\exists f \colon [\Gamma] \to [B]^{[A]} \qquad \exists g \colon [\Gamma] \to [A]}{\exists h \colon [\Gamma] \to [B]}.$$

Es ist $\lambda^{-1}f\colon [\Gamma] \times [A] \to [B]$. In der Kategorie der Mengen ginge die Setzung $h(x) := (\lambda^{-1}f)(x, g(x))$. Das gesuchte Abstraktum hierzu ist $h := (\lambda^{-1}f) \circ (\mathrm{id}, g)$.

Da die Negation $\neg A$ als Subjunktion $A \Rightarrow \bot$ definiert wird, brauchen wir diese nicht explizit zu betrachten. Dasselbe gilt für die Bijunktion $A \Leftrightarrow B$, die als $(A \Rightarrow B) \wedge (B \Rightarrow A)$ definiert wird.

Zuletzt verbleibt noch ex falso quodlibet $\bot \vdash A$ als ein weiterer Induktionsanfang zu bestätigen. Dazu muss ein Morphismus $f\colon 0 \to [A]$ existent sein, was aber gerade die Eigenschaft des Anfangsobjektes 0 darstellt. □

Weiterhin stellt sich heraus, dass die Schnittregel der schlichten Komposition von Morphismen entspricht. Genauer wird die Übersetzung des Schlusses

$$\frac{A \vdash B \qquad B \vdash C}{A \vdash C} \text{ gesichert durch } \frac{f\colon A \to B \qquad g\colon B \to C}{g \circ f\colon A \to C}.$$

Die übersetzte allgemeine Form der Schnittregel braucht nicht mehr bestätigt werden, dies ist bereits als abschließender Schritt in der Bestätigung der übersetzten Subjunktionsbeseitigung zu finden.

Die eigentliche BHK-Interpretation soll die Struktur von Beweisen der intuitionistischen Logik erklären. Dafür sollte man aber am besten ausgehend von einem Morphismus zum zugehörigen Beweis zurückgelangen können. Insbesondere sollte die Umkehrung von Satz 6.9 ebenfalls gelten. Es mag allerdings sein, dass zwischen den Objekten Morphismen bestehen, die keinem Beweis entsprechen. Betrachten wir Objekte wieder speziell als Mengen, kann dies ausgeschlossen werden, indem als Menge der erreichbaren Konstrukte, heißt formulierbaren Terme, die kleinste induktive Menge gewählt wird, die unter den verfügbaren Konstruktionsschritten abgeschlossen ist. Weil die Schritte zu verschiedenartigen Konstrukten führen, erhält man zudem eine frei erzeugte induktive Menge. Man kann nun per struktureller Rekursion über den Term eine Rückübersetzung zum ursprünglichen Beweis definieren. Die kleinste induktive Menge verhält sich demnach isomorph zur ursprünglichen Menge der formulierbaren Beweise.

Was ist dann aber gewonnen, wenn Beweise durch die Übersetzung lediglich in eine andere Gestalt gebracht werden? Zum einen ermöglicht die Übersetzung offenbar, Beweise mit Konzepten der Kategorientheorie in Verbindung zu bringen. Zum anderen wurden in der gerade gemachten Überlegung die Beweise zusätzlich in die Übersetzung miteinbezogen. Das Beschriebene System zur Konstruktion von Termen stellt sich später als der einfach getypte Lambda-Kalkül heraus, dem in der Informatik eine grundlegende Bedeutung zukommt. Der Lambda-Kalkül verhält sich so ähnlich wie ein Termersetzungssystem und formalisiert im Wesentlichen die Konstruktion und Applikation von Funktionen.

Von der BHK-Interpretation ausgehend gelangt man somit zum *Curry-Howard-Isomorphismus*, der die Übersetzung von Beweisen in die entsprechenden Terme des einfach getypten Lambda-Kalküls beschreibt. Aussagen werden hierbei in Typen übersetzt.

7. Diskrete Mathematik

7.1. Kombinatorik

7.1.1. Endliche Summen

Problemstellungen der Kombinatorik führen oftmals zu Summen. Ein Grund dafür liegt darin, dass die Lösungen bestimmter Differenzengleichungen als Summen darstellbar sind. In der Analysis sind die Reihen sehr bedeutsam, das sind Summen von unendlich vielen Summanden. Der Wert einer Reihe wird hierbei, sofern existent, als der Grenzwert der Folge der Partialsummen erklärt. Somit bauen Reihen konzeptuell auf den endlichen Summen auf. Ein weiteres wichtiges Werkzeug stellen die formalen Potenzreihen bereit, die ohne den Grenzwertbegriff auskommen, aber analoge Rechenregeln zu den eigentlichen Potenzreihen aufweisen. In der linearen Algebra dienen Summen ebenfalls zur allgemeinen Formulierung von Rechenregeln. Darin einbegriffen ist das Rechnen mit Polynomen. Auch in der Stochastik sind Summen dienlich, zum Beispiel zur Erklärung des Erwartungswertes einer Zufallsgröße.

Zur Einführung eine kleine Aufgabe. Man beobachtet

$$1+3=2^2,\quad 1+3+5=3^2,\quad 1+3+5+7=4^2,\quad 1+3+5+7+9=5^2.$$

Anscheinend besteht die Gesetzmäßigkeit

$$1+3+\ldots+(2n-1)=n^2.$$

In Worten ist die Summe der ersten ungeraden Zahlen eine Quadratzahl, genauer die Anzahl der Summanden zum Quadrat. Ein strenger Beweis dafür findet sich unschwer. Zuvor will ich aber die allgemeinen Gesetzmäßigkeiten für das Rechnen mit Summen diskutieren.

Statt Summen mit drei Pünktchen zu beschreiben, verwendet man meist die unmissverständliche Notation

$$\sum_{k=1}^{n} a_k := a_1 + a_2 + \ldots + a_n.$$

Beispielsweise gilt

$$\sum_{k=1}^{3}(2k-1) = (2\cdot 0-1)+(2\cdot 1-1)+(2\cdot 2-1) = 1+3+5.$$

Die drei Pünktchen treten allerdings noch in der Erklärung der Summe auf. Wir werden sie wieder gänzlich los, indem die Summe rekursiv festgelegt wird.

Definition 7.1 (Summe).
Für eine Folge $a\colon \mathbb{Z} \to \mathbb{R}$ ist die *Summe* der a_k für $k = m$ bis $k = n$ rekursiv definiert gemäß

$$\sum_{k=m}^{m-1} a_k := 0, \qquad \sum_{k=m}^{n} a_k := a_n + \sum_{k=m}^{n-1} a_k.$$

Es darf k, genannt der *Laufindex*, als gebundene Variable nur im Term a_k auftreten. Die Zahlen m, n nennt man die *Grenzen* der Summierung. Ist a lediglich auf $D \subseteq \mathbb{Z}$ definiert, können wir a schlicht per $a_k := 0$ für $k \notin D$ auf ganz $\mathbb{Z}$ erweitern. Dies bietet sich an, wenn die Summe stets definiert sein soll, unabhängig davon wie die Grenzen gewählt werden.

Satz 7.1. Es gilt $\sum_{k=m}^{n} ca_k = c \sum_{k=m}^{n} a_k$ für jede Konstante c.

Beweis. Induktion über n. Im Induktionsanfang $n = m - 1$ gilt

$$\sum_{k=m}^{m-1} ca_k \overset{\text{def}}{=} 0 = c \cdot 0 \overset{\text{def}}{=} c \sum_{k=m}^{m-1} a_k.$$

Der Induktionsschritt ist

$$\sum_{k=m}^{n} ca_k \overset{\text{def}}{=} ca_n + \sum_{k=m}^{n-1} ca_k \overset{\text{IV}}{=} ca_n + c \sum_{k=m}^{n-1} a_k = c\Big(a_n + \sum_{k=m}^{n-1} a_k\Big) \overset{\text{def}}{=} c \sum_{k=m}^{n} a_k. \qquad \square$$

Satz 7.2. Es gilt $\sum_{k=m}^{n}(a_k + b_k) = \sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k$.

Beweis. Induktion über n. Im Induktionsanfang $n = m - 1$ reduziert sich die Behauptung zu $0 = 0 + 0$. Der Induktionsschritt ist

$$\begin{aligned}\textstyle\sum_{k=m}^{n}(a_k + b_k) &\overset{\text{def}}{=} a_n + b_n + \textstyle\sum_{k=m}^{n-1}(a_k + b_k) \overset{\text{IV}}{=} a_n + b_n + \sum_{k=m}^{n-1} a_k + \sum_{k=m}^{n-1} b_k \\ &\overset{\text{def}}{=} \textstyle\sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k. \qquad \square\end{aligned}$$

Man kann den Raum $\mathrm{Abb}(\mathbb{Z}, \mathbb{R})$ auch als einen Vektorraum über dem Körper $\mathbb{R}$ deuten, wobei die Multiplikation mit einem Skalar $c \in \mathbb{R}$ und die vektorielle Addition punktweise definiert werden, das heißt,

$$(ca)_k := ca_k, \qquad (a + b)_k := a_k + b_k.$$

Dann ist $\sum_m^n a := \sum_{k=m}^n a_k$ eine lineare Abbildung, denn die Homogenität gilt laut Satz 7.1, und die Additivität laut Satz 7.2.

Der Differenzoperator

$$(\Delta a)_n := a_{n+1} - a_n$$

wird die nachfolgende Diskussion übersichtlicher halten. Es handelt sich ebenfalls um eine lineare Abbildung, das heißt, für jede Konstante c und alle Folgen a, b gilt $\Delta(ca) = c\Delta a$ und $\Delta(a + b) = \Delta a + \Delta b$, denn

$$\begin{aligned}(\Delta(ca))_n &= ca_{n+1} - ca_n = c(a_{n+1} - a_n) = c(\Delta a)_n,\\ (\Delta(a+b))_n &= (a_{n+1} + b_{n+1}) - (a_n + b_n) = a_{n+1} - a_n + b_{n+1} - b_n\\ &= (\Delta a)_n + (\Delta b)_n.\end{aligned}$$

Satz 7.3 (Teleskopsumme). Es gilt $\sum_{k=m}^{n-1}(\Delta a)_k = a_n - a_m$.

Beweis. Induktion über n. Im Induktionsanfang $n = m$ sieht man trivial, dass beide Seiten der Gleichung den Wert null haben. Der Induktionsschritt ist

$$\sum_{k=m}^{n}(\Delta a)_k \overset{\text{def}}{=} (\Delta a)_n + \sum_{k=m}^{n-1}(\Delta a)_k \overset{\text{IV}}{=} a_{n+1} - a_n + a_n - a_m = a_{n+1} - a_m. \qquad \square$$

Das Prinzip der Teleskopsummen bietet zwei wesentliche Vorteile. Der erste Vorteil besteht darin, dass man damit bequem nach Summenformeln fischen kann, wobei man deren Beweis mit geschenkt bekommt. Für $s_n := (n-1)/n$ gilt zum Beispiel

$$(\Delta s)_n = \frac{n}{n+1} - \frac{n-1}{n} = \frac{n^2 - (n+1)(n-1)}{(n+1)n} = \frac{1}{(n+1)n}.$$

Ergo findet sich die Summenformel

$$\sum_{k=1}^{n} \frac{1}{(k+1)k} = \frac{n}{n+1}.$$

Der zweite Vorteil besteht darin, dass sich der Beweis einer bereits ermittelten Summenformel dahingehend vereinfacht, dass nicht jedes mal ein Induktionsbeweis geführt werden muss, womit die unter Umständen schwierige Auffindung der wesentlichen Umformungen im Induktionsschritt erleichtert wird. Zum Beweis einer Summenformel der Form

$$\textstyle\sum_{k=1}^{n} a_k = s_{n+1} - s_1$$

genügt es nämlich, $\Delta s = a$ zu prüfen. Zur Summenformel

$$\sum_{k=1}^{n}(2k-1) = n^2$$

gilt beispielsweise $a_n = 2n-1$ und $s_n = (n-1)^2$. Es bestätigt sich kurzerhand

$$(\Delta s)_n = n^2 - (n-1)^2 = n^2 - n^2 + 2n - 1 = 2n - 1 = a_n.$$

Aus der nun bewiesenen Formel ziehen sich schnell weitere Schlüsse. Setzt man in die Gleichung

$$n^2 = \sum_{k=1}^{n}(2k-1) = 2\sum_{k=1}^{n} k - \sum_{k=1}^{n} 1$$

die Vereinfachung $\sum_{k=1}^{n} 1 = n$ ein und stellt sie daraufhin nach der verbleibenden Summe um, findet sich die Summenformel

$$\sum_{k=1}^{n} k = \frac{1}{2}(n^2+n) = \frac{n}{2}(n+1).$$

Die Werte dieser Summe werden die *Dreieckszahlen* genannt.[1]

Satz 7.4 (Aufteilung von Summen).
Für $m \le p \le n$ gilt $\sum_{k=m}^{n} a_k = \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n} a_k$.

Beweis. Induktion über n. Der Induktionsanfang liege in $n = p$. Die Gleichung ist in diesem erfüllt, weil $\sum_{k=p}^{p} a_k = a_p$ gilt. Der Induktionsschritt ist

$$\sum_{k=m}^{n} a_k \overset{\text{def}}{=} a_n + \sum_{k=m}^{n-1} a_k \overset{\text{IV}}{=} a_n + \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n-1} a_k \overset{\text{def}}{=} \sum_{k=m}^{p-1} a_k + \sum_{k=p}^{n} a_k. \square$$

Satz 7.5 (Indexverschiebung).
Für die Indexverschiebung der Distanz $d \in \mathbb{Z}$, kurz Indexshift, gilt $\sum_{k=m}^{n} a_k = \sum_{k=m+d}^{n+d} a_{k-d}$.

Beweis. Induktion über n. Im Induktionsanfang $n = m-1$ wird unmittelbar ersichtlich, dass beide Seiten der Gleichung im Wert null resultieren. Der Induktionsschritt gilt gemäß der Umformung

$$\sum_{k=m}^{n} a_k \overset{\text{def}}{=} a_n + \sum_{k=m}^{n-1} a_k \overset{\text{IV}}{=} a_{(n+d)-d} + \sum_{k=m+d}^{n+d-1} a_{k-d} \overset{\text{def}}{=} \sum_{k=m+d}^{n+d} a_{k-d}. \square$$

[1] Engl. *triangular numbers*. Eintrag A000217 in OEIS.

Ich will zusätzlich noch eine alternative Herleitung aufzeigen. Bei dieser wird der Indexbereich als Ungleichung beschrieben, was den Vorteil bietet, dass die Ungleichung einer Umformung unterzogen werden kann. Mit der Substitution $k = k' - d$ gelingt dahingehend die Überlegung

$$\sum_{k=m}^{n} a_k = \sum_{m \le k \le n} a_k = \sum_{m \le k'-d \le n} a_{k'-d} = \sum_{m+d \le k' \le n+d} a_{k'-d} = \sum_{k'=m+d}^{n+d} a_{k'-d}.$$

Satz 7.6 (Umkehrung der Reihenfolge).
Es gilt $\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} a_{n-k}$.

Beweis. Der Induktionsanfang bei $n = 0$ ist trivial. Beim Induktionsschritt macht man sich Satz 7.5 (Indexshift) und Satz 7.4 (Aufteilung) zunutze,

$$\sum_{k=0}^{n} a_{n-k} = a_{n-n} + \sum_{k=0}^{n-1} a_{n-k} = a_0 + \sum_{k=0}^{n-1} a_{n-(n-1-k)} = a_0 + \sum_{k=0}^{n-1} a_{k+1}$$
$$\overset{[k:=k-1]}{=} a_0 + \sum_{k=1}^{n} a_k = \sum_{k=0}^{0} a_k + \sum_{k=1}^{n} a_k = \sum_{k=0}^{n} a_k. \qquad \square$$

Dienlich ist die Umkehrung beispielsweise bei der folgenden klassischen Herleitung der Summenformel $\sum_{k=0}^{n} k = \frac{n}{2}(n+1)$. Aufbauend auf den bereits gezeigten Regeln stellt diese Herleitung ebenfalls einen strengen Beweis dar. Das doppelte der Summe betrachtet man hierbei als Addition der Summe zu sich selbst, wobei die Reihenfolge der zweiten Summe umgekehrt wird. Man erhält

$$2\sum_{k=0}^{n} k = \sum_{k=0}^{n} k + \sum_{k=0}^{n} k = \sum_{k=0}^{n} k + \sum_{k=0}^{n} (n-k)$$
$$= \sum_{k=0}^{n} (k + n - k) = \sum_{k=0}^{n} n = n \sum_{k=0}^{n} 1 = n(n+1).$$

Ganz allgemein dürfen die Summanden einer endlichen Summe beliebig umgeordnet werden; dies geht aus wiederholter Anwendung des Kommutativgesetzes und des Assoziativgesetzes hervor. Die Formalisierung dieser Tatsache geschieht vermittels einer frei wählbaren Permutation des Indexbereichs. Der zuvor diskutierte Sachverhalt stellt sich dahingehend als Spezialfall mit der auf der Menge $\{0, \ldots, n\}$ definierten Permutation $\pi(k) = n - k$ dar.

Satz 7.7 (Permutation der Reihenfolge).
Sei $M = \{k \in \mathbb{Z} \mid m \leq k \leq n\}$. Für jede Permutation $\pi\colon M \to M$ gilt

$$\sum_{k=m}^{n} a_k = \sum_{k=m}^{n} a_{\pi(k)}.$$

Beweis. Induktion über n. Sei ohne Beschränkung der Allgemeinheit $m = 1$. Im Induktionsanfang $n = 0$ und $n = 1$ ist die Gleichung offenkundig erfüllt.

Induktionsschritt. Induktionsvoraussetzung sei die Gültigkeit für M. Zu zeigen ist die Gültigkeit für $M \cup \{n+1\}$.

Sei t ein fester Parameter mit $1 \leq t \leq n+1$. Im Fall $\pi(t) = n+1$ geht man wie folgt vor. Man setze $\sigma(k) := \pi(k)$ für $1 \leq k \leq t-1$. Man setze $\sigma(k) := \pi(k+1)$ für $t \leq k \leq n$. Weil $n+1$ kein Wert von σ ist, muss σ eine Permutation $\sigma\colon M \to M$ sein. Ergo gilt

$$\begin{aligned}\sum_{k=1}^{n+1} a_{\pi(k)} &= \sum_{k=1}^{t-1} a_{\pi(k)} + a_{\pi(t)} + \sum_{k=t+1}^{n+1} a_{\pi(k)} = a_{\pi(t)} + \sum_{k=1}^{t-1} a_{\pi(k)} + \sum_{k=t}^{n} a_{\pi(k+1)} \\ &= a_{n+1} + \sum_{k=1}^{t-1} a_{\sigma(k)} + \sum_{k=t}^{n} a_{\sigma(k)} = a_{n+1} + \sum_{k=1}^{n} a_{\sigma(k)} \\ &\overset{\text{IV}}{=} a_{n+1} + \sum_{k=1}^{n} a_k = \sum_{k=1}^{n+1} a_k.\end{aligned}$$

Man beachte, dass in den beiden Randfällen $t = 1$ und $t = n+1$ die jeweilige Randsumme den Wert null hat und somit verschwindet. □

Dass die Reihenfolge der Summierung keine Rolle spielt, gibt Anlass zu einer abstrakten Fassung der Summierung selbst, die von vornherein auf die Festlegung einer bestimmten Reihenfolge verzichtet.

Satz 7.8. Sei M eine endliche Menge, sei $a\colon M \to \mathbb{R}$. Die Festlegung

$$\sum_{k \in M} a_k := \sum_{i=m}^{n} a_{f(i)}, \quad |M| = n - m + 1, \quad f\colon \{m, \ldots, n\} \to M \text{ bijektiv}$$

ist wohldefiniert, wobei f frei gewählt werden darf.

Beweis. Es gilt zu zeigen, dass die Summe auch wirklich unabhängig von der gewählten Bijektion ist. Sind f, g zwei solcher Bijektionen, so ist $\pi := g^{-1} \circ f$ ebenfalls

bijektiv, also eine Permutation des Indexbereichs. Vermittels Satz 7.7 gilt daher

$$\sum_{i=m}^{n} a_{f(i)} = \sum_{i=m}^{n} a_{g(\pi(i))} = \sum_{i=m}^{n} a_{g(i)}. \qquad \square$$

Im Weiteren ist man nun befähigt, die Begrenzung des Indexbereichs durch Ungleichungen streng zu definieren. Man legt fest

$$\sum_{m \le k \le n} a_k := \sum_{k \in M} a_k, \quad M := \{k \in \mathbb{Z} \mid m \le k \wedge k \le n\}.$$

Somit ist streng begründet, dass die Ungleichungen $m \le k$ und $k \le n$ einer Äquivalenzumformung unterzogen werden dürfen, denn von diesen bleibt die Menge M, und somit die Summe unberührt.

Ein letztes Detail darf aber nicht vergessen werden. Im Vorfeld der oder Anschluss an die Umformung der Ungleichungen wird ggf. eine Substitution der Laufvariable durchgeführt. Streng genommen muss noch abgeklärt werden, dass dies unter allen Umständen unverfänglich ist. Gewissheit schafft der

Satz 7.9 (Substitutionsregel). Ist $\varphi\colon M' \to M$ eine Bijektion, gilt

$$\sum_{k \in M} a_k = \sum_{k' \in M'} a_{\varphi(k')}.$$

Beweis. Zur Bijektion $f\colon \{1, \dots, |M|\} \to M$ existiert die Bijektion g mit $f = \varphi \circ g$, nämlich ist dies $g := \varphi^{-1} \circ f$. Infolge gilt

$$\sum_{k \in M} a_k = \sum_{i=1}^{|M|} a_{f(i)} = \sum_{i=1}^{|M|} a_{\varphi(g(i))} = \sum_{k' \in M'} a_{\varphi(k')}. \qquad \square$$

Für die allgemeine Form der Summierung gelten anloge Regeln. Genauer lassen sich die bereits bewiesenen Regeln direkt übertragen. Man prüft unschwer nach, dass gilt

$$\sum_{k \in M} c a_k = c \sum_{k \in M} a_k, \qquad \sum_{k \in M} (a_k + b_k) = \sum_{k \in M} a_k + \sum_{k \in M} b_k.$$

Zur Aufteilung sei $M = A \cup B$ mit $A \cap B = \emptyset$, also eine disjunkte Zerlegung von M in A, B. Dann gilt

$$\sum_{k \in M} a_k = \sum_{k \in A} a_k + \sum_{k \in B} a_k.$$

Bezüglich der Indikatorfunktion $1_A\colon M \to \{0,1\}$ zu $A \subseteq M$ gilt nun

$$\sum_{k\in M} 1_A(k)a_k = \sum_{k\in A} a_k,$$

denn $M = A \cup (M \setminus A)$ ist eine disjunkte Zerlegung, so dass gilt

$$\sum_{k\in M} 1_A(k)a_k = \sum_{k\in A} \underbrace{1_A(k)}_{=1} a_k + \sum_{k\in M\setminus A} \underbrace{1_A(k)}_{=0} a_k = \sum_{k\in A} a_k.$$

Ist $A \subseteq \mathbb{Z}$ endlich, dürfen wir daher definieren

$$\sum_{k} 1_A(k)a_k = \sum_{k\in\mathbb{Z}} 1_A(k)a_k := \sum_{k\in A} a_k,$$

da diese Summe in Wirklichkeit endlich ist, also niemals über unendlich viele Indizes läuft. Diese Sichtweise bietet unter Umständen Vorteile bei der Umformung von Summen.

7.1.2. Endliche Produkte

Endliche Produkte weisen analoge Rechenregeln zu den endlichen Summen auf, kommen allerdings weit weniger häufig vor als diese. Infolgedessen verlaufen die Beweise ebenso analog, weshalb ich darauf verichten will, diese weitläufig aufzuführen. Verbreitet sind endliche Produkte von Zahlen vor allem in der Kombinatorik, zum Beispiel in Form von Faktoriellen und Binomialkoeffizienten.

Für eine Folge $a\colon \mathbb{Z} \to \mathbb{R}$ definiert man rekursiv

$$\prod_{k=m}^{m-1} a_k := 1, \qquad \prod_{k=m}^{n} a_k := a_n \prod_{k=m}^{n-1} a_k.$$

Sei $c \in \mathbb{Z}_{\geq 0}$ eine Konstante. Es gilt

$$\prod_{k=m}^{n} a_k^c = \Big(\prod_{k=m}^{n} a_k\Big)^c, \qquad \prod_{k=m}^{n} (a_k b_k) = \Big(\prod_{k=m}^{n} a_k\Big)\Big(\prod_{k=m}^{n} b_k\Big).$$

Die Aufteilung von Produkten geht vonstatten gemäß

$$\prod_{k=m}^{n} a_k = \Big(\prod_{k=m}^{p-1} a_k\Big)\Big(\prod_{k=p}^{n} a_k\Big), \quad m \leq p \leq n+1.$$

Für jede Konstante $c \in \mathbb{R}$ gilt

$$\prod_{k=m}^{n} c = c^{n+1-m}, \qquad \prod_{k=m}^{n} (ca_k) = c^{n+1-m} \prod_{k=m}^{n} a_k.$$

Bei manchen Umformungen hilft das Teleskopprodukt

$$\prod_{k=m}^{n} \frac{a_{k+1}}{a_k} = \frac{a_{n+1}}{a_m}, \quad \text{sofern } \forall k \in \{m, \ldots, n\} \colon a_k \neq 0.$$

7.1.3. Anzahl der Abbildungen

Abstrakt gesehen thematisiert die abzählende Kombinatorik die Frage nach der Anzahl der Elemente endlicher Mengen. Die Mengenlehre hilft bei der Formalisierung, und infolgedessen beim strengen Beweis kombinatorischer Gesetzmäßigkeiten.

Zur Einführung stellt man häufig die Frage, wie viele mögliche Tupel eine Produktmenge enthält. Üblicherweise formuliert man diese unkomplizierte Aufgabe aber nicht in der Gestalt der Mengenlehre, sondern fragt, wie viele Zahlen eine Binärsequenz aus k Bits enthält, oder wie viele Zeichenketten der Länge k mit 26 Buchstaben formulierbar sind. Abstrakt gesehen geht es bei den Binärsequenzen um die Anzahl der Elemente von $\{0, 1\}^k$. Analog geht es bei den Zeichenketten um die Anzahl der Elemente von

$$\{\texttt{'a'}, \texttt{'b'}, \ldots, \texttt{'z'}\}^k.$$

Die Elemente sind jeweils Tupel der Länge k. Bezüglich B^k will ich $|B|$ so wie bei Stellenwertsystemen als die *Basis* bezeichnen. Dass $|A \times B| = |A| \cdot |B|$ für beliebige Mengen gilt, wurde bereits aufgezeigt. Demnach gilt

$$|B^k| = |B|^k = n^k, \quad n := |B|.$$

Mit den $n = 2$ Binärwerten sind dies 256 Binärsequenzen der Länge $k = 8$. Oder mit $n = 26$ Buchstaben sind es 676 Zeichenketten der Länge $k = 2$, und schon 17576 Zeichenketten der Länge $k = 3$.

In der Vergangenheit empfahlen Dienste häufig oder machten es obligatorisch, Passwörter dadurch sicherer zu machen, indem neben Buchstaben auch noch Ziffern und Sonderzeichen eingegeben werden sollten. Man empfahl also, die Basis n zu vergrößern, um die Länge gering zu halten. Nimmt man an, dass so ein Passwort rein zufällig ausgewürfelt wird, ist es bei bereits hinreichend großer Basis aber wesentlich sinnvoller, abermals den Exponenten k zu erhöhen. Bei Hinzunahme der

10 Ziffern zu den 26 Buchstaben bekommt man mit $n = 36$ für $k = 3$ eine Zahl von 46656 Zeichenketten. Unterscheidet man zudem zwischen kleinen und großen Buchstaben, sind es immerhin 238328. Mit $n = 26$ und $k = 4$ sind es aber bereits 456976.

In der professionellen Kryptografie nennt man n^k die *Schlüsselraumgröße*, die bei modernen symmetrischen Kryptosystemen mindestens 2^{128} beträgt. Und der Schlüssel wird am besten rein zufällig gewählt. Zur Kodierung in der Basis n muss man die Gleichung $2^{128} = n^k$ nach k umformen, das macht

$$k = 128\,\frac{\lg(2)}{\lg(n)}.$$

Das sind acht Blöcke je vier Hexadezimalziffern. Oder 22 Zeichen zur Basis 62. Es ist davon auszugehen, dass dies ebenso sehr ein sicheres Passwort darstellt.

Zur Erinnerung, ein Tupel aus B^k darf auch als eine endliche Folge interpretiert werden, das ist eine Abbildung $\{0, \ldots, k-1\} \to B$. Dahingehend ergibt sich die abstrakte Fassung

Satz 7.10 (Anzahl der Abbildungen).
Seien X, Y endliche Mengen mit $|X| = k$ und $|Y| = n$. Die Menge der Abbildungen $X \to Y$ enthält n^k Elemente.

Beweis. Induktion über k. Im Anfang $k = 0$ ist $X = \emptyset$. Es gibt genau eine Abbildung $\emptyset \to Y$, nämlich die leere Abbildung. Gleichermaßen ist $n^0 = 1$.

Zum Induktionsschritt. Induktionsvoraussetzung sei die Gültigkeit für $k - 1$. Es sei $|X| = k$ und $|Y| = n$. Gesucht ist die Anzahl der Möglichkeiten zur Festlegung der Abbildung $f\colon X \to Y$. Sei $x \in X$ fest. Für die Festlegung $f(x) = y$ bestehen nun genau n Möglichkeiten, nämlich so viele, wie es Elemente $y \in Y$ gibt. Die Festlegung der übrigen Werte geschieht gemäß der Einschränkung von f auf $X \setminus \{x\}$, und die Zahl dieser Abbildungen beträgt n^{k-1} gemäß der Induktionsvoraussetzung. Man hat also n mal n^{k-1} Möglichkeiten, das sind n^k. □

Gleichwohl die dargelegte Beweisführung recht genau erscheint, bezieht sie sich im Induktionsschritt an der wesentlichen Stelle auf rationale Überlegungen, die allerdings nicht aus bereits formal bewiesenen Regeln der Mengenlehre entstammen. Zur Schließung der Lücke muss man sich daran erinnern, dass

$$|Y^{A \cup B}| = |Y^A \times Y^B| = |Y^A| \cdot |Y^B|, \text{ sofern } A \cap B = \emptyset.$$

Zusammen mit $|Y^{\{x\}}| = n$ wird der Induktionsschritt damit präzisiert zu

$$|Y^X| = |Y^{(X \setminus \{x\}) \cup \{x\}}| = |Y^{X \setminus \{x\}}| \cdot |Y^{\{x\}}| \overset{\text{IV}}{=} n^{k-1} \cdot n = n^k.$$

7.1.4. Anzahl der Injektionen

Bei den Passwörtern durfte jedes Zeichen beliebig oft vorkommen. Stellt man die Aufgabe so, dass jedes Zeichen nur ein einziges Mal vorkommen darf, verringert sich die Anzahl der Möglichkeiten. Zum Beispiel wird gefragt nach der Zahl der Ketten, die aus zwei der vier Buchstaben a, b, c, d besteht. Man kann dazu ein Baumdiagramm aufstellen. Für den ersten Buchstaben bestehen vier Möglichkeiten, woraufhin für den zweiten Buchstaben jeweils noch drei Möglichkeiten übrig bleiben. Das macht $4 \cdot 3 = 12$ Möglichkeiten. Die Berechnung verläuft gemäß der

Definition 7.2 (Fallende Faktorielle).
Für $n \in \mathbb{Z}$ und $k \in \mathbb{Z}_{\geq 0}$ legt man rekursiv fest

$$n^{\underline{0}} := 1, \qquad n^{\underline{k}} := n \cdot (n-1)^{\underline{k-1}}.$$

Zu $n = 4$ Buchstaben und $k = 2$ Stellen ergibt sich

$$n^{\underline{k}} = 4^{\underline{2}} = 4 \cdot 3^{\underline{1}} = 4 \cdot 3 \cdot 2^{\underline{0}} = 4 \cdot 3 \cdot 1 = 12.$$

Per Induktion über k bestätigt man die Produktformel

$$n^{\underline{k}} = \prod_{i=0}^{k-1} (n-i) = \prod_{i=1}^{k} (n+1-i) = n \cdot (n-1) \cdot \ldots \cdot (n+1-k).$$

Der Sachverhalt kann nun wieder abstrakt gefasst werden.

Satz 7.11 (Anzahl der Injektionen).
Seien X, Y endliche Mengen, wobei $|X| = k$ und $|Y| = n$ gelte. Die Menge der Injektionen $X \to Y$ enthält $n^{\underline{k}}$ Elemente.

Beweis. Induktion über k. Im Anfang $k = 0$ ist $X = \emptyset$. Es gibt genau eine Injektion $\emptyset \to Y$, nämlich die leere Abbildung. Gleichermaßen gilt $n^{\underline{0}} = 1$.

Zum Schritt. Voraussetzung sei die Gültigkeit für $k-1$. Es sei $|X| = k$ und $|Y| = n$. Gesucht ist die Anzahl der Möglichkeiten zur Festlegung der Injektion $f\colon X \to Y$. Sei $x \in X$ fest. Für die Festlegung $f(x) = y$ bestehen genau n Möglichkeiten, nämlich so viele, wie es Elemente $y \in Y$ gibt. Bei der Festlegung der übrigen entfällt y aufgrund der Injektivität von f. Für die Festlegung betrachtet man f daher als Injektion

$$f\colon X \setminus \{x\} \to Y \setminus \{y\},$$

von denen es laut Voraussetzung $(n-1)^{\underline{k-1}}$ gibt. Es sind also n mal $(n-1)^{\underline{k-1}}$ Möglichkeiten, was definitionsgemäß gleich $n^{\underline{k}}$ ist. □

Im Spezialfall $n = k$ muss müssen die Abbildungen außerdem surjektiv, und damit bijektiv sein. Deren Anzahl wird gezählt durch die spezielle fallende Faktorielle $n^{\underline{n}}$, die man *Fakultät* nennt und $n!$ notiert. Spezialisiert man die Definition der fallenden Faktoriellen dahingehend, ergibt sich die

Definition 7.3 (Fakultät).
Für $n \in \mathbb{Z}_{\geq 0}$ legt man rekursiv fest

$$0! := 1, \qquad n! := n \cdot (n-1)!.$$

Speziell im Fall $X = Y$ sind die Bijektionen die Permutationen auf der Menge X. Somit ist $n!$ auch die Anzahl der Permutationen von n verschiedenen Objekten. In Alltagssprache formuliert lassen sich n unterschiedliche Gegenstände in $n!$ Reihenfolgen bringen.

7.1.5. Anzahl der Teilmengen

Aber wie verhält es sich, wenn die Reihenfolge keine Rolle spielt? Dazu sei eine Menge Y von $n = |Y|$ Objekten gegeben, aus der k Objekte ausgewählt werden. Diese Auswahl nennt man eine *Kombination*. Sie ist als Ansammlung unterschiedlicher Objekte, bei der es nicht auf die Reihenfolge ankommt, aber wesensmäßig nichts anderes als eine Menge. Gesucht ist demzufolge die Anzahl der Teilmengen von Y, die genau k Elemente enthalten. Die Lösung dieser Aufgabe wird berechnet durch den *Binomialkoeffizient* $\binom{n}{k}$, der zunächst einmal erklärt werden muss.

Definition 7.4 (Binomialkoeffizent).
Für $n \in \mathbb{Z}$ und $k \in \mathbb{Z}_{\geq 0}$ setzt man

$$\binom{n}{k} := \frac{1}{k!} n^{\underline{k}}.$$

Stellt man den Binomialkoeffizent als Funktion von (n, k) tabellarisch dar, erkennt man früher oder später, dass dieser einer einfachen Rekurrenz genügt. Nämlich liegt, insofern n die Zeilen und k die Spalten durchläuft, die Summe zweier Einträge unterhalb des zweiten Eintrags; die graphische Darstellung dieses Sachverhaltes wird *pascalsches Dreieck* genannt. Das heißt, es gilt der

Satz 7.12. Für $n \in \mathbb{Z}$ und $k \geq 1$ gilt

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Beweis. Es findet sich die Umformung

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \frac{(n-1)^{\underline{k-1}}}{(k-1)!} + \frac{(n-1)^{\underline{k}}}{k!} = \frac{k(n-1)^{\underline{k-1}}}{k!} + \frac{(n-1)^{\underline{k-1}}(n-k)}{k!}$$
$$= \frac{(n-1)^{\underline{k-1}}}{k!}(k+n-k) = \frac{n(n-1)^{\underline{k-1}}}{k!} = \frac{n^{\underline{k}}}{k!} = \binom{n}{k}. \square$$

Setzt man die beiden Anfangsbedingungen $\binom{n}{0} = 1$ und $\binom{n}{n} = 1$ zusätzlich voraus, wird $\binom{n}{k}$ für $0 \leq k \leq n$ durch die Rekurrenz charakterisiert. Genau genommen muss dafür zudem abgesichert werden, dass die Rekursion zu jedem Argument (n, k) durch die Anfangsbedingungen fundiert ist, was aber leicht ersichtlich wird. Da n bei der Rekurrenz in jedem Argument um eins verringert auftritt, wird aufgrund $0 \leq k \leq n$ irgendwann $k = 0$ oder $k = n$ sein müssen, weil sich der Abstand zwischen 0 und n immer weiter vermindert.

Satz 7.13 (Anzahl der Teilmengen).
Sei Y eine $|Y| = n$ Elemente enthaltende endliche Menge und $C_k(Y)$ die Menge der k-elementigen Teilmengen von Y. Es gilt $|C_k(Y)| = \binom{n}{k}$.

Beweis. Induktion über (n, k). Im Anfang ist $k = 0$ oder $k = n$. Der abstruse Fall $k = 0$ sucht nach Teilmengen ohne Elemente. Es existiert genau eine solche Menge, nämlich die leere Menge, womit $C_0(Y) = 1$ ist. Der Fall $k = n$ sucht nach Teilmengen, die so viele Elemente haben wie Y. Dies kann nur Y selbst sein, womit $C_n(Y) = 1$ gilt. Gleichermaßen gilt $\binom{n}{0} = 1$ und $\binom{n}{n} = 1$.

Induktionsvoraussetzung sei die Gültigkeit für $(n-1, k-1)$ und $(n-1, k)$. Man nimmt nun ein Element y aus Y heraus, womit darin $n-1$ verbleiben. Entscheidet man sich, y zur Teilmenge hinzuzufügen, verbleiben noch $k-1$ Elemente auszuwählen. Entscheidet man sich dagegen, verbleibt die Teilmenge unverändert, womit nach wie vor k Elemente auszuwählen sind. Die Anzahl der Möglichkeiten ist somit

$$|C_k(Y)| = |C_{k-1}(Y \setminus \{y\})| + |C_k(Y \setminus \{y\})| \overset{\text{IV}}{=} \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}. \square$$

Kombinatorische Aufgaben tendieren dazu, sich einer bildlichen Anschauung zu entziehen. Um dem ein wenig entgegenzuwirken, will ich den folgenden Sachverhalt aufführen, der die Binomialkoeffizienten mit monotonen Gitterwegen in Beziehung setzt. Als Beispiel listet Abb. 7.1 jeden der 10 möglichen Wege von Knoten $(0, 0)$ nach Knoten $(3, 2)$ auf.

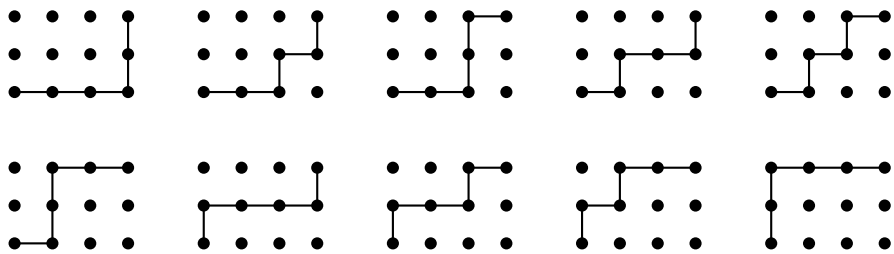


Abbildung 7.1.: Die 10 monotonen Gitterwege von $(0,0)$ nach $(3,2)$.

Satz 7.14. Ein Weg auf dem Gitter $\mathbb{Z} \times \mathbb{Z}$ heiße *monoton*, wenn von (x,y) aus lediglich der Schritt nach $(x+1,y)$ oder der Schritt nach $(x,y+1)$ gewährt ist. Die Anzahl der monotonen Gitterwege von $(0,0)$ nach (x,y) beträgt

$$\frac{(x+y)!}{x!y!} = \binom{x+y}{x} = \binom{x+y}{y}.$$

Beweis. Es bezeichne $f(x,y)$ die Anzahl der monotonen Wege von $(0,0)$ nach (x,y). Zum Erreichen eines Punktes auf den Achsen besteht immer nur ein möglicher Weg, womit $f(x,0) = 1$ und $f(0,y) = 1$ gelten muss. Ein nicht auf den Achsen befindlicher Punkt (x,y) kann von $(x-1,y)$ oder $(x,y-1)$ aus erreicht werden, was zur Rekurrenz

$$f(x,y) = f(x-1,y) + f(x,y-1)$$

führt. Die Tabellierung der Werte macht ersichtlich, dass sie ein gedrehtes pascalsches Dreieck erzeugt. Wir setzen daher $C(x+y,x) := f(x,y)$ und führen die Koordinatentransformation $x+y = n$ und $x = k$ aus. Die Rekurrenz nimmt damit die Form

$$C(n,k) = C(n-1,k-1) + C(n-1,k), \quad C(n,0) = 1, \quad C(n,n) = 1$$

an, die aber eindeutig den Binomialkoeffizienten charakterisiert. □

Ein analoger Sachverhalt besteht für Gitter in höheren Dimensionen, was zum Begriff des *Multinomialkoeffizienten* führt.

Man kann auch eine gruppentheoretische Sichtweise auf die Kombinationen einnehmen. Die Teilmengen von Y mit k Elementen sind wie Listen, die jedes der k Elemente genau einmal enthalten. Der Unterschied zu diesen Listen besteht genau

darin, dass es bei den Teilmengen nicht auf die Reihenfolge der Elemente ankommt. Um die Anzahl der Teilmengen zu zählen, müssen demzufolge Listen, die sich lediglich durch ihre Reihenfolge unterscheiden, als äquivalent angesehen werden. Ihre Äquivalenzklassen stellen somit quasi Listen dar, die die Reihenfolge ihrer Elemente vergessen haben.

Wir verwenden eine beliebige Menge X mit $|X| = k$ als Indexmenge, beispielsweise direkt $X := \{0, \dots, k-1\}$. Je zwei Listen sind damit gegeben durch Injektionen $f, g\colon X \to Y$. Die beiden Listen werden als äquivalent angesehen, sofern sie sich lediglich durch eine Permutation π ihrer Indizes unterscheiden, also

$$f \sim g :\Leftrightarrow \exists \pi \in S_k\colon f = g \circ \pi.$$

Es sei ferner $\mathrm{Inj}(X, Y)$ die Menge der Injektionen $X \to Y$ und $\mathrm{Inj}(X, Y)/S_k$ die Quotientenmenge bezüglich der Äquivalenzrelation bzw. der Gruppe S_k.

Satz 7.15. Es sei $C_k(Y)$ die Menge der k-elementigen Teilmengen von Y. Zwischen $\mathrm{Inj}(X, Y)/S_k$ und $C_k(Y)$ besteht eine kanonische Bijektion.

Beweis. Wir definieren diese Bijektion als

$$\varphi\colon \mathrm{Inj}(X, Y)/S_k \to C_k(Y), \quad \varphi([f]) := f(X),$$

wobei $[f] = f \circ S_k$ die Äquivalenzklasse des Repräsentanten f bezeichne. Die Abbildung φ ist wohldefiniert, denn für je zwei f, g mit $f \sim g$ ist eine Permutation π vorhanden, so dass gilt

$$f(X) = (g \circ \pi)(X) = g(\pi(X)) = g(X).$$

Zum Nachweis der Injektivität von φ muss $[f] = [g]$ aus $f(X) = g(X)$ gewonnen werden. Gesucht ist also eine Permutation π mit $f = g \circ \pi$. Weil g injektiv ist, existiert eine Linksinverse g^{-1}, so dass $\pi := g^{-1} \circ f$ gewählt werden kann. Es verbleibt die Gleichheit $f = g \circ g^{-1} \circ f$ zu bestätigen. Zwar ist g^{-1} im Allgemeinen keine Rechtsinverse von g, ihre Einschränkung auf $g(X)$ aber schon. Wegen $f(X) = g(X)$ hebt sich $g \circ g^{-1}$ daher auf $f(X)$ weg.

Zur Surjektivität von φ. Hier ist zu zeigen, dass es zu jeder Menge $B \in C_k(Y)$ eine Injektion $f\colon X \to Y$ mit $f(X) = B$ gibt. Weil X und B gleichmächtig sind, existiert eine Bijektion $f_0\colon X \to B$, womit man die gesuchte Injektion mit der Setzung $f(x) := f_0(x)$ erhält. □

Bezüglich $|X| = k$ tut sich nun die Umformung

$$|C_k(Y)| \overset{(1)}{=} |\mathrm{Inj}(X, Y)/S_k| \overset{(2)}{=} \frac{|\mathrm{Inj}(X, Y)|}{|S_k|} = \frac{n^{\underline{k}}}{k!} = \binom{n}{k}$$

Tabelle 7.1.: Der zwölffaltige Weg

	$f \in \mathrm{Abb}(X,Y)$	$f \in \mathrm{Inj}(X,Y)$	$f \in \mathrm{Sur}(X,Y)$
f	n^k	$n^{\underline{k}}$	$n!\left\{{k \atop n}\right\}$
$f \circ S_k$	$\binom{n+k-1}{k}$	$\binom{n}{k}$	$\binom{k-1}{k-n}$
$S_n \circ f$	$\sum_{i=0}^{n} \left\{{k \atop i}\right\}$	$[k \leq n]$	$\left\{{k \atop n}\right\}$
$S_n \circ f \circ S_k$	$p_n(n+k)$	$[k \leq n]$	$p_n(k)$

auf. Der Schritt (1) wurde bereits durch Satz 7.15 abgeklärt. Die Einsicht in (2) erhält man folgendermaßen. Für jede Gruppe G gilt die Bahnformel $|G| = |f \circ G| \cdot |G_f|$. Bei trivialer Fixgruppe G_f gilt $|G_f| = 1$, mithin $|f \circ G| = |G|$. Namentlich bei der symmetrischen Gruppe $G = S_k$ tritt dieser Umstand ein. Aus diesem Grund enthält jede Bahn $f \circ S_k$ gleich viele Elemente, $|S_k|$ an der Zahl. Weil die Bahnen außerdem paarweise disjunkt sind, ergibt sich deshalb die Faktorisierung

$$|\mathrm{Inj}(X,Y)| = |S_k| \cdot |\mathrm{Inj}(X,Y)/S_k|.$$

Der gemachte Gedankengang kann auch unter anderen Umständen vorgenommen werden, dergestalt dass statt den Injektionen die Surjektionen oder sämtliche Abbildungen betrachtet werden. Es entsteht der *zwölffaltige Weg*, – wenn man so will, ein kleines Periodensystem der Kombinatorik.

7.2. Zur elementaren Zahlentheorie

7.2.1. Kongruenzen

Die elementare Zahlentheorie beschäftigt sich mit der Teilbarkeit von Zahlen, den Resten, die bei der Ganzzahldivision übrig bleiben und der Auffindung ganzzahliger Lösungen von Gleichungen. Lange Zeit als ein Gebiet der reinen Mathematik angesehen, stellte sie sich als bedeutsam für die praktische Informatik und die Kryptologie heraus.

Die *modulare Arithmetik*, auch *Restklassenarithmetik* genannt, ist ein wichtiges Hilfsmittel der elementaren Zahlentheorie. Sie geht aus der gewöhnlichen Arithmetik hervor, indem zwei ganze Zahlen als gleich angesehen werden, wenn sie bei Division durch eine vorab gewählte feste Zahl denselben Rest lassen. Diese Form der Gleichheit nennt man *Kongruenz*, die feste Zahl den *Modul*. Kongruenz zweier Zahlen ist damit gleichbedeutend, dass die eine Zahl zur anderen um ein Vielfaches des Moduls verschoben liegt, was nichts anderes bedeutet, als dass der Modul die Differenz der beiden Zahlen teilt.

Man vermittelt das Konzept gern am Lauf von Uhrzeigern. Hierbei werden die Stunden 0, 12, 24, 36 usw. als gleich angesehen. Entsprechend werden die Stunden 1, 13, 25, 37 usw. als gleich angesehen. Es verbleiben nur noch 12 unterschiedliche Zeitpunkte, die vollen Stunden $0, 1, \ldots, 11$. Ein Uhrzeiger auf 11 Uhr trifft zwei Stunden später auf 1 Uhr. Anders ausgedrückt sind die Zahlen $11 + 2$ und 1 kongruent Modulo 12, was nach Carl Friedrich Gauß in der Form

$$11 + 2 \equiv 1 \pmod{12}$$

notiert wird.

Definition 7.5 (Kongruenz).
Zwei ganze Zahlen a, b heißen kongruent modulo m, wenn ihre Differenz $a - b$ durch m teilbar ist,

$$a \equiv b \pmod{m} \;:\Leftrightarrow\; \exists k \in \mathbb{Z}\colon a - b = km.$$

Statt »$(\mathrm{mod}\ m)$« schreibt man beim Rechnen meist kürzer »(m)«.

Satz 7.16. Die Kongruenz ist eine Äquivalenzrelation, das heißt, es gilt

$$\begin{aligned}
&a \equiv a \pmod{m}, && \text{(Reflexivität)}\\
&a \equiv b \Rightarrow b \equiv a \pmod{m}, && \text{(Symmetrie)}\\
&a \equiv b \wedge b \equiv c \Rightarrow a \equiv c \pmod{m}. && \text{(Transitivität)}
\end{aligned}$$

Beweis. Für die Reflexivität ist ein k mit $0 = a - a = km$ zu finden. Setze $k = 0$.

Bei der Symmetrie gibt es nach Voraussetzung ein k mit $a - b = km$. Dann ist $b - a = -km$. Setze $k' = -k$. Es gibt also k' mit $b - a = k'm$, somit gilt $b \equiv a$.

Bei der Transitivität gibt es nach Voraussetzung ein k mit $a - b = km$ und ein l mit $b - c = lm$. Das heißt, es gilt $km + lm = a - c$, ergo $a - c = (k + l)m$. Setze $k' = k + l$. Es gibt also k' mit $a - c = k'm$. Somit gilt $a \equiv c$. □

Satz 7.17. Sind a, b, c ganze Zahlen, dann gilt

$$\begin{aligned} a \equiv b \pmod{m} &\iff a + c \equiv b + c \pmod{m}, \\ a \equiv b \pmod{m} &\iff a - c \equiv b - c \pmod{m}. \end{aligned}$$

Beweis. Unter Beachtung von $(b + c) - (a + c) = b - a$ findet man

$$\begin{aligned} a \equiv b \pmod{m} &\iff (\exists k \in \mathbb{Z}\colon b - a = km) \\ &\iff (\exists k \in \mathbb{Z}\colon (b + c) - (a + c) = km) \\ &\iff a + c \equiv b + c \pmod{m}. \end{aligned}$$

Für die Subtraktion von c ist die Überlegung analog. □

Satz 7.18. Sind a, b, c ganze Zahlen, dann gilt

$$a \equiv b \pmod{m} \implies ac \equiv bc \pmod{m}.$$

Beweis. Unter der Voraussetzung $a \equiv b \pmod{m}$ gibt es ein k mit $b - a = km$. Es gilt

$$b - a = km \iff (b - a)c = kcm \iff bc - ac = k'm$$

mit $k' := kc$. Man hat also

$$(\exists k' \in \mathbb{Z}\colon bc - ac = k'm) \iff ac \equiv bc \pmod{m}. \quad \square$$

Satz 7.19. Gilt $a \equiv a' \pmod{m}$ und $b \equiv b' \pmod{m}$, dann gilt auch

$$\begin{aligned} a + b &\equiv a' + b' \pmod{m}, \\ a - b &\equiv a' - b' \pmod{m}, \\ ab &\equiv a'b' \pmod{m}. \end{aligned}$$

Beweis. Man findet

$$\left.\begin{aligned} a \equiv a' &\implies a + b \equiv a' + b \\ b \equiv b' &\implies a' + b \equiv a' + b' \end{aligned}\right\} \implies a + b \equiv a' + b \equiv a' + b' \pmod{m}.$$

Für die Subtraktion ist die Überlegung analog. Für die Multiplikation ebenfalls:

$$\left.\begin{aligned} a \equiv a' &\implies ab \equiv a'b \\ b \equiv b' &\implies a'b \equiv a'b' \end{aligned}\right\} \implies ab \equiv a'b \equiv a'b' \pmod{m}. \quad \square$$

Satz 7.20. Addition des Moduls führt auf eine kongruente Zahl:

$$a \equiv a + m \equiv a - m \pmod{m}.$$

Beweis. Es gilt

$$a \equiv a + m \pmod{m} \iff (\exists k \in \mathbb{Z}\colon km = (a + m) - a = m).$$

Setze $k = 1$. Bei

$$a \equiv a - m \pmod{m} \iff (\exists k \in \mathbb{Z}\colon km = (a - m) - a = -m)$$

setze $k = -1$. $\square$

7.2.2. Teilbarkeit

Die Beziehung »m teilt a« ist definiert als

$$m \mid a \ :\iff (\exists k \in \mathbb{Z}\colon a = km) \iff a \equiv 0 \pmod{m}.$$

Während die Kongruenz eine Äquivalenzrelation ist, stellt die Teilbarkeit eine Halbordnung dar, dergestalt dass sie die drei Axiome

$\forall a \in \mathbb{Z}\colon a \mid a,$	(Reflexivität)
$\forall a, b \in \mathbb{Z}\colon a \mid b \wedge b \mid a \Rightarrow a = b,$	(Antisymmetrie)
$\forall a, b, c \in \mathbb{Z}\colon a \mid b \wedge b \mid c \Rightarrow a \mid c.$	(Transitivität)

erfüllt. Teilt a die Zahl b, ist a also in gewisser Weise kleiner als b.

7.2.3. Restklassenringe

Wir könnten nun beginnen, mit der modularen Arithmetik interessante Probleme zu lösen. Zunächst möchte ich aber erläutern, wie die modulare Arithmetik mit dem Restklassenring zusammenhängt. Unter diesem Blickwinkel bekommen wir ein tieferes Verständnis und können Mittel der Ringtheorie und Gruppentheorie anwenden.

Zu einer ganzen Zahl a ist die Restklasse modulo m definiert als

$$[a]_m := \{x \mid x \equiv a \pmod{m}\}.$$

Eine alternative Schreibweise für $[a]_m$ ist $a + m\mathbb{Z}$. Weil die Kongruenz eine Äquivalenzrelation ist, handelt es sich bei den Restklassen um Äquivalenzklassen. Wir betrachten nun die Quotientenmenge

$$\mathbb{Z}/m\mathbb{Z} := \{[a]_m \mid a \in \mathbb{Z}\}.$$

Nun können wir die Addition und Multiplikation von Restklassen definieren.

Satz 7.21. Auf $\mathbb{Z}/m\mathbb{Z}$ sind die beiden Operationen

$$\begin{aligned}[a]_m + [b]_m &:= [a+b]_m,\\ [a]_m \cdot [b]_m &:= [ab]_m\end{aligned}$$

wohldefiniert.

Beweis. Zu zeigen ist, dass $a + b \equiv x + y$ gilt, sofern $a \equiv x$ und $b \equiv y$ ist. Gemäß Satz 7.17 gilt

$$\begin{aligned}a \equiv x &\iff a + b \equiv x + b,\\ b \equiv y &\iff x + b \equiv x + y.\end{aligned}$$

Aus den beiden Prämissen erhalten wir demzufolge $a+b \equiv x+b \equiv x+y$. Die Argumentation zur Multiplikation ist analog, wobei Satz 7.18 zur Anwendung kommt. □

Die Struktur $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ nennt man den *Restklassenring* zum Modul m.

Satz 7.22. Jeder Restklassenring ist ein kommutativer unitärer Ring.

Beweis. Bereits bewiesen wurde, dass die ganzen Zahlen einen kommutativen unitären Ring bilden. Aufgrund der Wohldefiniertheit der Addition und Multiplikation ist Kongruenz modulo m eine Kongruenzrelation. Satz 3.22 zeigt somit die Behauptung. □

Zudem ist die Quotientenabbildung

$$\varphi\colon \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}, \quad \varphi(a) := [a]_m.$$

ein Eins-erhaltender Ringhomomorphismus, wie aus dem Beweis von Satz 3.22 hervorgeht.

7.2.4. Euklidische Division

Satz 7.23 (Lemma zur euklidischen Division).
Zu je zwei ganzen Zahlen a, b mit $b \neq 0$ gibt es zwei eindeutig bestimmte Zahlen q, r mit $0 \leq r < |b|$, so dass $a = bq + r$. Man nennt q den *Quotient* und r den *Rest*.

Beweis der Existenz. Betrachten wir zunächst den Fall, bei dem $a \geq 0$ und $b > 0$ ist. Man kann dann sooft b von a abziehen, bis sich eine Zahl ≥ 0 und $< b$ ergibt. Formal bilden wir die Folge $r_k := a - bk$. Nun muss für irgendein $k \geq 0$ schließlich $0 \leq r_k < b$ sein. Damit ist $q = k$ und $r = r_k$ gefunden.

Sei nun $a < 0$. Wie bereits gezeigt gibt es q, r mit $-a = bq + r$. Für $r = 0$ haben wir dann mit $q' := -q$ und $r' := 0$ einen Quotient und einen Rest. Sei nun $r \neq 0$. Dann gilt

$$a = -bq - r = -(q+1)b + b - r.$$

Mit $q' := -(q+1)$ und $r' := b - r$ gibt es somit auch in diesem Fall einen Quotient und einen Rest. Der Rest erfüllt auch die gewünschte Ungleichung, denn aus $r < b$ ergibt sich $0 < r'$ und aus $0 < r$ ergibt sich $r' < b$.

Sei nun a beliebig und $b < 0$. Dann gibt es q, r mit $a = (-b)q + r$. Setze also $q' := -q$ und $r' := r$. Damit gilt $a = bq' + r'$, womit auch in diesem Fall ein Quotient und ein Rest gefunden ist. □

Beweis der Eindeutigkeit. Das Paar q, r erfülle $a = bq + r$ und q', r' erfülle ebenfalls $a = bq' + r'$. Dann gilt

$$bq + r = bq' + r', \iff b(q - q') = r' - r, \implies |b||q - q'| = |r' - r|.$$

Aus $0 < r < |b|$ und $0 < r' < |b|$ erhält man außerdem $|r' - r| < |b|$. Somit muss $|b||q - q'| < |b|$ sein, also $|q - q'| < 1$. Eine nichtnegative ganze Zahl kann aber nur dann kleiner als eins sein, wenn sie null ist. Damit hat man

$$|q - q'| = 0, \iff q - q' = 0, \iff q = q'.$$

Entsprechend folgt $r = r'$. □

Bei der euklidischen Division $a : b$ ist $(a \bmod b)$ eine geläufige Schreibweise für den Rest. Das Lemma zur euklidischen Division sagt uns, dass jede Restklasse von $\mathbb{Z}/m\mathbb{Z}$ einen kanonischen Repräsentant besitzt. Nämlich besitzt die Restklasse $[a]_m$ den kanonischen Repräsentant $r = (a \bmod m)$, denn $a = mq + r$ bedeutet dass $a - r$ durch m teilbar ist, also

$$a \equiv r \pmod{m}, \quad \text{bzw.} \quad [a]_m = [r]_m.$$

7.2.5. Rundung

Manche Formeln oder Rechnungen verlangen das Runden von Zahlen auf eine ganze Zahl oder auf eine bestimmte Zahl von Nachkommastellen. Im Algorithmus, der Vektorgrafiken als Rastergrafik aus Pixeln rendern soll, muss zum Beispiel zwangsläufig an irgendeiner Stelle gerundet werden. Außerdem ist Runden in vielen händischen Rechnungen allgegenwärtig, sei es in wissenschaftlichen, technischen, gewerblichen oder alltäglichen.

Als nächstes will ich daher erklären, wie Funktionen zum Runden formal definiert werden, und welche Eigenschaften sie besitzen. Als wichtige Grundfunktionen treten dabei die *Abrundungsfunktion*, engl. *floor*, und die *Aufrundungsfunktion*, engl. *ceil*, auf, die vereinzelt auch in der Kombinatorik und in der Zahlentheorie vorkommen.

Definition 7.6 (Floor). Für $x \in \mathbb{R}$ definiert man

$$y = \lfloor x \rfloor \ :\Leftrightarrow\ y \in \mathbb{Z} \wedge 0 \leq x - y < 1.$$

Definition 7.7 (Ceil). Für $x \in \mathbb{R}$ definiert man

$$y = \lceil x \rceil \ :\Leftrightarrow\ y \in \mathbb{Z} \wedge 0 \leq y - x < 1.$$

Satz 7.24. Für jede ganze Zahl k gilt $\lfloor k + x \rfloor = k + \lfloor x \rfloor$.

Beweis. Aufgrund der Prämisse $k \in \mathbb{Z}$ ist $y \in \mathbb{Z}$ äquivalent zu $y - k \in \mathbb{Z}$. Unter dieser Gegebenheit findet sich mit Def. 7.6 die äquivalente Umformung

$$\begin{aligned} y = \lfloor k + x \rfloor &\iff y \in \mathbb{Z} \wedge 0 \leq (k + x) - y < 1 \\ &\iff y - k \in \mathbb{Z} \wedge 0 \leq x - (y - k) < 1 \\ &\iff y - k = \lfloor x \rfloor \iff y = k + \lfloor x \rfloor. \end{aligned}$$

□

Satz 7.25. Für $0 \leq x < 1$ gilt $\lfloor x \rfloor = 0$.

Beweis. Dies folgt unmittelbar aus Def. 7.6. □

Satz 7.26. Es gilt

$$\lfloor x \rfloor = \max\{k \in \mathbb{Z} \mid k \leq x\} = \min\{k \in \mathbb{Z} \mid x < k+1\},$$
$$\lceil x \rceil = \min\{k \in \mathbb{Z} \mid x \leq k\} = \max\{k \in \mathbb{Z} \mid k-1 < x\}.$$

Beweis. Mit bezüglich $y = \lfloor x \rfloor$ entfalteten Def. 7.6, 3.29 lautet die Aussage

$$y \in \mathbb{Z} \wedge 0 \leq x - y < 1 \iff y \in \mathbb{Z} \wedge y \leq x \wedge (\forall k \in \mathbb{Z}\colon k \leq x \Rightarrow k \leq y).$$

Für die Implikation von links nach rechts ist im Wesentlichen

$$k \in \mathbb{Z}, y \in \mathbb{Z}, k \leq x, x < y + 1 \vdash k \leq y$$

zu zeigen. Man erhält zunächst $k < y + 1$ per Transitivgesetz. Wegen $k, y \in \mathbb{Z}$ folgt daraus $k \leq y$. Für die Implikation von rechts nach links ist im Wesentlichen

$$y \in \mathbb{Z}, (\forall k \in \mathbb{Z}\colon k \leq x \Rightarrow k \leq y) \vdash x < y + 1$$

zu zeigen. Angenommen, es wäre $y + 1 \leq x$. Die Allaussage wird mit $k := y + 1$ spezialisiert. Per Modus ponens erhält man den Widerspruch $y + 1 \leq y$. Ergo muss die Annahme falsch sein, was äquivalent zu $x < y + 1$ ist. Die restlichen Beweise verlaufen analog. □

Satz 7.27. Für $x \in \mathbb{R}$ und $n \in \mathbb{Z}_{\geq 1}$ gilt

$$\left\lfloor \frac{x}{n} \right\rfloor = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor.$$

Beweis. Sei $a := x - \lfloor x \rfloor$. Per Def. 7.6 gilt $0 \leq a < 1$. Laut dem Lemma zur euklidischen Division gilt außerdem $\lfloor x \rfloor = qn + r$ mit $q = \lfloor \frac{\lfloor x \rfloor}{n} \rfloor$ und $0 \leq r \leq n - 1$. Infolge gilt $0 \leq a + r < n$, also $0 \leq \frac{a+r}{n} < 1$ und somit $\lfloor \frac{a+r}{n} \rfloor = 0$ laut Satz 7.25. Es findet sich

$$\left\lfloor \frac{x}{n} \right\rfloor = \left\lfloor \frac{a + \lfloor x \rfloor}{n} \right\rfloor = \left\lfloor \frac{a + qn + r}{n} \right\rfloor \overset{(1)}{=} q + \left\lfloor \frac{a+r}{n} \right\rfloor \overset{(2)}{=} q = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor,$$

wobei (1) laut Satz 7.24 gilt, und (2) wie soeben ausgeführt. □

Satz 7.28. Für $x \in \mathbb{R}$ und $m, n \in \mathbb{Z}$ mit $n \geq 1$ gilt

$$\left\lfloor \frac{m+x}{n} \right\rfloor = \left\lfloor \frac{m+\lfloor x \rfloor}{n} \right\rfloor.$$

Beweis. Laut Satz 7.24 ist $m + \lfloor x \rfloor = \lfloor m + x \rfloor$. Die Aussage folgt nun als Korollar aus Satz 7.27. □

Das Runden einer Zahl auf eine ganze Zahl kann vermittels floor als

$$\operatorname{round}(x) := \operatorname{sgn}(x) \lfloor |x| + \tfrac{1}{2} \rfloor$$

beschrieben werden. Das Runden auf die n-te Nachkommastelle daraufhin als

$$\operatorname{round}(x, n) := \frac{\operatorname{round}(10^n x)}{10^n}.$$

8. Elemente der Stochastik

8.1. Grundbegriffe

8.1.1. Ereignisse

Die Wahrscheinlichkeitstheorie beschäftigt sich mit *Zufallsexperimenten*. Darunter versteht man ein Experiment mit zufälligem Ausgang, das, um der Wissenschaftlichkeit genüge zu tun, unter genau definierten Versuchsbedingungen durchgeführt wird. Der Ausgang führt immer zu einem *Ergebnis*. Alle erreichbaren Ergebnisse fasst man zur *Ergebnismenge* zusammen. Allgemeiner genügt es, wenn jedes Ergebnis in der Ergebnismenge liegt, wobei diese aber auch Elemente enthalten darf, die das Experiment niemals abwirft. Jede Teilmenge der Ergebnismenge nennt man ein *Ereignis*. Die Potenzmenge der Ergebnismenge heißt *Ereignisraum*, sie besteht aus allen denkbaren Ereignissen. Man sagt, ein Ereignis sei eingetreten, wenn das Ergebnis des Versuchs im diesem Ereignis liegt.

Zu beachten ist, dass wir dabei eine endliche oder höchstens abzählbar unendliche Ergebnismenge voraussetzen. Bei überabzählbaren Ergebnismengen kommt es zu Unwägbarkeiten, deren Klärung Gegenstand der Maßtheorie ist.

Ein schlichtes Experiment bietet der Wurf des Spielwüfels, eines mit Augenzahlen beschrifteten regelmäßigen Hexaeders. Die Ergebnismenge wird als

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

festgelegt. Betrachten wir die drei Ereignisse

$$A := \{2\}, \quad B := \{1, 2\}, \quad C := \{1, 3\}.$$

Ist $\omega = 2$ das Ergebnis des Versuchs, sind die Ereignisse A, B eingetreten. Zwei Ereignisse, die niemals gleichzeitig eintreten, heißen *disjunkt*. So sind die A, C disjunkt, weil ihre Schnittmenge leer ist.

8.1.2. Wahrscheinlichkeiten

Man kann nicht voraussagen, wie ein Experiment ausgehen wird. Das Wahrscheinlichkeitsmaß liefert dennoch ein Maß dafür, wie sicher der Eintritt eines Ereignis-

ses ist. Wahrscheinlichkeit wird tiefergründig verständlich, wenn dasselbe Zufallsexperiment abermals wiederholt wird. Wir zählen, wie häufig ein Elementarereignis eingetreten ist.

Es sei ein Versuch n mal durchgeführt worden, was zu den Ergebnissen a_i für $i = 1$ bis $i = n$ geführt hat. Wir definieren die *relative Häufigkeit* des Ereignisses A als die Zahl

$$r_{n,a}(A) := \frac{1}{n}|\{i \in \{1, \ldots, n\} \mid a_i \in A\}|.$$

Relative Häufigkeiten bieten bei hinreichend großem n eine Näherung für die Wahrscheinlichkeit. Zur Vermessung eines Würfels wird man diesen also möglichst oft werfen wollen. Man erhält so die relativen Häufigkeiten der Elementarereignisse, und damit näherungsweise auch ihre Wahrscheinlichkeiten. So lässt sich feststellen, ob ein Würfel gezinkt wurde.

Fassen wir a als Funktion $i \mapsto a_i$ auf, können wir schreiben

$$\{i \mid a_i \in A\} = \{i \mid i \in a^{-1}(A)\} = a^{-1}(A).$$

Für disjunkte Ereignisse A, B erhält man nun

$$\begin{aligned} r_{n,a}(A \cup B) &= \tfrac{1}{n}|a^{-1}(A \cup B)| = \tfrac{1}{n}|a^{-1}(A) \cup a^{-1}(B)| \\ &= \tfrac{1}{n}|a^{-1}(A)| + \tfrac{1}{n}|a^{-1}(B)| = r_{n,a}(A) + r_{n,a}(B). \end{aligned}$$

Weil Ω die Zielmenge von a ist, muss $a_i \in \Omega$ für jedes $i \in \Omega$ gelten, also $|a^{-1}(\Omega)| = n$. Somit gilt

$$r_{n,a}(\Omega) = 1,$$

unabhängig davon, wie groß die Zahl n ist. Und weil sich aus der Definition von a unmittelbar $0 \leq |a^{-1}(A)| \leq n$ ergibt, besteht der Sachverhalt

$$0 \leq r_{n,a}(A) \leq 1.$$

Weiterhin ergibt sich, dass sich die Häufigkeit monoton bezüglich der Inklusion verhält, das heißt, es besteht die Implikation

$$A \subseteq B \Rightarrow r_{n,a}(A) \leq r_{n,b}(B).$$

Laut Satz 3.12 folgt nämlich $a^{-1}(A) \subseteq a^{-1}(B)$ aus $A \subseteq B$. Daraufhin ergibt sich $|a^{-1}(A)| \leq |a^{-1}(B)|$.

Die Vorstellung von der Wahrscheinlichkeit besteht darin, dass sich die relative Häufigkeit ihr mit steigender Anzahl der Versuche immer weiter nähert. Dies bedeutet, es wird meistens $P(A) \approx r_{n,a}(A)$ für hinreichend großes n gelten. Die bisher erläuterten Eigenschaften von $r_{r,a}$ sollten diesbezüglich für P bestehen bleiben, da diese nicht von n abhängen.

Untermauert wird diese Überlegung durch das *Gesetz der großen Zahlen*. Es zu diskutieren, würde an dieser Stelle aber zu weit führen. Zum einem müsste dafür zunächst der Begriff des Grenzwertes eingeführt werden. Zum anderen ist der Grenzwertbegriff der Analysis nicht in direkter Weise auf den Sachverhalt anwendbar. Auf dessen Basis müssten zunächst die Konvergenzbegriffe für Zufallsgrößen erörtert werden.

Stattdessen wird der Begriff des Wahrscheinlichkeitsmaßes erst einmal axiomatisch eingeführt. Darauf aufbauend, was unter einer Zufallsgröße zu verstehen sei.

Definition 8.1 (Endlicher Wahrscheinlichkeitsraum).
Sei Ω eine nichtleere endliche Menge. Das Paar (Ω, P) nennt man einen *endlichen Wahrscheinlichkeitsraum*, wenn

$$P\colon \mathcal{P}(\Omega) \to \mathbb{R}, \quad P(A) := \sum_{\omega \in A} P(\{\omega\})$$

die Eigenschaften $\mathrm{Bild}(P) = [0, 1]$ und $P(\Omega) = 1$ besitzt.

Ein endlicher Wahrscheinlichkeitsraum kann alternativ auch durch die kolmogorowschen Axiome charakterisiert werden. Der Vorteil besteht darin, dass diese später auch unter allgemeineren Umständen gültig bleiben.

Satz 8.1 (Axiome von Kolmogorow).
Sei Ω eine nichtleere endliche Menge und $P\colon \mathcal{P}(\Omega) \to \mathbb{R}$. Es ist (Ω, P) genau dann ein endlicher Wahrscheinlichkeitsraum, wenn

$P(A) \geq 0$ für $A \in \Omega$,
$P(\Omega) = 1$,
$P(A \cup B) = P(A) + P(B)$ für $A, B \in \Omega$ mit $A \cap B = \emptyset$.

Beweis. Es sei (Ω, P) ein endlicher Wahrscheinlichkeitsraum. Mit $A \cap B = \emptyset$ gilt für die Indikatorfunktionen die Gleichung

$$1_{A \cup B}(\omega) = 1_A(\omega) + 1_B(\omega).$$

Somit gelingt die Rechnung

$$P(A \cup B) = \sum_{\omega \in A \cup B} P(\{\omega\}) = \sum_{\omega \in \Omega} P(\{\omega\}) 1_{A \cup B}(\omega)$$
$$= \sum_{\omega \in \Omega} P(\{\omega\})(1_A(\omega) + 1_B(\omega)) = \sum_{\omega \in \Omega} P(\{\omega\}) 1_A(\omega) + \sum_{\omega \in \Omega} P(\{\omega\}) 1_B(\omega)$$
$$= \sum_{\omega \in A} P(\{\omega\}) + \sum_{\omega \in B} P(\{\omega\}) = P(A) + P(B).$$

Umgekehrt erfülle (Ω, P) die kolmogorowschen Axiome. Da die elementaren Ereignisse $\{\omega\}$ paarweise disjunkt sind, ergibt sich

$$P(A) = P(\bigcup_{\omega \in A} \{\omega\}) = \sum_{\omega \in A} P(\{\omega\}).$$

Mit dieser Summenformel erhält man die Aussage $P(A) \leq 1$ nun als Korollar aus Satz 8.2 mit $A := A$ und $B := \Omega$. □

Satz 8.2. Gilt $A \subseteq B$, so ist $P(A) \leq P(B)$.

Beweis. Mit $A \subseteq B$ ist $1_A(\omega) \leq 1_B(\omega)$ für jedes $\omega \in \Omega$. Ergo gilt

$$P(A) = \sum_{\omega \in \Omega} P(\{\omega\}) 1_A(\omega) \leq \sum_{\omega \in \Omega} P(\{\omega\}) 1_B(\omega) = P(B). \square$$

Die Vorstellung von der Wahrscheinlichkeit als Chance oder Risiko mag aufgrund ihres Charakters als Unwägbarkeit schwierig fassbar sein. Der frequentistische Wahrscheinlichkeitsbegriff wird hingegen mit seiner Erklärung von Wahrscheinlichkeit als relative Häufigkeit bei großer Zahl der Versuche schon greifbarer. Mir fällt dazu eine Sichtweise ein, die sich dem zufälligen Charakter so weit entzieht wie es nur geht. Lässt man die große Zahl der Versuche parallel ablaufen, gibt die die Wahrscheinlichkeit den *Anteil* der Ergebnisse an, die in der jeweiligen Teilmenge enthalten sind. Werdem zum Beispiel eine Million Spielwürfel auf einmal geworfen, wird jede der Augenzahlen voraussichtlich näherungsweise mit dem Anteil $\frac{1}{6}$ vorkommen. Oder denkt man sich eine Population von unzählig vielen Individuen, und erhält jedes die gleiche Wahrscheinlichkeit $P(A)$, ein bestimmtes Merkmal A zu besitzen, wird man in der Population einen Anteil von näherungsweise $P(A)$ Individuen mit dem Merkmal A vorfinden.

Das Wahrscheinlichkeitsmaß P erfüllt einige Gesetzmäßigkeiten, die elementar aus den Axiomen ableitbar sind.

Satz 8.3. Allgemein gilt $P(A^c) = 1 - P(A)$.

Beweis. Da $\Omega = A \cup A^c$ eine Zerlegung von Ω in zwei disjunkte Mengen ist, erhält man $P(\Omega) = P(A) + P(A^c)$. Die Gleichung $P(\Omega) = 1$ wird daraufhin kurzum nach $P(A^c)$ umgeformt. □

Satz 8.4. Allgemein gilt $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Beweis. Wir erinnern uns daran, dass man für Indikatorfunktionen vermittels einer Wahrheitstafel den Sachverhalt

$$1_{A\cup B}(\omega) = 1_A(\omega) + 1_B(\omega) - 1_{A\cap B}(\omega)$$

bestätigen kann. Hiermit formt man wieder wie im Beweis von Satz 8.1 die Summenformel für $P(A \cup B)$ um. □

Alternativ kann die Identität folgendermaßen direkt aus dem dritten kolmogorowschen Axiom abgeleitet werden, ohne den Weg über die Summenformel zu gehen. Man überzeugt sich mühelos davon, dass A via

$$A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$$

als Vereinigung disjunkter Mengen dargestellt werden kann. Ergo gilt

$$P(A) = P(A \cap B) + P(A \cap B^c).$$

Analog bekommt man

$$A \cup B = (A \cup B) \cap (B \cup B^c) = ((A \cup B) \cap B) \cup ((A \cup B) \cap B^c)$$

als Vereinigung disjunkter Mengen. Der letzte Term vereinfacht sich mit dem Absorptionsgesetz $(A \cup B) \cap B = B$ und $(A \cup B) \cap B^c = A \cap B^c$ via Distributivgesetz, Komplementärgesetz und Neutralitätsgesetz. Demzufolge gilt

$$P(A \cup B) = P(B) + P(A \cap B^c).$$

Löst man die vorherige Gleichung nun nach $P(A \cap B^c)$ auf und setzt sie in diese ein, findet sich schließlich Satz 8.4.

Vermittels Satz 8.4 wird unschwer die Formel

$$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$

gewonnen. Umgekehrt kann Satz 8.4 unmittelbar aus ihr zurückgewonnen werden, indem einfach $C := \emptyset$ gesetzt wird. Diese Formeln folgen einem allgemeinen Prinzip, das allgemein für Vereinigungen endlich vieler Mengen gilt, – man spricht vom *Prinzip der Inklusion und Exklusion.*

8.1.3. Zufallsgrößen

Eine *Zufallsgröße* darf man sich als eine Funktion $X\colon \Omega \to \Omega'$ vorstellen, die eine kausale Verbindung zwischen den Ergebnismengen Ω, Ω' schafft. Ein Ergebnis $\omega \in \Omega$ führt zu $X(\omega)$. Ursächlich für ein $x \in \Omega'$ sind daher all die ω mit $x = X(\omega)$. Das heißt, ursächlich für das Elementarereignis $\{x\}$ ist dessen Urbild $X^{-1}(\{x\})$. Infolge muss die Wahrscheinlichkeit von $\{x\}$ die es Urbildes sein. Insofern definiert man auf Ω' das Wahrscheinlichkeitsmaß

$$P_X\colon \mathcal{P}(\Omega') \to [0,1], \quad P_X(A) := P(X^{-1}(A)).$$

Man nennt P_X die *Verteilung* von X. Mit der identischen Zufallsgröße

$$\mathrm{id}\colon \Omega \to \Omega, \quad \mathrm{id}(\omega) := \omega$$

versteht sich auch das ursprüngliche Maß P als die Verteilung $P = P_{\mathrm{id}}$.

Geläufig sind die Schreibweisen

$$P(X = x) := P(X^{-1}(\{x\})), \qquad \{X = x\} := X^{-1}(\{x\}),$$
$$P(X \in A) := P(X^{-1}(A)), \qquad \{X \in A\} := X^{-1}(A).$$

Es ist $\{X = x\}$ dasselbe wie $\{X \in \{x\}\}$. Ist P die Gleichverteilung auf Ω, ergibt sich

$$P(X \in A) = \frac{|\{X \in A\}|}{|\Omega|}.$$

Standardbeispiel. Wir werfen zwei Spielwürfel. Die Ergebnismenge sei

$$\Omega := \{1,\dots,6\} \times \{1,\dots,6\},$$

und jedes der 36 elementaren Ereignisse sei gleich wahrscheinlich, habe also die Wahrscheinlichkeit $\frac{1}{36}$. Es bezeichne ω_1 das Ergebnis des ersten, und ω_2 das des zweiten Wurfs. Wir betrachten die Zufallsgröße

$$X\colon \Omega \to \{2,\dots,12\}, \quad X(\omega_1,\omega_2) := \omega_1 + \omega_2.$$

Gesucht sei $P(X = 4)$. Man ermittelt

$$\{X = 4\} = \{(1,3),(2,2),(3,1)\}, \quad \text{ergo } P(X = 4) = \tfrac{3}{36}.$$

Allgemein zerfällt ein Ereignis A ja in seine disjunkten Elementarereignisse $\{x\}$, so dass $A = \bigcup_{x\in A}\{x\}$ gilt. Weil nun die Fasern $X^{-1}(\{x\})$ ebenfalls disjunkt sind, muss $P(X \in A)$ die Summe der $P(X = x)$ mit $x \in A$ sein. Das heißt, man rechnet

$$P(X \in A) = P(X^{-1}(\bigcup_{x\in A}\{x\})) = P(\bigcup_{x\in A} X^{-1}(\{x\})) = \sum_{x\in A} P(X = x).$$

Die Verteilung P_X ist demzufolge bereits eindeutig bestimmt, sobald $P(X = x)$ für jedes $x \in \Omega'$ vorliegt. Dies motiviert uns, die Funktion

$$p_X \colon \Omega \to [0, 1], \quad p_X(x) := P(X = x)$$

zu definieren, genannt die *Wahrscheinlichkeitsfunktion* der Zufallsgröße X.

8.2. Bedingte Wahrscheinlichkeiten

8.2.1. Mehrstufige Experimente

Es findet ein zweistufiges Experiment statt, welches sich sich aus einem ersten und einem zweiten Wurf eines Spielwürfels zusammensetzt. Bei jedem der Würfe bestehe eine Gleichverteilung. Zur Frage steht, wie wahrscheinlich das Ereignis $\{(6,6)\}$ ist. Ein Paar (ω_1, ω_2) fasse hierbei das Ergebnis ω_1 des ersten und ω_2 des zweiten Wurfs zusammen.

Die Wahrscheinlichkeit der ersten Sechs beträgt $\frac{1}{6}$, die der zweiten ebenfalls $\frac{1}{6}$. Sie multiplizieren sich zu $\frac{1}{36}$, richtig?

Es wäre doch möglich, dass zwischen den beiden Würfen eine, sagen wir, geisterhafte Beziehung besteht, dergestalt dass der zweite Wurf niemals in einer Sechs resultiert, sofern das Ergebnis des ersten eine war. Trotzdem sind die Wahrscheinlichkeiten bei jedem der Würfe für sich allein gesehen gleichverteilt. Dafür muss man nicht unbedingt die Wirklichkeit manipulieren. Das Phänomen ist bereits bei der Erzeugung von Zufallszahlen im Computer beobachtbar. War die erste Zufallszahl eine Sechs, braucht der Generator die zweite lediglich solange zu verwerfen, wie sie eine Sechs sein sollte. Umstände dieser Art stellen nicht nur ein Gedankenspiel dar, so dass wir uns notgedrungen mit ihnen auseinandersetzen müssen. Sie führen zum Begriff der *bedingten Wahrscheinlichkeit*.

Bisher wurde immer nur die Verteilung der Wahrscheinlichkeiten eines Würfels für sich allein betrachtet. Das war modelliert durch die Größe

$$X_0 \colon \Omega \to \Omega, \quad X_0(\omega) := \omega, \quad \Omega := \{1, \dots, 6\},$$

mit der Gleichverteilung P_0, so dass $P_0(X = 6) = \frac{1}{6}$.

Wir modellieren das zweistufige Experiment durch die Zufallsgröße

$$X \colon \Omega^2 \to \Omega^2, \quad X(\omega) := (X_1, X_2)(\omega) = (X_1(\omega), X_2(\omega)),$$

die sich mit $\omega = (\omega_1, \omega_2)$ aus den zwei Zufallsgrößen

$$\begin{aligned} X_1 &\colon \Omega^2 \to \Omega, \quad X_1(\omega_1, \omega_2) := \omega_1, \\ X_2 &\colon \Omega^2 \to \Omega, \quad X_2(\omega_1, \omega_2) := \omega_2 \end{aligned}$$

zusammensetzt. Es stellt $X_1(\omega)$ das Ergebnis des ersten und $X_2(\omega)$ das des zweiten Wurfs dar. Wie gewünscht gilt

$$(X_1(\omega), X_2(\omega)) = (X_0(\omega_1), X_0(\omega_2)) = (\omega_1, \omega_2).$$

Es bezeichne P die Verteilung auf Ω^2. Wir wissen hier allerdings lediglich

$$P(X_1 = \omega_1) = P_0(X_0 = \omega_1) = \tfrac{1}{6},$$
$$P(X_2 = \omega_2) = P_0(X_0 = \omega_2) = \tfrac{1}{6}.$$

Die Fehlannahme besteht nun darin, dass per se

$$P(\{X_1 = \omega_1\} \cap \{X_2 = \omega_2\}) = P(X_1 = \omega_1)P(X_2 = \omega_2)$$

gelten müsse. Ist diese Gleichung erfüllt, nennt man die Zufallsgrößen X_1, X_2 *unabhängig*. In der bisherigen Sichtweise, wo wir nur X_0 mit P_0 gesehen haben, war es uns nicht möglich, stochastische Abhängigkeit zu beschreiben. Man notiert allgemein

$$P(X = x, Y = y) := P(\{X = x\} \cap \{X = y\}) = P(X = x)P(Y = y \mid X = x).$$

Der letzte Faktor bezeichne hierbei die bedingte Wahrscheinlichkeit für das Ereignis $\{Y = y\}$, unter der Bedingung, dass $\{X = x\}$ bereits eingetreten ist.

Definition 8.2 (Bedingte Wahrscheinlichkeit).
Die *bedingte Wahrscheinlichkeit* für den Eintritt von A unter der Bedingung B ist für $P(B) \neq 0$ definiert gemäß

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)}.$$

Wir setzen speziell $B := \{X = x\}$ und $A := \{Y = y\}$ ein, das macht

$$P(Y = y \mid X = x) = \frac{P(X = x, Y = y)}{P(X = x)}.$$

Sind X, Y unabhängig, gilt also

$$P(Y = y \mid X = x) = P(Y = y).$$

Mit der geisterhaften Beziehung zwischen den Würfeln wäre allerdings

$$0 = P(X_2 = 6 \mid X_1 = 6) \neq P(X_1 = 6) = \tfrac{1}{6}.$$

Reflektiert man den Begriff der bedingten Wahrscheinlichkeit eingehend, fällt auf, dass dieser nicht zwangsläufig ein mehrstufiges Experiment voraussetzt. Zugleich kann ein zweistufiges Experiment auch als einstufiges mit Paaren als Ergebnissen interpretiert werden. Zur Verdeutlichung will ich die bedingte Wahrscheinlichkeit

aber nochmals bei einem eigentlich einstufigen Experiment diskutieren. Es wird einmalig ein gewöhnlicher Würfel geworfen, die Wahrscheinlichkeit für das Ereignis $\{4\}$ beträgt offenkundig $\frac{1}{6}$. Es sei nun B das Ereignis, dass eine gerade Zahl geworfen wurde. Diesbezüglich ergibt sich

$$P(\{4\} \mid B) = \frac{P(\{4\} \cap B)}{P(B)} = \frac{P(\{4\})}{1/2} = 2 \cdot \frac{1}{6} = \frac{1}{3}.$$

Dieser Umstand mag intuitiv begreiflich sein. Wenn bereits bekannt ist, dass eine gerade Zahl geworfen wurde, verbleiben nur noch drei Möglichkeiten für das Ergebnis, womit jedes der drei zugehörigen elementaren Ereignisse, die anscheinend gleichberechtigt sind, die Wahrscheinlichkeit $\frac{1}{3}$ besitzen muss.

Bei einem mehrstufigen Zufallsexperiment stellt das *Baumdiagramm* ein Hilfsmittel zur Schaffung von Übersicht dar. An ihm wird die erste und zweite Pfadregel ersichtlich gemacht. Wir betrachten dazu zunächst ein zweistufiges Experiment mit der Ergebnismenge $\Omega = \Omega_1 \times \Omega_2$. Ist $a \in \Omega$ mit $a = (a_1, a_2)$ das Ergebnis des Experiments, so ist $a_1 \in \Omega$ das erste Teilergebnis. Es stellt sich zunächst einmal die Frage, wie die Wahrscheinlichkeit zu diesem ausgedrückt wird, wo P eigentlich auf Ω definiert ist, nicht aber auf Ω_1. Sie muss gleich $P(\{a_1\} \times \Omega_2)$ sein, insofern das zweite Ergebnis noch unbekannt verbleibt, also keine Rolle spielen darf. Nun würde man analog sagen, das zweite Ergebnis besäße die Wahrscheinlichkeit $P(\Omega_1 \times \{a_2\})$, aber das kann so nicht allgemein stimmen, denn die Wahrscheinlichkeit zu a_2 ist unter Umständen vom Ausgang des ersten Teilexperiments abhängig. Zur Abkürzung sei $B := \{a_1\} \times \Omega_2$ und $A := \Omega_1 \times \{a_2\}$. Wenn die Wahrscheinlichkeit vom Ausgang zu a_2 gewünscht ist, bei der a_1 bereits eingetreten ist, handelt es sich nicht um $P(A)$, sondern um $P(A \mid B)$. Man rechnet diesbezüglich

$$P(\{a\}) = P(B \cap A) = P(B)P(A \mid B).$$

Die ist die *erste Pfadregel*. Sie besagt, dass die Wahrscheinlichkeit eines elementaren Ereignisses gleich dem Produkt der Wahrscheinlichkeiten entlang des Pfades ist, der zu diesem Ereignis führt.

Beispiel. In einer Urne befinden sich eine weiße und drei schwarze Kugeln. Es werden zwei Kugeln gezogen, wobei ihre Reihenfolge von Belang sei. Gesucht ist die Wahrscheinlichkeit der elementaren Ereignisse. Die Abb. 8.1 zeigt das Baumdiagramm zu diesem Experiment. Wir haben hier die Erbebnismenge

$$\Omega = \Omega_1 \times \Omega_2 = \{\bullet, \circ\} \times \{\bullet, \circ\} = \{(\bullet, \bullet), (\bullet, \circ), (\circ, \bullet), (\circ, \circ)\}.$$

Zu $B := \{\bullet\} \times \Omega_2$ und $A := \Omega_1 \times \{\circ\}$ ergibt sich

$$P(\{(\bullet, \circ)\}) = P(B)P(A \mid B) = \tfrac{3}{4} \cdot \tfrac{1}{3} = \tfrac{1}{4}.$$

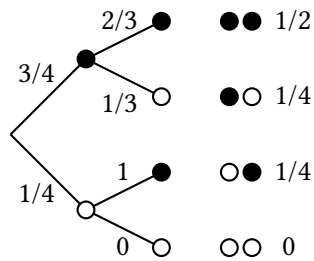


Abbildung 8.1.: Baumdiagramm zur Ziehung aus der Urne

Um die erste Pfadregel für Experimente mit mehr als zwei Stufen bestätigen zu können, müssen wir zunächst ein paar Vorbetrachtungen machen. Zunächst wird sich die Frage auftun, wie man bedingte Wahrscheinlichkeiten von bedingten Wahrscheinlichkeiten bildet. Bei fester Bedingung B bildet die bedingte Wahrscheinlichkeit wiederum ein Wahrscheinlichkeitsmaß,

$$P_B \colon \Omega \to \mathbb{R}, \quad P_B(A) := P(A \mid B).$$

Diesbezüglich kann man nun von $P_B(A \mid C)$ reden. Allerdings findet sich hierzu die Umformung

$$P_B(A \mid C) = \frac{P_B(A \cap C)}{P_B(C)} = \frac{\frac{P(A \cap B \cap C)}{P(B)}}{\frac{P(B \cap C)}{P(B)}} = \frac{P(A \cap B \cap C)}{P(B \cap C)} = P(A \mid B \cap C).$$

Das heißt, die Schachtelung von bedingten Wahrscheinlichkeiten stimmt mit der bedingten Wahrscheinlichkeit unter konjunktiver Verbindung der Bedingungen, ausgedrückt als Schnittmenge, überein. Für $P(A \mid B \cap C)$ ist auch die Schreibweise $P(A \mid B, C)$ geläufig.

8.2.2. Gesetz der totalen Wahrscheinlichkeit

Satz 8.5 (Gesetz der totalen Wahrscheinlichkeit).
Es sei $\mathcal{Z}$ eine Zerlegung der Ergebnismenge Ω in paarweise disjunkte Mengen $B \in \mathcal{Z}$. Dann gilt

$$P(A) = \sum_{B \in \mathcal{Z}} P(A \mid B) P(B).$$

Beweis. Es findet sich die Umformung

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P(A \cap \bigcup_{B \in \mathcal{Z}} B) = P(\bigcup_{B \in \mathcal{Z}} (A \cap B)) \\ &= \sum_{B \in \mathcal{Z}} P(A \cap B) = \sum_{B \in \mathcal{Z}} P(A \mid B) P(B). \end{aligned}$$ □

9. Programmverifikation

9.1. Programme

Zur Lösung mancher Probleme, Beantwortung mancher Fragen mag man eine mehr oder weniger festgelegte Methode, ein Verfahren anwenden. Zur Präzisierung legt man das Verfahren durch einen Algorithmus fest. Mit konkreten Mitteln, das Verfahren durchführen zu können, entsteht aus dem Algorithmus ein *Programm*. Sei dem zweiten Weltkrieg kamen immer leistungsfähigere Rechenmaschinen auf, die immer anspruchsvollere Aufgaben bewältigten. Die Programme, die auf den Maschinen abliefen, wurden mit der Zeit immer komplexer.

Reflektiert man eine Weile, kann man zur Sichtweise gelangen, dass Problemlösung mehr oder weniger allgemein algorithmisch stattfinden kann. Es scheint, als ob wir uns fast notgedrungen mit Programmen auseinandersetzen müssen.

Die frühen Computer führten meist numerische Rechnungen aus, die den Bedürfnissen von Ingenieuren und Naturwissenschaftlern entsprangen. Leider nicht immer für friedliche Zwecke. Mit der Entwicklung ging die Auffindung vieler numerischer Verfahren und Ausreizung bereits bekannter Verfahren einher. Dazu ein kleines Beispiel. Das Programm 9.1 zeigt die iterative Berechnung der Quadratwurzel $x = \sqrt{a}$ mit dem Heron-Verfahren

$$x_0 := \frac{a+1}{2}, \quad x_{n+1} := \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Die Folge (x_n) konvergiert rasch gegen x.

9.2. Operationelle Semantik

Hat man bereits viele Programme geschrieben, gelesen und sich die Vorgänge auf der Assembler-Ebene angeschaut, mag man ein intuitives Verständnis dafür bekommen haben, wie Programme ablaufen. Es hat sich allerdings als sehr förderlich herausgestellt, diese Abläufe formal zu präzisieren. Das gelingt zum Beispiel, indem eine tatsächliche oder emulierte Rechenmaschine festgelegt wird. Um aber nicht den Fokus auf das Wesentliche zu verlieren, abstrahiert man von den Details

Listing 9.1: Programm zur iterativen Berechnung von Quadratwurzeln

```
def sqrt(a, epsilon = 1E-12):
    x = 0.5*(a + 1)
    while True:
        x = 0.5*(x + a/x)
        if abs((x*x - a)/a) < epsilon:
            return x
```

Listing 9.2: Programm zur iterativen Berechnung der Potenz x^n

```
def power(x, n):
    y = 1
    while n != 0:
        y = y*x
        n = n - 1
    return y
```

Listing 9.3: Programm zur rekursiven Berechnung der Potenz x^n

```
def power(x, n):
    if n == 0:
        return 1
    else:
        return x*power(x, n - 1)
```

der Maschine, indem den syntaktischen Konstrukten der Programmiersprache direkt eine Semantik zugeordnet wird. Ich will näher auf die *operationelle Semantik* eingehen.

Es sei T das Symbol für einen Term und Zahl das Symbol für eine Zahl. Wir legen mit der Produktionsregel

$$\mathrm{T} \to \mathrm{Zahl} \mid (\mathrm{T} + \mathrm{T}) \quad \text{bzw.} \quad \mathrm{T} \to \mathrm{Zahl} \quad \text{und} \quad \mathrm{T} \to (\mathrm{T} + \mathrm{T})$$

eine sehr einfache Grammatik fest. Damit ist gemeint, dass ein Term entweder eine Zahl oder ein geklammerter Summenterm sein soll, dessen Summanden wiederum Terme sind. Ein aus dem Startsymbol T erzeugter Term sei grammatisch, sobald nur noch Terminalsymbole, das sind in diesem Fall konkrete Zahlen, vorkommen. In EBNF notiert, nimmt die Produktionsregel die Form

```
T ::= Zahl | '(' T '+' T ')';
```

an. Bezeichnen wir mit Z die Menge der Zahlen und mit T die Menge der Terme, können wir die Grammatik alternativ auch durch die Inferenzregeln

$$\frac{t \in Z}{t \in T}, \qquad \frac{t_1 \in T \quad t_2 \in T}{(t_1 + t_2) \in T}$$

festlegen.

Die Auswertung von Termen findet statt gemäß einer Relation $(\rightarrow) \subseteq T \times Z$, wobei wir $t \rightarrow v$ für $(t, v) \in (\rightarrow)$ schreiben. Damit ist gemeint, dass t zu v auswertet oder auswerten kann. Die Relation wird festgelegt durch die Regeln

$$\frac{v \in Z}{v \rightarrow v}, \qquad \frac{t_1 \rightarrow v_1 \qquad t_2 \rightarrow v_2}{(t_1 + t_2) \rightarrow v'},$$

wobei v' der der Wert von $v_1 + v_2$ sein soll. Zu bemerken ist, dass durch diese Regeln keine Auswertungsreihenfolge vorgegeben wird.

Erweitern wir die Sprache nun dergestalt, dass in Ausdrücken auch Variablen $x, y, z \in \mathrm{Var}$ enthalten sein dürfen, stellt sich sogleich die Frage, welcher Wert einer Variablen bei ihrer Auswertung zukommen soll. Dies hängt anscheinend davon ab, in welchem *Zustand* sich das Programm gerade befindet. Nennen wir S die Menge der Zustände, ist mit einem Zustand $s \in S$ eine Abbildung $s \colon \mathrm{Var} \rightarrow Z$ verbunden. Meist brauchen wir lediglich den Teilzustand betrachten, der gerade von Bedeutung ist. Für den Zustand s mit $s(x) = 7$ und $s(y) = 2$ schreibt man dann auch kurzum $s = \{x = 7, y = 2\}$. Die Auswertungsrelation stellt nun eine Teilmenge von $T \times S \times Z$ dar, wobei $\langle t, s\rangle \rightarrow v$ bedeuten soll, dass der Term t im Zustand s den Wert v annehmen kann. Die Auswertung ist entsprechend festgelegt durch die Regeln

$$\frac{v \in Z}{\langle v, s\rangle \rightarrow v}, \qquad \frac{x \in \mathrm{Var}}{\langle x, s\rangle \rightarrow s(x)}, \qquad \frac{\langle t_1, s\rangle \rightarrow v_1 \qquad \langle t_2, s\rangle \rightarrow v_2}{\langle (t_1 + t_2), s\rangle \rightarrow v_1 + v_2}.$$

Die Sprache wird nun erweitert um Ausdrücke gemäß

$$\mathrm{E} \rightarrow \texttt{false} \mid \texttt{true} \mid \mathrm{T} = \mathrm{T} \mid \mathrm{T} \leq \mathrm{T} \mid \neg \mathrm{E} \mid (\mathrm{E} \wedge \mathrm{E}) \mid (\mathrm{E} \vee \mathrm{E}).$$

Es sei E die Menge der auf diese Weise formierbaren Ausdrücke. Zu diesen wird ebenfalls eine Auswertungsrelation definiert,

$$(\rightarrow) \subseteq E \times S \times \mathrm{Bool}, \quad \mathrm{Bool} := \{\texttt{false}, \texttt{true}\}.$$

Wir legen die Auswertung intuitiv fest als

$$\frac{}{\langle \texttt{false}, s\rangle \rightarrow \texttt{false}}, \qquad \frac{}{\langle \texttt{true}, s\rangle \rightarrow \texttt{true}}, \qquad \frac{\langle t_1, s\rangle \rightarrow v_1 \quad \langle t_2, s\rangle \rightarrow v_2}{\langle t_1 = t_2, s\rangle \rightarrow v'}$$

und so weiter, wobei v' der Wert **true** sein soll, wenn v_1, v_2 übereinstimmen, sonst **false**. Die Auswertung der Junktoren geschieht gemäß

$$\frac{\langle e, s\rangle \rightarrow v}{\langle \neg e, s\rangle \rightarrow v'}, \qquad \frac{\langle e_1, s\rangle \rightarrow v_1 \quad \langle e_2, s\rangle \rightarrow v_2}{\langle (e_1 \wedge e_2), s\rangle \rightarrow v'}$$

und so weiter, wobei v' jeweils der Wahrheitswert sein soll, der sich aus der Wahrheitstafel des jeweiligen Junktors ergibt.

Die Produktionsregel

$$\mathrm{P} \to \textbf{skip} \mid \mathrm{Var} := \mathrm{T} \mid \mathrm{P}; \mathrm{P} \mid \textbf{if}\ \mathrm{E}\ \textbf{then}\ \mathrm{P}\ \textbf{else}\ \mathrm{P}\ \textbf{end} \mid \textbf{while}\ \mathrm{E}\ \textbf{do}\ \mathrm{P}\ \textbf{end}$$

beschreibt die Syntax einer kleinen imperativen Programmiersprache, die im Weiteren als Studienobjekt dienen wird. Wir definieren zunächst ein Semantik, die die intuitive Vorstellung vom Ablauf der Anweisungen und Kontrollstrukturen widerspiegelt. Dies geschieht durch Festlegung der Übergangsrelation $(\to) \subseteq P \times S \times S$. Mit $\langle p, s\rangle \to s'$ soll gemeint sein, dass der Zustand s mit dem Durchlauf des Programms p in den Zustand s' übergeht oder zumindest übergehen kann.

Zur Anweisung **skip** muss nicht viel gesagt werden, sie wird einfach übersprungen, der Zustand bleibt unberührt. Ihre Regel lautet daher

$$\frac{}{\langle \textbf{skip}, s\rangle \to s}.$$

Der Übergang zur Zuweisung verläuft gemäß

$$\frac{\langle t, s\rangle \to v}{\langle x := t, s\rangle \to s[x := v]}.$$

Zu einer Sequenz von Programmen verläuft der Übergang gemäß

$$\frac{\langle p_1, s\rangle \to s' \quad \langle p_2, s'\rangle \to s''}{\langle p_1; p_2, s\rangle \to s''}.$$

Bei if-Anweisungen findet der Übergang nach den Regeln

$$\frac{\langle e, s\rangle \to \textbf{true} \quad \langle p_1, s\rangle \to s'}{\langle \textbf{if}\ e\ \textbf{then}\ p_1\ \textbf{else}\ p_2\ \textbf{end}\rangle \to s'}, \qquad \frac{\langle e, s\rangle \to \textbf{false} \quad \langle p_2, s\rangle \to s'}{\langle \textbf{if}\ e\ \textbf{then}\ p_1\ \textbf{else}\ p_2\ \textbf{end}\rangle \to s'}$$

statt. Schließlich muss noch die Semantik für Schleifen festgelegt werden. Wertet die Schleifenbedingung zu **false** aus, wird der Schleifenkörper schlicht übersprungen, ohne eine Zustandsänderung herbeizuführen,

$$\frac{\langle e, s\rangle \to \textbf{false}}{\langle \textbf{while}\ e\ \textbf{do}\ p\ \textbf{end}, s\rangle \to s}.$$

Wertet die Schleifenbedingung zu **true** aus, liegt die Regel nicht mehr gänzlich auf der Hand. Die Überlegung ist folgende. Wurde bereits geklärt, dass der erste Durchlauf den Zustand s in s' überführt und die restliche Schleife von s' in s'' überführt, dann muss die Schleife den Zustand s in s'' überführen. Man gelangt zur Regel

$$\frac{\langle e, s\rangle \to \textbf{true} \quad \langle p, s\rangle \to s' \quad \langle \textbf{while}\ e\ \textbf{do}\ p\ \textbf{end}, s'\rangle \to s''}{\langle \textbf{while}\ e\ \textbf{do}\ p\ \textbf{end}, s\rangle \to s''}.$$

9.3. Der Hoare-Kalkül

Ein Hoare-Tripel $\{A\}\, p\, \{B\}$ soll die logische Aussage ausdrücken, dass nach dem Durchlaufen des Programms p die Aussage B gilt, sofern zuvor die Aussage A galt. Hierbei bezeichnet man A als *Vorbedingung* und B als *Nachbedingung*. Das Tripel macht allerdings keine Aussage darüber, ob das Programm p tatsächlich terminiert. Ist das Tripel erfüllt, nennt man das Programm in Bezug auf die Vor- und Nachbedingung *partiell korrekt*. Terminiert das Programm zudem, spricht man von *totaler Korrektheit*.

Bei A, B soll es sich um prädikatenlogische Formeln handeln, mit der Besonderheit, dass in ihr die Terme und Ausdrücke der Programmiersprache auftreten dürfen. Dazu kann formal die *Zusicherungssprache* festgelegt werden, aus der die Zusicherungen A, B entstammen sollen.

Die Gültigkeit eines Hoare-Tripels wird definiert gemäß

$$(\vDash \{A\}\, p\, \{B\}) \;:\Leftrightarrow\; \forall s \in S\colon (s \vDash A) \Rightarrow \forall s' \in S\colon (\langle p, s\rangle \to s') \Rightarrow (s' \vDash B).$$

Interessanterweise kommt darin recht direkt die Kripke-Semantik einer Notwendigkeit zum Vorschein. Die Zustände nehmen hierbei die Rolle der Welten ein, die Übergangsrelation je Programm p die Rolle der Zugänglichkeitsrelation. Demnach liegt eine multimodale Logik vor, da je Programm p ein Operator $\Box_p$ existiert, der üblicherweise $[p]$ geschrieben wird. Setzen wir diesbezüglich

$$(s \vDash [p]B) \;:\Leftrightarrow\; \forall s' \in S\colon (\langle p, s\rangle \to s') \Rightarrow (s' \vDash B),$$

findet sich die Äquivalenz

$$(\vDash \{A\}\, p\, \{B\}) \Leftrightarrow (\vDash A \Rightarrow [p]B),$$

die eine jähe Beziehung des Hoare-Kalküls zur ihr, der *dynamischen Logik* herstellt, auf die ich an dieser Stelle aber nicht näher eingehen will. Trotzdem mag der Leser sie später in Erinnerung rufen, um ein übersichtlicheres Bild von den Sachverhalten zu bekommen.

Die Regeln zur Herleitung der Tripel. Für Zuweisungen gilt

$$\frac{}{\vdash \{A[x := t]\}\; x := t\; \{A\}}.$$

Zur Gültigkeit dieser Regel. Es gelte $s \vDash A[x := t]$. Weiterhin wird $\langle x := t, s\rangle \to s'$ angenommen. Zu zeigen ist $s' \vDash A$. Insofern die Auswertung von Termen deterministisch stattfindet, existiert nun genau ein v mit $\langle t, s\rangle \to v$. Mithin liegt der Sachverhalt $s' = s[x := v]$ vor. Letztlich gilt die Äquivalenz

$$(s \vDash A[x := t]) \Leftrightarrow (s[x := v] \vDash A),$$

da man zum gleichen Resultat gelangt, wenn x in A direkt mit v belegt wird, oder zunächst gegen t ersetzt und daraufhin t den Wert v erhält. Formal bestätigt man dies per struktureller Induktion über den Aufbau von A.

Für eine Sequenz von Programmen gilt

$$\frac{\vdash \{A\}p_1\{B\} \qquad \vdash \{B\}p_2\{C\}}{\vdash \{A\}p_1; p_2\{C\}}.$$

Zur Gültigkeit dieser Regel. Es gelte $s \vDash A$. Angenommen, es kann s'' erreicht werden, also $\langle s, p_1; p_2\rangle \to s''$. Dann muss ein Zustand s' mit $\langle p_1, s\rangle \to s'$ und $\langle p_2, s'\rangle \to s''$ existieren. Sollte s' nicht eindeutig bestimmt sein, betrachten wir s' einfach fest, aber beliebig. Zu zeigen ist $s'' \vDash C$. Vermittels der ersten Prämisse erhält man $s' \vDash B$, vermittels der zweiten daraufhin $s'' \vDash C$.

Die Regel für Verzweigungen lautet

$$\frac{\vdash \{A \wedge B\}\, p_1\, \{C\} \qquad \vdash \{A \wedge \neg B\}\, p_2\, \{C\}}{\vdash \{A\}\ \texttt{if}\ B\ \texttt{then}\ p_1\ \texttt{else}\ p_2\ \texttt{end}\ \{C\}}.$$

Zur Gültigkeit dieser Regel. Sei s der Zustand vor, und s' der Zustand nach Durchlaufen der Verzweigung. Es gelte $s \vDash A$, zu zeigen ist $s' \vDash C$. Fallunterscheidung. Im Fall $\langle B, s\rangle \to \texttt{true}$ gilt $s \vDash B$, also $s \vDash A \wedge B$. Es wird p_1 ausgeführt, also muss $\langle p_1, s\rangle \to s'$ gelten. Laut der ersten Prämisse gilt dann $s' \vDash C$. Im Fall $\langle B, s\rangle \to \texttt{false}$ gilt $s \vDash \neg B$, also $s \vDash A \wedge \neg B$. Es wird p_2 ausgeführt, also muss $\langle p_2, s\rangle \to s'$ gelten. Laut der zweiten Prämisse gilt dann $s' \vDash C$.

Die Regel für Schleifen lautet

$$\frac{\vdash \{I \wedge B\}\, p\, \{A\}}{\vdash \{I\}\ \texttt{while}\ B\ \texttt{do}\ p\ \texttt{end}\ \{I \wedge \neg B\}}.$$

Zur Gültigkeit dieser Regel. Sei s der Zustand vor, und s'' der Zustand nach dem Durchlaufen der Schleife. Es gelte $s \vDash I$. Induktion über die Anzahl der Durchläufe. Im Induktionsanfang wird findet kein Durchlauf von p statt, womit $\langle B, s\rangle \to \texttt{false}$ gilt, also $s \vDash \neg B$. Da dem Zustand keine Änderung widerfährt, gilt $s'' = s$, also $s'' \vDash I \wedge \neg B$. Nun zum Induktionsschritt. Da p durchlaufen wird, gilt $\langle B, s\rangle \to \texttt{true}$, also $s \vDash B$, ergo $s \vDash I \wedge B$. Mit dem Durchlauf von p existiert s' mit $\langle p, s\rangle \to s'$. Laut der Prämisse gilt $s' \vDash I$. Mit dem Durchlaufen der restlichen Schleife geht s' in s'' über. Vermittels der Induktionsvoraussetzung erhält man $s'' \vDash I \wedge \neg B$.

Eine Aussage I, welche sowohl vor als auch nach dem Durchlauf des Schleifenrumpfs und somit der Schleife als Ganzes gültig ist, bezeichnet man als *Schleifeninvariante*. Die Auffindung einer zielführenden Schleifeninvariante ist nicht immer einfach.

Listing 9.4: Schnelles Potenzieren

```
def power(x, n):
    y = 1
    while n > 1:
        if n%2 == 1: y = y*x
        n //= 2; x = x*x
    if n == 1: y = y*x
    return y
```

Listing 9.5: Mit Hilfsvariablen

```
def power(x, n):
    y = 1; a = x; i = n
    while i > 1:
        if i%2 == 1: y = y*a
        i //= 2; a = a*a
    if i == 1: y = y*a
    return y
```

Gilt ein Tripel als bestätigt, kann man intuitiv die Vorbedingung beliebig verstärken und die Nachbedingung beliebig abschwächen. Man erhält ein Tripel mit weniger Aussagekraft, das aber für die Beschreibung des Sachverhaltes genügen mag. Diese Überlegung wird formalisiert durch die Regel

$$\frac{\vdash A' \Rightarrow A \qquad \vdash \{A\}\, p\, \{B\} \qquad \vdash B \Rightarrow B'}{\vdash \{A'\}\, p\, \{B'\}}.$$

Zur Gültigkeit dieser Regel. Aus $s \vDash A'$ und $\langle p, s\rangle \longrightarrow s'$ ist $s' \vDash B$ abzuleiten. Vermittels $\vDash A' \Rightarrow A$ erhält man zunächst $s \vDash A$, vermittels $\vDash \{A\}\, p\, \{B\}$ daraufhin $s' \vDash B$, und vermittels $\vDash B \Rightarrow B'$ daraufhin schließlich $s' \vDash B'$.

Die Regel setzt sich zusammen aus den beiden Spezialfällen

$$\frac{\vdash A' \Rightarrow A \qquad \vdash \{A\}\, p\, \{B\}}{\vdash \{A'\}\, p\, \{B\}}, \qquad \frac{\vdash \{A\}\, p\, \{B\} \qquad \vdash B \Rightarrow B'}{\vdash \{A\}\, p\, \{B'\}}.$$

Beispiel. Das Listing 9.4 zeigt den Algorithmus zum schnellen Potenzieren. Wie 9.2 berechnet dieser die Potenz x^n, benötigt dafür aber weniger Multiplikationen. Soll beispielsweise x^{100} berechnet werden, kann dies auf $(x^{50})^2$ zurückgeführt werden. Dafür dass am Ende einmal quadriert wird, wurden 50 Multiplikationen eingespart. Der Algorithmus stellt ein allgemeines Verfahren hierfür dar.

Allerdings ist nicht mehr auf den ersten Blick ersichtlich, ob der Algorithmus korrekt arbeitet. Zur Analyse werden zunächst die Hilfsvariablen a, i eingeführt, damit x, n konstant bleiben, siehe Listing 9.5. Nach dem Durchlaufen des Programms soll die Nachbedingung $y = x^n$ bestehen. Zunächst findet sich das Tripel

$$\{y = x^n \wedge i = 0 \vee y \cdot a = x^n \wedge i = 1\}$$
```
if i == 1: y = y*a
```
$$\{y = x^n\}.$$

Im Fall $i = 1$ reduziert sich die Vorbedingung nämlich zu $y \cdot a = x^n$, die bedingte Anweisung wird ausgeführt und das Resultat ist $y = x^n$ laut Zuweisungsregel. Im Fall $i \neq 1$ muss $i = 0$ sein, denn nach der Schleife wird $0 \leq i \leq 1$ gelten. In diesem Fall reduziert sich die Vorbedingung zu $y = x^n$, die bedingte Anweisung wird nicht ausgeführt, das Resultat ist somit ebenfalls $y = x^n$.

Nun kommt ein kritischer Schritt, dessen Auffindung ohne Übung schwierig sein mag. Vor der Schleife gilt $a^i = x^n$, was etwas mit der Schleifeninvariante zu tun haben könnte. Betrachten wir nun die Vorbedingung

$$y = x^n \wedge i = 0 \vee y \cdot a = x^n \wedge i = 1.$$

Ist $a^i = x^n$ so modifizierbar, dass diese gilt? Ja, nämlich zu $y \cdot a^i = x^n$, denn

$$\text{im Fall } i = 0 \text{ gilt } y \cdot a^i = y \cdot a^0 = y,$$
$$\text{im Fall } i = 1 \text{ gilt } y \cdot a^i = y \cdot a^1 = y \cdot a.$$

Die Gleichung $y \cdot a^i = x^n$ gilt ebenfalls vor der Schleife, denn dort ist $y = 1$. Zudem kommen in der Gleichung alle drei Variablen vor, die sich in der Schleife verändern. Tatsächlich stellt diese Gleichung eine zielführende Schleifeninvariante dar.

Nun zum Schleifenrumpf. Das Tripel zur ersten Anweisung ist

$\{y \cdot a^i = x^n\}$
`if i%2 == 1: y = y*a`
$\{y \cdot a^{i-1} = x^n \wedge i \bmod 2 = 1 \vee y \cdot a^i = x^n \wedge i \bmod 2 = 0\}$.

Die lange Nachbedingung kann man kompakter schreiben als

$$y \cdot a^{i - i \bmod 2} = x^n.$$

Außerdem gilt $i - i \bmod 2 = 2\lfloor \frac{i}{2} \rfloor$, wobei $\lfloor \frac{i}{2} \rfloor$ mit dem Programmterm `i//2` gleichbedeutend ist. Mit dieser Vorbereitung stellt die restliche Argumentation nur mehr eine Leichtigkeit dar. Gemäß der Zuweisungsregel sind die Tripel

$\{y \cdot a^{2\lfloor i/2 \rfloor} = x^n\}$		$\{y \cdot a^{2i} = x^n\}$
`i //= 2;`	und	`a = a*a;`
$\{y \cdot a^{2i} = x^n\}$		$\{y \cdot a^i = x^n\}$

erfüllt. Damit gilt die Invariante als bewiesen. Weil nichts weiter zu beweisen verbleibt, gilt auch der gesamte Algorithmus als bewiesen.

9.4. Zum Kalkül der schwächsten Vorbedingung

Zu einem Programm p mag man zu einer Nachbedingung B mehr als eine Vorbedingung A finden, so dass $\{A\}\,p\,\{B\}$ gilt. Dahingehend gelangt man zur Überlegung, ob man A nicht zu einer Vorbedingung A' mit $A \Rightarrow A'$ abschwächen kann, so dass $\{A'\}\,p\,\{B\}$ erfüllt bleibt. Wir nennen A_0 *schwächste liberale Vorbedingung*, wenn für jede andere Vorbedingung A die Äquivalenz

$$\{A\}\,p\,\{B\} \Leftrightarrow (A \Rightarrow A_0)$$

besteht. Man schreibt $\mathrm{wlp}(p, B) := A_0$, wobei wlp für *weakest liberal precondition* steht. Wie bei Erörterung der Semantik des Hoare-Tripels bereits festgestellt wurde, handelt es sich bei $\mathrm{wlp}(p, B)$ um nichts anderes als die modale Aussage $[p]B$.

Satz 9.1. Es gilt $[p_1; p_2]B \equiv [p_1][p_2]B$.

Beweis. Zur Abkürzung sei $\varphi_1 :\equiv (\langle p_1, s\rangle \to s')$ sowie $\varphi_2 :\equiv (\langle p_2, s'\rangle \to s'')$ und $\psi :\equiv (s'' \vDash B)$. Es findet sich die äquivalente Umformung

$$\begin{aligned}
&(s \vDash [p_1][p_2]B) \iff (\forall s' \colon \varphi_1 \Rightarrow \forall s'' \colon \varphi_2 \Rightarrow \psi)\\
\iff &(\forall s' \colon \forall s'' \colon \varphi_1 \wedge \varphi_2 \Rightarrow \psi) \iff (\forall s'' \colon \forall s' \colon \varphi_1 \wedge \varphi_2 \Rightarrow \psi)\\
\iff &(\forall s'' \colon (\exists s' \colon \varphi_1 \wedge \varphi_2) \Rightarrow \psi) \iff (s \vDash [p_1; p_2]B).
\end{aligned}$$

Der letzte Schritt gelingt hierbei aufgrund der Äquivalenz

$$(\exists s' \colon \varphi_1 \wedge \varphi_2) \Leftrightarrow (\langle p_1; p_2, s\rangle \to s''). \qquad \square$$

Satz 9.2. Es gilt $[x := t]B \equiv B[x := t]$.

Beweis. Es gelte $s \vDash [x := t]B$, womit $s' \vDash B$ für jedes s' mit $\langle x := t, s\rangle \to s'$ gelten muss. Es existiert aber genau ein solches s', und für dieses gilt $s' = s[x := v]$ mit $\langle t, s\rangle \to v$. Diesbezüglich ist $s' \vDash B$ gleichbedeutend mit $s \vDash B[x := t]$. Diese Argumentation kann auch unschwer umgekehrt werden. $\square$

Der Umstand, dass man einfach die Symbolik verdrehen kann, ist natürlich eine glückliche Fügung in Form einer syntaktischen Koinzidenz. Wählt man eine andere Notation, etwa $\square_{x:=t}B \equiv B[t/x]$, geht sie verloren.

Satz 9.3. Aus $\vDash A$ folgt $\vDash [p]A$.

Beweis. Unter der Annahme $\langle p, s\rangle \to s'$ muss $s' \vDash A$ bestätigt werden. Laut der Prämisse ist $s' \vDash A$ aber per se erfüllt. $\square$

Satz 9.4. Es gilt $[p](A \Rightarrow B) \Rightarrow ([p]A \Rightarrow [p]B)$.

Beweis. Unter der Annahme $\langle p, s\rangle \rightarrow s'$ muss $s' \vDash B$ bestätigt werden. Vermittels der ersten Prämisse erhält man $s' \vDash A$ aus der Annahme. Vermittels der zweiten Prämisse erhält man $s' \vDash A \Rightarrow B$ aus der Annahme, was definitionsgemäß zu $(s' \vDash A) \Rightarrow (s' \vDash B)$ äquivalent ist. Per Modus ponens folgt somit $s' \vDash B$. □

Satz 9.5. Es gilt $[p](A \wedge B) \equiv [p]A \wedge [p]B$.

Beweis. Dies ist ein allgemeiner Umstand der Modallogik K, die laut Satz 9.3 und Satz 9.4 je fixem Programm p vorliegt. □

Satz 9.6. Es gilt $[\texttt{if}\ B\ \texttt{then}\ p_1\ \texttt{else}\ p_2\ \texttt{end}]C \equiv (B \Rightarrow [p_1]C) \wedge (\neg B \Rightarrow [p_2]C)$.

Beweis. Es gelte die linke Seite der Äquivalenz. Zum ersten Teil der Konjunktion. Aus der Annahme von $s \vDash B$ und $\langle p_1, s\rangle \rightarrow s'$ ist $s' \vDash C$ abzuleiten. Wegen $\langle B, s\rangle \rightarrow \texttt{true}$ erhält man

$$\langle \texttt{if}\ B\ \texttt{then}\ p_1\ \texttt{else}\ p_2\ \texttt{end}, s\rangle \rightarrow s',$$

woraus $s' \vDash C$ vermittels der linken Seite folgt. Die Argumentation zum zweiten Teil der Konjunktion verläuft analog.

Es gelte die rechte Seite der Äquivalenz. Es ist $s' \vDash C$ abzuleiten. Fallunterscheidung. Im Fall $\langle B, s\rangle \rightarrow \texttt{true}$ gilt $s \vDash B$ und außerdem $\langle p_1, s\rangle \rightarrow s'$. Mit diesen erhält man $s' \vDash C$ vermittels des ersten Teils der Konjunktion. Im Fall $\langle B, s\rangle \rightarrow \texttt{false}$ verläuft die Argumentation analog. □

Ich möchte bemerken, dass in klassischer Logik die allgemeine Äquivalenz

$$(B \Rightarrow A_1) \wedge (\neg B \Rightarrow A_2) \Leftrightarrow (B \wedge A_1) \vee (\neg B \wedge A_2)$$

besteht.

10. Typentheorie

Mit der Zeit hielt die Typentheorie Einzug in die Grundlagen der Mathematik und gesellte sich neben die Mengenlehre. Viele Theorembeweisassistenten arbeiten mit einer Typentheorie, die eine Logik höherer Stufe kodiert. Um nicht nur oberflächlich mit diesen Assistenten arbeiten zu können, sondern auch ein gewissen Verständnis zu erhalten, wie sie funktionieren, müssen wir uns notgedrungen mit diesen Systemen auseinandersetzen. Darüber hinaus bestehen enge Bezüge zu den Typsystemen in der Informatik.

Wir müssen hier zwischen zwei Begriffen unterscheiden. Die *Typentheorien* sind ideale Typsysteme, denen vom Wesen her ein logisches System innewohnt. *Die Typentheorie*, eine Teildisziplin der mathematischen Logik, studiert diese Typsysteme und klärt ihre Konsistenz ab, wobei ein Bezug zur Beweistheorie hergestellt wird.

10.1. Abhängige Typentheorie

10.1.1. Begrifflichkeiten

Ich will in den Darlegungen den sorgfältigen Ausführungen von [16], [12] und [23] folgen. So wie es mehrere Systeme der Mengenlehre gibt, haben sich auch mehrere Systeme der abhängigen Typentheorie herausgebildet. Ich werde mich im Folgenden auf den *induktiven Konstruktionenkalkül* beziehen.

Im formalen System der Mengenlehre gibt es erstens Terme, die für Mengen stehen, zweitens Formeln, die für Aussagen stehen und drittens Beweise. Das sind drei syntaktische Systeme, die voneinander getrennt definiert werden. In der einfachen Typentheorie gibt es gleichermaßen Terme, Typen und Beweise. Dagegen fallen Terme, Typen und Beweise in der abhängigen Typentheorie, sofern sie reiner Art ist, syntaktisch zum *Ausdruck* zusammen.

Dafür gibt es aber drei unterschiedliche Formen von Urteilen. Dies sind

1. das Urteil Γ ctx, laut dem Γ ein wohlgeformter Kontext sei,
2. das Urteil $\Gamma \vdash a : A$, laut dem der Ausdruck a im Kontext Γ vom Typ A sei,
3. das Urteil $a \equiv b$, laut dem die Ausdrücke a, b urteilsmäßig gleich seien.

Ein Kontext Γ ist eine Liste der Form

$$\Gamma = [x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n],$$

die kodiert, dass die jeweilige Variable x_i vom Typ A_i sei. Die x_i stehen hierbei für paarweise verschiedene Variablen, nicht aber für Metavariablen, für die zusammengesetzte Ausdrücke eingesetzt werden dürften. Anders als in den herkömmlichen Sequenzen muss der Kontext eine Liste sein, da es auf die Reihenfolge der Elemente ankommt, denn im Ausdruck A_j dürfen nur die Variablen x_i mit $i < j$ auftauchen.

Es darf auch $n = 0$ sein, was bedeuten soll, dass der leere Kontext $\Gamma = []$ ebenfalls ein wohlgeformter ist. Die Schreibweise $\Gamma, x : A$ steht analog zu den herkömmlichen Sequenzen für die Konkatenation der Listen Γ und $[x : A]$.

10.1.2. Formuierungsregeln

Von den Schlussregeln der abhängigen Typentheorie gruppieren sich zunächst einmal die *Formierungsregeln*, die der Aufstellung wohlgeformter Typen dienen. Dem Formalismus nach wird ausschließlich mit auf diese Weise geformten Typen gearbeitet.

Jedem Ausdruck kommt ein *Typ* zu, auch den Typen selbst. Hierfür dient eine Abfolge von Konstantensymbolen

$$U_0, U_1, U_2, \ldots,$$

die wir *Typuniversen* nennen wollen. Statt U_i ist auch die längere Schreibweise Type_i geläufig. Hat ein Ausdruck A den Typ U_i für ein i, so wollen wir A selbst als einen *Typ* betrachten. Ein Ausdruck, der kein Typ ist, heißt *Term*. Es verhält sich so ähnlich wie mit einer Mengenlehre, in der Klassen die einzige Sorte sind. Die Abgrenzung der Mengen von den echten Klassen geschieht dahingehend nicht syntaktisch, sondern kommt inhaltlich aus den Regeln des formalen Systems zum Vorschein.

Da stellt sich die Frage, warum nicht ein Universum, nennen wir es U, genügt. Da diesem selbst ein Typ zukommen müsste, würden wir $U : U$ festsetzen, wie es Per Martin-Löf in der ursprünglichen Fassung seiner Typentheorie tat. Allerdings entdeckte Jean-Yves Girard kurz darauf, das das System damit inkonsistent wird, wir also einen Widerspruch ableiten können. Dieser Umstand, den man nun *girardsches Paradoxon* nennt, verhält sich analog zum russellschen Paradoxon. [22]

Für Typuniversen U_i bestehen die Regeln

$$\frac{\Gamma \text{ ctx}}{\Gamma \vdash U_i : U_{i+1}} \text{ univ-intro}, \qquad \frac{\Gamma \vdash A : U_i}{\Gamma \vdash A : U_{i+1}} \text{ univ-cumul}.$$

Kontexte werden eingeführt, erweitert gemäß

$$\frac{}{[] \text{ ctx}} \text{ ctx-empty}, \qquad \frac{\Gamma \vdash A : U_i}{\Gamma, x : A \text{ ctx}} \text{ ctx-ext},$$

mit der Nebenbedingung, dass x nicht bereits in Γ vorkommt.

Eine Grundsequenz wird eingeführt via

$$\frac{[x_1 \colon A_1, \ldots, x_n \colon A_n] \text{ ctx}}{x_1 \colon A_1, \ldots, x_n \colon A_n \vdash x_i \colon A_i} \text{ var}, \quad \text{oder kurz} \quad \frac{\Gamma \text{ ctx}}{\Gamma \vdash x \colon A}((x \colon A) \in \Gamma).$$

Der Typ abhängiger Funktionen wird formiert gemäß

$$\frac{\Gamma \vdash A \colon U_i \qquad \Gamma, x \colon A \vdash B \colon U_i}{\Gamma \vdash \Pi(x \colon A).\, B \colon U_i} \text{ Pi-form.}$$

Beim Typ Prop $:= U_0$ genügt es aber bereits, wenn B vom Typ U_0 ist,

$$\frac{\Gamma, x \colon A \vdash B \colon U_0}{\Gamma \vdash \Pi(x \colon A).\, B \colon U_0} \text{ Pi-form0.}$$

Der Typ $A \to A$ der Funktionen von A zu A wird als $\Pi(x \colon A).\, A$ dargestellt. Wir würden diesen Typ nun gerne formieren. Hierbei muss aber klargestellt werden, was A sein soll. Formiert werden kann zum Beispiel der Typ $\Pi(A \colon U_0).\, A \to A$, womit $\Pi(A \colon U_0).\, (A \to A)$ gemeint ist. Zuvor wird kurz festgestellt, dass die Einführung von Grundsequenzen der Form $A \colon U_i \vdash A \colon U_i$ eine zulässige Regel ist. Es finden sich die Bäume:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{[] \text{ ctx}}\,\text{ctx-empty}}{\vdash U_i \colon U_{i+1}}\,\text{univ-intro}}{A \colon U_i \text{ ctx}}\,\text{ctx-ext}}{A \colon U_i \vdash A \colon U_i}\,\text{var}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{A \colon U_0 \vdash A \colon U_0}}{A \colon U_0, x \colon A \text{ ctx}}\,\text{ctx-ext}}{A \colon U_0, x \colon A \vdash A \colon U_0}\,\text{var}}{A \colon U_0 \vdash \Pi(x \colon A).\, A \colon U_0}\,\text{Pi-form0}}{\vdash \Pi(A \colon U_0).\, \Pi(x \colon A).\, A \colon U_0}\,\text{Pi-form0}$$

Für $i \geq 1$ liegt der Typ allerdings in einem höhergestuften Universum:

$$\dfrac{\dfrac{}{\vdash U_i \colon U_{i+1}} \qquad \dfrac{\dfrac{\dfrac{}{A \colon U_i \vdash A \colon U_i} \qquad \dfrac{}{A \colon U_i, x \colon A \vdash A \colon U_i}}{A \colon U_i \vdash \Pi(x \colon A).\, A \colon U_i}\,\text{Pi-form}}{A \colon U_i \vdash \Pi(x \colon A).\, A \colon U_{i+1}}\,\text{univ-cumul}}{\vdash \Pi(A \colon U_i).\, \Pi(x \colon A).\, A \colon U_{i+1}}\,\text{Pi-form}$$

Man beachte, dass A hier, anders als in der Regel Pi-form, lediglich für eine Typvariable steht, also atomar ist. Dies wird durch die Kontexte klar, die nur Variablen enthalten dürfen.

10.1.3. Einführungs- und Beseitigungsregeln

Die Regeln zur Einführung, Beseitigung abhängiger Funktionen lauten

$$\frac{\Gamma, x\colon A \vdash b\colon B}{\Gamma \vdash (x \mapsto b)\colon \Pi(x\colon A).\, B}\ \text{Pi-intro}, \qquad \frac{\Gamma \vdash f\colon \Pi(x\colon A).\, B \qquad \Gamma \vdash a\colon A}{\Gamma \vdash f(a)\colon B[x := a]}\ \text{Pi-elim}.$$

Hierbei sind a, b, A, B Ausdrücke.

Speziell für von x unabhängiges B ergibt sich daraus entsprechend

$$\frac{\Gamma, x\colon A \vdash b\colon B}{\Gamma \vdash (x \mapsto b)\colon A \to B}, \qquad \frac{\Gamma \vdash f\colon A \to B \qquad \Gamma \vdash a\colon A}{\Gamma \vdash f(a)\colon B}.$$

Wie bei jedem lambda-Kalkül ist auch $\lambda x.\, b$ statt $x \mapsto b$ geläufig, wobei dies Kurzschreibweisen für $\lambda(x\colon A).\, b$ und $(x\colon A) \mapsto b$ sind. Die Klammern dürfen hierbei entfallen, ich will sie aber zur besseren Lesbarkeit beibehalten.

Beispielsweise ist $\Pi(A\colon U_i).\, A \to A$ bewohnt, denn es findet sich der Baum:

$$\dfrac{\dfrac{\dfrac{}{A\colon U_i, x\colon A \vdash x\colon A}}{A\colon U_i \vdash (x \mapsto x)\colon A \to A}\ \text{Pi-intro}}{\vdash (A \mapsto x \mapsto x)\colon \Pi(A\colon U_i).\, A \to A}\ \text{Pi-intro}$$

10.1.4. Bezug zur Logik

Den Typ $\Pi(x\colon A).\, B$ deuten wir als die logische Aussage $\forall(x\colon A).\, B$. Weiterhin deuten wir den Typ $A \to B$ als die Aussage $A \Rightarrow B$.

Für $A, B\colon U_0$ kann man die Konjunktion, in analoger Weise wie das schon in der Logik zweiter Stufe möglich ist, kodieren als

$$A \wedge B := \Pi(X\colon U_0).\, (A \to B \to X) \to X.$$

Dahingehend ist unschwer feststellbar, dass die drei Funktionen

$$\begin{aligned}
&\mathrm{conj}\colon \Pi(A\colon U_0).\, \Pi(B\colon U_0).\, A \to B \to A \wedge B,\\
&\mathrm{proj}_1\colon \Pi(A\colon U_0).\, \Pi(B\colon U_0).\, A \wedge B \to A,\\
&\mathrm{proj}_2\colon \Pi(A\colon U_0).\, \Pi(B\colon U_0).\, A \wedge B \to B
\end{aligned}$$

konstruierbar sind. Nämlich findet sich

$$\begin{aligned}
&\mathrm{conj} := A \mapsto B \mapsto (a\colon A) \mapsto (b\colon B) \mapsto X \mapsto (f\colon A \to B \to X) \mapsto f(a)(b),\\
&\mathrm{proj}_1 := A \mapsto B \mapsto (c\colon A \wedge B) \mapsto c(A)((a\colon A) \mapsto (b\colon B) \mapsto a),\\
&\mathrm{proj}_2 := A \mapsto B \mapsto (c\colon A \wedge B) \mapsto c(B)((a\colon A) \mapsto (b\colon B) \mapsto b).
\end{aligned}$$

Sollte nicht klar sein, wie diese Terme zustande kommen, kann man zunächst die Herleitungsbäume für die logischen Aussagen erstellen und daraufhin die Beweisterme ergänzen.

11. Maschinengestütztes Beweisen

11.1. Terme und Typen

11.1.1. Zur Aussagenlogik

Tabelle 11.1 zeigt die Funktionen, mit denen Terme für die Konjunktion und die Disjunktion konstruiert werden können.

Ein kurzes Beispiel. Mit der Herleitung

$$\dfrac{\dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \quad \dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash A}}{\Gamma \vdash B \wedge A} \quad \rightsquigarrow \quad \dfrac{\dfrac{\Gamma \vdash h : A \wedge B}{\Gamma \vdash \mathtt{proj2}(h) : B} \quad \dfrac{\Gamma \vdash h : A \wedge B}{\Gamma \vdash \mathtt{proj1}(h) : A}}{\Gamma \vdash \mathtt{conj}(\mathtt{proj2}(h))(\mathtt{proj1}(h)) : B \wedge A}$$

kommt man bezüglich $\Gamma := [h : A \wedge B]$ zum Term

$$(h \mapsto \mathtt{conj}(\mathtt{proj2}(h))(\mathtt{proj1}(h))) : A \wedge B \to B \wedge A.$$

Der Quelltext hierzu:

```
Theorem conj_commutes A B:
  A /\ B -> B /\ A.
Proof.
  exact (fun h => conj (proj2 h) (proj1 h)).
Qed.
```

Tabelle 11.1.: Funktionen für die Konjunktion und die Disjunktion

Junktor	Einführung	Beseitigung
Konjunktion	$\mathtt{conj}\ A\ B : A \to B \to A \wedge B$	$\mathtt{proj1}\ A\ B : A \wedge B \to A$ $\mathtt{proj2}\ A\ B : A \wedge B \to B$
Disjunktion	$\mathtt{or\_introl}\ A\ B : A \to A \vee B$ $\mathtt{or\_intror}\ A\ B : A \to B \vee A$	$\mathtt{or\_elim}\ A\ B : (A \to C) \to (A \to C) \to (A \vee B \to C)$

11.2. Taktiken

Statt Terme zu schreiben, nutzt man lieber Taktiken. Sie bieten den Vorteil, dass die Schlussregeln mit ihnen rückwerts angewendet werden können. Man gibt dabei das zu beweisende Theorem als *Ziel* vor. Vermittels den Taktiken, die zu den Schlussregeln des natürlichen Schließens gehören, kann das Ziel auf ein oder mehrere *Unterziele* zurückgeführt werden. Das geht bis zum Erreichen von Grundsequenzen so weiter.

Mit $\Gamma := [A\colon \mathrm{Prop}, B\colon \mathrm{Prop}]$ gilt

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\Gamma, g\colon \neg B \vdash g\colon \neg B} \quad \dfrac{\dfrac{}{\Gamma, f\colon A \to B \vdash f\colon A \to B} \quad \dfrac{}{\Gamma, a\colon A \vdash a\colon A}\ \text{exact } a}{\Gamma, f\colon A \to B, a\colon A \vdash f(a)\colon B}\ \text{apply } f}{\Gamma, f\colon A \to B, g\colon \neg B, a\colon A \vdash g(f(a))\colon \bot}\ \text{apply } g}{\Gamma, f\colon A \to B, g\colon \neg B \vdash a \mapsto g(f(a))\colon \neg A}\ \text{intro } a}{\Gamma, f\colon A \to B \vdash g \mapsto a \mapsto g(f(a))\colon \neg B \to \neg A}\ \text{intro } g}{\Gamma \vdash f \mapsto g \mapsto a \mapsto g(f(a))\colon (A \to B) \to (\neg B \to \neg A)}\ \text{intro } f.$$

Man beachte, dass der konstruierte Term $f \mapsto g \mapsto a \mapsto g(f(a))$ am Anfang noch unbekannt ist. Bekannt sind im jeweiligen Schritt lediglich die zur Verfügung stehenden Variablen und der Typ, dessen Term zu konstruieren ist. Erst nachdem man sich rückwärts von der Wurzel aus zu den Blättern hin durchgearbeitet hat, kann man den Weg zur Wurzel hin zurücklaufen und dabei schrittweise den Terme konstruieren. In der Praxis gibt man sich in der Regel zufrieden, bei den Blättern angekommen zu sein. Der tatsächliche Verlauf sieht also so aus:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\Gamma, g\colon \neg B \vdash g\colon \neg B} \quad \dfrac{\dfrac{}{\Gamma, f\colon A \to B \vdash f\colon A \to B} \quad \dfrac{}{\Gamma, a\colon A \vdash ?\colon A}\ \text{exact } a}{\Gamma, f\colon A \to B, a\colon A \vdash ?\colon B}\ \text{apply } f}{\Gamma, f\colon A \to B, g\colon \neg B, a\colon A \vdash ?\colon \bot}\ \text{apply } g}{\Gamma, f\colon A \to B, g\colon \neg B \vdash ?\colon \neg A}\ \text{intro } a}{\Gamma, f\colon A \to B \vdash ?\colon \neg B \to \neg A}\ \text{intro } g}{\Gamma \vdash ?\colon (A \to B) \to (\neg B \to \neg A)}\ \text{intro } f$$

Der Quelltext hierzu:

```
Theorem contraposition (A B: Prop):
  (A -> B) -> (~B -> ~A).
Proof.
  intro f. intro g. intro a.
  apply g. apply f. exact a.
Qed.
```

A. Formelsammlung

A.1. Logik

Natürliches Schließen

	Einführung	Beseitigung
$\wedge$	$\dfrac{\Gamma \vdash A \qquad \Gamma' \vdash B}{\Gamma, \Gamma' \vdash A \wedge B}$	$\dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad \dfrac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}$
$\vee$	$\dfrac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad \dfrac{\Gamma \vdash B}{\Gamma \vdash A \vee B}$	$\dfrac{\Gamma \vdash A \vee B \qquad \Gamma', A \vdash C \qquad \Gamma'', B \vdash C}{\Gamma, \Gamma', \Gamma'' \vdash C}$
$\Rightarrow$	$\dfrac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B}$	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash B}$
$\neg$	$\dfrac{\Gamma, A \vdash \bot}{\Gamma \vdash \neg A}$	$\dfrac{\Gamma \vdash \neg A \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash \bot}$
$\Leftrightarrow$	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash B \Rightarrow A}{\Gamma, \Gamma' \vdash A \Leftrightarrow B}$	$\dfrac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash A \Rightarrow B} \quad \dfrac{\Gamma \vdash A \Leftrightarrow B}{\Gamma \vdash B \Rightarrow A}$
$\forall$	$\dfrac{\Gamma \vdash A}{\Gamma \vdash \forall x \colon A}\,(x \notin \mathrm{FV}(\Gamma))$	$\dfrac{\Gamma \vdash \forall x \colon A}{\Gamma \vdash A[x := t]}$
$\exists$	$\dfrac{\Gamma \vdash A[x := t]}{\Gamma \vdash \exists x \colon A}$	$\dfrac{\Gamma \vdash \exists x \colon A \qquad \Gamma', A \vdash B}{\Gamma, \Gamma' \vdash B}\,(x \notin \mathrm{FV}(\Gamma, \Gamma', B))$

Name	Kürzel	Regel	Schema	Logik
Ex falso quodlibet	EFQ	$\dfrac{\Gamma \vdash \bot}{\Gamma \vdash A}$	$\bot \Rightarrow A$	intuitionistisch
Satz vom ausgeschlossenen Dritten	LEM	$\dfrac{}{\vdash A \vee \neg A}$	$A \vee \neg A$	klassisch
Beseitigung der Doppelnegation	DNE	$\dfrac{\Gamma \vdash \neg\neg A}{\Gamma \vdash A}$	$\neg\neg A \Rightarrow A$	klassisch

Zulässige Regeln der klassischen Logik

Name	Regel	Schema
Modus ponens	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash B}$	$(A \Rightarrow B) \wedge A \Rightarrow B$
Schnittregel	$\dfrac{\Gamma, A \vdash B \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash B}$	$(A \Rightarrow B) \wedge A \Rightarrow B$
Kettenschluss	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash B \Rightarrow C}{\Gamma, \Gamma' \vdash A \Rightarrow C}$	$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$
Modus tollens	$\dfrac{\Gamma \vdash A \Rightarrow B \qquad \Gamma' \vdash \neg B}{\Gamma, \Gamma' \vdash \neg A}$	$(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A$
Kontraposition	$\dfrac{\Gamma \vdash A \Rightarrow B}{\Gamma \vdash \neg B \Rightarrow \neg A}$	$(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$
umgekehrte Kontraposition	$\dfrac{\Gamma \vdash \neg B \Rightarrow \neg A}{\Gamma \vdash A \Rightarrow B}$	$(\neg B \Rightarrow \neg A) \Rightarrow (A \Rightarrow B)$
Modus tollendo ponens	$\dfrac{\Gamma \vdash A \vee B \qquad \Gamma' \vdash \neg A}{\Gamma, \Gamma' \vdash B}$	$(A \vee B) \wedge \neg A \Rightarrow B$
Modus ponendo tollens	$\dfrac{\Gamma \vdash \neg(A \wedge B) \qquad \Gamma' \vdash A}{\Gamma, \Gamma' \vdash \neg B}$	$\neg(A \wedge B) \wedge A \Rightarrow \neg B$
Resolution	$\dfrac{\Gamma \vdash A \vee B \qquad \Gamma' \vdash \neg A \vee C}{\Gamma, \Gamma' \vdash B \vee C}$	$(A \vee B) \wedge (\neg A \vee C) \Rightarrow B \vee C$
Exportation	$\dfrac{\Gamma \vdash A \wedge B \Rightarrow C}{\Gamma \vdash A \Rightarrow (B \Rightarrow C)}$	$(A \wedge B \Rightarrow C) \Rightarrow (A \Rightarrow (B \Rightarrow C))$
Umkehrung der Exportation	$\dfrac{\Gamma \vdash A \Rightarrow (B \Rightarrow C)}{\Gamma \vdash A \wedge B \Rightarrow C}$	$(A \Rightarrow (B \Rightarrow C)) \Rightarrow (A \wedge B \Rightarrow C)$
Absorption	$\dfrac{\Gamma \vdash A \Rightarrow B}{\Gamma \vdash A \Rightarrow A \wedge B}$	$(A \Rightarrow B) \Leftrightarrow (A \Rightarrow A \wedge B)$
Abschwächung	$\dfrac{\Gamma \vdash B}{\Gamma, A \vdash B}$	$B \Rightarrow (A \Rightarrow B)$
Ersetzungsregel	$\dfrac{\Gamma \vdash A \Leftrightarrow B \qquad \Gamma' \vdash C(A)}{\Gamma, \Gamma' \vdash C(B)}$	$(A \Leftrightarrow B) \wedge C(A) \Rightarrow C(B)$
Reductio ad absurdum	$\dfrac{\Gamma, \neg A \vdash \neg B \qquad \Gamma', \neg A \vdash B}{\Gamma, \Gamma' \vdash A}$	$(\neg A \Rightarrow B) \wedge (\neg A \Rightarrow \neg B) \Rightarrow A$

A.2. Mengenlehre

Mengenbildungen

$u \in \mathcal{U} :\Leftrightarrow \exists C\colon u \in C$	$u \in \{x \mid A(x)\} :\Leftrightarrow u \in \mathcal{U} \wedge A(u)$
$A \subseteq B :\Leftrightarrow \forall x \in A\colon x \in B$	$u \in \{x \in M \mid A(x)\} :\Leftrightarrow u \in M \wedge A(u)$
$\mathcal{P}(M) := \{A \mid A \subseteq M\}$	$u \in \{x_1, \ldots, x_n\} :\Leftrightarrow u = x_1 \vee \ldots \vee u = x_n$
$A \cap B := \{x \mid x \in A \wedge x \in B\}$	$\bigcap \mathcal{A} := \{x \mid \forall A \in \mathcal{A}\colon x \in A\}$
$A \cup B := \{x \mid x \in A \vee x \in B\}$	$\bigcup \mathcal{A} := \{x \mid \exists A \in \mathcal{A}\colon x \in A\}$
$A \setminus B := \{x \mid x \in A \wedge x \notin B\}$	$\bigcap_{i\in I} A_i := \{x \mid \forall i \in I\colon x \in A_i\}$
$A \triangle B := (A \setminus B) \cup (B \setminus A)$	$\bigcup_{i\in I} A_i := \{x \mid \exists i \in I\colon x \in A_i\}$
$A \sqcup B := (\{1\} \times A) \cup (\{2\} \times B)$	$A \times B := \{t \mid \exists x \in A\colon \exists y \in B\colon t = (x, y)\}$
$\bigsqcup_{i\in I} A_i := \bigcup_{i\in I}(\{i\} \times A_i)$	$\prod_{i\in I} A_i := \{f\colon I \to \bigcup_i A_i \mid \forall i\colon f(i) \in A_i\}$

Boolesche Algebra

Schnitt	Vereinigung	Bezeichnung
$A \cap \emptyset = \emptyset$	$A \cup G = G$	Extremalgesetze
$A \cap A^c = \emptyset$	$A \cup A^c = G$	Komplementärgesetze
$A \cap A = A$	$A \cup A = A$	Idempotenzgesetze
$A \cap G = A$	$A \cup \emptyset = A$	Neutralitätsgesetze
$A \cap B = B \cap A$	$A \cup B = B \cup A$	Kommutativgesetze
$A \cap (B \cap C) = (A \cap B) \cap C$	$A \cup (B \cup C) = (A \cup B) \cup C$	Assoziativgesetze
$(A \cap B)^c = A^c \cup B^c$	$(A \cup B)^c = A^c \cap B^c$	De morgansche Gesetze
$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$	Absorptionsgesetze

Hierbei soll $A, B, C \subseteq G$ und $A^c := G \setminus A$ sein, G die Grundmenge.

Distributivgesetze

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

Literaturverzeichnis

[1] Gerhard Gentzen: *Untersuchungen über das logische Schließen.* In: *Mathematische Zeitschrift.* Band 39, 1935, S. 176–210, S. 405–431.

[2] Gerhard Gentzen: *Die Widerspruchsfreiheit der reinen Zahlentheorie.* In: *Mathematische Annalen.* Band 112, 1936, S. 493–565.

[3] Gerhard Gentzen: *Die gegenwärtige Lage in der mathematischen Grundlagenforschung. Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie.* In: *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften.* Heft 4, S. Hirzel, Leipzig 1938.

[4] Eckart Menzler-Trott: *Gentzens Problem. Mathematische Logik im nationalsozialistischen Deutschland.* Birkhäuser, Basel 2001.

[5] Ingebrigt Johansson: *Der Minimalkalkül, ein reduzierter intuitionistischer Formalismus.* In: *Compositio Mathematica.* Band 4, 1937, S. 119–136.

[6] Hannes Diener, Maarten McKubre-Jordens: *Classifying Material Implications over Minimal Logic.* In: *Archive for Mathematical Logic.* Band 59, 2020, S. 905–924. doi:10.1007/s00153-020-00722-x.

[7] Jan von Plato: *Elements of Logical Reasoning.* Cambridge University Press, Cambridge 2013.

[8] Jan von Plato: *Gentzen's Logic.* In: *Handbook of The History of Logic.* Band 5, North-Holland, 2009.

[9] Francis Jeffry Pelletier, Allen P. Hazen: *A History of Natural Deduction.* In: *Handbook of The History of Logic.* Band 11, North-Holland, 2012.

[10] Francis Jeffry Pelletier, Allen Hazen: *Natural Deduction Systems in Logic.* In: *The Stanford Encyclopedia of Philosophy.*

[11] Andrzej Indrzejczak: *Natural Deduction.* In: *The Internet Encyclopedia of Philosophy.*

[12] Samuel Mimram: *Program = Proof*. Laboratoire d'Informatique de l'Ecole polytechnique, Palaiseau 2020.

[13] Dirk W. Hoffmann: *Grenzen der Mathematik*. Springer, Berlin 2011.

[14] Heinz-Dieter Ebbinghaus, Jörg Flum, Wolfgang Thomas: *Einführung in die mathematische Logik*. Springer Spektrum, 1978, 6. Auflage 2018. doi:10.1007/978-3-662-58029-5.

[15] Open Logic Project: *The Open Logic Text*. Complete Build, Oktober 2022.

[16] Jeremy Avigad: *Mathematical Logic and Computation*. Cambridge University Press, 2023.

[17] Jeremy Avigad: *Foundations*. September 2021. arXiv:2009.09541v4.

[18] Heinrich Wansing (Hrsg.): *Dag Prawitz on Proofs and Meaning*. In: *Outstanding Contributions to Logic*. Springer, 2015. doi:10.1007/978-3-319-11041-7.

[19] Herbert B. Enderton: *A Mathematical Introduction to Logic*. Academic Press, New York 1972, 2. Auflage 2001.

[20] Paul Taylor: *Practical Foundations of Mathematics*. Cambridge University Press, 1999.

[21] Patrick Blackburn, Johan van Benthem: *Modal Logic: A Semantic Perspective*. In: *Studies in Logic and Practical Reasoning*. Band 3, 2007, S. 1–84.

[22] Thierry Coquand: *Type Theory*. In: *The Stanford Encyclopedia of Philosophy*.

[23] *Homotopy Type Theory. Univalent Foundations of Mathematics*. The Univalent Foundations Program, 2013.

[24] Christine Paulin-Mohring: *Introduction to the Calculus of Inductive Constructions*. In: Bruno Woltzenlogel Paleo, David Delahaye: *All about Proofs, Proofs for All*. In: *Studies in Logic*. Band 55, 2015.

[25] Jonathan P. Seldin: *Coquand's Calculus of Constructions: A Mathematical Foundation for a Proof Development System*. In: *Formal Aspects of Computing*. Band 4, 1992, S. 425–441. doi:10.1007/BF01211392.

[26] Oliver Deiser: *Grundbegriffe der Mathematik. Sprache, Zahlen und erste Erkundungen*. März 2021, letzte Version Oktober 2022.

[27] Georg Cantor: *Beiträge zur Begründung der transfiniten Mengenlehre.* In: *Mathematische Annalen.* Band 46, 1895, S. 481.

[28] Abraham Adolf Fraenkel: *Einleitung in die Mengenlehre.* Springer, Berlin 1919, 3. Auflage 1928.

[29] Oliver Deiser: *Einführung in die Mengenlehre.* Springer, 2002, 3. Auflage 2010.

[30] Heinz-Dieter Ebbinghaus: *Einführung in die Mengenlehre.* Springer Spektrum, 5. Auflage 2021.

[31] Herbert B. Enderton: *Elements of Set Theory.* Academic Press, New York 1977.

[32] Michael A. Shulman: *Set theory for category theory.* Oktober 2008. arXiv:0810.1279.

[33] Thomas Jech: *Set Theory: The Third Millennium Edition, revised and expanded.* Springer, 2002. doi:10.1007/3-540-44761-X.

[34] Ben Siraphob: *A non-trivial trivial theorem: doing classical mathematics in Coq.* Blogpost, 27. Juni 2021.

[35] Stefan Müller-Stach: *Richard Dedekind: Was sind und was sollen die Zahlen? Stetigkeit und Irrationale Zahlen.* Springer, 2017.

[36] Paul Lorenzen: *Die Definition durch vollständige Induktion.* In: *Monatshefte für Mathematik und Physik.* Band 47, 1939, S. 356–358.

[37] Paul R. Halmos: *Naive Set Theory.* Springer, 1974.

[38] Klaus Mainzer: *Natürliche, ganze und rationale Zahlen.* In: Heinz-Dieter Ebbinghaus u. a.: *Zahlen.* Springer, 1983, 3. Auflage 1992.

[39] Christoph Lamm: *Karl Grandjot und der Dedekindsche Rekursionssatz.* In: *Mitteilungen der DMV.* Band 24, 2016, S. 37–45.

[40] Tobias Glosauer: *Elementar(st)e Gruppentheorie.* Springer, 2016.

[41] Norbert Henze: *Stochastik für Einsteiger.* Springer, 1997, 12. Auflage 2018.

[42] Glynn Winskel: *The Formal Semantics of Programming Languages.* The MIT Press, Cambridge (Massachusetts) 1993.

[43] David Harel, Dexter Kozen, Jerzy Tiuryn: *Dynamic Logic.* The MIT Press, Cambridge (Massachusetts) 2000.

Index