Notizen zur Vorlesung „Mathematische Logik“

Markus Junker

Sommersemester 2023

Inhaltsverzeichnis

1 Einstimmung, Geschichte, Motivation

> *Denn die Pioniere der Mathematik hatten sich von gewissen Grundlagen brauchbare Vorstellungen gemacht, aus denen sich Schlüsse, Rechnungsarten, Resultate ergaben, deren bemächtigten sich die Physiker, um neue Ergebnisse zu erhalten, und endlich kamen die Techniker, nahmen oft bloß die Resultate, setzten neue Rechnungen darauf und es entstanden die Maschinen. Und plötzlich, nachdem alles in schönste Existenz gebracht war, kamen die Mathematiker – jene, die ganz innen herumgrübeln, – darauf, daß etwas in den Grundlagen der ganzen Sache absolut nicht in Ordnung zu bringen sei; tatsächlich, sie sahen zuunterst nach und fanden, daß das ganze Gebäude in der Luft stehe. Aber die Maschinen liefen! Man muß daraufhin annehmen, daß unser Dasein bleicher Spuk ist; wir leben es, aber eigentlich nur auf Grund eines Irrtums, ohne den es nicht entstanden wäre. Es gibt heute keine zweite Möglichkeit so phantastischen Gefühls wie die des Mathematikers.* [1]

Die Logik ist eine eigene, ursprünglich eher philosophische Wissenschaftsdisziplin, historisch gerne als „die Lehre vom korrekten Schließen“ beschrieben. Betrieben wurde sie vor allem als Propädeutik, also als eine Grundlage wissenschaftlichen Arbeitens. Es ist klar, dass die Mathematik mit ihrem Anspruch, mathematische Aussagen nicht nur zu finden, sondern auch zu beweisen, in prominenter Weise auf einem Verständnis von korrekten Schlüssen beruht.

Traditionell wird zwischen *Form* und *Inhalt* von Aussagen unterschieden. Die logische Korrektheit eines Schlusses sollte sich allein aus der Form der beteiligten Aussagen ergeben und nicht aus den jeweiligen spezifischen Inhalten, weshalb oft von *formaler Logik* gesprochen wird (was also noch nichts mit Formeln zu tun hat).

Als Begründer der formalen Logik gilt Aristoteles (384–322). Er hat ein (vollständiges) System von *Syllogismen* genannten logischen Schlussweisen aufgestellt, in denen grob gesprochen aus dem Verhältnis zweier Begriffe/Eigenschaften P und Q einerseits und Q und R andererseits auf das Verhältnis zwischen P und R geschlossen wird, beispielsweise:

> Alle, die Logik, können, sind schlau.
> Es gibt Studierende, die nicht schlau sind.
> Also gibt es Studierende, die keine Logik können.

Etwa ein Jahrhundert später findet sich bei der Stoa (z. B. Chrysippos, ca. 280–208) große Teile dessen, was heute klassische Aussagenlogik genannt wird und dem mathematischen Arbeiten zu Grunde liegt. In den Schlüssen der Stoa werden zwei Aussagen wahrheitswertfunktional verbunden und aus der Wahrheit oder Falschheit der einen Aussage auf die Wahrheit oder Falschheit der anderen geschlossen, beispielsweise:

> Wenn Joe Biden jünger als 75 Jahre ist, dann wurde er nach 1945 geboren.
> Joe Biden wurde vor 1945 geboren.
> Also ist Joe Biden nicht jünger als 75 Jahre.

Aus moderner Sicht gab es dann zweitausend Jahre lang keine entscheidenden Neuerungen. Allerdings entwickelt Ramon Llull (1232–1316) die Idee einer „logischen Maschine“, mit der durch systematische Kombinationen von Begriffen alle wahren Aussagen gefunden werden können sollen. Diese Idee wird von Gottfried Wilhelm Leibniz (1646–1716) aufgegriffen, der an

[1] Aus: Robert Musil „Der mathematische Mensch“, 1913.
Zitiert nach: Mitteilungen der Deutschen Mathematiker-Vereinigung, Band 20, Heft 1, 2012.
`https://www.degruyter.com/document/doi/10.1515/dmvm-2012-0020/pdf`

einem „Alphabets des menschlichen Denkens“ arbeitet, „dessen ‚Buchstabenkombinationen‘ alle menschlichen Begriffe mechanisch auf Grundbegriffe zurückführt, mit denen man alle wahren Sätze mechanisch erhält“. [2]

Einen Neuanfang für die Logik gab es mit der Entdeckung von George Boole (1815–1864), dass man Syllogistik und Ausssagenlogik algebraisieren kann, also so formalisieren kann, dass man mit den entsprechenden Rechengesetzen die Gültigkeit logischer Schlussweisen „ausrechnen“ kann. Die Strukturen, in denen diese Rechnungen ausgeführt werden, wurden daher *Boole'sche Algebren* genannt. Mit Boole wird die formale Logik zu einer *formalisierten Logik* und die theoretische Grundlage von außermathematischen Anwendungen universeller Rechenmaschinen wie der *analytic engine* von Charles Babbage (1791–1871) ist gelegt.

Gottlob Frege (1848–1925) mit der „Begriffsschrift“ von 1879 und unabhängig davon Charles Sanders Peirce (1839–1914) haben die formalisierte Logik dann neu geschaffen bzw. ausgedehnt auf das, was heute *Prädikatenlogik* heißt. In der Prädikatenlogik kann man die gesamte Mathematik formalisieren und erstmals einen präzisen Definition davon angeben, was ein mathematischer Beweis sein sollte. Auch Giuseppe Peano (1858–1932) hat wichtige Beiträge geleistet.

Unabhängig davon gab es im 19. Jahrhundert eine Reihe von mathematischen Entwicklungen, die gezeigt haben, dass man solche soliden Grundlagen für das mathematische Arbeiten braucht und nicht allein auf Intuition oder nachprüfbare Ergebnisse vertrauen kann. Stichworte hierfür sind: Kontrolle über die Berechnung und Vertauschbarkeit von Grenzwerten; Entdeckung stetiger, nicht differenzierbarer Funktionen oder raumfüllernder stetiger Kurven; vor allem aber die Entwicklung der Mengenlehre durch Georg Cantor (1845–1918), in der man mit den „Paradoxien des Unendlichen“ umgehen muss. Der letztlich entscheidende Moment war allerdings der Versuch von Gottlob Frege, eine axiomatische Grundlegung der Arithmetik zu schaffen. Aus seinen Axiomen ließ sich die Zermelo-Russsell'sche Antinomie (Ernst Zermelo 1971–1953, Bertrand Russell 1872–1970) herleiten: die Existenz der in sich widersprüchlichen Menge $\{M \mid M \notin M\}$, also einer Menge, deren Elemente genau alle Mengen sind, die sich nicht selbst enthalten.

Darüber, welche Konsequenzen daraus zu ziehen wären und welche Methoden in der Mathematik zugelassen sein sollten, gab es Kontroversen. David Hilbert (1862–1943) wollte mit unumstrittenen mathematischen Methoden zeigen, dass umstrittene Methoden nicht zu Widersprüchen führten. Dazu hat er logische Formeln und formale Beweise selbst als mathematische Objekte aufgefasst und angefangen, Sätze über sie zu beweisen. Dies ist die eigentliche Geburtsstunde der Mathematischen Logik als mathematischer Teildisziplin.

Kurt Gödel (1906–1978) hat mit seinen Theoremen (Vollständigkeitssatz, Unvollständigskeitssätze) die Möglichkeiten und Grenzen des Hilbert'schen Programms ausgelotet. Sie gelten als Hauptergebnisse der Mathematischen Logik. In diesem Zusammenhang wurden auch grundlegende Fragen über die Möglichkeiten und Grenzen der Berechenbarkeit gestellt und beantwortet und damit die Theoretische Informatik begründet (wichtige Namen hierfür: Emil Post 1897–1954, Alonzo Church 1903–1995, Stephen Kleene 1909–1994, Alan Turing 1912–1954).

Durchgesetzt als Grundlage für die mathematische Praxis hat sich einerseits die klassische zweiwertige Logik und andererseits das im Wesentlichen von Zermelo entwickelte mengentheoretische Axiomensystem ZFC. Mit der Möglichkeit, dass dieses widersprüchlich sein könnte, hat man sich arrangiert, und vertraut darauf, dass der Kern der Mathematik davon nicht berührt wäre.

Aus den Fragen und Techniken im Umfeld des Hilbert'schen Programms und der Gödel'schen Sätze haben sich eigenständige mathematische Teilgebiete entwickelt: Mengenlehre, Beweistheo-

[2] Zitat nach https://de.wikipedia.org/wiki/Gottfried_Wilhelm_Leibniz

rie, Rekursions- oder Berechenbarkeitstheorie und Modelltheorie (Alfred Tarski, 1901–1983). Sie werden zur Mathematischen Logik gezählt, haben sich aber von den Fragen der Grundlegung der Mathematik weitgehend entfernt. Gemeinsam ist ihnen die logische Sprache, also das Nutzen der Möglichkeit, Mathematik zu formalisieren.

Zunächst einmal unabhängig von dem bisher aufgezeigten Entwicklungsstrang steht die Modallogik. Bereits Aristoteles hat mit Modalitäten versehene Aussagen wie „*Notwendigerweise sind alle Griechen sterblich*“ oder „*Möglicherweise sind alle Griechen sterblich*“ untersucht (neben modal unbestimmten Aussagen wie „*Alle Griechen sind sterblich*“) und ein System modaler Syllogismen etnwickelt Bis heute gibt es allerdings kein allgemein akzeptiertes Verständis von Aristoteles' Modallogik.

Die moderne, axiomatische Modallogik wurde von Clarence Irving Lewis (1883–1964) ins Leben gerufen aus der Motivation heraus, eine Logik mit einer Implikation zu finden, die näher am intuitiven Folgerungsbegriff ist als die „materiale Implikation“ der klassischen Aussagenlogik. Durch die „Semantik möglicher Welten“ von Saul Kripke (1940–2022) hat die Modallogik einen großen Aufschwung genommen und insbesondere in der Informatik breite Anwendungen gefunden. Abgesehen davon, dass die Modallogik eine schöne Theorie mit interessanten Ergebnissen bietet, ist sie für die Mathematische Logik unter anderem deshalb interessant, weil sie die Modellisierung von Alternativen zur klassischen Aussagenlogik wie dem Intuitionismus zulässt, der von L. E. J. Brouwer (1881–1966) als Ausweg aus der oben beschriebenen Grundlagenkrise der Mathematik entwickelt wurde, sich in der Mathematik aber nicht durchsetzen konnte. Auch die Ergebnisse im Umfeld des Zweiten Gödel'schen Unvollständigkeitssatzes lassen sich in modallogischer Sprache schön beschreiben, weshalb die Vorlesung mit einem kleinen Einblick in die Modallogik schließt.

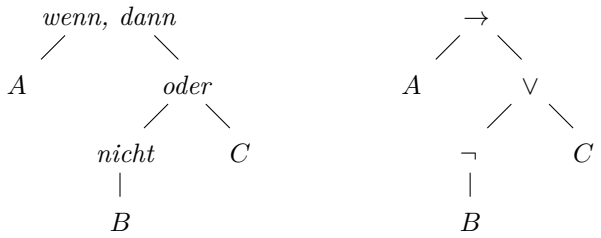
2 „Modell“-Logik: die Aussagenlogik

Die (klassische zweiwertige) Aussagenlogik betrachtet das Verhalten von Aussagen – also von Sätzen, die wahr oder falsch sein können – bezüglich dieser beiden Wahrheitswerte. Man interessiert sich insbesondere für „wahrheitswertfunktionale“ Zusammensetzungen von Teilaussagen.

Beispiel: *Wenn f eine reelle Funktion mit kompaktem Definitionsbereich ist, dann ist f nicht stetig oder f ist beschränkt.*

Teilaussagen:	Abkürzung:
f ist eine reelle Funktion mit kompaktem Definitionsbereich.	A
f ist stetig.	B
f ist bechränkt.	C
Zusammensetzungen:	
wenn ... dann ...	$\to$
nicht ...	$\neg$
... oder ...	$\vee$

Die Struktur des Gesamtsatzes versteht man am besten als *Baum*:



Es gibt mehrere verbreitete Möglichkeiten, solch einen Baum als *Symbolfolge/Zeichenkette* zu schreiben:

- Die *Infix-Notation* mit verschiedenen Klammerungsvarianten; hier $(A \to (\neg B \vee C))$
- die klammerfreie *Polnische Notation* $\to A \vee \neg BC$
- bzw. die *Umgekehrte Polnische Notation* $AB\neg C\vee \to$.

Die formalisierte Aussagenlogik führt für diese Analyse eine formale Sprache ein, für die man zunächst das *Alphabet* angeben muss, also welche Zeichen/Symbole vorkommen können, und dann die *Grammatik*, also die Regeln, nach denen aussagenlogische Formeln gebildet werden.

2.1 Syntax

Alphabet

Das Alphabet besteht aus den beiden Klammern (und) , den beiden *Aussagenkonstanten Verum* $\top$ und *Falsum* $\bot$, den fünf *Junktoren* (*Negation-*, *Konjunktions-*, *Disjunktions-*, *Implikations-* und *Äquivalenzjunktor*)

$$\neg \qquad \wedge \qquad \vee \qquad \to \qquad \leftrightarrow$$

sowie abzählbar unendlich vielen *Aussagenvariablen*[3]

$$A_0 \qquad A_1 \qquad A_2 \qquad A_3 \qquad A_4 \qquad A_5 \qquad \dots$$

[3] Jede Aussagenvariable, also z. B. auch $A_{87232837}$, zählt als *ein* Symbol.

Die Zeichen $\neg, \wedge, \vee$ werden in der Mathematischen Logik ausschließlich in den aussagenlogischen Formeln benutzt (sogenannte *objektsprachliche Zeichen*) und nicht im mathematischen Reden über Formeln (als *metasprachliche Zeichen*). Das ist keine große Einschränkung, da man die kurzen Wörter „nicht“, „und“, „oder“ nicht unbedingt durch Zeichen abkürzen muss.

Ich benutzte $A_i, A_j, \ldots$ als Variablen für Aussagenvariablen.

Grammatik

Konzeptionell ist es am besten, aussagenlogische Formeln als Bäume zu definieren. Platzsparender und verbreiteter ist allerdings die Schreibweise in Infix-Notation. Für Beweise praktisch ist bisweilen die Polnische Notation. Ich werde bei Bedarf frei zwischen den Schreibweisen wechseln und gebe die Regeln in diesen drei Varianten an.

Definition 2.1 *Aussagenlogische Formeln sind* etikettierte, orientierte Wurzelbäume[4] *bzw. in Infix- und Polnischer Notation* Zeichenfolgen*, die durch endliche Anwendung der folgenden Regeln gewonnen werden können:*

- $\top$ *und* $\bot$ *und jede Aussagenvariable bilden jeweils eine aussagenlogische Formel* (als Wurzel eines einelementigen Baums bzw. als einelementige Zeichenfolge)*;*
- *Für jede aussagenlogische Formel* φ *ist*

  $$\begin{array}{c}\neg\\ |\\ \varphi\end{array}$$

  bzw. in Infix- und Polnischer Notation $\neg\varphi$ *eine aussagenlogische Formel.*
  (Gemeint ist damit der Baum, der $\neg$ als Wurzel hat, unterhalb derer der Baum φ hängt, bzw. die Zeichenfolge aus dem Symbol $\neg$ gefolgt von der Zeichenkette, für die φ steht.)
- *Für alle aussagenlogischen Formeln* φ_1 *und* φ_2 *sind*

  $$\begin{array}{cccccccc} \wedge & & \vee & & \to & & \leftrightarrow & \\ / & \backslash & / & \backslash & / & \backslash & / & \backslash \\ \varphi_1 & \varphi_2 & \varphi_1 & \varphi_2 & \varphi_1 & \varphi_2 & \varphi_1 & \varphi_2 \end{array}$$

  bzw. in Infix-Notation

  $$(\varphi_1 \wedge \varphi_2) \qquad (\varphi_1 \vee \varphi_2) \qquad (\varphi_1 \to \varphi_2) \qquad (\varphi_1 \leftrightarrow \varphi_2)$$

  bzw. in Polnischer Notation

  $$\wedge\varphi_1\varphi_2 \qquad \vee\,\varphi_1\varphi_2 \qquad \to\,\varphi_1\varphi_2 \qquad \leftrightarrow\,\varphi_1\varphi_2$$

  aussagenlogische Formeln.

Die Anzahl der Formeln, die durch einen Junktor zu einer neuen Formel verbunden werden, heißt Stelligkeit *des Junktors:* $\neg$ *ist* einstelliger Junktor*,* $\wedge, \vee, \to, \leftrightarrow$ *sind* zweistellige Junktoren*.*

[4] *Bäume* sind zusammenhängende, zykelfreie Graphen; *Wurzelbäume* haben zusätzliche eine ausgezeichnete Ecke. *Etikettiert* bedeutet hier, dass jeder Ecke des Graphen ein Symbol aus dem Alphabet zugeordnet ist. *Orientiert* bedeutet, dass die Nachbarn einer Ecke auf feste Weise angeordnet sind.
Ausführlicherer Definitionen dieser Begriffe findet man z. B. in meinem Skript „Logik für Studierende der Informatik“; für diese Vorlesung reicht allerdings das intuitive Verständnis.

Induktion über den Aufbau der Formeln

Formeln sind *induktiv* (man sagt auch: *rekursiv*) aufgebaut. Es sollte klar sein, dass die Definition als Baum die Eindeutigkeit des Aufbaus sicherstellt: Aus dem Baum kann man die sukzessiven Konstruktionsschritte der Formel ablesen.

Definitionen und Beweise müssen häufig diesen Aufbauprozess nachvollziehen; man nenne dies Definitionen bzw. Beweise „per Induktion über den Aufbau der Formel“. Definition 2.2 und der Beweist zu Satz 2.3 liefern erste Beispiele. Formeln als Wurzelbäume haben eine *Höhe*, nämlich den maximalen Abstand einer Ecke von der Wurzel. Man kann die Induktion über den Aufbau der Formel als „normale“ Induktion über die Höhe von Formeln verstehen.

Definition 2.2 Teilformeln *einer aussagenlogischen Formel φ sind diejenigen aussagenlogischen Formeln, die im Aufbauprozess von φ vorkommen. Präzise ist die Menge der Teilformeln einer aussagenlogischen Formel φ per Induktion über den Aufbau der Formeln wie folgt definiert:*

- $\text{Teilformeln}(A_i) = \{A_i\}$, $\text{Teilformeln}(\top) = \{\top\}$ *und* $\text{Teilformeln}(\bot) = \{\bot\}$.
- $\text{Teilformeln}(\neg\psi) = \{\neg\psi\} \cup \text{Teilformeln}(\psi)$
- *für jeden zweistelligen Junktor $* \in \{\wedge, \vee, \to, \leftrightarrow\}$ ist*
  $\text{Teilformeln}(\,(\psi_1 * \psi_2)\,) = \{\,(\psi_1 * \psi_2)\,\} \cup \text{Teilformeln}(\psi_1) \cup \text{Teilformeln}(\psi_2)$.

Man schreibt $\varphi(A_{i_1}, \ldots, A_{i_n})$ für eine aussagenlogische Formel φ mit der Eigenschaft, dass alle Aussagenvariablen, die Teilformeln von φ sind, sich unter $A_{i_1}, \ldots, A_{i_n}$ befinden.

Satz 2.3 (Eindeutige Lesbarkeit) *Aussagenlogische Formeln in Infix- und in Polnischer Notation sind* eindeutig lesbar*, d. h. es gibt genau eine aussagenlogische Formel als Baum, dem die gegebene Infix- bzw. Polnischer Notation entspricht.*

Beweis für die Polnische Notation: Man braucht dafür folgendes Lemma.

Lemma 2.4 *Kein echtes Anfangsstück einer aussagenlogischen Formel in Polnischer Notation ist selbst eine aussagenlogischen Formel in Polnischer Notation.*

Man spart sich etwas Schreibarbeit, wenn man dieses Lemma zusammen mit der eindeutigen Lesabrkeit per Induktion über den Aufbau der Formeln beweist:

- Für $\varphi = A_i, \top, \bot$ ist die eindeutige Lesbarkeit offensichtlich.
  Echtes Anfangsstück ist jeweils nur die leere Zeichenfolge, die keine Formel ist.
- Für $\varphi = \neg\psi$ ist klar, dass φ nur durch Anwenden der Regel für den Negationsjunktor entstanden sein kann. Also ist ψ eine aussagenlogische Formel, die per Induktion eindeutig lesbar ist. Damit ist auch φ eindeutig lesbar.
  Echte Anfangsstücke sind die leere Zeichenfolge und Folgen der Form $\neg\psi'$ für ein echtes Anfangsstück ψ' von ψ. Nach Induktion ist ψ' keine aussagenlogische Formel. Dann ist auch $\neg\psi'$ keine aussagenlogische Formel, da sie als solche nur durch Anwenden der Regel für den Negationsjunktor aus ψ' entstanden sein könnte.
- Für $\varphi = *\psi_1\psi_2$ mit $* \in \{\wedge, \vee, \to, \leftrightarrow\}$ ist klar, dass φ nur durch Anwenden der Regel für den Junktor $*$ entstanden sein kann. Angenommen es gilt auch $\varphi = *\varphi_1\varphi_2$. Wenn $\psi_1 \neq \varphi_1$, dann ist entweder φ_1 ein echtes Anfangsstück von ψ_1 oder umgekehrt, was per Induktion nicht möglich ist.

Echte Anfangsstücke von φ sind neben der leeren Symbolfolge, die keine Formel ist, Zeichenfolgen $*\psi_1'$ und $*\psi_1\psi_2'$ mit echten Anfangsstücken ψ_i' von ψ_i. Wäre solch ein Anfangsstück von der Form $*\varphi_1\varphi_2$ mit aussagenlogischen Formeln φ_i, dann wäre entweder φ_1 echtes Anfangsstück von ψ_1, oder $\varphi_1 = \psi_1$ und φ_2 echtes Anfangsstück von ψ_2, oder ψ_1 echtes Anfangsstück von φ_1, was alles nach Induktion nicht geht. □

Beweis für die Infixnotation:

Für die eindeutige Lesbarkeit sind in diesem Fall die Klammern verantwortlich, daher braucht es einige Vorbereitungen über Klammerungen. Für eine aussagenlogische Formel φ sei k_φ die Anzahl von Klammern in φ und $\breve{\varphi}$ die *Klammerung* von φ, das ist die Zeichenfolge der Länge k_φ, die man aus φ erhält, indem man alle Zeichen außer den Klammern entfernt. Außerdem sei die *Klammerungstiefe* $\mathrm{Kt}(\breve{\varphi}, i)$ an der Stelle $i \in \{0, \ldots, k_\varphi\}$ die Anzahl der öffnenden minus die Anzahl der schließenden Klammern unter den ersten i Symbolen von $\breve{\varphi}$.

Lemma 2.5 *Für jede aussagenlogische Formel φ gilt:*

$$\mathrm{Kt}(\breve{\varphi}, i) \geqslant 0 \text{ für alle } i = 0, \ldots, k_\varphi \quad \text{und} \quad \mathrm{Kt}(\breve{\varphi}, i) = 0 \iff \big[\, i = 0 \text{ oder } i = k_\varphi \,\big]$$

Beweis per Induktion über den Aufbau der Formeln:

- Für $\varphi = A_i, \top, \bot$ ergibt die die leere Klammerung, die offenbar alle Bedingungen erfüllt.
- Im Negationsschritt ändert sich die Klammerung nicht.
- Sei also $\varphi = (\varphi_1 * \varphi_2)$. Klar ist $\mathrm{Kt}(\breve{\varphi}, 0) = 0 = \mathrm{Kt}(\breve{\varphi}, k_\varphi)$ und dass $\mathrm{Kt}(\breve{\varphi}, i) = 1 + \mathrm{Kt}(\breve{\varphi}_1, i-1)$ für $i = 1, \ldots, k_{\varphi_1} + 1$ gilt und $\mathrm{Kt}(\breve{\varphi}, i) = 1 + \mathrm{Kt}(\breve{\varphi}_2, i - k_{\varphi_1} - 1)$ für $i = k_{\varphi_1} + 1, \ldots, k_\varphi - 1)$. Per Induktion folgen damit problemlos die Behauptungen. □

Damit kann man nun per Induktion über den Aufbau der Formeln simultan beweisen, dass alle aussagenlogischen Formeln φ in Infix-Notation eindeutig lesbar sind und kein echtes Anfangsstück selbst wieder eine aussagenlogischen Formel in Infix-Notation ist:

- Für $\varphi = A_i$, $\top$ oder $\bot$ ist dies offensichtlich.
- Für $\varphi = \neg\varphi'$ folgt die Behauptungen per Induktion, da einerseits φ' per Induktion eindeutig lesbar ist, also auch φ, und andererseits ein echtes Anfangsstück von der Form $\neg\alpha$ für ein echtes Anfangsstück α von φ' wäre.
- Für $\varphi = (\varphi_1 * \varphi_2)$ ist diese Zerlegung eindeutig, denn falls auch $\varphi = (\psi_1 \circ \psi_2)$, ist entweder φ_1 ein Anfangsstück ψ_1 oder umgekehrt. Per Induktion gilt also $\varphi_1 = \psi_1$. Da φ_1, φ_2 per Induktion eindeutig lesbar sind, folgt dies auch für φ. Dass kein echtes Anfangsstück von φ eine aussagenlogischen Formel in Infix-Notation sein kann, ergibt sich aus dem Klammerungslemma, weil erst mit der letzten schließenden Klammer wieder die Klammerungstiefe 0 erreicht ist, was bei einer aussagenlogischen Formel der Fall sein muss. □

2.2 Semantik

Wahrheitswertverlauf

Die beiden Wahrheitswerte „wahr“ und „falsch“ werden mit 1 und 0 identifiziert. Jedem n-stelligen Junktor $*$ ist eine Wahrheitswertfunktion $\tilde{*} : \{0,1\}^n \to \{0,1\}$ zugeordnet. Es ist hierfür praktisch, $\top$ und $\bot$ als nullstellige Junktoren aufzufassen.

- $\tilde{\top}$ ist die konstante Funktion 1 und $\tilde{\bot}$ die konstante Funktion 0.

- $\tilde{\neg}(x) = 1 - x \qquad \tilde{\wedge}(x,y) = \min\{x,y\} \qquad \tilde{\vee}(x,y) = \max\{x,y\}$
- $\tilde{\rightarrow}(x,y) = \begin{cases} 1 & x \leqslant y \\ 0 & x > y \end{cases}$ und $\tilde{\leftrightarrow}(x,y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases}$

Definition 2.6 *Jeder aussagenlogischen Formel φ wird eine Funktion $\tilde{\varphi} : \{0,1\}^{\mathbb{N}} \to \{0,1\}$ zugeordnet:*

- *Wenn $\varphi = A_j$ ist $\tilde{\varphi}((x_i)_{i\in\mathbb{N}}) := x_j$.*
- *Wenn (in Polnischer Notation) $\varphi = *\varphi_1 \dots \varphi_k$ mit k-stelligem Junktor $*$, dann ist*

$$\tilde{\varphi}((x_i)_{i\in\mathbb{N}}) := \tilde{*}\big(\tilde{\varphi}_1((x_i)_{i\in\mathbb{N}}), \dots, \tilde{\varphi}_k((x_i)_{i\in\mathbb{N}})\big)$$

$\tilde{\varphi}$ heißt Wahrheitswertverlauf *der Formel φ.*

Es ist klar, dass der Wahrheitswertverlauf von φ nur von den tatsächlich in φ vorkommenden Aussagenvariablen abhängt. Manchmal bezeichnet man als Wahrheitswertverlauf daher die Einschränkung von $\tilde{\varphi}$ auf diejenigen Variablen, die den in φ vorkommenden Aussagenvariablen entsprechen.

Praktisch rechnet man den Wahrheitswertverlauf einer aussagenlogischen Formel in dieser eingeschränkten Form in einer *Wahrheitstafel* aus, z. B. für $(A_0 \to (\neg A_1 \vee A_3))$:

A_0	A_1	A_3	$\neg A_1$	$(\neg A_1 \vee A_3)$	$(A_0 \to (\neg A_1 \vee A_3))$
0	0	0	1	1	1
0	0	1	1	1	1
0	1	0	0	0	1
0	1	1	0	1	1
1	0	0	1	1	1
1	0	1	1	1	1
1	1	0	0	0	0
1	1	1	0	1	1

Der Berechnungsaufwand einer Wahrheitstafel wächst exponentiell mit der Anzahl der vorkommenden Aussagenvariablen. Für viele Aussagenvariablen ist dies nicht mehr praktikabel. Es gibt einige Verfahren, die in vielen Fällen günstiger sind, um den Wahrheitswertverlauf zu bestimmen bzw. damit verbundene Fragen zu beantworten (z.B. Erfüllbarkeit, siehe nächsten Unterabschnitt). Manche sind im Skript „Logik für Studierende der Informatik“ ausgeführt. Die generelle Frage, ob es dafür einen schnellen Algorithmus gibt, ist ein großes offenes Problem der Mathematik bzw. Theoretischen Informatik (im Wesentlichen das sogenannte P=NP-Problem).

Aussagen über aussagenlogische Formeln

Definition 2.7

- *Eine aussagenlogische Formel φ ist eine* Tautologie*, wenn $\tilde{\varphi}$ konstant 1 ist.*
  Dafür schreibt man $\vdash \varphi$.
- *Eine aussagenlogische Formel φ heißt* erfüllbar*, wenn $\tilde{\varphi}$ nicht konstant 0 ist.*
- *Zwei aussagenlogische Formeln φ und ψ heißen* logisch äquivalent *zueinander, wenn $\tilde{\varphi} = \tilde{\psi}$.*
  Dafür schreibe ich $\varphi \sim \psi$.
- *Eine aussagenlogische Formel ψ* folgt logisch *aus einer Menge aussagenlogischer Formeln $\{\varphi_i \mid i \in I\}$ – oder wird von ihr* impliziert *– wenn $\inf\{\tilde{\varphi}_i \mid i \in I\} \leqslant \tilde{\psi}$.*
  Dafür schreibt man $\{\varphi_i \mid i \in I\} \vdash \psi$; im Falle einer endlichen Menge auch $\varphi_1, \dots, \varphi_n \vdash \psi$.

Im letzten Punkt wird das Infimum in der partiell geordneten Menge aller Funktionen $\{0,1\}^{\mathbb{N}} \to \{0,1\}$ als punktweises Minimum gebildet.

$\vdash$ und $\sim$ sind metasprachliche Zeichen, also nicht Teil des aussagenlogischen Alphabets!

Lemma 2.8

- $\vdash \varphi \iff \varphi \sim \top \iff \top \vdash \varphi \iff \emptyset \vdash \varphi$.
- φ *ist erfüllbar* $\iff \nvdash \neg\varphi \iff \varphi \not\sim \bot$.
- $\varphi \sim \psi \iff \vdash (\varphi \leftrightarrow \psi) \iff [\varphi \vdash \psi$ *und* $\psi \vdash \varphi]$
- $\varphi_1, \ldots, \varphi_n \vdash \psi \iff \vdash (((\ldots((\varphi_1 \wedge \varphi_2) \wedge \varphi_3) \wedge \ldots) \wedge \varphi_n) \to \psi)$

Dieses Lemma ist mit Absicht in fast reiner Formelsprache geschrieben, um zwei Dinge deutlich zu machen: Erstens erhöht reine Formelprache nicht unbedingt die Lesbarkeit. Mehr Text wäre hier hilfreicher. Zweitens muss man drei verschiedene Zeichenebenen unterscheiden: Im ersten Punkt kommen die objektsprachlichen aussagenlogischen Formeln φ und $\top$ vor; es kommen die metasprachlichen Zeichen $\vdash$ und $\sim$ vor, die Aussagen über aussagenlogische Formel machen; schließlich gibt es das metametasprachliche Zeichen $\Leftrightarrow$, das die Äquivalenz von Aussagen über aussagenlogische Formel behauptet. Insbesondere tauchen in der dritten Zeile drei verschiedene Äquivalenzzeichen auf: $\leftrightarrow$, $\sim$ und $\Leftrightarrow$. Man sollte sich unbedingt die Funktionsweisen und Bedeutungen dieser drei Zeichen in Abgrenzung zueinander klar machen.

Beweis Der Beweis dieser Aussagen kann direkt erfolgen, ohne über den Aufbau der Formeln zu gehen. Zum Beispiel bedeuten die vier Aussagen der ersten Zeile nach Definition:

$$\tilde{\varphi} = 1 \qquad \tilde{\varphi} = \tilde{\top} \qquad \inf\{\tilde{\top}\} \leqslant \tilde{\varphi} \qquad \inf \emptyset \leqslant \tilde{\varphi}$$

Die Äquivalenz dieser Aussagen sieht man, da (ebenfalls per Definition) $\tilde{\top}$ die konstante Funktion 1 ist und außerdem die konstante Funktion 1 größtes Element im Raum der Funktionen $\{0,1\}^{\mathbb{N}} \to \{0,1\}$ ist, und daher auch Infimum der leeren Menge.

Die Äquivalenzen der zweiten und dritten Zeile sieht man ähnlich ein. Zum letzten Punkt:

$\varphi_1, \ldots, \varphi_n \vdash \psi$ ist per Definition gleichwertig mit $\inf\{\tilde{\varphi}_1, \ldots, \tilde{\varphi}_n\} \leqslant \tilde{\psi}$, also mit

$$\begin{aligned} &\min\{\tilde{\varphi}_1((x_i)_{i\in\mathbb{N}}), \ldots, \tilde{\varphi}_n((x_i)_{i\in\mathbb{N}})\} \leqslant \tilde{\psi}((x_i)_{i\in\mathbb{N}}) \quad \text{für jedes } (x_i)_{i\in\mathbb{N}} \\ \text{bzw.} \quad &\tilde{\to}\big(\min\{\tilde{\varphi}_1((x_i)_{i\in\mathbb{N}}), \ldots, \tilde{\varphi}_n((x_i)_{i\in\mathbb{N}})\}, \tilde{\psi}((x_i)_{i\in\mathbb{N}})\big) = 1 \quad \text{für jedes } (x_i)_{i\in\mathbb{N}} \end{aligned}$$

Da man das Minimum als sukzessives paarweises Minimum schreiben kann, also

$$\begin{gathered} \min\big\{\tilde{\varphi}_1((x_i)_{i\in\mathbb{N}}), \ldots, \tilde{\varphi}_n((x_i)_{i\in\mathbb{N}})\big\} = \\ \min\big\{\ldots \min\big\{\min\big\{\tilde{\varphi}_1((x_i)_{i\in\mathbb{N}}), \tilde{\varphi}_2((x_i)_{i\in\mathbb{N}})\big\}, \tilde{\varphi}_3((x_i)_{i\in\mathbb{N}})\big\} \ldots, \tilde{\varphi}_n((x_i)_{i\in\mathbb{N}})\big\} \end{gathered}$$

und das Minimum die Wahrheitswertfunktion ist, die dem Junktor $\wedge$ zugeordnet ist, folgt schließlich die Gleichwertigkeit mit der Bedingung, dass der Wahrheitswertverlauf der aussagenlogischen Formel $((\ldots((\varphi_1 \wedge \varphi_2) \wedge \varphi_3) \wedge \cdots \wedge \varphi_n) \to \psi)$ konstant 1 ist. $\square$

2.3 Logische Gesetze

Substitutionsprinzipien

Lemma 2.9 *Wenn man in einer aussagenlogischen Formel φ ein Vorkommen einer Teilformel φ' durch eine aussagenlogische Formel ψ' ersetzt, erhält man wieder eine aussagenlogische Formel ψ.*

Man kann dieses Lemma beweisen (von φ' ausgehend per Induktion über den Aufbau von Formeln, die φ' als Teilformel enthalten); der Beweis bringt allerdings m. E. keinen Zugewinn gegenüber der anschaulichen Einsichtigkeit: In der Darstellung als Bäume bedeutet es, dass ein Teilbaum durch einen anderen Baum ersetzt wird.

In der aussagenlogischen Formel $\varphi = (\neg(A_0 \to \bot) \vee (A_2 \leftrightarrow (A_0 \to \bot)))$ kann man zum Beispiel das rechte Vorkommen der Formel $\varphi' = (A_0 \to \bot)$ durch die Formel $\psi' = \neg(A_3 \wedge A_0)$ ersetzen und erhält $\psi = (\neg(A_0 \to \bot) \vee (A_2 \leftrightarrow \neg(A_3 \wedge A_0)))$. Ersetzt man das gleiche Vorkommen der Teilformel durch $\psi'' = A_0$, erhält man $(\neg(A_0 \to \bot) \vee (A_2 \leftrightarrow A_0))$.

Satz 2.10 (Prinzip der äquivalenten Substitution)
Ersetzt man in einer aussagenlogischen Formel φ ein Vorkommen einer Teilformel durch eine dazu logische äquivalente Formel, erhält man eine zu φ logisch äquivalente Formel.

Mit den Notationen des Lemmas: Falls $\varphi' \sim \psi'$, ist auch $\varphi \sim \psi$.

Zum Beispiel folgt aus diesem Prinzip, dass wegen $A_0 \sim \neg\neg A_0$ auch $(A_0 \wedge A_1) \sim (\neg\neg A_0 \wedge A_1)$.

Das Prinzip gilt, weil in der Auswertung der Formeln φ und ψ den Teilformeln φ' und ψ' jeweils der gleiche Wahrheitswert zugeordnet wird und der weitere Auswertungsprozess derselbe ist. Wie der Wahrheitswert jeweils zustand kommt, ist unerheblich. Auch hierfür gebe ich keinen detaillierteren Beweis.

Aus dem Prinzip der äquivalenten Substitution folgen auch die *Kongruenzeigenschaft* der logischen Äquivalenz bzgl. aller Junktoren: Aus $\varphi_1 \sim \psi_1$ und $\varphi_2 \sim \psi_2$ folgt zum Beispiel $(\varphi_1 \wedge \varphi_2) \sim (\psi_1 \wedge \psi_2)$, und analog für die anderen Junktoren.

Ersetzt man in einer aussagenlogischen Formel φ alle Vorkommen einer Aussagenvariablen A_i durch eine aussagenlogische Formel ψ, erhält man eine aussagenlogische Formel $\varphi[\frac{\psi}{A_i}]$.

Zum Beispiel ist

$$((A_1 \vee (A_2 \to \neg A_1))\,[\frac{\neg(A_3 \leftrightarrow A_4)}{A_1}] \;=\; ((\neg(A_3 \leftrightarrow A_4) \vee (A_2 \to \neg\neg(A_3 \leftrightarrow A_4)))$$
$$((A_1 \vee (A_2 \to \neg A_1))\,[\frac{\neg A_1}{A_1}] \;=\; ((\neg A_1 \vee (A_2 \to \neg\neg A_1))$$

Satz 2.11 (Prinzip der uniformen Substitution)
Uniforme Substitution erhält logische Äquivalenz: Wenn $\varphi_1 \sim \varphi_2$, dann $\varphi_1\,[\frac{\psi}{A_i}] \sim \varphi_2\,[\frac{\psi}{A_i}]$.

Zum Beispiel folgt aus diesem Prinzip, dass wegen $A_0 \sim \neg\neg A_0$ auch $(A_0 \wedge A_1) \sim \neg\neg(A_0 \wedge A_1)$.

Das Prinzip gilt, weil für jeden Wahrheitswert für A_0 in der Auswertung der Formeln φ_1 und φ_2 der gleiche Wahrheitswert herauskommt. Also kommt auch nach der Ersetzung für jeden Wahrheitswert, den ψ annehmen kann, am Ende der gleiche Wahrheitswert heraus. Ein detaillierterer Beweis unterbleibt auch hier.

Elementare logische Gesetze

Satz 2.12 *Es gelten die folgenden* elementaren logischen Gesetze *für alle aussagenlogischen Formeln φ, ψ, χ:*

$\top$-/$\bot$-Gesetze

$\neg\top \sim \bot$	$\neg\bot \sim \top$	*Verum-Falsum-Dualität*
$(\varphi \wedge \top) \sim \varphi$	$(\varphi \vee \bot) \sim \varphi$	*neutrale Elemente von $\wedge$ bzw. $\vee$*
$(\varphi \wedge \bot) \sim \bot$	$(\varphi \vee \top) \sim \top$	*absorbierende Elemente von $\wedge$ bzw. $\vee$*

$\neg$-Gesetze

$\neg\neg\varphi \sim \varphi$	*Doppelnegationsgesetz*
$(\varphi \vee \neg\varphi) \sim \top$	*Prinzip des ausgeschlossenen Dritten*
$(\varphi \wedge \neg\varphi) \sim \bot$	*Prinzip des ausgeschlossenen Widerspruchs*

$\wedge$-/$\vee$-Gesetze

$(\varphi \wedge \varphi) \sim \varphi$	$(\varphi \vee \varphi) \sim \varphi$
$(\varphi \wedge \psi) \sim (\psi \wedge \varphi)$	$(\varphi \vee \psi) \sim (\psi \vee \varphi)$
$((\varphi \wedge \psi) \wedge \chi) \sim (\varphi \wedge (\psi \wedge \chi))$	$((\varphi \vee \psi) \vee \chi) \sim (\varphi \vee (\psi \vee \chi))$
$((\varphi \wedge \psi) \vee \chi) \sim ((\varphi \vee \chi) \wedge (\psi \vee \chi))$	$((\varphi \vee \psi) \wedge \chi) \sim ((\varphi \wedge \chi) \vee (\psi \wedge \chi))$

Gesetze von *de Morgan*

$\neg(\varphi \wedge \psi) \sim (\neg\varphi \vee \neg\psi)$	$\neg(\varphi \vee \psi) \sim (\neg\varphi \wedge \neg\psi)$

Definition von $\rightarrow$ bzw. $\leftrightarrow$

$(\varphi \rightarrow \psi) \sim (\neg\varphi \vee \psi)$	$(\varphi \leftrightarrow \psi) \sim ((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$

Die $\wedge$-/$\vee$-Gesetze heißen der Reihe nach: *Idempotenz*, *Kommutativität* und *Assoziativität von $\wedge$ bzw. $\vee$* sowie *Distributivität von $\vee$ über $\wedge$* bzw. umgekehrt.

Beweis Nachprüfen mit Wahrheitstafeln! □

Diese elementaren logischen Gesetze reichen zusammen mit dem Prinzip der äquivalenten Substitution aus, um alle andern logischen Gesetze herzuleiten. Trotzdem ist es nützlich, einige weitere Gesetze zu kennen:

Satz 2.13 *Es gelten für alle aussagenlogischen Formeln φ, ψ, χ:*

Definition von $\neg$:	$(\varphi \rightarrow \bot) \sim \neg\varphi$	
Absorptionsgesetze:	$((\varphi \wedge \psi) \vee \varphi) \sim \varphi$	$((\varphi \vee \psi) \wedge \varphi) \sim \varphi$
Kontraposition:	$(\varphi \rightarrow \psi) \sim (\neg\psi \rightarrow \neg\varphi)$	$(\varphi \leftrightarrow \psi) \sim (\neg\psi \leftrightarrow \neg\varphi)$
„Currying“:	$((\varphi \wedge \psi) \rightarrow \chi) \sim (\varphi \rightarrow (\psi \rightarrow \chi))$	
Transitivität von $\rightarrow$:	$(\varphi \rightarrow \psi), (\psi \rightarrow \chi) \vdash (\varphi \rightarrow \chi)$	
mit $\varphi = \top$: *modus ponens*	$\psi, (\psi \rightarrow \chi) \vdash \chi$	
mit $\chi = \bot$: *modus tollens*	$(\varphi \rightarrow \psi), \neg\psi \vdash \neg\varphi$	
Links-Distributivität	$(\varphi \rightarrow (\psi \wedge \chi)) \sim ((\varphi \rightarrow \psi) \wedge (\varphi \rightarrow \chi))$	
von $\rightarrow$ über $\wedge$ bzw. $\vee$:	$(\varphi \rightarrow (\psi \vee \chi)) \sim ((\varphi \rightarrow \psi) \vee (\varphi \rightarrow \chi))$	
Rechts-Antidistributivität	$((\varphi \wedge \psi) \rightarrow \chi) \sim ((\varphi \rightarrow \chi) \vee (\psi \rightarrow \chi))$	
von $\rightarrow$ über $\wedge$ bzw. $\vee$:	$((\varphi \vee \psi) \rightarrow \chi) \sim ((\varphi \rightarrow \chi) \wedge (\psi \rightarrow \chi))$	

Disjunktive Normalform

Da iterierte Konjunktionen und Disjunktionen häufig vorkommen, ist es nützlich, folgende Abkürzungen einzuführen:

$$(\varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_n) \quad \text{oder} \quad \bigwedge_{i=1}^{n} \varphi_i \quad \text{für die Formel} \quad ((\cdot\cdot\,(\varphi_1 \wedge \varphi_2) \wedge \cdots) \wedge \varphi_n)$$

$$(\varphi_1 \vee \varphi_2 \vee \cdots \vee \varphi_n) \quad \text{oder} \quad \bigvee_{i=1}^{n} \varphi_i \quad \text{für die Formel} \quad ((\cdot\cdot\,(\varphi_1 \vee \varphi_2) \vee \cdots) \vee \varphi_n)$$

Insbesondere ist mit den in der Mathematik üblichen Konventionen $\bigwedge_{i=1}^{1} \varphi_i = \bigvee_{i=1}^{1} \varphi_i = \varphi_1$ sowie $\bigwedge_{i=1}^{0} \varphi_i = \top$ und $\bigvee_{i=1}^{0} \varphi_i = \bot$ (nämlich das jeweilige neutrale Element).

Die Notationen $\bigwedge_{i=1}^{n} \varphi_i$ und $\bigvee_{i=1}^{n} \varphi_i$ funktionieren als Abkürzungen nur mit der fest gegebenen Reihenfolge der Formeln. Interessiert man sich für aussagenlogische Formeln nur bis auf logische Äquivalenz, kann man auch Mengennotationen wie $\bigwedge\{\varphi_i \mid i \in I\}$ bzw. $\bigvee\{\varphi_i \mid i \in I\}$ nutzen.

Definition 2.14 *Eine Formel ist in* disjunktiver Normalform (DNF)*, wenn sie eine Disjunktion von Konjunktionen von Aussagenvariablen bzw. negierten Aussagenvariablen ist, d. h. von der Form*

$$((L_{11} \wedge \cdots \wedge L_{1n_1}) \vee \cdots \vee (L_{k1} \wedge \cdots \wedge L_{kn_k})) \quad \textit{bzw.} \quad \bigvee_{i=1}^{k} \bigwedge_{j=1}^{n_i} L_{ij},$$

eine Formel ist in konjunktiver Normalform (KNF) *wenn sie eine Konjunktion von Disjunktionen von Aussagenvariablen bzw. negierten Aussagenvariablen ist, d. h. von der Form*

$$((L_{11} \vee \cdots \vee L_{1n_1}) \wedge \cdots \wedge (L_{k1} \vee \cdots \vee L_{kn_k})) \quad \textit{bzw.} \quad \bigwedge_{i=1}^{k} \bigvee_{j=1}^{n_i} L_{ij},$$

wobei die „Literale“ L_{ij} jeweils Aussagenvariablen A_l oder negierte Aussagenvariablen $\neg A_l$ sind und $k, n_1, \ldots, n_k \in \mathbb{N}$.

Satz 2.15 *Jede aussagenlogische Formel φ ist logisch äquivalent zu einer aussagenlogischen Formel in disjunktiver Normalform und logisch äquivalent zu einer aussagenlogischen Formel in konjunktiver Normalform.*

Beweis 1: Man kodiert die Wahrheitstafel von φ in einer aussagenlogischen Formel. Ohne Einschränkung seien $A_1, \ldots, A_n$ die in φ vorkommenden Aussagenvariablen. Für jede Abbildung $\beta : \{1, \ldots, n\} \to \{0, 1\}$ sei

$$[\neg_\beta A_i] := \begin{cases} A_i & \text{falls } \beta(i) = 1 \\ \neg A_i & \text{falls } \beta(i) = 0 \end{cases}$$

Sei nun $\varphi_\beta := \bigwedge_{i=1}^{n} [\neg_\beta A_i]$. Dann gilt

$$\tilde{\varphi_\beta}((x_i)_{i\in\mathbb{N}}) = 1 \iff \beta = (x_i)_{i\in\mathbb{N}} \restriction_{\{1,\ldots,n\}}$$

Es folgt die Existenz einer disjunktiven Normalform:

$$\varphi \;\sim\; \bigvee \Big\{ \bigwedge_{i=1}^{n} [\neg_\beta A_i] \;\Big|\; \beta = (x_i)_{i\in\mathbb{N}} \restriction_{\{1,\ldots,n\}} \text{ mit } \tilde{\varphi}((x_i)_{i\in\mathbb{N}}) = 1 \Big\}$$

Eine konjunktive Normalform bekommt man als

$$\varphi \;\sim\; \bigwedge \Big\{ \bigvee_{i}^{n} \neg[\neg_\beta A_i] \;\Big|\; \beta = (x_i)_{i\in\mathbb{N}} \restriction_{\{1,\ldots,n\}} \text{ mit } \tilde{\varphi}((x_i)_{i\in\mathbb{N}}) = 0 \Big\}$$

wobei man ggf. noch doppelte Negationen eliminieren muss. □

Formeln, die man durch diesen Beweis erhält, heißen *kanonische disjunktive* bzw. *konjunktive Normalform*: In jeder Konjunktion der DNF bzw. Disjunktion des KNF kommt jede in φ vorkommende Aussagenvariable vor. Sie ist nicht eindeutig, kann aber eindeutig gemacht werden, wenn man noch eine Reihenfolge der vorkommenden β's festlegt (zusätzlich zu den implizit festgelegten Reihenfolge der Aussagenvariablen und der in Notation festgelegten Klammerungsreihenfolge). Kanonische KNF einer Tautologie ist $\top$ und kanonische DNF einer negierten Tautologie ist $\bot$. Ohne diese beiden Zeichen gäbe es hierfür keine kanonische KNF bzw. DNF.

Der Beweis zeigt noch mehr:

Folgerung 2.16 *Jeder von nur endlich vielen Variablen abhängige Wahrheitswertverlauf ist von der Form $\tilde{\varphi}$ für eine aussagenlogische Formel φ, die so gewählt werden kann, dass keine anderen Junktoren außer $\wedge, \vee, \neg, \top, \bot$ vorkommen.*

Definition 2.17 *Eine Menge von Junktoren, für die die voranstehende Folgerung gilt, heißt* vollständiges Junktorensystem*.*

Satz 2.18 *$\{\wedge, \neg\}$ und $\{\vee, \neg\}$ sind vollständige Junktorensysteme.*

Beweis $\{\wedge, \neg\}$ ist wegen Folgerung 2.16, der Substitutionsprinzipien und $\bot \sim (A_0 \wedge \neg A_0)$, $\top \sim \neg\bot$ und $(A_0 \vee A_1) \sim \neg(\neg A_0 \wedge \neg A_1)$ ein vollständiges Junktorensystem.

Wegen $(A_0 \wedge A_1) \sim \neg(\neg A_0 \vee \neg A_1)$ ist auch $\{\vee, \neg\}$ ein vollständiges Junktorensystem. □

Beweis 2 von Satz 2.15: Man beweist über den Aufbau der Formeln, dass sich jede Formel in DNF bringen lässt.

- $A_i, \top, \bot$ sind bereits in DNF.
- $\varphi = \neg\psi$: Ohne Einschränktung ist $\psi = \bigvee_{i=1}^{k} \bigwedge_{j=1}^{n_i} (\neg) A_{ij}$ per Induktion bereits in DNF. Dann ist $\varphi \sim \bigwedge_{i=1}^{k} \bigvee_{j=1}^{n_i} \neg(\neg) A_{ij}$. Durch Anwenden des Distributivgesetzes von $\wedge$ über $\vee$ und Eliminieren von evtl. Doppelnegationen erreicht man wieder DNF.
- $\varphi = (\varphi_1 \vee \varphi_2)$ ist bereits in DNF, wenn φ_1 und φ_2 in DNF gebracht sind.
- Die anderen Junktoren lassen sich auf Negation und Disjunktion zurückführen:

$$\begin{aligned}(\varphi_1 \wedge \varphi_2) &\sim \neg(\neg\varphi_1 \vee \neg\varphi_2)\\ (\varphi_1 \to \varphi_2) &\sim (\neg\varphi_1 \vee \varphi_2)\\ (\varphi_1 \leftrightarrow \varphi_2) &\sim (\neg(\varphi_1 \vee \varphi_2) \vee \neg(\neg\varphi_1 \vee \neg\varphi_2))\end{aligned}$$ □

Man sieht an diesem Beispiel, dass Beweise relativ kurz werden können, wenn man ein geschicktes vollständiges Junktorensystem betrachtet. Voraussetzung ist allerdings, dass es in der zu beweisenden Aussage nur um aussagenlogische Formeln bis auf logische Äquivalenz geht.

Ein Kalkül

Satz 2.19 *Alle logischen Äquivalenzen ergeben sich aus sukzessiver Anwendung der elementaren logischen Gesetze aus Satz 2.12 zusammen mit äquivalenter Substitution.*

Dieser Satz folgt sofort aus dem folgenden Satz, dessen Beweis einen weiteren Beweis von Satz 2.15 liefert.

Satz 2.20 *Jede aussagenlogische Formel lässt sich durch die elementaren logischen Gesetze aus Satz 2.12 zusammen mit äquivalenter Substitution in eindeutige kanonische disjunktive Normalform überführen.*

Beweis: Die Umformung erfolgt durch das unten skizzierte Verfahren. Mit den Kommutativitäts- und Assoziativitätsgesetzen für $\wedge$ und $\vee$ kann innerhalb von Konjunktionen und Disjunktionen stets beliebig umsortiert und geklammert werden. Nötige Umformungen dieser Art sind im Verfahren unten nicht explizit aufgeführt. Durch sie kann insbesondere am Ende aus einer kanonischen DNF die wie auch immer festgelegte eindeutige kanonische DNF erreicht werden.

1. Ersetze alle Junktoren $\leftrightarrow$ und dann alle Junktoren $\rightarrow$ durch ihre Definition.
2. Vereinfache mit den $\top$-/$\bot$-Gesetzen alle Vorkommen von $\top$ und $\bot$, bis nur noch $\top$ oder $\bot$ übrig bleibt oder kein $\top$ oder $\bot$ mehr vorkommt.
3. Ziehe mit den *de Morgan*'schen Gesetzen alle Negationen „nach innen“ bis unmittelbar vor Aussagenvariablen oder weitere Negationszeichen. Eliminiere mit der Doppelnegationsregel ggf. vorkommende Doppelnegationen.
4. Wende solange das Distributivgesetz von $\wedge$ über $\vee$ an, bis die Formel in DNF ist.
5. Ersetze gleichzeitiges Vorkommen von A_i und $\neg A_i$ in einer Konjunktion durch $\bot$ (Prinzip des ausgeschlossenen Widerspruchs) und führe wieder Schritt 2 durch. Dabei bleibt die Formel in DNF.
6. Ersetze ggf. eine Konjunktion K durch $(K \wedge \top)$ und dann $(K \wedge (A_i \vee \neg A_i))$. Mit Schritt 4 erhält man wieder DNF, wobei nun in allen Konjunktionen alle Aussagenvariablen vorkommen.
7. Eliminiere mit dem Idempotenz-Gesetz ggf. doppelt vorkommende (negierten) Aussagenvariablen in einer Konjunktion und ggf. doppelt vorkommende Konjunktionen. □

Den in Definition 2.7 eingeführten Zugang zu den Konzepten „Tautologie“, „erfüllbare Formel“, „logische Äquivalenz“ und „logische Folgerung“ über eine *Auswertung* der logischen Formeln nennt man häufig einen *semantischen Zugang*, weil man die Zuweisung von Wahrheitswerten als eine Zuweisung von „Bedeutung“ auffassen kann. Man könnte solch eine Zuweisung konkretisieren durch ein *Modell*: Für eine aussagenlogische Formel weist solch ein Modell einerseits jeder Aussagenvariablen eine konkrete Aussage zu und definiert andererseits einen Kontext, in dem jede dieser Aussagen entweder stimmt oder nicht stimmt. Die aussagenlogische Formel wird in dem Modell dadurch zu einer konkreten Aussage mit einer Bedeutung.

Satz 2.19 ermöglicht einen anderen, sogenannten *syntaktischen Zugang* zum Konzept der logischen Äquivalenz (den man über die Zusammenhänge in Lemma 2.8 auch auf die Konzepte Tautologie, Erfüllbarkeit und logische Folgerung ausdehnen könnte). Die elementaren logischen Gesetze kann man als Regeln begreifen, die nur etwas über die Manipulation von Zeichenfolgen aussagen. Solche Regelsysteme heißen auch *Kalküle*. Man kann einen Kalkül definieren, ohne irgendeine Bedeutung oder Funktionsweise der manipulierten Zeichenfolgen im Sinn zu haben. Die Aussage von Satz 2.12 nennt man üblicherweise die *Korrektheit* des Kalküls der elementaren logischen Gesetze und die Aussage von Satz 2.19 die *Vollständigkeit* des Kalküls.

Solch ein Kalkül wie in Satz 2.12, also eine Sammlung logischer Gesetze, aus denen alles ableitbar ist, ist natürlich weit davon entfernt, in irgendeiner Weise eindeutig zu sein. Die Bezeichnung

„elementares logisches Gesetz“ gilt also nur für diese Vorlesung und ist keine allgemein akzeptierte Klassifizierung.

Die beiden unterschiedlichen Herangehensweisen– semnatisch und syntaktisch – werden in der Prädikatenlogik noch wichtig sein!

3 Prädikatenlogik

3.1 Syntax und Semantik

Wenn man etwa im angeordneten Körper $\mathbb{R}$ der reellen Zahlen eine Aussage wie „*Das Produkt zweier positiver Zahlen ist positiv*“ formalisieren möchte, kann man dies in der Aussagenlogik nur durch eine Aussagenvariable A_i tun. Die „innere Struktur“ der Aussage ist darin nicht darstellbar; folglich sind auch logische Schlüsse, die auf dieser Struktur beruhen, aussagenlogisch nicht nachvollziehbar. Daher braucht man eine Erweiterung. Ausreichend, um alle Mathematik darstellen zu können (wenn auch nicht immer auf direkte Weise) ist die *Prädikatenlogik*, genauer die *Prädikatenlogik erster Stufe*.

Alphabet

Anders als in der Aussagenlogik gibt es für jede Anwendungsssituation eine eigene prädikatenlogische *Sprache*.

Definition 3.1 *Eine* (erststufige) Sprache $\mathscr{L}$ *besteht aus einer Menge* $\mathscr{L}_F$ *von* Funktionszeichen *und einer dazu disjunkten Menge* $\mathscr{L}_R$ *von* Relationszeichen *zusammen mit einer Abbildung* $s : \mathscr{L}_F \cup \mathscr{L}_R \to \mathbb{N}$*, die jedem Zeichen seine* Stelligkeit *zuordnet.*

0-stellige Funktionszeichen heißen auch *Konstanten(zeichen)*; Relationszeichen (insbesondere 1-stellige) werden auch *Prädikate/Prädikatszeichen* genannt. Manche andere Autoren verwenden *Signatur* anstelle von *Sprache*.

Als Funktionszeichen verwende ich meistens kleine, als Relationszeichen meistens große lateinische Buchstaben (ggf. mit Indizes oder anderen „Akzidenzien“), sowie geläufige Symbole wie z. B. $+$, $\circ$, $\cup$ bzw. $<$, $\sim$, $\subseteq$. Die Buchstaben v, w, x, y, z sind üblicherweise für Variablen reserviert und werden nicht als Funktionszeichen verwendet.

Prädikatenlogische Formeln nutzen neben dem in der Sprache festgelegten variablen Anteil des Alphabets auch die folgenden, fest vorgegebenen Symbole:

Individuenvariablen:	$v_0 \quad v_1 \quad v_2 \quad \cdots$
Quantorenzeichen:	Existenzquantor $\exists$ und Allquantor $\forall$
Junktoren:	$\neg \quad \wedge \quad \vee \quad \to \quad \leftrightarrow$
Aussagenkonstanten:	$\top \quad \bot$
Gleichheitszeichen:	$\doteq$
Klammern:	$(\quad)$

Es ist hilfreich, das objektsprachliche Gleichheitszeichen von dem (normalen) metasprachlichen Gleichheitszeichen $=$ zu unterscheiden, ähnlich wie der objektsprachliche Junktor $\to$ von dem metasprachlichen Implikationspfeil $\Rightarrow$ unterschieden ist. Ich benutze daher die Ziegler'sche Konvention, $\doteq$ als objektsprachliches Gleichheitszeichen zu benutzen.

$\mathscr{L}$-Strukturen

Die Prädikatenlogik ist wesentlich komplexer als die Aussagenlogik. Für ihr Verständnis ist es hilfreich, Syntax und Semantik gleichzeitig einzuführen. Die Auswertung von prädikatenlogischen Formeln einer gegebenen Sprache $\mathscr{L}$ erfolgt in sogenannten $\mathscr{L}$-Strukturen:

Definition 3.2 *Für eine erststufige Sprache $\mathscr{L}$ ist eine $\mathscr{L}$-*Struktur *$\mathscr{M}$ gegeben durch:*

- *eine nicht-leere Menge M, genannt* Universum, Träger- *oder* Grundmenge *der Struktur,*
- Interpretationen *der Zeichen aus $\mathscr{L}$ in $\mathscr{M}$, d. h. einer Funktion $f^{\mathscr{M}} : M^{s(f)} \to M$ für jedes Funktionszeichen $f \in \mathscr{L}_F$ und einer Relation $R^{\mathscr{M}} \subseteq M^{s(R)}$ für jedes Relationszeichen $R \in \mathscr{L}_R$.*

Für jede Menge M ist $M^0 = \{\emptyset\}$, also $|M^0| = 1$. Eine 0-stellige Funktion $\alpha : M^0 \to M$ ist daher konstant und kann mit der „Konstanten“ $\alpha(\emptyset) \in M$ identifiziert werden. Eine 0-stellige Relation $\Omega \subseteq M^0$ ist entweder $\emptyset$ oder $\{\emptyset\}$, was man als Wahrheitswerte $|\Omega| \in \{0, 1\}$ auffassen kann, weshalb Ω als Aussagenvariable verstanden werden kann.

Beispiel 3.3 Die Sprache $\mathscr{L} = \{f_0, f_1, c_0, c_1, R\}$ besteht aus zweistelligen Funktionszeichen f_0 und f_1, nullstelligen Funktionszeichen c_0 und c_1 und einem zweistelligen Relationszeichen R.

Eine mögliche $\mathscr{L}$-Struktur ist der angeordnete Körper der reellen Zahlen, also $M = \mathbb{R}$ mit $f_0^{\mathscr{M}}$ und $f_1^{\mathscr{M}}$ als Addition $+_{\mathbb{R}}$ bzw. Multiplikation $\cdot_{\mathbb{R}}$ auf $\mathbb{R}$, mit den Zahlen $c_0^{\mathscr{M}} = 0 \in \mathbb{R}$ und $c_1^{\mathscr{M}} = 1 \in \mathbb{R}$ als Interpretation der Konstantenzeichen und mit $R^{\mathscr{M}}$ als der Ordnungsrelation $\leqslant_{\mathbb{R}}$ auf $\mathbb{R}$. Man schreibt für diese Struktur kurz auch $\mathscr{M} = (\mathbb{R}; +_{\mathbb{R}}, \cdot_{\mathbb{R}}, 0, 1, \leqslant_{\mathbb{R}})$.

Eine weitere $\mathscr{L}$-Struktur ist $\mathscr{N} = (\mathbb{N}; +_{\mathbb{N}}, \cdot_{\mathbb{N}}, 0, 1, \leqslant_{\mathbb{N}})$, also der angeordnete Halbring der natürlichen Zahlen.

Im Extremfall ist $O = \{\emptyset\}$. Möchte man darauf eine $\mathscr{L}$-Struktur $\mathscr{O}$ definieren, dann gibt es für die Interpretationen der Funktionszeichen jeweils nur eine Möglichkeit (die konstante Funktion mit Wert $\emptyset$) und für $R^{\mathscr{O}}$ zwei Möglichkeiten (die leere Relation $\emptyset \subseteq O^2$ oder die volle Relation $\{(\emptyset, \emptyset)\} = O^2$).

Die Strukturen bzw. Interpretationen müssen nicht irgendwie „natürlich“ sein. Auch $\mathscr{P}$ mit $P = \mathbb{R}$ und $c_0^{\mathscr{P}} = c_1^{\mathscr{P}} = -\sqrt{\pi}$, der Funktion $f_0^{\mathscr{P}} : (x, y) \mapsto |y|$, falls $x \leqslant y$, und $(x, y) \mapsto \sin(xy)^3 - 2\ln(x - y)$ sonst, der Funktion $f_1^{\mathscr{P}}$, die konstant 0 ist, und der leeren Relation $\leqslant^{\mathscr{P}} = \emptyset$ ist eine vollwertige $\mathscr{L}$-Struktur.

Wenn ich prädikatenlogisch über z. B. den angeordneten Körper der reellen Zahlen sprechen möchte, werde ich der Einfachheit halber in Zukunft häufig die üblichen Symbole als Sprache nehmen, also $\mathscr{L}' = \{+, \cdot, 0, 1, \leqslant\}$ mit den offensichtlichen Stelligkeiten. In der $\mathscr{L}'$-Struktur $\mathscr{M}' = (\mathbb{R}; +_{\mathbb{R}}, \cdot_{\mathbb{R}}, 0_{\mathbb{R}}, 1_{\mathbb{R}}, \leqslant_{\mathbb{R}})$ ist dann also z. B. $+$ das Zeichen, das durch die Addition auf $\mathbb{R}$ interpretiert wird, die ich dann $+_{\mathbb{R}}$ schreibe. Während diese Schreibweise häufig einsichtig ist, weil die Addition $+_{\mathbb{R}}$ auf $\mathbb{R}$ als mathematisches Objekt ja tatsächlich etwas anderes ist als die Addition $+_{\mathbb{N}}$ auf $\mathbb{N}$, wirkt sie bei Konstanten etwas merkwürdig, weil im üblichen Verständnis $0_{\mathbb{R}} = 0_{\mathbb{N}}$. Manche Autoren variieren daher lieber die Zeichen und nehmen etwa $\dot{0}$, $\underline{0}$ oder $\mathbf{0}$ als Symbol in der Sprache und 0 als dessen Interpretation. Keine Lösung ist hier ideal. Wichtig ist, dass man die Konzepte von Zeichen und Interpretation versteht und gedanklich auseinanderhält. Dazu hilft es, dies auch notationell zu tun.

Insbesondere sollte verstanden sein, dass in der Definition einer $\mathscr{L}$-Struktur das Zeichen nicht seine Interpretation vorgibt. Man kann auch eine $\mathscr{L}$-Struktur auf den reellen Zahlen definieren,

in der das Zeichen $\leqslant$ durch die umgekehrte Relation $\geqslant_{\mathbb{R}}$ interpretiert wird. Das ist für Anwendungen natürlich nicht sinnvoll, aber wichtig für einige noch kommende Definitionen, in denen es um die Auswertung von $\mathscr{L}$-Formeln in beliebigen $\mathscr{L}$-Strukturen geht.

$\mathscr{L}$-Terme

Die prädikatenlogischen Formeln werden in drei Schritten eingeführt: Im ersten Schritt werden *Terme* definiert, die in einer Struktur für einzelne Element stehen. Ich führe wieder die Betrachtungsweise als Bäume und die Schreibweise in Polnischer Notation ein.

Definition 3.4 *Sei $\mathscr{L}$ eine prädikatenlogische Sprache.*

- *Jede Individuenvariable (als einelementiger Baum oder als einelementige Zeichenfolge) ist ein $\mathscr{L}$-Term.*
- *Für jedes Funktionszeichen $f \in \mathscr{L}_F$ und alle $\mathscr{L}$-Terme $\tau_1, \dots, \tau_{s(f)}$ ist der Baum*

$$\begin{array}{c} f \\ /\,|\,\cdot\cdot\,\diagdown \\ \tau_1\ \ \tau_2\ \dots\ \tau_{s(f)} \end{array}$$

bzw. in Polnischer Notation die Zeichenfolge $f\tau_1\tau_2\dots\tau_{s(f)}$ ein $\mathscr{L}$-Term.

Ich benutze kleine griechische Buchstaben wie σ und τ als Namen/Variablen für $\mathscr{L}$-Terme und bisweilen x, y, z (ggf. mit Varianten) als Variablen für Individuenvariablen (neben $v_i, v_j, \dots$).

In der Sprache $\mathscr{L}$ aus Beispiel 3.3 sind zum Beispiel die folgenden Zeichenfolgen $\mathscr{L}$-Terme in Polnischer Notation:

$$v_3 \qquad c_0 \qquad f_1v_3c_0 \qquad f_0c_1c_1 \qquad f_0v_3f_1v_3c_0 \qquad f_0f_0v_3f_1v_3c_0f_0c_1c_1$$

Geschachtelte Terme werden in dieser kondensierten Schreibweise schnell unlesbar, kommen in der Vorlesung aber kaum vor. Der besseren Lesbarkeit halber nutzt man bei zweistelligen Funktionszeichen wie $+$ auch die Infixnotation, also z. B. $v_3 + c_0$ statt $+v_3c_0$, braucht bei geschachtelten Termen dann aber zusätzlich Klammern.

Lemma 3.5 *$\mathscr{L}$-Terme in Polnischer Notation sind eindeutig lesbar, d. h. es gibt genau einen $\mathscr{L}$-Term als Baum, der diese Darstellung hat.*

Beweis Ähnlich wie bei der Aussagenlogik. □

$\mathscr{L}$-Terme sollen Elemente von $\mathscr{L}$-Strukturen bezeichnen. Um sie auswerten zu können, muss man allerdings auch mit den Individuenvariablen umgehen. Dazu hilft die nächste Definition:

Definition 3.6 *Sei $\mathscr{M}$ eine $\mathscr{L}$-Struktur. Eine* Belegung *(genauer: eine Belegung der Individuenvariablen mit Elementen aus M) ist eine Abbildung $\beta : \mathbb{N} \to M$.*

Definition 3.7 *Sei $\mathscr{M}$ eine $\mathscr{L}$-Struktur und β eine Belegung. Jedem $\mathscr{L}$-Term τ induktiv wird ein Element $\beta(\tau) \in M$ zugeordnet:* [5]

- $\beta(v_i) := \beta(i)$ *für jede Individuenvariable v_i*

[5] Notiz an mich: wäre $\tau^{(\mathscr{M},\beta)}$ besser?

- $\beta(f\tau_1 \dots \tau_{s(f)}) := f^{\mathscr{M}}\big(\beta(\tau_1), \dots, \beta(\tau_{s(f)})\big)$

$\beta(\tau)$ heißt *Auswertung von τ in $\mathscr{M}$ unter β*. Eine andere übliche Schreibweise dafür ist $\tau^{\mathscr{M}}[\beta]$.

Definition 3.7 ist ein Beispiel einer *Definition über den Aufbau der Terme*. Im Grunde ist ein Term ein geschachtelter Funktionsausdruck und er wird in einer Struktur ausgewertet, indem man sukzessive die interpretierenden Funktionen anwendet. Die Auswertung des $\mathscr{L}$-Terms $f_0f_0v_3f_1v_3c_0f_0c_1c_1$ im angeordneten Körper $\mathbb{R}$ als $\mathscr{L}$-Struktur wie in Beispiel 3.3, etwa für $\beta(3) = 5$, ist $+_{\mathbb{R}}(+_{\mathbb{R}}(5, \cdot_{\mathbb{R}}(5, 0)), +_{\mathbb{R}}(1, 1)) = +_{\mathbb{R}}(+_{\mathbb{R}}(5, 0), 2) = +_{\mathbb{R}}(5, 2) = 7$.

Lemma 3.8 *Die Auswertung von $\mathscr{L}$-Termen hängt nur von den vorkommenden Individuenvariablen ab: Sei $\mathscr{L}$ eine Sprache, τ in $\mathscr{L}$-Term, $\mathscr{M}$ eine $\mathscr{L}$-Struktur und β, β' zwei Belegungen mit $\beta(i) = \beta'(i)$ für alle i, sodass v_i in τ vorkommt. Dann ist $\beta(\tau) = \beta'(\tau)$.*

Beweis: Es ist offensichtlich, dass dieses Lemma gilt. Der Beweis soll aber dennoch als Beispiel eines *Beweises über den Aufbau der Terme* ausgeführt werden:

- Wenn $\tau = v_i$, dann kommt v_i in τ vor und nach Annahme ist $\beta(\tau) = \beta(i) = \beta'(i) = \beta'(\tau)$.
- Wenn $\tau = f\tau_1 \dots \tau_{s(f)}$, dann kommen alle in einem τ_i vorkommenden Individuenvariablen auch in τ vor, also gilt per Induktion $\beta(\tau_i) = \beta'(\tau_i)$ für $i = 1, \dots, s(f)$, und es folgt $\beta(\tau) = f^{\mathscr{M}}\big(\beta(\tau_1), \dots, \beta(\tau_{s(f)}) = f^{\mathscr{M}}\big(\beta'(\tau_1), \dots, \beta'(\tau_{s(f)}) = \beta'(\tau)$. □

Ein $\mathscr{L}$-Term, in dem keine Individuenvariablen vorkommen, wird auch *geschlossener $\mathscr{L}$-Term* genannt. Geschlossene $\mathscr{L}$-Term τ können also in einer $\mathscr{L}$-Struktur $\mathscr{M}$ unabhängig von einer Belegung durch ein Element $\tau^{\mathscr{M}} \in M$ ausgewertet werden.

Atomare $\mathscr{L}$-Formeln

Im zweiten Schritt werden *atomare Formeln* definiert, also Formeln, die nicht aus anderen Formeln zusammengesetzt sind. Sie werden in einer Struktur durch einen Wahrheitswert ausgewertet, stimmen also oder stimmen nicht.

Prädikatenlogische Formeln sollte man sich wieder als Bäume vorstellen. Die atomaren Formeln sind dann wieder einelementige Bäume, wobei die „Etiketten“ nun nicht mehr einzelen Symbole wie in der Aussagenlogik sind, sondern bereits komplexerer Ausdrücke sein können. Diese Ausdrücke könnte man selbst wieder als Bäume auffassen. Üblicher ist, sie als Zeichenfolgen zu schreiben in einer eingebürgerten gemischten Schreibweise aus teils Infix-, teils Polnischer Notation. Ich gebe hier nur die Definition in dieser Darstellung.

Definition 3.9 *Die folgenden Zeichenfolgen heißen* atomare $\mathscr{L}$-Formeln*:*

- $\top$ *und* $\bot$
- $\tau_1 \doteq \tau_2$ *für beliebige $\mathscr{L}$-Terme τ_1, τ_2*
- $R\tau_1 \dots \tau_{s(R)}$ *für Relationszeichen $R \in \mathscr{L}_R$ und beliebige $\mathscr{L}$-Terme $\tau_1, \dots, \tau_{s(R)}$.*

Der besseren Lesbarkeit halber nutzt man bei üblichen zweistelligen Relationszeichen wie $<$ auch die Infixnotation, also z. B. $\tau_1 < \tau_2$ statt $< \tau_1\tau_2$.

Man kann $\top, \bot$ und $\doteq$ in dem Sinn als „konstante“ 0- bzw. 2-stellige Relationszeichen auffassen, dass sie in jeder Struktur $\mathscr{M}$ eine feste Interpretation haben, nämlich $\top^{\mathscr{M}} = \{\emptyset\} = M^0$, $\bot^{\mathscr{M}} = \emptyset \subseteq M^0$ und $\doteq^{\mathscr{M}} = \{(m, m) \mid m \in M\} \subseteq M^2$. In dieser Auffassung sind alle atomaren Formeln von der Form $R\tau_1 \dots \tau_{s(R)}$ mit $R \in \mathscr{L}_R^+ := \mathscr{L}_R \cup \{\top, \bot, \doteq\}$. Dadurch vereinfachen sich die manche Beweise etwas.

Definition 3.10 *Sei $\mathscr{M}$ eine $\mathscr{L}$-Struktur und β eine Belegung. Für jede atomare $\mathscr{L}$-Formel φ ist auf die folgende Weise festgelegt, ob die Formel in $\mathscr{M}$ unter β gilt oder nicht. Dafür schreibt man $(\mathscr{M},\beta)\vDash\varphi$ bzw. $(\mathscr{M},\beta)\nvDash\varphi$:*

- *Stets ist $(\mathscr{M},\beta)\vDash\top$ und $(\mathscr{M},\beta)\nvDash\bot$.*
- $(\mathscr{M},\beta)\vDash\tau_1\doteq\tau_2:\Longleftrightarrow\ \beta(\tau_1)=\beta(\tau_2)$.
- $(\mathscr{M},\beta)\vDash R\tau_1\dots\tau_{s(R)}:\Longleftrightarrow\ \big(\beta(\tau_1),\dots,\beta(\tau_{s(R)})\big)\in R^{\mathscr{M}}$.

Grob gesprochen gilt eine Formel in einer Struktur also, wenn die intendierte Bedeutung stimmt, wobei die Intension des Gleichheitszeichens $\doteq$ die Gleichheit ist, die Intension eines Relationszeichens R bzw. Funktionszeichens f dessen Interpretation, also die Relation $R^{\mathscr{M}}$ bzw. die Funktion $f^{\mathscr{M}}$ ist, usw. Man schafft sich also lediglich eine Möglichkeit, mit Symbolen formelhaft das auszudrücken, was man ausdrücken möchte. Die spezifisch mathematische Schwierigkeit hierbei ist, dass man sowieso schon immer mit Formeln arbeitet. Der Unterschied ist, dass nun eine präzise Definition vorliegen hat für etwas, was vorher nur *ad hoc* und in sehr variabler Weise genutzt wurde. Außerdem sollte man sich klar machen, dass man Mathematik auch ohne Formeln betreiben kann (und dass dies jahrhundertelang auch weitgehend so praktiziert wurde).

Beliebige $\mathscr{L}$-Formeln

Im dritten und letzten Schritt werden beliebige *Formeln* definiert, die aus atomaren Formeln zusammengesetzt sind. Sie sollte man sich am besten wieder als Bäume vorstellen, wobei ich wieder nur die Darstellung in der gemischten Notation definieren werde.

Definition 3.11 *$\mathscr{L}$-Formeln sind wie folgt per Induktion definiert:*

- *Jede atomare $\mathscr{L}$-Formel ist eine $\mathscr{L}$-Formel.*
- *Wenn φ eine $\mathscr{L}$-Formel ist, dann auch* $\neg\varphi$.
- *Wenn φ,ψ $\mathscr{L}$-Formeln sind, dann auch* $(\varphi\wedge\psi)$ $(\varphi\vee\psi)$ $(\varphi\to\psi)$ $(\varphi\leftrightarrow\psi)$.
- *Wenn φ eine $\mathscr{L}$-Formel ist, dann auch für jede Individuenvariable v_i:* $\exists v_i\varphi$ *und* $\forall v_i\varphi$

Definition 3.12 *Sei $\mathscr{M}$ eine $\mathscr{L}$-Struktur und β eine Belegung. Dann richtet sich die Gültigkeit der mit Junktoren zusammengesetzten $\mathscr{L}$-Formeln nach der Wahrheitswertfunktionalität der Junktoren, das heißt es gilt*

$$\begin{array}{lll}
(\mathscr{M},\beta)\vDash\neg\varphi & \text{genau dann, wenn} & (\mathscr{M},\beta)\nvDash\varphi\\
(\mathscr{M},\beta)\vDash(\varphi\wedge\psi) & \text{genau dann, wenn} & \big[(\mathscr{M},\beta)\vDash\varphi \quad \text{und} \quad (\mathscr{M},\beta)\vDash\psi\big]\\
(\mathscr{M},\beta)\vDash(\varphi\vee\psi) & \text{genau dann, wenn} & \big[(\mathscr{M},\beta)\vDash\varphi \quad \text{oder} \quad (\mathscr{M},\beta)\vDash\psi\big]\\
(\mathscr{M},\beta)\vDash(\varphi\to\psi) & \text{genau dann, wenn} & \big[(\mathscr{M},\beta)\vDash\varphi\Longrightarrow(\mathscr{M},\beta)\vDash\psi\big]\\
(\mathscr{M},\beta)\vDash(\varphi\leftrightarrow\psi) & \text{genau dann, wenn} & \big[(\mathscr{M},\beta)\vDash\varphi\Longleftrightarrow(\mathscr{M},\beta)\vDash\psi\big]
\end{array}$$

Für die mit Quantoren zusammengesetzten $\mathscr{L}$-Formeln gilt

$$\begin{array}{ll}
(\mathscr{M},\beta)\vDash\exists v_j\,\varphi & \text{genau dann, wenn es ein } m\in M \text{ gibt mit } (\mathscr{M},\beta\tfrac{m}{v_j})\vDash\varphi\\
(\mathscr{M},\beta)\vDash\forall v_j\,\varphi & \text{genau dann, wenn für alle } m\in M \text{ gilt } (\mathscr{M},\beta\tfrac{m}{v_j})\vDash\varphi
\end{array}$$

wobei jeweils $\beta':=\beta\frac{m}{v_j}$ die Belegung ist mit $\beta'(j)=m$ und $\beta'(i)=\beta(i)$ für alle $i\neq j$.

Bemerkung: Man könnte auch definieren, dass die Belegung β jeder $\mathscr{L}$-Formel φ einen Wahrheitswert $\beta(\varphi)\in\{0,1\}$ zuordnet, nämlich 1 genau dann, wenn $(\mathscr{M},\beta)\vDash\varphi$. Dann gilt für die Junktorenschritte entsprechend zur Aussagenlogik $\beta(*\varphi_1\dots\varphi_k)=\tilde{*}\big(\beta(\varphi_1),\dots,\beta(\varphi_k)\big)$ mit $*\in\{\neg,\wedge,\vee,\to,\leftrightarrow\}$, Polnischer Notation und $k\in\{1,2\}$ der Stelligkeit des Junktors.

Freie Variable und $\mathscr{L}$-Aussagen

Es ist wiederum klar, dass die Frage, ob $(\mathscr{M}, \beta) \vDash \varphi$ gilt, nur von den in φ vorkommenden Individuenvariablen abhängt. Tatsächlich hängt sie nur von den nicht durch einen Quantor gebundenen Variablen ab. Weil wie im Beispiel $(Pv_1 \wedge \exists v_1\, Pv_1)$ eine Variable mehrfach vorkommen kann – teils im Bereich von Quantoren, teils nicht – muss man häufig genauer die einzelnen *Vorkommen von Variablen* betrachten.

Zunächst sind einige Sprechweisen nützlich: Die Quantorenzeichen $\exists$ und $\forall$ kommen nur in Verbindung mit einer Individuenvariable v_i vor. Man nennt die Zeichenfolge $\exists v_i$ bzw. $\forall v_i$ einen *Quantor*; das Zeichen v_i in dieser Verbindung wird also als Teil des Quantors angesehen. Wenn man daher von den „in einer Formel vorkommenden Individuenvariablen“ spricht, sind damit *nicht* diese als Teil eines Quantors auftretenden Zeichen gemeint. In Zukunft werde ich zudem häufig kurz von „Variablen“ sprechen, wenn Individuenvariablen gemeint sind.

Definition 3.13

- *Eine* Teilformel *einer $\mathscr{L}$-Formel φ ist analog zur Aussagenlogik wiederum eine Formel, die im induktiven Aufbau von φ vorkommt.*
- *Der* Wirkungsbereich *eines Vorkommens eines Quantors $\exists v_i$ bzw. $\forall v_i$ in einer $\mathscr{L}$-Formel φ ist diejenige Teilformel ψ, vor der der Quantor steht, so dass also $\exists v_i\, \psi$ bzw. $\forall v_i\, \psi$ Teilformel von φ ist. Als Baum gesehen ist es der unterhalb des Vorkommens des Quantors liegende Teilbaum.*
- *Ein Vorkommen einer Individuenvariable v_i wird durch das Vorkommen eines Quantors $\exists v_i$ bzw. $\forall v_i$* gebunden*, wenn es im Wirkungsbereich des Quantors liegt und in diesem nicht bereits durch ein anderes Vorkommen eines Quantors gebunden ist.*
- *Eine Individuenvariable v_i ist* frei *in einer $\mathscr{L}$-Formel φ, wenn es ein nicht durch einen Quantor gebundenes Vorkommen von v_i in φ gibt.*
- *Man schreibt $\varphi(v_{i_1}, \ldots, v_{i_n})$ dafür, dass die freien Individuenvariablen von φ sich unter $v_{i_1}, \ldots, v_{i_n}$ befinden. Zudem soll $v_{i_j} \neq v_{i_k}$ für $i_j \neq i_k$ gelten.*
- *Eine $\mathscr{L}$-Formel ohne freie Individuenvariable heißt* $\mathscr{L}$-Aussage *(auch* geschlossene $\mathscr{L}$-Formel *oder* $\mathscr{L}$-Satz*).*

Sei $\mathscr{L} = \{P, Q\}$ mit einstelligen Relationszeichen P und Q. In der $\mathscr{L}$-Formel

$$(\exists v_1(\exists v_1 \forall v_1(Qv_1 \wedge Pv_1) \rightarrow Pv_1) \wedge Qv_1)$$

werden das erste und das zweite Vorkommen (von links gelesen) von v_1 durch den Quantor $\forall v_1$ gebunden, dessen Wirkungsbereich die Teilformel $(Qv_1 \wedge Pv_1)$ ist. Das dritte Vorkommen von v_1 durch das erste Vorkommen von $\exists v_1$ gebunden, dessen Wirkungsbereich die Teilformel $(\exists v_1 \forall v_1(Qv_1 \wedge Pv_1) \rightarrow Pv_1)$ ist. Durch das zweite Vorkommen von $\exists v_1$ wird kein Vorkommen einer Variablen gebunden, denn die Vorkommen der Variable v_1 in seinem Wirkungsbereich $\forall v_1(Qv_1 \wedge Pv_1)$ sind bereits durch den Allquantor gebunden. Das letzte Vorkommen von v_1 ist durch keinen Quantor gebunden, v_1 ist also freie Variable der Formel.

Lemma 3.14 *Die Auswertung von $\mathscr{L}$-Formeln hängt nur von den freien Variablen ab:*
Sei $\mathscr{L}$ eine Sprache, φ eine $\mathscr{L}$-Formel, $\mathscr{M}$ eine $\mathscr{L}$-Struktur und β, β' zwei Belegungen mit $\beta(i) = \beta'(i)$ für alle i, so dass v_i frei in φ ist. Dann gilt $(\mathscr{M}, \beta) \vDash \varphi \iff (\mathscr{M}, \beta') \vDash \varphi$.

Beweis per Induktion über den Aufbau der Formeln:

- Für atomare $\mathscr{L}$-Formeln φ gilt die Aussage nach Lemma 3.8, da alle in atomaren $\mathscr{L}$-Formeln vorkommenden Individuenvariablen darin frei sind.
- Für die Junktorenschritte $\varphi = *\varphi_1 \dots \varphi_k$ mit $* \in \{\neg, \wedge, \vee, \to, \leftrightarrow\}$ etc. gilt die Aussage per Induktion, da jede in einem φ_i frei vorkommende Variable auch in φ frei ist.
- Sei nun $\varphi = \exists v_i\, \psi$ oder $\varphi = \forall v_i\, \psi$. Die freien Variablen von ψ sind die freien Variablen von φ und eventuell v_i. Für jedes $m \in M$ gilt also nach Induktion $(\mathscr{M}, \beta\frac{m}{v_i}) \vDash \psi \iff (\mathscr{M}, \beta'\frac{m}{v_i})$. Es folgt in beiden Fällen $(\mathscr{M}, \beta) \vDash \varphi \iff (\mathscr{M}, \beta') \vDash \varphi$. □

$\mathscr{L}$-Aussagen φ kann man also in $\mathscr{L}$-Strukturen $\mathscr{M}$ unabhängig von Belegungen auswerten. Man schreibt dann $\mathscr{M} \vDash \varphi$ bzw. $\mathscr{M} \nvDash \varphi$ und sagt, „φ gilt in $\mathscr{M}$“, „φ stimmt in $\mathscr{M}$“, „φ trifft in $\mathscr{M}$ zu“, „φ ist wahr in $\mathscr{M}$“, „$\mathscr{M}$ erfüllt φ“ oder „$\mathscr{M}$ ist Modell von φ“ bzw. die entsprechenden verneinten Aussagen.

Aussagenlogik als Teil der Prädikatenlogik

Syntaktisch ist zunächst jede aussagenlogische Formel $\varphi(A_0, \dots, A_{n-1})$ eine $\mathscr{L}$-Formel für Sprachen $\mathscr{L}$, die die Aussagenvariablen $A_0, \dots, A_{n-1}$ enthalten.

Die Auswertungsprozesse passen wie folgt zusammen: Eine $\mathscr{L}$-Struktur $\mathscr{M}$ ordnet jeder in $\mathscr{L}$ vorkommenden Aussagenvariable A_i einen Wahrheitswert zu, nämlich 1, wenn $\mathscr{M} \vDash A_i$, und 0, wenn $\mathscr{M} \nvDash A_i$. Umgekehrt kann man jede Verteilung der Wahrheitswerte in einer entsprechend definierten $\mathscr{L}$-Struktur realisieren. Eine Wahrheitswertverteilung $(x_i)_{i\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ kann man daher als Abstraktion von $\mathscr{L}'$-Strukturen für $\mathscr{L}' = \{A_0, A_1, A_2, \dots\}$ ansehen, indem man die Grundmenge „vergisst“.

Aufgrund der übereinstimmenden Definitionen der Auswertungsprozesse ist klar:
Wenn $(x_i)_{i\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ mit $\mathscr{M}$ in dem Sinne übereinstimmt, dass $x_i = 1 \iff \mathscr{M} \vDash A_i$ für alle in $\mathscr{L}$ vorkommenden Aussagenvariablen A_i, dann ist $\tilde{\varphi}((x_i)_{i\in\mathbb{N}}) = 1 \iff \mathscr{M} \vDash \varphi$.

3.2 Aussagen über prädikatenlogische Formeln

Sobald man einen Auswertungsmöglichkeit für Formeln hat, kann man entsprechend wie in der Aussagenlogik die Konzepte „Tautologie“, „Erfüllbarkeit“, „logische Äquivalenz“ und „logische Folgerung“ definieren. Der Begriff „Tautologie“ wird in der Prädikatenlogik allerdings eine andere Bedeutung bekommen und daher durch „Allgemeingültigkeit“ ersetzt. Gleiches gilt für das Zeichen $\vdash$, das durch $\vDash$ ersetzt wird.

Definition 3.15

- *Eine $\mathscr{L}$-Formel φ ist* allgemeingültig*, wenn $(\mathscr{M}, \beta) \vDash \varphi$ für alle $\mathscr{L}$-Strukturen $\mathscr{M}$ und alle Belegungen β. Dafür schreibt man $\vDash \varphi$.*
- *Eine $\mathscr{L}$-Formel φ heißt* erfüllbar*, wenn es eine $\mathscr{L}$-Struktur $\mathscr{M}$ und eine Belegung β mit $(\mathscr{M}, \beta) \vDash \varphi$ gibt.*
- *Zwei $\mathscr{L}$-Formeln φ und ψ heißen* logisch äquivalent *zueinander, wenn $(\mathscr{M}, \beta) \vDash \varphi \iff (\mathscr{M}, \beta) \vDash \psi$ für alle $\mathscr{L}$-Strukturen $\mathscr{M}$ und alle Belegungen β. Dafür schreibe ich $\varphi \sim \psi$.*
- *Eine $\mathscr{L}$-Formel ψ* folgt logisch aus *oder* wird impliziert von *einer Menge von $\mathscr{L}$-Formeln $\{\varphi_i \mid i \in I\}$, wenn für alle $\mathscr{L}$-Strukturen $\mathscr{M}$ und alle Belegungen β Folgendes gilt: Wenn $(\mathscr{M}, \beta) \vDash \varphi_i$ für alle $i \in I$, dann auch $(\mathscr{M}, \beta) \vDash \psi$. Dafür schreibt man $\{\varphi_i \mid i \in I\} \vDash \psi$; im Falle einer endlichen Menge auch $\varphi_1, \dots, \varphi_n \vDash \psi$.*

Analog zu Lemma 2.8 sieht man:

Lemma 3.16 *Die folgenden Aussagen sind zeilenweise äquivalent:*

(a)	$\vDash \varphi$	(b)	$\varphi \sim \top$	(c)	$\top \vDash \varphi$	(d)	$\emptyset \vDash \varphi$
(a)	φ *ist erfüllbar*	(b)	$\nvDash \neg\varphi$	(c)	$\varphi \not\sim \bot$		
(a)	$\varphi \sim \psi$	(b)	$\vDash (\varphi \leftrightarrow \psi)$	(c)	$[\varphi \vDash \psi$ *und* $\psi \vDash \varphi]$		
(a)	$\varphi_1, \dots, \varphi_n \vDash \psi$	(b)	$\vDash ((\varphi_1 \wedge \dots \wedge \varphi_n) \to \psi)$				

Außerdem ist per Definition klar, dass man sich in der folgenden Weise immer auf $\mathscr{L}$-Aussagen zurückziehen kann:

Lemma 3.17 *Für $\mathscr{L}$-Formeln $\varphi(v_{i_1}, \dots, v_{i_n}), \psi(v_{i_1}, \dots, v_{i_n})$ und $\varphi_i(v_{i_1}, \dots, v_{i_n})$ gilt:*

- *φ ist genau dann allgemeingültig, wenn $\forall v_{i_1} \dots \forall v_{i_n}\, \varphi$ allgemeingültig ist.*
- *φ ist genau dann erfüllbar, wenn $\exists v_{i_1} \dots \exists v_{i_n}\, \varphi$ erfüllbar ist.*
- *φ und ψ sind genau dann logisch äquivalent zueinander, wenn $\forall v_{i_1} \dots \forall v_{i_n}\, (\varphi \leftrightarrow \psi)$ allgemeingültig ist.*
- *ψ folgt genau dann logisch aus $\varphi_1, \dots, \varphi_n$, wenn $\forall v_{i_1} \dots \forall v_{i_n}\, ((\varphi_1 \wedge \dots \wedge \varphi_n) \to \psi)$ allgemeingültig ist.*

Achtung: Im Allgemeinen sind φ und ψ nicht gemnau dann logisch äquivalent zueinander, wenn $\forall v_{i_1} \dots \forall v_{i_n}\, \varphi$ und $\forall v_{i_1} \dots \forall v_{i_n}\, \psi$ logisch äquivalent zueinander sind. Zum Beispiel sind für ein einstelliges Relationszeichen P die beiden $\{P\}$-Formeln $\forall v_1 \forall v_2\, Pv_1$ und $\forall v_1 \forall v_2\, Pv_2$ logisch äquivalent zueinander (und beide zu $\forall v_0\, Pv_0$), aber Pv_1 und Pv_2 sind nicht logisch äquivalent zueinander: Man kann z. B. eine zwei-elementige $\{P\}$-Struktur $\mathscr{M}$ mit $M = \{a, b\}$ und $P^{\mathscr{M}} = \{a\}$ definieren. Für Belegungen mit $\beta(1) = a$ und $\beta(2) = b$ gilt $(\mathscr{M}, \beta) \vDash Pv_1$, aber $(\mathscr{M}, \beta) \nvDash Pv_2$. Analog für die logische Folgerung!

Abhängigkeit von der Sprache

Falls $\mathscr{L} \subseteq \mathscr{L}'$, ist jede $\mathscr{L}$-Formel auch eine $\mathscr{L}'$-Formel. Man sieht leicht, dass die gerade definierten Konzepte nicht von der Sprache abhängen, d. h. dass zum Beispiel eine (nicht) allgemeingültige $\mathscr{L}$-Formel auch als $\mathscr{L}'$-Formel (nicht) allgemeingültig ist

Denn zum einen wird aus jeder $\mathscr{L}'$-Struktur $\mathscr{M}'$ eine $\mathscr{L}$-Struktur $\mathscr{M}$, indem man die Interpretation der Zeichen aus $\mathscr{L}' \setminus \mathscr{L}$ „ vergisst“, und es ist klar, dass dann für $\mathscr{L}$-Formeln φ und beliebige Belegungen β gilt: $(\mathscr{M}', \beta) \vDash \varphi \iff (\mathscr{M}, \beta) \vDash \varphi$. Umgekehrt kann man jede $\mathscr{L}$-Struktur zu einer $\mathscr{L}'$-Struktur machen, indem man die zusätzlichen Zeichen beliebig interpretiert, ohne dass sich an der Gültigket einer $\mathscr{L}$-Formel unter einer Belegung etwas ändern würde.

3.3 Logische Gesetze

Substitutionsprinzipien

Analog zur Aussagenlogik ist offensichtlich, dass man aus einer $\mathscr{L}$-Formel wieder eine $\mathscr{L}$-Formel erhält, wenn man eine Teilformel durch eine andere $\mathscr{L}$-Formel ersetzt. Eine Spezialfall hiervon ist es, Aussagenvariablen durch $\mathscr{L}$-Formeln zu ersetzen. Zunächst kann man die beiden Substitutionsprinzipien der Aussagenlogik übertragen:

Satz 3.18
(a) Äquivalente Substitution: Wenn ψ aus einer $\mathscr{L}$-Formel φ dadurch entsteht, dass eine Teilformel φ' von φ durch eine zu φ' logisch äquivalente Formel ψ' ersetzt wird, dann ist ψ logisch äquivalent zu φ.

(b) Uniforme Substitution: Wenn eine aussagenlogische Tautologie $\varphi(A_1, \ldots, A_n)$ und $\mathscr{L}$-Formeln $\varphi_1, \ldots, \varphi_n$ gegeben sind, dann ist die $\mathscr{L}$-Formel $\varphi\left[\frac{\varphi_1}{A_1} \ldots \frac{\varphi_n}{A_n}\right]$ allgemeingültig, die aus φ durch simultanes Ersetzen aller Vorkommen von A_i durch φ_i entsteht.

Allgemeiner gilt für alle $\mathscr{L}$-Strukturen $\mathcal{M}$ und Belegungen β:

$$(\mathcal{M}, \varphi) \vDash \varphi\left[\begin{smallmatrix}\varphi_1 \\ A_1\end{smallmatrix} \ldots \begin{smallmatrix}\varphi_n \\ A_n\end{smallmatrix}\right] \iff \tilde{\varphi}\big((x_i)_{i\in\mathbb{N}}\big) = 1$$

für alle Wahrheitswertverteilungen $(x_i)_{i\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ mit $x_i = 1 \iff (\mathcal{M}, \beta) \vDash \varphi_i$.

Beweis wie in der Aussagenlogik, da es im Auswertungsprozess nur auf den Wahrheitswert von $(\mathscr{M}, \beta) \vDash \varphi'$ bzw. $(\mathscr{M}, \beta) \vDash A_i$ ankommt. □

Definition 3.19 *Allgemeingültige $\mathscr{L}$-Formeln, die durch uniforme Substitution aus aussagenlogischen Tautologien entstehen, heißen* $\mathscr{L}$-Tautologien.

$(\exists v_1\, Pv_1 \vee \neg\, \exists v_1\, Pv_1)$ ist also eine aus z. B. $(A_1 \vee \neg A_1)$ entstandene $\{P\}$-Tautologie; dagegen ist $(\exists v_1\, Pv_1 \leftrightarrow \exists v_2\, Pv_2)$ eine allgemeingültige $\{P\}$-Formel, die keine $\{P\}$-Tautologie ist.

Für $\mathscr{L}$-Terme τ und σ sei $\tau[\frac{\sigma}{v_i}]$ der $\mathscr{L}$-Term, den man erhält, indem man jedes Vorkommen von v_i in τ durch σ ersetzt. Es ist klar, dass dies wieder ein $\mathscr{L}$-Term ist.

Lemma 3.20 (Substitutionslemma für Terme)
Seien σ und τ $\mathscr{L}$-Terme und β eine Belegung in einer $\mathscr{L}$-Struktur $\mathcal{M}$. Dann gilt

$$\beta\big(\tau\left[\tfrac{\sigma}{v_i}\right]\big) \;=\; \beta\tfrac{\beta(\sigma)}{v_i}(\tau)$$

Beweis Klar durch Nachdenken darüber, was die Formel des Lemmas aussagt. □

Wenn man das Substitutionslemma für Terme zum Substitutionslemma für Formeln ausweiten möchte, muss man ausschließen, dass sich bei der Substitution die Quantifizierungsstruktur ändert. Dazu definiert man zum einen $\varphi[\frac{\sigma}{v_i}]$ als die $\mathscr{L}$-Formel, die aus der $\mathscr{L}$-Formel φ dadurch entsteht, dass man jedes *freie Vorkommen* von v_i in φ durch den $\mathscr{L}$-Term σ ersetzt. Zum andern braucht man die folgende Definition:

Definition 3.21 *Die Variable v_i ist* frei für einen $\mathscr{L}$-Term σ in einer $\mathscr{L}$-Formel φ, *wenn bei der Ersetzung $\varphi[\frac{\sigma}{v_i}]$ kein Vorkommen einer Individuenvariablen v_j in σ durch einen Quantor von φ gebunden wird.*

Formal über den Aufbau der Formeln: *v_i ist frei für σ in φ, falls entweder v_i nicht frei in φ ist oder falls v_i frei in φ ist und einer der folgenden Fälle gilt:*

- *φ ist atomar;*
- *$\varphi = \neg\psi$ und v_i ist frei für σ in ψ;*
- *$\varphi = (\psi * \psi')$ mit beliebigem zweistelligen Junktor $*$ und v_i ist frei für σ in ψ und in ψ';*
- *$\varphi = \exists v_j\, \psi$ oder $\varphi = \forall v_j\, \psi$, v_i ist frei für σ in ψ und v_j kommt in σ nicht vor.*

Lemma 3.22 (Substitutionslemma)
Sei φ eine $\mathscr{L}$-Formel, σ ein $\mathscr{L}$-Term und $\mathscr{M}$ eine $\mathscr{L}$-Struktur mit Belegeung β. Dann gilt für jede Variable v_i, die frei für σ in φ ist:

$$(\mathcal{M}, \beta) \vDash \varphi\left[\tfrac{\sigma}{v_i}\right] \iff \big(\mathcal{M}, \beta\tfrac{\beta(\sigma)}{v_i}\big) \vDash \varphi$$

Beweis Beweis per Induktion über den Aufbau der Formeln. Für atomare $\mathscr{L}$-Formeln $\varphi = R\tau_1 \dots \tau_n$ (inklusive $\top$, $\bot$ und $\tau_1 \doteq \tau_2$) hat man:

$$\begin{aligned}
(\mathscr{M},\beta) \vDash R\tau_1\dots\tau_n\ \left[\tfrac{\sigma}{v_i}\right] &\iff (\mathscr{M},\beta) \vDash R\ \tau_1[\tfrac{\sigma}{v_i}]\ \dots\ \tau_n[\tfrac{\sigma}{v_i}] \\
&\iff \big(\beta(\tau_1[\tfrac{\sigma}{v_i}]),\dots,\beta(\tau_n[\tfrac{\sigma}{v_i}])\big)\ \in R^{\mathscr{M}} \\
\textit{Substitutionslemma für Terme} &\iff \big(\beta\tfrac{\beta(\sigma)}{v_i}(\tau_1),\dots,\beta\tfrac{\beta(\sigma)}{v_i}(\tau_n)\big) \in R^{\mathscr{M}} \\
&\iff \big(\mathscr{M},\beta\tfrac{\beta(\sigma)}{v_i}\big) \vDash R\tau_1\dots\tau_n
\end{aligned}$$

Die Junktorenschritte sind unproblematisch, wenn man sich klar macht, dass $(\varphi \wedge \psi)[\frac{\sigma}{v_i}] = (\varphi[\frac{\sigma}{v_i}] \wedge \psi[\frac{\sigma}{v_i}])$ etc. gilt.

Sei nun also $\varphi = \exists v_j\,\psi$ (für den Allquantor funktioniert die entsprechende Argumentation). Wegen der Annahme, dass v_i frei für σ in φ ist, kommt v_j in σ nicht vor. Wenn v_i nicht frei in φ ist, ist $\varphi[\frac{\sigma}{v_i}] = \varphi$ und die Aussage des Substitutionslemmas folgt mit Lemma 3.14.

Sei also v_i frei in φ. Dann muss $i \neq j$ sein und:

$$\begin{aligned}
(\mathscr{M},\beta) \vDash \exists v_j\,\psi[\tfrac{\sigma}{v_i}] &\iff \text{es gibt ein } m \in M \text{ mit } \big(\mathscr{M},\beta\tfrac{m}{v_j}\big) \vDash \psi[\tfrac{\sigma}{v_i}] \\
\textit{nach Induktionsvoraussetzung} &\iff \text{— " —} \quad \big(\mathscr{M},\beta\tfrac{m}{v_j}\tfrac{\beta\frac{m}{v_j}(\sigma)}{v_i}\big) \vDash \psi \\
\textit{da } v_j \textit{ in } \sigma \textit{ nicht vorkommt} &\iff \text{— " —} \quad \big(\mathscr{M},\beta\tfrac{m}{v_j}\tfrac{\beta(\sigma)}{v_i}\big) \vDash \psi \\
\textit{da } i \neq j &\iff \text{— " —} \quad \big(\mathscr{M},\beta\tfrac{\beta(\sigma)}{v_i}\tfrac{m}{v_j}\big) \vDash \psi \\
&\iff \big(\mathscr{M},\beta\tfrac{\beta(\sigma)}{v_i}\big) \vDash \exists v_j\,\psi
\end{aligned}$$ □

Gleichheits- und Quantorengesetze

Satz 3.23 *Die folgenden $\mathscr{L}$-Aussagen sind allgemeingültig (für beliebige Funktionszeichen $f \in \mathscr{L}_F$ und Relationszeichen $R \in \mathscr{L}_R$). Sie heißen $\mathscr{L}$-*Gleichheitsgesetze.

Reflexivität:	$\forall v_i\ v_i \doteq v_i$
Symmetrie:	$\forall v_i \forall v_j\,(v_i \doteq v_j \to v_j \doteq v_i)$
Transitivität:	$\forall v_i \forall v_j \forall v_k\,((v_i \doteq v_j \wedge v_j \doteq v_k) \to v_i \doteq v_k)$
Kongruenz:	*mit* $n = s(f)$ *bzw.* $n = s(R)$
	$\forall v_{i_1} \cdots \forall v_{i_{2n}}((v_{i_1} \doteq v_{i_{n+1}} \wedge \cdots \wedge v_{i_n} \doteq v_{i_{2n}}) \to f v_{i_1} \dots v_{i_n} \doteq f v_{i_{n+1}} \dots v_{i_{2n}})$
	$\forall v_{i_1} \cdots \forall v_{i_{2n}}((v_{i_1} \doteq v_{i_{n+1}} \wedge \cdots \wedge v_{i_n} \doteq v_{i_{2n}}) \to (R v_{i_1} \dots v_{i_n} \leftrightarrow R v_{i_{n+1}} \dots v_{i_{2n}}))$

Beweis: Klar □

Satz 3.24 *Es gelten die logischen Äquivalenzen und Folgerungen:*

„Unnötige Quantoren“:

Wenn v_i nicht frei in φ ist, dann $\quad \exists v_i\,\varphi \sim \forall v_i\,\varphi \sim \varphi$

Umbenennung gebundener Variablen:

Wenn v_i frei für v_j in φ ist und v_j nicht frei in φ ist, dann

$\exists v_i\,\varphi \sim \exists v_j\ \varphi[\frac{v_j}{v_i}]$ $\qquad$ $\forall v_i\,\varphi \sim \forall v_j\ \varphi[\frac{v_j}{v_i}]$

Verhältnis von Quantoren untereinander:

$\forall v_i\,\varphi \vDash \exists v_i\,\varphi$ $\qquad$ $\exists v_i \forall v_j\,\varphi \vDash \forall v_j \exists v_i\,\varphi$

$\exists v_i \exists v_j\,\varphi \sim \exists v_j \exists v_i\,\varphi$ $\qquad$ $\forall v_i \forall v_j\,\varphi \sim \forall v_j \forall v_i\,\varphi$

Dualität / verallgemeinerte Regeln von *de Morgan*:

$$\exists v_i \neg \varphi \;\sim\; \neg \forall v_i \varphi \qquad\qquad \forall v_i \neg \varphi \;\sim\; \neg \exists v_i \varphi$$

Vertauschungen von Quantoren mit Junktoren:

$$\exists v_i (\varphi \vee \psi) \;\sim\; (\exists v_i \varphi \vee \exists v_i \psi) \qquad \forall v_i (\varphi \wedge \psi) \;\sim\; (\forall v_i \varphi \wedge \forall v_i \psi)$$

$$\exists v_i (\varphi \wedge \psi) \;\vDash\; (\exists v_i \varphi \wedge \exists v_i \psi) \qquad (\forall v_i \varphi \vee \forall v_i \psi) \;\vDash\; \forall v_i (\varphi \vee \psi)$$

Falls v_i nicht frei in φ ist, dann *Falls v_i nicht frei in φ ist, dann*

$$\exists v_i (\varphi \wedge \psi) \;\sim\; (\varphi \wedge \exists v_i \psi) \qquad \forall v_i (\varphi \vee \psi) \;\sim\; (\varphi \vee \forall v_i \psi)$$

$$\exists v_i (\varphi \vee \psi) \;\sim\; (\varphi \vee \exists v_i \psi) \qquad \forall v_i (\varphi \wedge \psi) \;\sim\; (\varphi \wedge \forall v_i \psi)$$

„Definition der Quantoren“: *Falls v_i frei für τ in φ ist, dann*

$$\varphi[\tfrac{\tau}{v_i}] \;\vDash\; \exists v_i \varphi \qquad\qquad \forall v_i \varphi \;\vDash\; \varphi[\tfrac{\tau}{v_i}]$$

Die Umkehrungen der sechs Implikationen $\vDash$ gelten im Allgemeinen nicht!

Beweis: Die Beweise funktionieren in der Regel durch direkte Anwendung der Definition. Beispielhaft sei das Gesetz für die Umbenennung gebundener Variablen gezeigt:

$$\begin{aligned}
(\mathcal{M},\beta) \vDash \exists v_j\ \varphi[\tfrac{v_j}{v_i}] &\iff \text{es gibt ein } m \in M \text{ mit } (\mathcal{M}, \beta\tfrac{m}{v_j}) \vDash \varphi[\tfrac{v_j}{v_i}] \\
\text{\textit{Substitutionslemma:}}\ &\iff \text{es gibt ein } m \in M \text{ mit } (\mathcal{M}, \beta\tfrac{m}{v_j}\tfrac{\beta\frac{m}{v_j}(v_j)}{v_i}) \vDash \varphi \\
&\iff \text{es gibt ein } m \in M \text{ mit } (\mathcal{M}, \beta\tfrac{m}{v_j}\tfrac{m}{v_i}) \vDash \varphi \\
v_j \text{ \textit{nicht frei in} } \varphi\ &\iff \text{es gibt ein } m \in M \text{ mit } (\mathcal{M}, \beta\tfrac{m}{v_i}) \vDash \varphi \\
&\iff (\mathcal{M},\beta) \vDash \exists v_i \varphi
\end{aligned}$$

□

Man kann sich auch bei jeder Bedingung überlegen, dass sie notwendig ist:

- Wenn v_j frei in φ ist, führt die Umbenennung z. B. in $\varphi = \forall v_i\, v_i \doteq v_j$, zu der nicht logisch äquivalenten Formel $\forall v_j\, v_j \doteq v_j$.
- Wenn v_i nicht frei für v_j in φ ist, führt die Umbenennung z. B. in $\varphi = \forall v_i \forall v_j\, v_i \doteq v_j$, zu der nicht logisch äquivalenten Formel $\forall v_j \forall v_j\, v_j \doteq v_j$.

Ableitungsregeln für Quantoren

Satz 3.25 *Angenommen $(\varphi \to \psi)$ ist eine allgemeingültige $\mathscr{L}$-Formel.*

Modus Ponens*:* *Wenn φ ebenfalls allgemeingültig ist, dann ist auch ψ allgemeingültig.*

$\exists$-Einführungsregel*: Wenn v_i nicht frei in ψ ist, dann ist auch $(\exists v_i \varphi \to \psi)$ allgemeingültig.*

$\forall$-Einführungsregel*: Wenn v_i nicht frei in φ ist, dann ist auch $(\varphi \to \forall v_i \psi)$ allgemeingültig.*

Wichtig ist insbesondere der Speziallfall $\varphi = \top$ mit $(\varphi \to \psi) \sim \psi$ und $(\varphi \to \forall v_i \psi) \sim \forall v_i \psi$.

Beweis Die Gültigkeit des *modus ponens* ist klar, da mit $(\mathscr{M},\beta) \vDash (\varphi \to \psi)$ und $(\mathscr{M},\beta) \vDash \varphi$ auch $(\mathscr{M},\beta) \vDash \psi$ gilt.

Für die $\exists$-Einführungsregel nimmt man an, dass $(\mathscr{M},\beta) \vDash \exists v_i \varphi$ gilt., d. h. dass es ein $m \in M$ gibt mit $(\mathscr{M}, \beta\frac{m}{v_i}) \vDash \varphi$. Aus der Allgemeingültigkeit von $(\varphi \to \psi)$ und *modus ponens* folgt $(\mathscr{M}, \beta\frac{m}{v_i}) \vDash \psi$, und da v_i nicht frei in ψ vorkommt, gilt auch $(\mathscr{M},\beta) \vDash \psi$.

Die $\forall$-Einführungsregel folgt daraus mit Dualität und den üblichen Substitutionen. □

Aus einer Implikation wie $\forall v_i \varphi \vDash \exists v_i \varphi$ bzw. $\vDash (\forall v_i \varphi \to \exists v_i \varphi)$ ergibt sich mit *modus ponens* eine Ableitungsregel: Wenn $\forall v_i \varphi$ allgemeingültig ist, dann auch $\exists v_i \varphi$. Die Umkehrung gilt aber

im Allgemeinen nicht! Beispielsweise besagt die ∀-Einführungsregel *nicht*, dass $((\varphi \to \psi) \to (\varphi \to \forall v_i\, \psi))$ allgemeingütig ist, wenn v_i nicht frei in φ ist. Man kann den *modus ponens* also nicht umkehren: Wenn aus der Allgemeingültigkeit von φ die Allgemeingültigkeit von ψ folgt, impliziert dies nicht die Allgemeingültigkeit von $(\varphi \to \psi)$.

Vollständige Junktoren-Quantoren-Systeme und Normalformen

Aus den Substitutionsprinzipien folgt zunächst, dass ein aussgenlogisch vollständiges Junktorensystem auch prädikatenlogisch gesehen ein vollständiges Junktorensysteme ist d. h. bis auf logische Äquivalenz kann auf die anderen Junktoren verzichtet werden. Aus den Quantorengesetzen ergibt sich zusätzlich:

Folgerung 3.26 $\{\exists\}$ *und* $\{\forall\}$ *sind* vollständige Quantorensysteme*, d. h. jede $\mathscr{L}$-Formel ist logisch äquivalent zu einer $\mathscr{L}$-Formel, in der das jeweils andere Quantorenzeichen nicht vorkommt.*

$\{\neg, \wedge, \exists\}$ *und* $\{\neg, \wedge, \forall\}$ *sind Beispiele* vollständiger Junktoren-Quantoren-Systeme*, d. h. jede $\mathscr{L}$-Formel ist logisch äquivalent zu einer $\mathscr{L}$-Formel, in der jeweils keine anderen Junktoren und Quantorenzeichen vorkommen.*

Definition 3.27 *Eine $\mathscr{L}$-Formel ist in* pränexer Normalform*, falls sie die Form*

$$\mathsf{Q}_1 v_{i_1}\, \ldots\, \mathsf{Q}_\mathsf{n} v_{i_n}\, \psi$$

hat mit $\mathsf{Q}_\mathsf{i} \in \{\exists, \forall\}$ und quantorenfreiem *ψ (d. h. ψ enthält keine Quantoren).*

Satz 3.28 *Jede $\mathscr{L}$-Formel ist zu einer $\mathscr{L}$-Formel in pränexer Normalform äquivalent. Zusätzlich kann der quantorenfreie Teil in DNF gebracht werden.*

Beweisskizze: Ohne Einschränkung kommen keine anderen Junktoren als $\neg, \wedge$ und $\vee$ vor. Gebundene Variablen werden so umbenannt, dass zu jedem Quantorzeichen eine eigene Variable gehört, die auch nicht frei in der Formel vorkommen. Negationen werden mit den Regeln von de Morgan „nach innen“ gezogen, Quantoren mit den Vertauschbarkeitsregeln mit $\wedge$ und $\vee$ nach außen.

Wenn man die Junktoren noch wie in der Aussagenlogik sortiert und Doppelnegationen eliminiert, bekommt man den quantorenfreien Teil in DNF. □

4 Der Vollständigkeitssatz

Definition 4.1 *Eine $\mathscr{L}$-*Theorie *ist eine Menge von $\mathscr{L}$-Aussagen.*

Eine $\mathscr{L}$-Theorie T heißt konsistent *(oder* erfüllbar*), wenn sie ein* Modell *hat, also eine $\mathscr{L}$-Struktur $\mathscr{M}$ so, dass $\mathscr{M} \vDash \varphi$ für alle $\varphi \in T$ (wofür man kurz $\mathscr{M} \vDash T$ schreibt).*

Eine $\mathscr{L}$-Theorie T heißt vollständig*, wenn sie konsistent ist und $T \vDash \varphi$ oder $T \vDash \neg\varphi$ für alle $\mathscr{L}$-Aussagen φ gilt.*

Lemma 4.2 **(a)** *T ist genau dann inkonsistent, wenn $T \vDash \bot$.*

(b) *T ist genau dann vollständig, wenn für alle $\mathscr{L}$-Aussagen φ* entweder *$T \vDash \varphi$* oder *$T \vDash \neg\varphi$.*

Beweis: (a) $T \vDash \bot \iff$ für alle $\mathscr{L}$-Strukturen $\mathscr{M}$ gilt $\mathscr{M} \nvDash T$ oder $\mathscr{M} \vDash \bot$. Letzteres gilt aber nie.

(b) Da stets $T \vDash \top$, impliziert die Bedingung insbesondere $T \nvDash \neg\top \sim \bot$, also ist T nach (a) konsistent. Umgekehrt kann für eine konsistente $\mathscr{L}$-Theorie T nicht $T \vDash \varphi$ und $T \vDash \neg\varphi$ gelten, weil dann auch $\mathscr{M} \vDash \varphi$ und $\mathscr{M} \vDash \neg\varphi$ für die Modelle $\mathscr{M}$ von T gelten müsste. □

Definition 4.3 (a) *Ein* $\mathscr{L}$-Kalkül $\mathbb{K}$ *besteht aus einer Menge von Ableitungsregeln. Ein* $\mathbb{K}$-Beweis *(oder* $\mathbb{K}$-Ableitung*) ist eine Folge von* $\mathscr{L}$*-Formeln* $\varphi_0 \dots \varphi_n$*, wobei jedes* φ_i *durch eine Ableitungsregel von* $\mathbb{K}$ *aus Formeln aus* $\{\varphi_0, \dots, \varphi_{i_1}\}$ *folgt.*[6]

Eine $\mathscr{L}$*-Theorie* T beweist φ in $\mathbb{K}$ *(oder:* φ ist $\mathbb{K}$-beweisbar aus T*), wenn es einen* $\mathbb{K}$*-Beweis* $\varphi_0 \dots \varphi_n$ *mit* $\varphi = \varphi_n$ *gibt und die Regeln von* $\mathbb{K}$ *die Verwendung von* $\mathscr{L}$*-Aussagen aus* T *im Beweis ohne Bedingungen erlauben. Dafür schreibt man* $T \vdash_{\mathbb{K}} \varphi$*.*

(b) *Eine* $\mathscr{L}$*-Theorie* T *heißt* $\mathbb{K}$-widersprüchlich*, falls* $T \vdash_{\mathbb{K}} \bot$*, sonst* $\mathbb{K}$-widerspruchsfrei.

(c) $\mathbb{K}$ *heißt* korrekt *oder* sound*, falls aus* $T \vdash_{\mathbb{K}} \varphi$ *stets* $T \vDash \varphi$ *folgt, und* vollständig*, falls umgekehrt aus* $T \vDash \varphi$ *stets* $T \vdash_{\mathbb{K}} \varphi$ *folgt.*

Satz 4.4 (Vollständigkeitssatz, Gödel 1929)
Es gibt für jede prädikatenlogische Sprache vollständige, korrekte Kalküle.

Gödel hat diesen Satz natürlich durch die Angabe konkreter Kalküle gezeigt. Der Beweis des Satzes – für einen anderen konkreten Kalkül – wird das gesamte Kapitel in Anspruch nehmen. Der konkrete Kalkül – sowohl bei Gödel wie auch hier – hat noch eine weitere schöne Eigenschaft (maschinelle Auswertbarkeit), die in einem späteren Kapitel eine Rolle spielen wird, und auch aus dem Beweis wird sich ein wichtiger Satz ergeben (Satz von Löwenheim-Skolem). Ein wichtiges Corollar kann direkt aus dem Satz gewonnen werden:

Satz 4.5 (Kompaktheitssatz der Prädikatenlogik) *Für jede* $\mathscr{L}$*-Theorie* T *gilt, dass sie genau dann konsistent ist, wenn jede endliche Teiltheorie* $T_0 \subseteq T$ *konsistent ist.*

Beweis: „⇒“ ist klar, da jedes Modell $\mathscr{M} \vDash T$ auch Modell von $T_0 \subseteq T$ ist.

„⇐“: Wenn T inkonsistent ist und $\mathbb{K}$ ein vollständiger, korrekter Kalkül, gilt nach dem Vollständigkeitssatz $T \vdash_{\mathbb{K}} \bot$. In einem entsprechenden $\mathbb{K}$-Beweis kommen aber nur endlich viele $\mathscr{L}$-Aussagen aus T vor, die also eine endliche $\mathbb{K}$-widersprüchliche Teilmenge T_0 bilden, die wiederum nach dem Vollständigkeitssatz auch inkonsistent ist. □

Definition 4.6 *Der hier betrachteten* $\mathscr{L}$*-Kalküle* $\mathbb{K}_{\mathscr{L}}$ *bestehen aus folgenden Regeln:*

Es gilt ohne Voraussetzungen $T \vdash_{\mathbb{K}_{\mathscr{L}}} \varphi$ *für die folgenden* $\mathscr{L}$*-Formeln* φ („Axiome“):

[Tautologie]	*alle* $\mathscr{L}$*-Tautologien*
[$\doteq$-Axiom]	*alle* $\mathscr{L}$*-Gleichheitsgesetze*
[$\forall$-Axiom]	*alle* $\mathscr{L}$*-Formeln der Form* $(\forall v_i\, \psi \to \psi[\frac{\tau}{v_i}])$*, wobei* v_i *frei für* τ *in* ψ *ist*
[Prämisse]	*alle* $\mathscr{L}$*-Aussagen in* T

und es gelten die folgenden „eigentlichen Ableitungsregeln“:

[modus ponens]	*wenn* $T \vdash_{\mathbb{K}_{\mathscr{L}}} (\varphi \to \psi)$ *und* $T \vdash_{\mathbb{K}_{\mathscr{L}}} \varphi$*, dann* $T \vdash_{\mathbb{K}_{\mathscr{L}}} \psi$
[$\to$-Einführung]	*wenn* $T \cup \{\varphi\} \vdash_{\mathbb{K}_{\mathscr{L}}} \psi$ *für eine* $\mathscr{L}$*-Aussage* φ*, dann* $T \vdash_{\mathbb{K}_{\mathscr{L}}} (\varphi \to \psi)$
[$\forall$-Einführung]	*wenn* $T \vdash_{\mathbb{K}_{\mathscr{L}}} \varphi$*, dann* $T \vdash_{\mathbb{K}_{\mathscr{L}}} \forall v_i\, \varphi$

[6] Es gibt andere Kalküle, deren Beweise aus komplizierteren Objekten als nur einzelnen Formeln bestehen.

Vorbemerkung 1: Ich arbeite in diesem Kalkül der Kürze wegen nur mit Allquantoren $\forall v_i$ und nur mit dem Junktor $\to$. Wenn man $\bot$ als 0-stelligen Junktor auffasst, ist $\{\to, \bot\}$ ein vollständiges Junktorensystem. In meiner Darstellung der Prädikatenlogik ist daher $\{\forall, \to\}$ ein vollständiges Junktoren-Quantoren-Sytem, weil $\bot$ als atomare $\mathscr{L}$-Formel vorhanden ist. Ich nutze allerdings der besseren Verständlichkeit wegen andere Zeichen als Abkürzung: $\exists v_i$ soll abkürzend für $\neg\forall v_i\neg$ stehen und $\neg\varphi$ abkürzend für $(\varphi \to \bot)$. [7]

Die $\mathscr{L}$-Gleichheitsgesetze, in denen eigentlich $\wedge$ vorkommt, sollen in $\mathbb{K}_{\mathscr{L}}$ in der folgenden, insgesamt logisch äquivalenten Version verwendet werden:

[Transitivität] $\forall v_i\, \forall v_j\, \forall v_k\, (\, v_i \doteq v_j \to (v_j \doteq v_k \to v_i \doteq v_k))$

[Kongruenz] mit $n = s(f)$ bzw. $n = s(R)$

$\forall v_{i_1} \dots \forall v_{i_{2n}} (v_{i_1} \doteq v_{i_{n+1}} \to (\dots \to (v_{i_n} \doteq v_{i_{2n}} \to f v_{i_1} \dots v_{i_n} \doteq f v_{i_{n+1}} \dots v_{i_{2n}}) \dots))$

$\forall v_{i_1} \dots \forall v_{i_{2n}} (v_{i_1} \doteq v_{i_{n+1}} \to (\dots \to (v_{i_n} \doteq v_{i_{2n}} \to (R v_{i_1} \dots v_{i_n} \to R v_{i_{n+1}} \dots v_{i_{2n}})) \dots))$

Vorbemerkung 2: Ich schreibe der Kürze halber $\vdash_{\mathscr{L}}$ statt $\vdash_{\mathbb{K}_{\mathscr{L}}}$. Es ist klar, dass ein $\mathbb{K}_{\mathscr{L}}$-Beweis einer $\mathscr{L}$-Formel φ auch ein $\mathbb{K}_{\mathscr{L}'}$-Beweis für jede größere Sprache $\mathscr{L}' \supseteq \mathscr{L}$ ist. Umgekehrt ist aber aus der Definition heraus zunächst völlig offen, ob sich aus einem $\mathbb{K}_{\mathscr{L}'}$-Beweis von φ auch ein $\mathbb{K}_{\mathscr{L}}$-Beweis gewinnen lässt. Dies ergibt sich erst als eine Folgerung aus dem Vollständigkeitssatz für die Kalküle $\mathbb{K}_{\mathscr{L}}$. Da aber im Laufe seines Beweises die betrachtete Sprache erweitert werden muss, ist es notwendig, sich über die jeweils verwendete Sprache im Klaren zu sein.

Beispiele für Ableitungen im Kalkül $\mathbb{K}_{\mathscr{L}}$ wird es im Laufe des Beweises des Vollständigkeitssatzes geben. Der Beweis des folgenden vorbereitenden Lemmas liefert einen ersten Eindruck:

Lemma 4.7 *Sei T eine $\mathscr{L}$-Theorie und φ eine $\mathscr{L}$-Formel bzw. in Teil (a) eine $\mathscr{L}$-Aussage.*

(a) *Wenn $T \cup \{\neg\varphi\}$ $\mathbb{K}$-widersprüchlich ist, gilt $T \vdash_{\mathscr{L}} \varphi$.*

(b) *Wenn $T \vdash_{\mathscr{L}} \varphi$ und $T \vdash_{\mathscr{L}} \neg\varphi$, ist T $\mathbb{K}_{\mathscr{L}}$-widersprüchlich.*

Beweis:

(a)	[Voraussetzung]	$T \cup \{\neg\varphi\} \vdash_{\mathscr{L}} \bot$	
	[$\to$-Einführung]	$T \vdash_{\mathscr{L}} (\neg\varphi \to \bot)$	aus Zeile (1)
	[Tautologie]	$T \vdash_{\mathscr{L}} ((\neg\varphi \to \bot) \to \varphi)$	
	[modus ponens]	$T \vdash_{\mathscr{L}} \varphi$	aus Zeilen (2) und (3)
(b)	[Voraussetzung 1]	$T \vdash_{\mathscr{L}} \varphi$	
	[Voraussetzung 2]	$T \vdash_{\mathscr{L}} \neg\varphi$	
	[Tautologie]	$T \vdash_{\mathscr{L}} (\varphi \to (\neg\varphi \to \bot))$	
	[modus ponens]	$T \vdash_{\mathscr{L}} (\neg\varphi \to \bot)$	aus Zeilen (1) und (3)
	[modus ponens]	$T \vdash_{\mathscr{L}} \bot$	aus Zeilen (2) und (4)

□

Satz 4.8 (Vollständigkeitssatz für $\mathbb{K}_{\mathscr{L}}$) *Eine $\mathscr{L}$-Theorie T beweist genau dann eine $\mathscr{L}$-Formel φ in $\mathbb{K}_{\mathscr{L}}$, wenn φ aus T logisch folgt:*

$$T \vdash_{\mathscr{L}} \varphi \iff T \vDash \varphi$$

Mit $\varphi = \bot$ gilt insbesondere: T ist genau dann $\mathbb{K}_{\mathscr{L}}$-widerspruchsfrei, wenn T konsistent ist; mit $T = \emptyset$ gilt insbesondere: φ ist genau dann $\mathbb{K}_{\mathscr{L}}$-beweisbar, wenn φ allgemeingültig ist.

[7] Man kann den Kalkül auf die anderen Zeichen ausdehnen, indem man eine zusätzliche Regel einführt, welche es erlaubt, innerhalb jeder Formel die Zeichenfolge $\neg\forall v_i\neg$ durch $\exists v_i$ zu ersetzen. Für die anderen Junktoren reicht sogar die Tautologie-Regel für $\mathscr{L}$-Formeln mit beliebigen Junktoren. Der Beweis ist allerdings etwas subtil, weil das Prinzip der äquivalenten Substitution nicht in den Kalkül eingebaut ist, also keine nachträgliche Ersetzung innerhalb einer Formel erlaubt.

Beweis von „⇒“, also der Korrektheit des Kalküls: Dies folgt aus den im vorigen Kapitel beschriebenen Gesetzen und Ableitungsregeln, wobei man sich im Beweis der Ableitungsregeln auf Modelle von T beschränkt. Nicht explizit aufgeführt ist die $\to$-Einführungsregel, deren Korrektheit aber leicht einzusehen ist. □

Der Beweis der Vollständigkeit des Kalküls wird den Rest des Kapitels in Anspruch nehmen.

Schritt 1: $\mathscr{L}$-Aussagen reichen

Lemma 4.9 *Sei C eine Menge von Konstanten und $\mathscr{L}_C = \mathscr{L} \cup C$. Sei T eine $\mathscr{L}$-Theorie, $\varphi(v_1, \dots, v_n)$ eine $\mathscr{L}$-Formel und seien $c_1, \dots, c_n \in C$ paarweise verschieden. Dann gilt:*

(a) $T \models \varphi \iff T \models \forall v_1 \dots \forall v_n \, \varphi \iff T \models \varphi\,[\frac{c_1}{v_1}] \cdots [\frac{c_n}{v_n}]$

(b) $T \vdash_{\mathscr{L}} \varphi \iff T \vdash_{\mathscr{L}} \forall v_1 \dots \forall v_n \, \varphi \iff T \vdash_{\mathscr{L}_C} \varphi\,[\frac{c_1}{v_1}] \cdots [\frac{c_n}{v_n}] \iff T \vdash_{\mathscr{L}_C} \varphi$

Beweis: (a) ist klar nach den Definitionen im vorherigen Kapitel.

(b) Aus $T \vdash_{\mathscr{L}} \varphi$ folgt mit sukzessiver [∀-Einführung] $T \vdash_{\mathscr{L}} \forall v_1 \dots \forall v_n \, \varphi$.

Aus $T \vdash_{\mathscr{L}} \forall v_1 \dots \forall v_n \, \varphi$ folgt zunächst $T \vdash_{\mathscr{L}_C} \forall v_1 \dots \forall v_n \, \varphi$, da ein $\mathbb{K}_{\mathscr{L}}$-Beweis auch ein Beweis in $\mathscr{L}_C$ ist. Mit dem [∀-Axiom] $T \vdash_{\mathscr{L}_C} (\forall v_1 \dots \forall v_n \, \varphi \to \forall v_2 \dots \forall v_n \, \varphi[\frac{c_1}{v_1}])$ und [modus ponens] erhält man daraus $T \vdash_{\mathscr{L}_C} \forall v_2 \dots \forall v_n \, \varphi[\frac{c_1}{v_1}]$ und sukzessive $T \vdash_{\mathscr{L}_C} \varphi[\frac{c_1}{v_1}] \cdots [\frac{c_n}{v_n}]$.

Angenommen $T \vdash_{\mathscr{L}_C} \varphi[\frac{c_1}{v_1}] \cdots [\frac{c_n}{v_n}]$. Sei k so groß, dass die Variablen $v_{k+1}, \dots, v_{k+n}$ in dem entsprechenden $\mathbb{K}_{\mathscr{L}_C}$-Beweis nicht vorkommen. Dann kann man aus diesem $\mathbb{K}_{\mathscr{L}_C}$-Beweis in folgender Weise einen $\mathbb{K}_{\mathscr{L}_C}$-Beweis für $\varphi[\frac{v_{k+1}}{v_1}] \cdots [\frac{v_{k+n}}{v_n}]$ machen: Man ersetzt zunächst jedes Vorkommen von c_i durch v_{k+1}. In allen Fällen, in denen nicht Konstanten in besonderer Weise ins Spiel kommen, bleiben dabei Ableitungsregeln erhalten. Die beiden Sonderfälle sind das $\doteq$-Axiom für 0-stellige Funktionszeichen und die ∀-Axiome: (1) Bei den ∀-Axiomen muss nur sichergestellt sein, dass die ersetzte Variable frei für den ersetzenden Term ist. Dies ist bei vollständig neuen Variablen sicher der Fall. (2) Das ggf. vorkommenden $\doteq$-Axiom $c_i \doteq c_i$ muss man durch eine kurze Ableitung ersetzen, nämlich durch [$\doteq$-Axiom] $\forall v_0 \, v_0 \doteq v_0$, [∀-Axiom] $(\forall v_0 \, v_0 \doteq v_0 \to v_{k+i} \doteq v_{k+i})$ und [modus ponens]. Also gilt $T \vdash_{\mathscr{L}_C} \varphi[\frac{v_{k+1}}{v_1}] \cdots [\frac{v_{k+1}}{v_n}]$, und da in dem Beweis keine der Konstanten c_i mehr vorkommen sogar $T \vdash_{\mathscr{L}} \varphi[\frac{v_{k+1}}{v_1}] \cdots [\frac{v_{k+n}}{v_n}]$.

Mit sukzessiver [∀-Einführung] folgt daraus $T \vdash_{\mathscr{L}} \forall v_{1+k} \dots \forall v_{n+k} \, \varphi[\frac{v_{1+k}}{v_1}] \cdots [\frac{v_{n+k}}{v_n}]$. Darin kann man wie oben über die passenden [∀-Axiome] v_i für v_{i+k} „re-substituieren“ (weil v_{i+k} für die freien Vorkommen von v_i eingesetzt wurde, ist jedes Vorkommen von v_{i+k} auch wieder frei für v_i) und auf $T \vdash_{\mathscr{L}} \varphi$ schließen. [8]

Damit hat man in einem Ringschluss die Äquivalenz der ersten drei Aussagen gezeigt. Da $\mathscr{L}$ beliebig ist, geht auch $\mathscr{L} = \mathscr{L}_C$, was die Äquivalenz mit $T \vdash_{\mathscr{L}_C} \varphi$ zeigt. □

Man kann sich im Beweis des Vollständigkeitssatzes folglich auf $\mathscr{L}$-Aussagen φ beschränken.

Lemma 4.10 *Für den Beweis der anderen Richtung, also der Vollständigkeit des Kalküls, reicht es zu zeigen, dass $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreie Theorien T konsistent sind.*

Beweis: Angenommen $T \nvdash_{\mathscr{L}} \varphi$ für eine $\mathscr{L}$-Aussage φ. Dann ist nach Lemma 4.7 $T \cup \{\neg\varphi\}$ $\mathbb{K}_{\mathscr{L}}$-widerspruchsfrei, hat also nach Annahme ein Modell $\mathscr{M}$. Dieses Modell zeigt $T \not\models \varphi$. □

[8] Man muss den Umweg über die neuen Variablen v_{k+i} gehen und kann nicht direkt die Konstanten c_i durch v_i ersetzen, weil man in dem nicht näher bekannten $\mathbb{K}_{\mathscr{L}_C}$-Beweis keine Kontrolle darüber hat, ob dann in allen Anwendungen von ∀-Axiomen die „frei für“-Bedingung erfüllt wäre.

Schritt 2: Henkin-Konstanten

Die Grundidee für die Konstruktion eines Modells kann man in der Konstruktion des Quotientenkörpers $\mathbb{Q}$ aus $\mathbb{Z}$ sehen: Hier nimmt man als Elemente Äquivalenzklassen von Termen, also Äquivalenzklassen von Bruchzahlen $\frac{a}{b}$. Im Allgemeinen reicht das aber nicht: Wenn man z. B. einen algebraisch abgeschlossenen Körper konstruieren will, braucht man auch Nullstellen von Polynomen. In analoger Weise muss man hier „Lösungen“ für gewisse Formeln „erfinden“. Das sind die sogenannten Henkin-Konstanten.

Definition 4.11 *Sei $\mathscr{L}_0 := \mathscr{L}$, dann induktiv $\mathscr{L}_{n+1} := \mathscr{L}_n \cup \{c_\varphi \mid \varphi(v_i)\ \mathscr{L}_n$-Formel$\}$ und schließlich $\mathscr{L}_H := \bigcup_{n\in\mathbb{N}} \mathscr{L}_n$, wobei $c_\varphi \notin \mathscr{L}_n$ jeweils neue Konstanten sind. Weiter sei $T_0 := T$,*

$$T_{n+1} := T_n \cup \left\{(\exists v_i\, {}^*\varphi \to \varphi[\tfrac{c_\varphi}{v_i}]) \;\middle|\; \varphi(v_i)\ \mathscr{L}_n\text{-Formel}\right\}$$

und $T_H := \bigcup_{n\in\mathbb{N}} T_n$. Dabei soll $*$ bedeuten, dass im Falle $\varphi = \neg\psi$ an Stelle der $\mathscr{L}$-Formel $(\exists v_i\, \varphi \to \varphi[\frac{c_\varphi}{v_i}]) = (\neg\forall v_i \neg\neg\psi \to \neg\psi[\frac{c_\varphi}{v_i}])$ die $\mathscr{L}$-Formel $(\neg\forall v_i\psi \to \neg\psi[\frac{c_\varphi}{v_i}])$ steht.

T_H heißt *Henkin-Theorie* von T, die Konstanten c_φ heißen *Henkin-Konstanten* und die Formeln $(\exists v_i\, {}^*\varphi \to \varphi[\frac{c_\varphi}{v_i}])$ *Henkin-Axiome*.

Lemma 4.12
Die Henkin-Theorie einer $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreien $\mathscr{L}$-Theorie ist $\mathbb{K}_{\mathscr{L}_H}$-widerspruchsfrei.

Beweis: Zunächst ist eine $\mathbb{K}_{\mathscr{L}}$-wiederspruchsfreie $\mathscr{L}$-Theorie T auch $\mathbb{K}_{\mathscr{L}_H}$-widerspruchsfrei, weil man Konstanten in einem $\mathbb{K}_{\mathscr{L}_H}$-Beweis von $\bot$ nach Lemma 4.9 eliminieren kann.

Die Vereinigung T' einer aufsteigenden Kette $T'_0 \subseteq T'_1 \subseteq T'_2 \subseteq \cdots$ von $\mathbb{K}_{\mathscr{L}_H}$-widerspruchsfreien $\mathscr{L}_H$-Theorien ist wieder $\mathbb{K}_{\mathscr{L}_H}$-widerspruchsfrei, denn in einem $\mathbb{K}_{\mathscr{L}_H}$-Beweis von $\bot$ aus T' kommen nur endlich viele $\mathscr{L}_H$-Aussagen aus T' vor, die bereits in einem T'_n liegen, das dann $\mathbb{K}_{\mathscr{L}_H}$-widersprüchlich wäre. Also gibt es nach dem Zorn'schen Lemma eine maximale $\mathbb{K}_{\mathscr{L}_H}$-widerspruchsfreie, durch Hinzunahme von Henkin-Axiomen aus T entstandene Theorie T''.

Angenommen ein Henkin-Axiom $(\exists v_i^*\varphi \to \varphi[\frac{c_\varphi}{v_i}])$ fehlt in T''. Wegen der Maximalität gilt mit [$\to$-Einführung] $T'' \vdash_{\mathscr{L}_H} ((\exists v_i^*\varphi \to \varphi[\frac{c_\varphi}{v_i}]) \to \bot)$.

Fall 1 $\varphi \neq \neg\psi$: Da $(((A_0 \to A_1) \to \bot) \to A_0)$ und $(((A_0 \to A_1) \to \bot) \to \neg A_1)$ Tautologien sind, erhält man mit [modus ponens] daraus $T'' \vdash_{\mathscr{L}_H} \exists v_i\varphi = \neg\forall v_i\neg\varphi$ und $T'' \vdash_{\mathscr{L}_H} \neg\varphi[\frac{c_\varphi}{v_i}]$. Aus letzterem folgt mit Lemma 4.9 $T'' \vdash_{\mathscr{L}_H} \forall v_i\neg\varphi$, also ist T'' nach Lemma 4.7 $\mathbb{K}_{\mathscr{L}_H}$-widersprüchlich: Widerspruch!

Fall 2 $\varphi = \neg\psi$: Wie in Fall 1 erhält man $T'' \vdash_{\mathscr{L}_H} \neg\forall v_i\psi$ und $T'' \vdash_{\mathscr{L}_H} \neg\neg\psi[\frac{c_\varphi}{v_i}]$. Mit der Tautologie $(\neg\neg A_0 \to A_0)$ bekommt man daraus zunächst $T'' \vdash_{\mathscr{L}_H} \psi[\frac{c_\varphi}{v_i}]$, um dann wieder mit Lemma 4.9 $T'' \vdash_{\mathscr{L}_H} \forall v_i\psi$ und einen Widerspruch zu erhalten. □

Man kann sich für den Beweis des Vollständigkeitssatzes also darauf beschränken, die Konsistenz einer $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreien Henkin-$\mathscr{L}$-Theorie zu zeigen.

Schritt 3: Deduktive Vervollständigung

Definition 4.13 *Eine $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreie $\mathscr{L}$-Theorie T heißt* deduktiv vollständig*, wenn $T \vdash_{\mathscr{L}} \varphi$ oder $T \vdash_{\mathscr{L}} \neg\varphi$ für alle $\mathscr{L}$-Aussagen φ gilt.*

Lemma 4.14 *Jede $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreie $\mathscr{L}$-Theorie lässt sich zu einer deduktiv vollständigen $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreien $\mathscr{L}$-Theorie erweitern.*

Beweis: Wie im Beweis von Lemma 4.12 gibt es eine maximale $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreie Erweiterung T der gegebenen $\mathscr{L}$-Theorie. Angenommen $T \nvdash_{\mathscr{L}} \varphi$. Dann ist insbesondere $\varphi \notin T$. Wegen der Maximalität von T ist also $T \cup \{\varphi\}$ $\mathbb{K}_{\mathscr{L}}$-widersprüchlich, d. h. es gilt $T \cup \{\varphi\} \vdash_{\mathscr{L}} \bot$ bzw. mit [$\to$-Einführung] $T \vdash_{\mathscr{L}} (\varphi \to \bot) = \neg\varphi$. Also ist T deduktiv vollständig. □

Folgerung 4.15 *Man kann jede Henkin-Theorie zu einer deduktiv vollständigen Henkin-Theorie erweitern.*

Beweis: Die Eigenschaft Henkin-Theorie zu sein geht bei einer Erweiterung der Theorie in derselben Sprache nicht verloren, da keine neuen Formeln hinzukommen! □

Man kann sich für den Beweis des Vollständigkeitssatzes also darauf beschränken, die Konsistenz einer deduktiv vollständigen $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreien Henkin-Theorie zu zeigen.

Schritt 4: Die Termstruktur

Sei nun T eine deduktiv vollständige $\mathbb{K}_{\mathscr{L}}$-widerspruchsfreien Henkin-$\mathscr{L}$-Theorie. Auf der Menge der $\mathscr{L}$-Terme ist eine binäre Relation definiert durch

$$\tau_1 \approx \tau_2 \quad :\Longleftrightarrow \quad T \vdash_{\mathscr{L}} \tau_1 \doteq \tau_2$$

Lemma 4.16 $\approx$ *ist eine Äquivalenzrelation.*

Beweis: Dies folgt aus der $\mathbb{K}_{\mathscr{L}}$-Beweisbarkeit der ersten drei Gleichheitsgesetze. Ich zeige beispielhaft die Transitivität, und zwar ausführlich als Beispiel für einen $\mathbb{K}_{\mathscr{L}}$-Beweis. Ohne Einschränkung sollen die Variablen v_0, v_1, v_2 in τ_0, τ_1, τ_2 nicht vorkommen. Dann lassen sich aus T in $\mathbb{K}_{\mathscr{L}}$ beweisen:

[Voraussetzung]	$\tau_0 \doteq \tau_1$
[Voraussetzung]	$\tau_1 \doteq \tau_2$
[$\doteq$-Axiom]	$\forall v_0 \forall v_1 \forall v_2 (v_0 \doteq v_1 \to (v_1 \doteq v_2 \to v_0 \doteq v_2))$
[$\forall$-Axiom]	$(\forall v_0 \forall v_1 \forall v_2 (v_0 \doteq v_1 \to (v_1 \doteq v_2 \to v_0 \doteq v_2)) \to \forall v_1 \forall v_2 (\tau_0 \doteq v_1 \to (v_1 \doteq v_2 \to \tau_0 \doteq v_2)))$
[modus ponens]	$\forall v_1 \forall v_2 (\tau_0 \doteq v_1 \to (v_1 \doteq v_2 \to \tau_0 \doteq v_2))$
[$\forall$-Axiom]	$(\forall v_1 \forall v_2 (\tau_0 \doteq v_1 \to (v_1 \doteq v_2 \to \tau_0 \doteq v_2)) \to \forall v_2 (\tau_0 \doteq \tau_1 \to (\tau_1 \doteq v_2 \to \tau_0 \doteq v_2)))$
[modus ponens]	$\forall v_2 (\tau_0 \doteq \tau_1 \to (\tau_1 \doteq v_2 \to \tau_0 \doteq v_2))$
[$\forall$-Axiom]	$(\forall v_2 (\tau_0 \doteq \tau_1 \to (\tau_1 \doteq v_2 \to \tau_0 \doteq \tau_2)) \to (\tau_0 \doteq \tau_1 \to (\tau_1 \doteq \tau_2 \to \tau_0 \doteq \tau_2)))$
[modus ponens]	$(\tau_0 \doteq \tau_1 \to (\tau_1 \doteq \tau_2 \to \tau_0 \doteq \tau_2))$
[modus ponens]	$(\tau_1 \doteq \tau_2 \to \tau_0 \doteq \tau_2)$
[modus ponens]	$\tau_0 \doteq \tau_2$

□

Die $\approx$-Äquivalenzklasse eines $\mathscr{L}$-Terms τ schreibe ich $\widetilde{\tau}$.

Lemma 4.17 *Auf der Menge M der $\approx$-Äquivalenzklassen von $\mathscr{L}$-Termen wird eine $\mathscr{L}$-Struktur $\mathscr{M}$ definiert durch*

$$f^{\mathscr{M}}\left(\widetilde{\tau_1}, \ldots, \widetilde{\tau_{s(f)}}\right) \quad := \quad \widetilde{f\tau_1 \ldots \tau_{s(f)}}$$
$$\left(\widetilde{\tau_1}, \ldots, \widetilde{\tau_{s(R)}}\right) \in R^{\mathscr{M}} \quad :\Longleftrightarrow \quad T \vdash_{\mathscr{L}} R\tau_1 \ldots \tau_{s(R)}$$

Beweis: Die Wohldefiniertheit folgt analog zum vorherigen Lemma aus den $\doteq$-Axiomen, wobei nun die Kongruenz-Gleichheitsgesetze benötigt werden. □

Schritt 5: $\mathscr{M}$ ist das gesuchte Modell

Da man dies nur induktiv über den Aufbau der Formeln beweisen kann, braucht man eine Belegung. Naheliegenderweise nimmt man

$$\tilde{\beta} : \mathbb{N} \to M,\ i \mapsto \widetilde{v_i}$$

Lemma 4.18 *Für jeden $\mathscr{L}$-Term τ gilt $\tilde{\beta}(\tau) = \tilde{\tau}$.*

Beweis per Induktion über den Aufbau der Terme: Im Induktionsschritt ist

$$\tilde{\beta}(f\tau_1 \dots \tau_{s(f)}) = f^{\mathscr{M}}(\tilde{\beta}(\tau_1), \dots, \tilde{\beta}(\tau_{s(f)})) = f^{\mathscr{M}}(\widetilde{\tau_1}, \dots, \widetilde{\tau_{s(f)}}) = \widetilde{f\tau_1 \dots \tau_{s(f)}} \qquad \square$$

Satz 4.19 *Für jede $\mathscr{L}$-Formel φ gilt*

$$(\mathscr{M}, \tilde{\beta}) \vDash \varphi \iff T \vdash_{\mathscr{L}} \varphi$$

Insbesondere ist $\mathscr{M}$ ein Modell von T und T ist somit konsistent.

Wir brauchen für den Beweis noch ein technisches Lemma:

Lemma 4.20 *Für jeden $\mathscr{L}$-Term τ gibt es eine Konstante in $\widetilde{\tau}$.*

Beweis: Sei v_n eine Variable, die in τ nicht vorkommt, und c die Henkin-Konstante der $\mathscr{L}$-Formel $\tau \doteq v_n$, so dass also das Henkin-Axiom $(\neg\forall v_n \neg \tau \doteq v_n \to \tau \doteq c)$ in T liegt. Es reicht $T \vdash_{\mathscr{L}} \neg\forall v_n \neg\tau \doteq v_n$ zu zeigen, denn dann folgt mit [modus ponens] $T \vdash_{\mathscr{L}} \tau \doteq c$, d. h. $c \in \widetilde{\tau}$. Andernfalls gilt wegen der deduktiven Vollständigkeit $T \vdash_{\mathscr{L}} \forall v_n \neg\tau \doteq v_n$, woraus mit dem passenden [$\forall$-Axiom] und [modus ponens] $T \vdash_{\mathscr{L}} \neg\tau \doteq \tau$ folgt, was aber im Widerspruch zu Lemma 4.16 und der $\mathbb{K}_{\mathscr{L}}$-Widerspruchsfreiheit von T steht! $\square$

Beweis von Satz 4.19 per Induktion über den Aufbau der $\mathscr{L}$-Formeln:

$T \vdash_{\mathscr{L}} \top$ gilt wegen [Tautologie] und $T \nvdash_{\mathscr{L}} \bot$, da T nach Voraussetzung $\mathbb{K}$-widerspruchsfrei ist. Nach Definition von $\approx$ und $\mathscr{M}$ und Lemma 4.18 sieht man außerdem:

$$\begin{aligned}
(\mathscr{M}, \tilde{\beta}) \vDash\ & \tau_1 \doteq \tau_2 && \iff \tilde{\beta}(\tau_1) = \tilde{\beta}(\tau_2) \iff \widetilde{\tau_1} = \widetilde{\tau_2} \iff T \vdash_{\mathscr{L}} \tau_1 \doteq \tau_2 \\
(\mathscr{M}, \tilde{\beta}) \vDash\ & R\tau_1 \dots, \tau_n && \iff \big(\tilde{\beta}(\tau_1), \dots, \tilde{\beta}(\tau_n)\big) \in R^{\mathscr{M}} \\
& && \iff \quad (\widetilde{\tau_1}, \dots, \widetilde{\tau_n}) \quad \in R^{\mathscr{M}} \quad \iff T \vdash_{\mathscr{L}} R\tau_1 \dots, \tau_n
\end{aligned}$$

Für die Implikation hat man per Induktion und deduktiver Vollständigkeit von T zunächst:

$$\begin{aligned}
(\mathcal{M}, \tilde{\beta}) \vDash (\varphi \to \psi) \quad & \iff \quad (\mathcal{M}, \tilde{\beta}) \nvDash \varphi \quad \text{oder} \quad (\mathcal{M}, \tilde{\beta}) \vDash \psi \\
& \iff \quad T \nvdash_{\mathscr{L}} \varphi \quad \text{oder} \quad T \vdash_{\mathscr{L}} \psi \\
& \iff \quad T \vdash_{\mathscr{L}} \neg\varphi \quad \text{oder} \quad T \vdash_{\mathscr{L}} \psi
\end{aligned}$$

Aus den Tautologien $(\neg A_0 \to (A_0 \to A_1)$ und $(A_1 \to (A_0 \to A_1)$ bekommt man mit [modus ponens] rechts in beiden Fällen $T \vdash_{\mathscr{L}} (\varphi \to \psi)$. Gilt umgekehrt $T \vdash_{\mathscr{L}} (\varphi \to \psi)$, folgt aus [modus ponens] $T \nvdash_{\mathscr{L}} \varphi$ oder $T \vdash_{\mathscr{L}} \psi$, was die gesucht Äquivalenz abschließt.

Im Quantorenschritt gilt zunächst:

$$\begin{aligned}
(\mathcal{M}, \tilde{\beta}) \vDash \forall v_i \varphi \iff & \text{ für alle } m \in M \text{ gilt } (\mathcal{M}, \tilde{\beta}\tfrac{m}{v_i}) \vDash \varphi \\
\textit{Lemma 4.20} \iff & \text{ für alle Konstanten } c \in \mathscr{L} \text{ gilt } (\mathcal{M}, \tilde{\beta}\tfrac{\tilde{c}}{v_i}) \vDash \varphi \\
\textit{Substitutionslemma und Lemma 4.18} \iff & \text{ für alle Konstanten } c \in \mathscr{L} \text{ gilt } (\mathcal{M}, \tilde{\beta}) \vDash \varphi[\tfrac{c}{v_i}] \\
\textit{Induktion} \iff & \text{ für alle Konstanten } c \in \mathscr{L} \text{ gilt } T \vdash_{\mathscr{L}} \varphi[\tfrac{c}{v_i}]
\end{aligned}$$

Nun muss man noch zeigen, dass „für alle Konstanten $T \vdash_{\mathscr{L}} \varphi[\frac{c}{v_i}]$“ äquivalent zu $T \vdash_{\mathscr{L}} \forall v_i \varphi$ ist. Die Richtung „$\Leftarrow$“ gilt allgemein, denn aus $T \vdash_{\mathscr{L}} \forall v_i \varphi$ folgt $T \vdash_{\mathscr{L}} \varphi[\frac{c}{v_i}]$ für jede Konstante c durch das passende [$\forall$-Axiom] und [modus ponens].

Die Umkehrrichtung „$\Rightarrow$“ gilt dagegen nur in der speziellen Situation hier: Aus $T \nvdash_{\mathscr{L}} \forall v_i \varphi$ folgt mit deduktiver Vollständigkeit $T \vdash_{\mathscr{L}} \neg\forall v_i \varphi$, Mit dem Henkin-Axiom $(\neg\forall v_i \varphi \to \neg\varphi[\frac{c_{\neg\varphi}}{v_i}]) \in T$ und [modus ponens] bekommt man daraus $T \vdash_{\mathscr{L}} \neg\varphi[\frac{c_{\neg\varphi}}{v_i}]$, also mit erneut deduktiver Vollständigkeit $T \nvdash_{\mathscr{L}} \varphi[\frac{c_{\neg\varphi}}{v_i}]$. □

Das im Beweis des Vollständigkeitssatzes konstruierte Modell hat nicht mehr Elemente als es $\mathscr{L}$-Terme gibt. Man kann sich überlegen (mehr dazu im Kapitel zur Mengenlehre), dass es nicht mehr als abzählbare viele $\mathscr{L}$-Terme gibt, wenn die Sprache $\mathscr{L}$ abzählbar ist, und höchstens so viele $\mathscr{L}$-Terme wie die Mächtigkeit von $\mathscr{L}$, wenn $\mathscr{L}$ überabzählbar ist. Insgesamt ergibt dies:

Satz 4.21 (Löwenheim-Skolem [9])
Jede konsistente $\mathscr{L}$-Theorie hat ein Modell $\mathscr{M}$ mit $|M| \leqslant \max\{|\mathbb{N}|, |\mathscr{L}|\}$.

5 Ergänzungen

5.1 Prädikatenlogik zweiter Stufe

Hier nur ein kurzer, informeller Ausblick:

Für eine Prädikatenlogik zweiter Stufe erweitert man die Syntax der Prädikatenlogik erster Stufe in der Sprache $\mathscr{L}$ um Relationsvariablen V_i^n der Stelligkeit n. Zusätzliche atomare Formeln sind dann von der Form $V_i^n \tau_1 \dots \tau_n$ für $\mathscr{L}$-Terme τ_i. Außerdem sind zweitstufige Quantifizierungen $\forall V_i^n \varphi$ und $\exists V_i^n \varphi$ bei der Formelbildung erlaubt. (Möglicherweise will man noch zweitstufige Funktions- und Relationszeichen, das ist aber für die folgenden Anmerkungen nicht wichtig.)

Belegungen in $\mathscr{L}$-Strukturen $\mathscr{M}$ werden auf die Relationsvariablen ausgedehnt, indem jedem V_i^n eine Teilmenge von M^n zugeordnet wird. Die Auswertung einer zweistufigen $\mathscr{L}$-Formel erfolgt dann in der naheliegenden Weise.

Man kann nun (für beliebiges $\mathscr{L}$) verschiedenste zweistufige $\mathscr{L}$-Formeln angeben kann, die nur in endlichen Strukturen erfüllt sind. Man sagt etwa, dass jede partielle Ordnung ein maximales Element hat:

$$\varphi = \forall V_0^2 \big(\big(\forall v_0\, V_0^2 v_0 v_0 \wedge \forall v_0 \forall v_1 ((V_0^2 v_0 v_1 \wedge V_0^2 v_1 v_0) \to v_0 \doteq v_1) \wedge \\ \forall v_0 \forall v_1 \forall v_2 ((V_0^2 v_0 v_1 \wedge V_0^2 v_1 v_2) \to V_0^2 v_0 v_2)\big) \to \exists v_0 \forall v_1 (v_0 \doteq v_1 \vee \neg V_0^2 v_0 v_1)\big)$$

oder dass jede injektive Abbildung auch surjektiv ist.

Mit den (erststufigen) $\mathscr{L}$-Formeln φ_n, die ausdrücken, dass es mindestens n Elemente gibt, hat man eine endlich konsistente, aber nicht konsistente Formelmenge $\{\varphi\} \cup \{\varphi_n \mid n \in \mathbb{N}\}$. Also gilt der Kompaktheitssatz *nicht* für die zweitstufige Prädikatenlogik, und damit kann es auch keinen praktikablen [10] vollständigen, korrekten Kalkül für die zweitstufige Prädikatenlogik geben, weil der Kompaktheitssatz eine Folgerung daraus wäre.

[9] In der Modelltheorie beweist man eine stärkere Version des Satzes.

[10] Man kann natürlich den trivialen Kalkül hinschreiben mit dem Axiom „$\vdash \varphi$, falls φ allgemeingültig“. Dieser Kalkül ist per Definition korrekt und vollständig, aber sinnlos. Was man möchte, sind „maschinell ausführbare“ Kalküle, was aber erst präzisiert werden muss.

Auch der Satz von Löwenheim-Skolem gilt zweitstufig nicht: Man drückt z.B. aus, dass man eine nicht-triviale dichte strikte Ordnung wie $(\mathbb{Q}, <)$ hat, und dass jede beschränkte Teilmenge ein Supremum hat:

$$\begin{aligned}
&\exists V_0^2\big(\ \exists v_0 \exists v_1 \neg v_0 \doteq v_1 \wedge \forall v_0 \neg V_0^2 v_0 v_0 && \textit{nicht-trivial und irreflexiv}\\
&\quad \wedge \forall v_0 \forall v_1 \forall v_2 ((V_0^2 v_0 v_1 \wedge V_0^2 v_1 v_2) \to V_0^2 v_0 v_2) && \textit{transitiv}\\
&\quad \wedge \forall v_0 \forall v_1 (V_0^2 v_0 v_1 \vee v_0 \doteq v_1 \vee V_0^2 v_1 v_0) && \textit{trichotomisch}\\
&\quad \wedge \forall v_0 \forall v_2 (V_0^2 v_0 v_2 \to \exists v_1 (V_0^2 v_0 v_1 \wedge V_0^2 v_1 v_2)) && \textit{dicht}\\
&\quad \wedge \forall V_1^1\big(\ \exists v_0 (V_1^1 v_0 \wedge \exists v_1 (V_0^2 v_0 v_1 \wedge \neg V_1^1 v_1)) && \textit{nicht leere und beschränkte Teilmenge}\\
&\qquad \to\ \exists v_2 (\forall v_3 (V_1^1 v_3 \to (v_3 \doteq v_2 \vee V_0^2 v_3 v_2)) && \textit{hat kleinste obere Schranke}\\
&\qquad\quad \wedge \forall v_4 (\forall v_3 (V_1^1 v_3 \to (v_3 \doteq v_2 \vee V_0^2 v_3 v_4)) \to (v_2 \doteq v_4 \vee V_0^2 v_2 v_4)))\big)\big)
\end{aligned}$$

Ein Modell dieser zweitstufigen $\mathscr{L}$-Theorie ist immer überabzählbar.

Erwähnt sei der Satz von Lindström, der die Prädikatenlogik erster Stufe dadurch charakterisiert, dass der Kompaktheitssatz und der Satz von Löwenheim-Skolem gelten (mehr dazu in Ebbinghaus, Flum, Thomas „Einführung in die mathematische Logik“).

5.2 Boole'sche Algebren

Sei $\mathscr{L}_{BA} = \{\sqcap, \sqcup, c, 0, 1\}$ mit zweistelligen Funktionszeichen $\sqcap$ und $\sqcup$, einem einstelligen Funktionszeichen c und Konstanten 0 und 1. Statt $c(x)$ schreibt man in diesem Zusammenhang üblicherweise x^c mit höherer Bindungsstärke als $\sqcap$ und $\sqcup$.

Definition 5.1 *Eine* Boole'sche Algebra *ist ein Modell der $\mathscr{L}_{BA}$-Theorie, die Folgendes ausdrückt:*

(a) $\sqcap$ *und* $\sqcup$ *sind idempotent, kommutativ und assoziativ.*

(b) $\sqcap$ *ist distributiv über* $\sqcup$ *und umgekehrt.*

(c) 1 *ist neutrales Element von* $\sqcap$ *und* 0 *neutrales Element von* $\sqcup$.

(d) *Es gelten die* Komplementgesetze $x \sqcap x^c = 0$ *und* $x \sqcup x^c = 1$.

Dies ist kein minimales Axiomensystem, denn man kann z. B. auf die Idempotenz und Assoziativität verzichten. Traditionell werden allerdings viel mehr Axiome aufgeführt, etwa die *Absorptionsgesetze* $x \sqcup (x \sqcap y) = x$ und $x \sqcap (x \sqcup y) = x$, die *de Morgan'schen Regeln* $(x \sqcap y)^c = x^c \sqcup y^c$ und $(x \sqcup y)^c = x^c \sqcap y^c$, die *Extremalgesetze* $x \sqcap 0 = 0$ und $x \sqcup 1 = 1$ sowie die *„Doppelnegationsregel“* $x^{cc} = x$.

Eine $\{\sqcap, \sqcup\}$-Struktur, die (a) erfüllt, heißt *Verband* bzw. *distributiver Verband*, wenn sie zusätzlich (b) erfüllt. Eine $\{\sqcap, \sqcup, 0, 1\}$-Struktur, die (a) und (c) erfüllt, heißt *beschränkter Verband*. Eigenschaft (d) wird dann *komplementär* genannt, so dass Boole'sche Algebren auch als *komplementäre beschränkte distributive Verbände* charakterisierbar sind.

In einer Boole'schen Algebra $\mathscr{B}$ ist durch

$$x \sqsubseteq y :\iff x \sqcap y = x \ \big(\iff x \sqcup y = y\big)$$

eine partielle Ordnung definiert. Man stellt fest, dass $x \sqcap y$ ein Infimum und $x \sqcup y$ ein Supremum von x und y bezüglich dieser Ordnung ist und 0 kleinstes sowie 1 größtes Element.

Sei $\mathscr{L}$ eine prädikatenlogische Sprache. Logische Äquivalenz ist per Definition klarerweise eine Äquivalenzrelation auf der Menge der $\mathscr{L}$-Formeln. Ich schreibe $[\![\varphi]\!]$ für die Äquivalenzklasse von

φ. Aufgrund der Substitutionsregeln sind die folgenden Operationen wohldefiniert:

$$\begin{aligned} [\![\varphi]\!] \wedge [\![\psi]\!] &:= [\![(\varphi \wedge \psi)]\!] \\ [\![\varphi]\!] \vee [\![\psi]\!] &:= [\![(\varphi \vee \psi)]\!] \\ \neg[\![\varphi]\!] &:= [\![\neg\varphi]\!] \end{aligned}$$

Satz 5.2 *Mit den eben definierten Operationen sowie den Äquivalenzklassen $[\![\bot]\!]$ und $[\![\top]\!]$ als Interpretationen von 0 bzw. 1 sind die folgenden Beispiele Boole'scher Algebren:*

- *Die Algebra $\mathscr{A}$ der Äquivalenzklassen aussagenlogischer Formeln.*
- *Die Algebra $\mathscr{A}_n$ der Äquivalenzklassen aussagenlogischer Formeln $\varphi(A_0, \ldots, A_{n-1})$, in denen höchstens die Aussagenvariablen $A_0, \ldots, A_{n-1}$ vorkommen.*
- *Die Algebra $\mathscr{F}(\mathscr{L})$ der Äquivalenzklassen von $\mathscr{L}$-Formeln.*
- *Die Algebra $\mathscr{F}_n(\mathscr{L})$ der Äquivalenzklassen von $\mathscr{L}$-Formeln $\varphi(v_0, \ldots, v_{n-1})$, in denen höchstens die freien Variablen $v_0, \ldots, v_{n-1}$ vorkommen.*

Beweis folgt aus den in Kapitel 2 bzw. 3 bewiesenen logischen Gesetzen. □

Solche Boole'sche Algebren heißen *Tarski-Lindenbaum-Algebren*. Die partielle Ordnung ist darin durch die Implikation gegeben:

$$\varphi \sqsubseteq \psi \iff \vDash (\varphi \to \psi)$$

$\mathscr{A}_n$ ist eine endliche Boole'sche Algebra, die anderen Beispiele sind unendlich.

Dualität

Definition 5.3 *Gegeben eine Boole'sche Algebra $\mathscr{B} = (B; \sqcap^{\mathscr{B}}, \sqcup^{\mathscr{B}}, c^{\mathscr{B}}, 0^{\mathscr{B}}, 1^{\mathscr{B}})$, ist die* duale Boole'sche Algebra *$\mathscr{B}^*$ definiert als $\mathscr{B}^* = (B; \sqcup^{\mathscr{B}}, \sqcap^{\mathscr{B}}, c^{\mathscr{B}}, 1^{\mathscr{B}}, 0^{\mathscr{B}})$.*

Die partielle Ordnung $\sqsubseteq^{\mathscr{B}^*}$ ist also die umgedrehte Ordnung $\sqsupseteq^{\mathscr{B}}$ der Ordnung von $\mathscr{B}$.

Satz 5.4 *$\mathscr{B}$ und $\mathscr{B}^*$ sind isomorph zueinander via $b \mapsto b^c$.*

Beweis: Nachrechnen! □

Dieser Isomorphismus wird auch als *Antiautomorphimsus* von $\mathscr{B}$ bezeichnet. Er erklärt die Symmetrie der logischen Gesetze bezüglich $\wedge/\vee$ und $\top/\bot$. Außerdem überführt er im Falle der Tarski-Lindenbaum-Algebren $\forall$ in $\exists$ und umgekehrt (wegen $\neg\forall v_i\varphi \sim \exists v_i\neg\varphi$ bzw. $\neg\exists v_i\varphi \sim \forall v_i\neg\varphi$).

Die Dualität lässt sich in verschiedener Weise als logisches Prinzip übersetzen. Wenn φ eine $\mathscr{L}$-Formel ist, in der die Junktoren $\to$ und $\leftrightarrow$ nicht vorkommen, sei φ^* die $\mathscr{L}$-Formel, die aus φ entsteht, indem alle Vorkommen von $\wedge$ durch $\vee$, von $\top$ durch $\bot$ und von $\forall$ durch $\exists$ ersetzt werden und jeweils (simultan) umgekehrt. Dann gilt:

- $\varphi \sim \psi$ genau dann, wenn $\varphi^* \sim \psi^*$.
- Insbesondere: φ ist genau dann allgemeingültig, wenn φ^* nicht erfüllbar ist.

Sei zudem $\overline{\varphi}$ die $\mathscr{L}$-Formel, die aus φ entsteht, indem alle atomaren Teilformeln simultan durch ihre Negation ersetzt werden. Dann gilt:

- $\neg\varphi \sim \overline{\varphi^*}$
- Insbesondere: φ ist genau dann allgemeingültig, wenn $\neg\overline{\varphi^*}$ allgemeingültig ist.

Der Stone'sche Darstellungssatz

Für jede beliebige Menge M ist die *Potenzmengenalgebren* $(\mathfrak{P}(M); \cap, \cup, \complement, \emptyset, M)$ ein weiteres Beispiel einer Boole'schen Algebra.

Definition 5.5 *Ein Homomomorphismus zwischen Boole'schen Algebren $\mathscr{B}$ und $\mathscr{B}'$ (oder ein $\mathscr{L}_{BA}$-Homomorphismus) ist eine Abbildung $h : B \to B'$, die mit allen Interpretationnen der Funktionszeichen verträglich ist, für die also $h(a \sqcap^{\mathscr{B}} b) = h(a) \sqcap^{\mathscr{B}'} h(b)$, $h(a \sqcup^{\mathscr{B}} b) = h(a) \sqcup^{\mathscr{B}'} h(b)$, $h(a^{c^{\mathscr{B}}}) = h(a)^{c^{\mathscr{B}'}}$, $h(0^{\mathscr{B}}) = 0^{\mathscr{B}'}$ und $h(1^{\mathscr{B}}) = 1^{\mathscr{B}'}$ gilt.* [11]

Die Menge der Wahrheitswerte $\{0, 1\}$ bildet eine Boole'sche Algebra, indem man sie wahlweise mit $\mathfrak{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$ oder mit $\mathscr{A}_0 = \{[\![\bot]\!], [\![\top]\!]\}$ identifiziert. Eine $\mathscr{L}$-Struktur $\mathscr{M}$ zusammen mit einer Belegung β definiert einen Homomorphismus

$$h^{(\mathscr{M},\beta)} : \mathscr{F}(\mathscr{L}) \to \{0, 1\}, \quad h^{(\mathscr{M},\beta)}(\varphi) = 1 \iff (\mathscr{M}, \beta) \vDash \varphi$$

Analog definiert in der Aussagenlogik eine Wahrheitswertverteilung (d. h. Belegung der Aussagenvariablen mit Wahrheitswerten) $\omega : \{A_i \mid i \in \mathbb{N}\} \to \{0, 1\}$ einen Homomorphismus $h^\omega : \mathscr{A} \to \{0, 1\}$. Im Fall der endlichen Algebren $\mathscr{A}_n$ ist h^ω bereits festgelegt durch die eingeschränkte „partielle Belegung" $\{A_0, \ldots, A_{n-1}\} \to \{0, 1\}$. Bis auf logische Äquivalenz ist einen aussagenlogische Formel durch ihren Wahrheitswertverlauf bestimmt, also insbesondere durch die Menge der partiellen Belegungen, die die Formel wahr machen. Dies definiert einen Isomorphismus Boole'scher Algebren

$$\begin{aligned} \mathscr{A}_n &\to \mathfrak{P}\big(\{\{A_0, \ldots, A_{n-1}\} \to \{0, 1\}\}\big) \\ \varphi &\mapsto \{\omega \mid h^\omega(\varphi) = 1\} \end{aligned}$$

Da es 2^n solche partielle Belegungen gibt, ist $|\mathscr{A}_n| = 2^{2^n}$. $\mathscr{A}_n$ ist übrigens die *freie von n Elementen erzeugte Boole'sche Algebra*, also in einem präzisne Sinn die größte Boole'sche Algebra, die von n Elementen erzeugt wird.

Man kann diese Überlegung auf beliebige Boole'sche Algebren ausdehnen, bekommt im Allgemeinen aber keinen Isomosphimus, sondern nur eine Einbettung:

Satz 5.6 (Darstellungssatz von Stone)
Jede Boole'sche Algebra ist Unteralgebra einer Potenzmengenalgebra.

Beweis: Man betrachtet die Menge $\mathrm{Hom}(\mathscr{B}, \{0, 1\})$ der Homomorphismen Boole'scher Algebren von $\mathscr{B}$ nach $\{0, 1\}$. Dann definiert man

$$\begin{aligned} S : \mathscr{B} &\to \mathfrak{P}\big(\mathrm{Hom}(\mathscr{B}, \{0, 1\})\big) \\ b &\mapsto \big\{h : \mathscr{B} \to \{0, 1\} \;\big|\; h(b) = 1\big\} \end{aligned}$$

Man rechnet recht problemlos nach, dass S ein Homomorphismus Boole'scher Algebren ist. (Einzig die Verträglichkeit mit der Komplementabbildung ist nicht ganz offensichtlich. Man muss zeigen, dass ein Homomorphismus h von b und b^c genau eines auf 0 und eines auf 1 abbildet, was aus $x \sqcup x^c = 1$ und $x \sqcap x^c = 0$ folgt.)

Die Injektivität ist im Falle der Tarski-Lindenbaum-Algebren leicht: Falls $S(\varphi) = S(\psi)$, gilt insbesondere $h^{(\mathscr{M},\beta)}(\varphi) = h^{(\mathscr{M},\beta)}(\psi)$ für alle $\mathscr{M}$ und β, d. h. φ und ψ sind logisch äquivalent zueinander bzw. $[\![\varphi]\!] = [\![\psi]\!]$ in $\mathscr{F}(\mathscr{L})$.

[11] Auf analoge Weise definiert man $\mathscr{L}$-Homomorphismen zwischen beliebigen $\mathscr{L}$-Strukturen $\mathscr{M}$ und $\mathscr{M}'$ für die Funktionszeichen in $\mathscr{L}$. Für Relationszeichen $R \in \mathscr{L}_R$ fordert man $(a_1, \ldots, a_n) \in R^{\mathscr{M}} \Rightarrow (h(a_1), \ldots, h(a_n)) \in R^{\mathscr{M}'}$. Mit der stärkeren Bedingungen „$\Leftrightarrow$" ergibt sich ein *starker $\mathscr{L}$-Homomorphismus*.

Im Allgemeinen ist die Injektivität schwieriger: Man muss für je zwei Elemente $b_0 \neq b_1$ einen Homomorphismus $h : \mathscr{B} \to \{0,1\}$ mit $h(b_0) \neq h(b_1)$ finden. Dazu analysiert man Kerne von Homomorphismen: $\mathrm{Kern}(h) := \{b \in B \mid h(b) = 0\}$ ist (a) abgeschlossen unter $\sqcup$, (b) nach unten abgeschlossen unter $\sqsubseteq$, d. h. $a \sqsubseteq b \in \mathrm{Kern}(h)$ impliziert $a \in \mathrm{Kern}(h)$, und (c) enthält für jedes $b \in B$ entweder b oder b^c. Teilmengen von B mit den beiden Abgeschlossenheitseigenschaften heißen *Ideale*, mit zusätzlich der dritten Eigenschaft *maximale Ideale*. Jedes maximale Ideal I ist auch Kern eines Homomorphismus, nämlich $h_I : \mathscr{B} \to \{0,1\}$ mit $h_I(b) = 0 \iff b \in I$.

Wenn o. B. d. A. $b_1 \not\sqsubseteq b_0$ kann man nun mit dem Zorn'schen Lemma ein maximales Ideal I mit $b_0 \in I$ und $b_1 \notin I$ konstruieren; h_I ist dann der gesuchte Homomorphimus. □

Statt den Kern eines Homomorphismus kann man ebenso gut auch den *dualen Kern* $\mathrm{Kern}^*(h) := \{b \in B \mid h(b) = 1\}$ betrachten. Er ist (a) abgeschlossen unter $\sqcap$, (b) nach oben abgeschlossen unter $\sqsubseteq$, d. h. $a \sqsubseteq b$ und $a \in \mathrm{Kern}(h)$ impliziert $b \in \mathrm{Kern}(h)$, und (c) enthält für jedes $b \in B$ entweder b oder b^c. Teilmengen von B mit diesen beiden Abgeschlossenheitseigenschaften heißen *Filter*, mit zusätzlich der dritten Eigenschaft *Ultrafilter*. Jeder Ultrafilter U ist dualer Kern eines Homomorphismus, nämlich $h_U : \mathscr{B} \to \{0,1\}$ mit $h_U(b) = 1 \iff b \in U$.

6 Mengenlehre

6.1 Einführung

Mengen sind von Georg Cantor in die Mathematik eingeführte mathematische Objekte, die seitdem in der Mathematik omnipräsent geworden sind. Ein sorgloser Umgang mit Mengen kann allerdings leicht zu Widersprüchen führen, daher braucht man klare und widerspruchsfreie Regeln für den Umgang mit Mengen, typischerweise durch ein Axiomensystem. Dazu gibt viele alternative Möglichkeiten; in der Praxis durchgesetzt hat sich das System ZFC, benannt nach den Logikern Ernst *Zermelo* und Abraham *Fraenkel* sowie dem Auswahlaxiom, engl. *Choice*.

ZFC ist eine $\mathscr{L}_{\mathrm{ML}}$-Theorie in der *Sprache der Mengenlehre* $\mathscr{L}_{\mathrm{ML}} = \{\in\}$ mit dem zweistelligen Relationszeichen $\in$. In der Praxis werden allerdings weitere vertraute Symbole benutzt wie $\subseteq, \cap, \cup, \emptyset, \mathfrak{P}$ bzw. Schreibweisen wie $\{a_1, \ldots, a_n\}$, die man entweder als Abkürzungen verstehen muss oder die durch sogenannte definitorische Erweiterungen eingeführt werden. Diese Symbole und Schreibweisen werde ich ohne große Erläuterungen verwenden.

Mengen vs. Mengen

Man hofft, dass ZFC eine konsistente $\mathscr{L}_{\mathrm{ML}}$-Theorie ist, kann dies aber nach dem Zweiten Gödel'schen Unvollständigkeitssatz nicht beweisen. Man nimmt also an, dass es $\mathscr{L}_{\mathrm{ML}}$-Strukturen gibt, die Modelle von ZFC sind. Die Elemente des Universums M eines solchen Modells $\mathscr{M}$ der Mengenlehre heißen dann *Mengen*. Andererseits ist M selbst (nach der allgemeinen Definition von $\mathscr{L}$-Strukturen) eine nicht näher spezifizierte Menge. In ZFC schließt man aus, dass Mengen Elemente von sich selbst sind und insbesondere, dass es eine Menge aller Mengen gibt. Man hat daher zwei Arten oder zumindest zwei Betrachtungsweisen von Mengen: „$\mathscr{M}$-Mengen" relativ zu Modellen $\mathscr{M}$ von ZFC und „Meta-Mengen", aus denen die Universen der Modelle von ZFC stammen. Dies ist eine unbefriedigende und tendenziell verwirrende Situation. Es gibt mehrere Ansätze, dieses Problematik aufzulösen:

- Man kann eine Hierarchie von Modellen von ZFC annehmen, so dass das Universum M eines Modells $\mathscr{M} \vDash \mathsf{ZFC}$ Element eines Modells einer höheren Hierarchiestufe ist.

- Man nimmt an, dass alle Mengen zusammen etwas bilden, was man „Klasse“ nennt, und dass man diese Klasse, obwohl sie keine Menge ist, dennoch in einem verallgemeinerten Sinn als eine Art $\mathscr{L}_{\mathrm{ML}}$-Struktur $\mathbb{V}$ auffassen kann, die ZFC erfüllt.
  Ich übernehme diesen Ansatz und schreibe $\mathbb{V} \vDash \varphi$ dafür, dass eine $\mathscr{L}_{\mathrm{ML}}$-Formel φ im allgemeinen mengentheoretischen Universum gilt. Per Definition ist also $\mathbb{V} \vDash \mathsf{ZFC}$ sein.
- Eine Art Mischung aus beiden Ansätzen ist die *von Neumann–Gödel–Bernays*'sche Mengenlehre, die ZFC erweitert und axiomatisch mit zwei Hierarchiestufen arbeitet: Mengen und (echten) Klassen.

Universeller Anspruch

Man hat festgestellt, dass man alle bisherige Mathematik in ZFC modellisieren kann. Konkret bedeutet dies, dass man z. B. in jedem Modell $\mathscr{M} \vDash \mathsf{ZFC}$ eine $\mathscr{M}$-Menge $N_{\mathscr{M}}$ findet und darauf Operationen $+$ und $\cdot$ definieren kann sowie eine Relation $\leqslant$, so dass sich $(N_{\mathscr{M}}; +, \cdot, \leqslant)$ wie die natürlichen Zahlen mit Addition, Multiplikation und Anordnung verhält. Ähnlich findet man die reellen und komplexen Zahlen in jedem Modell der Mengenlehre.

Man kann also (als Erfahrungstatsache, nicht als Theorem) alle Mathematik innerhalb von ZFC betreiben. Manche Mathematiker:innen tun so oder sind sogar der Meinung, dass es nur das mengentheoretische Universum $\mathbb{V}$ gibt und alle mathematischen Objekte Elemente oder Teilklassen von $\mathbb{V}$ sind.

Naive Mengenlehre

Von Cantor stammt die Beschreibung:

> *Unter einer „Menge“ verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unserer Anschauung oder unseres Denkens (welche die „Elemente“ von M genannt werden) zu einem Ganzen.*

Cantor war bereits bewusst, dass nicht jede solche Zusammenfassung als Menge betrachtet werden darf, ohne dass es zu Widersprüchen kommt. Wollte man das versuchen, bekäme man die $\mathscr{L}_{\mathrm{ML}}$-Theorie der „naiven Mengenlehre“ mit folgenden Axiomen:

[Extensionalität] $\forall v_0 \forall v_1 \big(\forall v_2 (v_2 \in v_0 \leftrightarrow v_2 \in v_1) \to v_0 \doteq v_1\big)$

[Komprehension] für jede $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi(v_0, v_1, \ldots, v_n)$ das Axiom:
$$\forall v_1 \ldots \forall v_n \exists v_{n+1} \forall v_0 \big(v_0 \in v_{n+1} \leftrightarrow \varphi(v_0, v_1, \ldots, v_n)\big)$$

[Komprehension] ist ein sogenanntes *Axiomenschema*, d. h. nicht ein Einzelaxiom, sondern abhängig von einem Parameter, hier φ. Es besagt in üblicher Schreibweise, dass für jede $\mathscr{L}_{\mathrm{ML}}$-Formel φ und alle $a_1, \ldots, a_n$ die Menge $\{x \mid \varphi(x, a_1, \ldots, a_n) \text{ gilt}\}$ existiert.

Naive Mengenlehre ist inkonsistent, denn mit $\varphi(v_0) = \neg v_0 \in v_0$ bekommt man die *Russell-Zermelo'sche Antinomie*:
$$\exists v_1 \forall v_0 \big(v_0 \in v_1 \leftrightarrow \neg v_0 \in v_0\big))$$
Mit $v_0 = v_1$ ergibt dies die inkonsistente $\mathscr{L}_{\mathrm{ML}}$-Formel $v_1 \in v_1 \leftrightarrow \neg v_1 \in v_1$.

Alle Axiomatisierungen der Mengenlehre versuchen, solche Widersprüche zu vermeiden, aber möglichst viel der naiven Mengenlehre zu bewahren, indem das Komprehensionsschema geeignet eingeschränkt wird. In ZFC ist Komprehension daher nur für gewisse $\mathscr{L}_{\mathrm{ML}}$-Formeln φ erlaubt.

Andere Heransgehensweisen arbeiten mit Hierarchisierungen von Mengen und schränken den Quantor $\forall v_0$ im Komprehensionsaxiom auf Mengen einer niedrigeren Hierarchiestufe als v_1 ein.

Von dieser letzten Heransgehensweise übernimmt man für ZFC zumindest terminologisch gerne die zweite Hierarchiestufe: Alle Mengen, die die $\mathscr{L}_{\mathrm{ML}}$-Formel φ im Komprehensionsaxiom erfüllen, bilden zusammen eine *Klasse*, und wenn diese Klasse keine Menge ist, eine *echte Klasse*. Die Zusammenstellung $\{m \text{ Menge} \mid m \notin m\}$ ist also echte Klasse und bewirkt damit keinen Widerspruch, weil sie sich nicht unter den betrachteten Mengen m befindet.

Mengen ohne Urelemente

Das Cantor'sche, vor-axiomatische Verständnis von Mengen besteht darin, dass man irgendwelche gegebenen Objekte hat wie z. B. die natürlichen Zahlen $0, 1, 2, \dots$, aus denen man Mengen wie $\{1, 2, 3\}$ oder $\{2, 3, 5, 7, \dots\}$ oder $\mathbb{N}$ bilden kann, und dann weiter $\mathfrak{P}(\mathbb{N})$ oder $\{1, 2, 3\} \times \mathbb{N}$ oder die Menge der Funktionen von $\{2, 3, 5, 7, \dots\} \to \mathbb{N}$, etc.

Innerhalb eines Modells von ZFC oder innerhalb von $\mathbb{V}$ gibt es dagegen *nur* Mengen, d. h. jedes Element einer Menge ist auch stets selbst eine Menge. Elemente, die selbst keine Mengen sind – sogenannte *Urelemente* – könnte man zwar auch in die Mengenlehre einbauen. Dies wird aber eher selten getan, weil man mit Urelementen beweisbar keine stärkere Ausdruckskraft gewinnt, die Theorie aber komplizierter wird.

6.2 Das Axiomensystem ZFC

Wie üblich soll $\exists! v_i$ eine Kurzschreibweise für den Quantor „*es gibt genau ein* v_i" sein, d. h. für eine $\mathscr{L}$-Formel φ steht $\exists! v_i\, \varphi(v_i)$ als Abkürzung für $\exists v_i(\varphi \wedge \forall v_j(\varphi[\frac{v_j}{v_i}] \to v_i \doteq v_j))$ mit $j \neq i$. Der besseren Lesbarkeit halber werde ich für $\varphi[\frac{v_j}{v_i}]$ zukümnftig $\varphi(y_j)$ schreiben und außerdem v, w, x, y, z und Varianten davon als Variablen für Individuenvariablen benutzen.

Definition 6.1 ZFC *ist definiert als die aus den im Folgenden beschriebenen acht Axiomen* [Extensionalität] *bis* [Auswahl] *und zwei Axiomentschemata* [Aussonderung] *und* [Ersetzung] *bestehende* $\mathscr{L}_{\mathrm{ML}}$*-Theorie.*

Die Axiome in ZFC sind nicht unabhängig, zum Beispiel könnte man auf [Paarmenge], [Aussonderung] und [Leere Menge] verzichten. Man findet daher in der Literatur verschiedene Varianten.

[Extensionalität] $\forall y_0 \forall y_1 \big(\forall x(x \in y_0 \leftrightarrow x \in y_1) \to y_0 \doteq y_1\big)$

d. h. zwei Mengen sind genau dann gleich, wenn sie die gleichen Elemente haben.

Fünf Spezialfälle der Komprehension:

[Paarmenge] $\forall y_1 \forall y_2 \exists w \forall x \big(x \in w \leftrightarrow (x \doteq y_1 \vee x \doteq y_2)\big)$

[Potenzmenge] $\forall y \exists w \forall x \big(x \in w \leftrightarrow \forall x'(x' \in x \to x' \in y)\big)$

[Vereinigung] $\forall y \exists w \forall x \big(x \in w \leftrightarrow \exists y'(x \in y' \wedge y' \in y)\big)$

d. h. zu je zwei Mengen y_1, y_2 existiert die Paarmenge $\{y_1, y_2\}$ und zu jeder Menge y die Potenzmenge $\mathfrak{P}(y)$ und die Vereinigungsmenge $\bigcup y$, die die Vereinigung der Elemente von y ist. Alle diese Mengen sind nach dem Extensionalitätsprinzip eindeutig.

Folgerung: Mit dem Paarmengen- und dem „großen" Vereinigungsaxiom bekommt man zu zwei Mengen y_1 und y_2 auch deren „kleine" Vereinigung $y_1 \cup y_2 := \bigcup\{y_1, y_2\}$.

[Aussonderung] Für jede $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi(x, y_1, \ldots, y_n)$

$$\forall y \forall y_1 \ldots \forall y_n \exists w \forall x \big(x \in w \leftrightarrow (x \in y \wedge \varphi(x, y_1, \ldots, y_n))\big)$$

[Ersetzung] Für jede $\mathscr{L}_{\mathrm{ML}}$-Formel $\psi(v, v', y_1, \ldots, y_n)$

$$\forall y \forall y_1 \ldots \forall y_n \big(\forall v \exists! v' \, \psi(v, v', y_1, \ldots, y_n) \to \exists w \forall x \big(x \in w \leftrightarrow \exists v (v \in y \wedge \psi(v, x, y_1, \ldots, y_n))\big)\big)$$

d. h. für feste Parameter $y_1, \ldots, y_n$ gibt es für jede Menge y die Teilmenge aller Elemente von y, die φ erfüllen, und das Bild $f[y]$ von y unter der durch ψ auf y definierten Funktion. Alle diese Mengen sind wiederum nach dem Extensionalitätsprinzip eindeutig.

Folgerung: Mit dem Aussonderungsaxiom bekommt man zu zwei Mengen y und y_1 auch deren Schnitt $y \cap y_1 := \{x \in y \mid x \in y_1\}$ und Differenz $y \setminus y_1 := \{x \in y \mid x \notin y_1\}$.

Aus dem Aussonderungsaxiom folgt, dass es in ZFC nicht die Menge aller Mengen geben kann, denn sonst könnte man daraus die Russell-Zermelo'sche Antinomie ableiten.

Zwei Axiome, die die Existenz bestimmter Mengen garantieren:

[Leere Menge] $\exists w \forall x \, \neg x \in w$

d. h. es gibt die leere Menge $\emptyset$, die nach dem Extensionalitätsprinzip eindeutig ist. [12]

[Unendlichkeit] $\exists w \big(\emptyset \in w \wedge \forall x (x \in w \to x \cup \{x\} \in w)\big)$

d. h. es gibt (nicht eindeutige) Mengen U, die die leere Menge als Element enthalten und mit jedem ihrer Elemente x auch $x \cup \{x\}$. Solch eine Menge U enthält damit mindestens $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \ldots$, also in einem intuitiven Sinn unendlich viele Elemente. Dazu später mehr.

Dass für jedes x auch $x \cup \{x\}$ eine Menge ist, sieht man so: $\{x\}$ bekommt man aus dem Paarmengenaxiom mit $x = y_1 = y_2$. Dann bildet man die Paarmenge $\{x, \{x\}\}$, deren Vereinigung gerade $x \cup \{x\}$ ist.

Schließlich zwei speziellere Axiome:

[Fundierung] $\forall y \, (\neg y \doteq \emptyset \to \exists x \, (x \in y \wedge x \cap y \doteq \emptyset)$

Das Fundierungsaxiom schließt merkwürdige Mengen aus, zum Beispiel Mengen m mit $m \in m$. Denn da m das einzige Element von $\{m\}$ ist, muss nach dem Fundierungsaxiom $m \cap \{m\} = \emptyset$ gelten, was genau dann der Fall ist, wenn $m \notin m$.

Das Fundierungsaxiom sieht zunächst merkwürdig aus. Es ist aber insofern unproblematisch, als man aus einem Modell von ZFC ohne Fundierung ein Modell von ZFC gewinnen kann, indem man sich auf die Mengen beschränkt, die das Fundierungsaxiom erfüllen. Außerdem kann man beweisen, dass jede nicht-fundierte Menge zu einer fundierten in Bijektion steht. In diesem Sinn können in nicht-fundierten Modellen auch keine Phänomene auftreten, die nicht schon in fundierten Modellen vorkommen. Die eigentliche Bedeutung des Fundierungsaxioms ist, dass es keine unendlichen absteigenden Ketten von Mengen $\cdots \in m_n \in \cdots \in m_1 \in m_0$ gibt.

Für das nächste Axiom braucht man das Konzept des *geordneten Paars* (y_1, y_2) von zwei Mengen y_1 und y_2. Die entscheidende Eigenschaft des geordneten Paars ist

$$(y_1, y_2) = (z_1, z_2) \iff y_1 = z_1 \text{ und } y_2 = z_2$$

Definition 6.2 *Für Mengen y_1 und y_2 wird das* geordnete Paar (y_1, y_2) *definiert als das* Kuratowski-Paar $\big\{\{y_1\}, \{y_1, y_2\}\big\}$ *und das* Kartesische Produkt $y_1 \times y_2$ *als* $\{(x_1, x_2) \mid x_i \in y_i\}$.

[12] Die leere Menge könnte man auch mit dem Aussonderungsaxiom aus einer beliebigen Menge gewinnen.

Lemma 6.3 *Das Kuratowski-Paar erfüllt die oben beschriebene Anforderung an ein geordnetes Paar. Die Extistenz des Kartesischen Produkts als Menge folgt aus den Axiomen von* ZFC*.*

Beweis: Übung □

[Auswahl] $\forall y\big(\neg\emptyset \in y \to \exists w \forall x\big(x \in y \to \exists! x'\big(\,(x,x') \in w \wedge x' \in x\big)\big)\big)$

Wenn $(x,x') \in w$, ist $x' \in \bigcup\bigcup w$. Die Bedingung rechts von der ersten Implikation sagt daher, dass $w \cap (y \times \bigcup\bigcup w)$ der Graph einer auf y definierten „Auswahlfunktion“ f mit der Eigenschaft $f(x) \in x$ für alle $x \in y$ ist.

Das Auswahlaxiom hat äquivalente Formulierungen, die es extrem plausibel erscheinen lassen, aber auch Konsequenzen, die paradoxal anmuten (z. B. das Banach-Tarski-Paradoxon). Es wurde nicht von Anfang an von allen akzeptiert und hat dadurch einen etwas mystischen Status erlangt. Manche Mathematiker:innen meinen daher, sie müssten (im Gegensatz zu den anderen Axiomen) die Verwendung des Auswahlaxioms besonders erwähnen. Allerdings stimmt dies nicht immer: Nicht immer, wenn in der Mathematik etwas ausgewählt wird, braucht man das Auswahlaxiom, und nicht immer, wenn man das Auswahlaxiom braucht, wird explizit etwas ausgewählt. Ich habe das Auswahlaxiom im Beweis des Vollständigkeitssatzes und im Beweis des Darstellungssatzes von Stone mehrfach in Form des Zorn'schen Lemmas verwendet.

Vielleicht ist folgende Beobachtung zu den Spezialfällen der Komprehension nützlich: Wenn man sich Mengen so hierarchisiert vorstellt, dass alle Mengen der Stufe $n+1$ ausschließlich Elemente der Stufe n enthalten, dann machen Paarmengen- und Potenzmengenaxiom aus Mengen der Stufe n Mengen der Stufe $n+1$, das Vereinigungsaxiom macht aus einer Menge der Stufe $n+1$ eine Menge der Stufe n und das Aussonderungsaxiom behält die Stufe bei.

Relationen und Funktionen

Ein für die Mathematik grundlegendes Konzept ist das der Abbildung oder Funktion. Sowohl im Ersetzungs- als auch im Auswahlaxiom ist es implizit vorhanden. Mit Hilfe der geordneten Paare kann man es explizit machen.

Definition 6.4 *Eine* Relation r zwischen Mengen m_1 und m_2 *ist ein Tripel* [13] (r, m_1, m_2)*, wobei* $r \subseteq m_1 \times m_2$. Definitionsbereich *und* Bildbereich *der Relation sind definiert als*

$$\begin{aligned} \mathrm{dom}(r) &:= \{x \mid \textit{es gibt } y \textit{ mit } (x,y) \in r\} \subseteq m_1 \\ \mathrm{im}(r) &:= \{y \mid \textit{es gibt } x \textit{ mit } (x,y) \in r\} \subseteq m_2 \end{aligned}$$

Während Definitions- und Bildbereich einer Relation durch r festgelegt sind, sind dies der *Quellbereich* m_1 und der *Zielbereich* m_2 nicht. Wenn $m_1 = \mathrm{dom}(r)$ und $m_2 = \mathrm{im}(r)$ oder wenn der genaue Quell- und Zielbereich irrelevant sind, spricht man auch kurz von r als der Relation.

Lemma 6.5 *Für jede Relation sind* $\mathrm{dom}(r)$ *und* $\mathrm{im}(r)$ *Mengen.*

Beweis: Wenn $(a,b) \in r$, sind $a, b \in \bigcup\bigcup r$. Man erhält dann $\mathrm{dom}(r)$ und $\mathrm{im}(r)$ mit dem Aussonderungsaxiom. □

Definition 6.6 *Eine* Funktion f von einer Menge m_1 in eine Menge m_2*,* $f : m_1 \to m_2$*, ist eine linkstotale und rechtseindeutige Relation* (f, m_1, m_2)*. Dabei bedeutet „linkstotal“, dass*

[13] Das Tripel ist auf der Metaebene, also hier nicht als Menge in ZFC gedacht.

$m_1 = \mathrm{dom}(f)$ *und „rechtseindeutig“, dass es für jedes* $a \in \mathrm{dom}(f)$ *genau ein* $b \in m_2$ *mit* $(a, b) \in f$ *gibt.*

Relationen und Funktionen werden also in ZFC durch ihre Graphen modellisiert. Durch den Graph einer Relation allein wird allerdings Quell- und Zielbereich nicht festgelegt, durch den Graph einer Funktion nicht der Zielbereich (was aber zum Beispiel für die Frage, ob die Funktion surjektiv ist, relevant ist).

Jede rechtseindeutige Menge f von Paaren ist Graph einer Funktion $f : \mathrm{dom}(f) \to \bigcup\bigcup f$. Wenn der Zielbereich nicht relevant ist, kann man ihn auch offen lassen und eine Funktion als $f : \mathrm{dom}(f) \to \mathbb{V}$ angeben.

Ist $f : m \to \mathbb{V}$ eine Funktion, schreibt man wie gewohnt $f(a)$ für das Bild eines Elements $a \in m$ unter f, also für das eindeutige b mit $(a, b) \in f$. Für eine Teilmenge $m' \subseteq m$ schreibt man $f[m']$ für die Menge $\{f(a) \mid a \in m'\}$. Da nicht ausgeschlossen ist, dass ein Element $a \in m$ auch Teilmenge $a \subseteq m$ ist, kann $f(a) \neq f[a]$ sein. Zum Beispiel ist stets $f[\emptyset] = \emptyset$, wenn aber $\emptyset \in \mathrm{dom}(f)$, kann $f(\emptyset)$ etwas Beliebiges sein.

Man kann die Konzepte von Relation und Funktion auch auf Klassen ausdehnen. Eine „klassengroße Relation“ ist dann eine definierbare Klasse von Paaren, etwa die Interpretation von $\in$ in $\mathbb{V}$, also die Klasse $\{(a, b) \mid a, b \text{ Mengen mit } a \in b\}$. Eine „klassengroße Funktion“ ist entsprechend eine definierbare Klasse von Paaren, die linkstotal und rechtseindeutig ist, zum Beispiel die Klasse $\{(a, \mathfrak{P}(a)) \mid a \text{ Menge}\}$, die der Zuordnung $a \mapsto \mathfrak{P}(a)$ entspricht.

Wenn $F : \mathbb{V} \to \mathbb{V}$ eine klassengroße Funktion ist und m eine Menge, dann ist die Einschränkung $F|_m$ von F auf m eine (mengengroße) Funktion, denn mit dem Ersetzungsaxiom ist auch $F[m]$ eine Menge und man bekommt $F|_m$ durch Aussonderung aus der Menge $m \times F[m]$.

Die Art und Weise, wie geordnete Paare und damit Relationen und Funktionen in der Mengenlehre definiert werden, ist kunstfertig, aber künstlich. Mit dieser Definition ist im Allgemeinen $(a, (b, c)) \neq ((a, b), c)$ und damit $a \times (b \times c) \neq (a \times b) \times c$. Wenn man geordnete n-Tupel braucht, muss man sie induktiv z. B. durch Linksklammerung definieren: $(a_1, \ldots, a_n) := ((\ldots((a_1, a_2), a_3), \ldots), a_n)$. Die naheliegende Verallgemeinerung des Kuratowski-Paares funktioniert dagegen nicht: $\{\{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}\}$ hat nicht die Eigenschaft eines Tripels.

Relativierte Quantoren

Folgende abkürzende Schreibweisen sind nützlich:

$$\exists x \in y\, \varphi \quad \text{für} \quad \exists x\, (x \in y \wedge \varphi)$$
$$\forall x \in y\, \varphi \quad \text{für} \quad \forall x\, (x \in y \to \varphi)$$

Die Dualitätsregeln gelten auch für diese *relativierten Quantoren*: $\neg\exists x \in y\, \varphi \sim \forall x \in y\, \neg\varphi$, etc.

6.3 Ordinalzahlen

Ordinalzahlen sind Zahlen, mit denen man die Elemente einer Menge abzählt (in der Grammatik die Zahlwörter *erstens, zweitens, drittens,...*); Kardinalzahlen sind dagegen Zahlen, die die Größe einer Menge messen (mit *null, eins, zwei, drei, ...*). Diese Konzepte sollen nun auf unendliche Mengen ausgedehnt werden. Die grundlegende Idee dafür ist die *Wohlordnung*, die das „Eins-Weiterzählen“ ermöglicht, indem sie immer ein kleinstes der noch nicht abgezählten Elemente anbietet.

Definition 6.7 *(a) Eine (totale oder lineare)* Ordnung auf m *ist eine reflexive, transitive, antisymmetrische totale Relation (r, m, m); eine (totale oder lineare)* strikte Ordnung auf m *ist eine irreflexive, transitive, asymmetrische totale Relation (r, m, m). Dabei bedeutet „total“ hier, dass $(a, b) \in r$ oder $(b, a) \in r$ für alle $a, b \in m$ mit $a \neq b$ gilt.*

(b) Eine Wohlordnung *ist eine (evtl. strikte) Ordnung einer Menge m, bei der jede nicht-leere Teilmenge von m ein kleinstes Element hat. m heißt dann* wohlgeordnet *durch die Ordnung.*

(c) Eine Menge m heißt transitiv*, wenn jedes Element von m auch Teilmenge von m ist. Anders ausgedrückt: Aus $a \in b \in m$ folgt $a \in m$.*

Eine andere Art, die Transitivität einer Menge m auszudrücken, ist $\bigcup m \subseteq m$.

Klar ist: Wenn r eine Wohlordnung auf m ist und m' eine Teilmenge von m, dann ist die Einschränkung von r auf m', also $(r \cap (m' \times m'), m', m')$, eine Wohlordnung auf m'.

Definition 6.8 *Eine* Ordinalzahl *ist eine transitive Menge, auf der $\in$ eine Wohlordnung ist.*

Beispiele für Ordinalzahlen sind die im Unendlichkeitsaxiom entstehenden Mengen: $\emptyset$, $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$, ...

Man kann sich überlegen, dass es eine $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi_{\mathbb{ORD}}(x)$ gibt, die genau von den Ordinalzahlen erfüllt wird. Die Ordinalzahlen bilden also eine Klasse $\mathbb{ORD}$. Einfacher findet man $\varphi_{\mathbb{ORD}}$ mit dem folgenden Lemma:

Lemma 6.9 *Für eine transitive Menge m sind äquivalent:*
(a) m ist eine Ordinalzahl.
(b) $\in$ definiert eine totale Ordnung auf m.
(c) $\in$ definiert eine totale Relation auf m, d. h. für alle $a, b \in m$ gilt: $a \in b$, $a = b$ oder $b \in a$.

Beweis: (a) $\Rightarrow$ (b) $\Rightarrow$ (c) ist klar. Es gelte also (c). Das Fundierungsaxiom verbietet Zykel $a_0 \in a_1 \in \cdots \in a_n = a_0$, weil die Menge $\{a_1, \ldots, a_n\}$ dann dem Axiom widerspräche.

Daher definiert $\in$ auf m eine Relation, die irreflexiv ist (nicht $a \in a$), asymmetrisch (nicht $a \in b$ und $b \in a$) und mit der zusätzlichen Voraussetzung auch transitiv: Denn mit $a, b, c \in m$ und $a \in b$ und $b \in c$ gibt es sowohl im Fall $a = c$ den Zykel $a \in b \in a$ als auch im Fall $c \in a$ den Zykel $a \in b \in c \in a$. Also bleibt nur die Möglichkeit $a \in c$.

Sie nun noch $\emptyset \neq m' \subseteq m$. Erneut mit dem Fundierungsaxiom gibt es ein $a \in m'$ mit $a \cap m' = \emptyset$. Für ein Element $a \neq b \in m'$ kann dann nicht $b \in a$ gelten, also muss $a \in b$ gelten. □

Lemma 6.10 *Seien $\alpha \neq \beta$ Ordinalzahlen.*

(a) Die echten Anfangsstücke von α bezüglich der durch $\in$ definierten Wohlordnung sind genau die Elemente von α und sind selbst Ordinalzahlen.

(b) Es gilt entweder $\alpha \in \beta$ oder $\beta \in \alpha$.

(c) $\alpha \in \beta \iff \alpha \subsetneq \beta$.

Beweis: (a) Sei η ein echtes Anfangsstück von α. Dann hat $\alpha \setminus \eta$ ein minimales Element b. Also ist $\eta = \{x \in \alpha \mid x \in b\} = b \cap \alpha = b$, da nach Transitivität $b \subseteq \alpha$. Insbesondere wird b durch $\in$ ebenfalls wohlgeordnet. Wenn $d \in c \in b$, dann sind $d, c \in \alpha$ wegen der Transitivität von α. Es folgt $d \in b$, weil $\in$ eine transitive Relation auf α definiert. Also ist auch b transitiv und damit einer Ordinalzahl.

Ist umgekehrt $b \in \alpha$ gegeben, dann ist $b \subseteq \alpha$ und das Argument für die Transivitität von b zeigt, dass b ein Anfangsstück von α ist.

(b) Sei $\gamma := \alpha \cap \beta$. Wenn $d \in c \in \gamma$, dann gilt wegen der Transitivität von α und β auch $d \in \alpha \cap \beta$. Also ist γ ein Anfangsstück sowohl von α als auch von β. Wenn $\gamma = \alpha$, ist $\alpha \in \beta$, und wenn $\gamma = \beta$ ist $\beta \in \alpha$. Bleibt der Fall, dass γ echtes Anfangsstück von α und von β ist. Dann gilt nach (a) aber $\gamma \in \alpha$ und $\gamma \in \beta$, also $\gamma \in \alpha \cap \beta = \gamma$: Widerspruch!

(c) $\alpha \in \beta \Rightarrow \alpha \subsetneq \beta$ folgt aus der Transitivität. Gilt umgekehrt $\alpha \subsetneq \beta$ und nicht $\alpha \in \beta$, folgt nach (b) $\beta \in \alpha$, also $\beta \subsetneq \alpha$ und Widerspruch. □

Folgerung 6.11 *Jede transitive Menge von Ordinalzahlen ist eine Ordinalzahl.*

Beweis: Lemma 6.10 (b) sagt, dass die Bedingung von Lemma 6.9 erfüllt ist. □

Folgerung 6.12 *Die Enthaltensrelation $\in$ (äquivalent: die Teilmengerelation $\subset$) definiert eine klassengroße Wohlordnung auf $\mathbb{ORD}$. Es folgt, dass $\mathbb{ORD}$ eine echte Klasse ist.*

Beweis: $\mathbb{ORD}$ ist eine transitive Klasse, da nach Lemma 6.10 (a) Elemente von Ordinalzahlen selbst Ordinalzahlen sind. Der Beweis von (c)⇒(b) in Lemma 6.9 überträgt sich auf Klassen; mit Lemma 6.10 (b) folgt daraus, dass $\in$ eine klassengroße Ordnung auf $\mathbb{ORD}$ definert.

Dass dies eine Wohlordnung ist, sieht man ebenfalls mit Lemma 6.9. Wenn man ohne das Fundierungsaxiom für Klassen auskommen will, argumentiert man so: Wenn $\mathbb{T}$ eine nicht-leere Teilklasse von $\mathbb{ORD}$ ist mit $\alpha \in \mathbb{T}$, dann enthält α ein $\in$-minimales Element β. Für beliebiges $\gamma \in \mathbb{T}$ ist nun entweder $\gamma \in \alpha$ oder $\gamma = \alpha$ oder $\alpha \in \gamma$. In den ersten beiden Fällen folgt $\beta \in \gamma$ aus der Wahl von β, im dritten Fall aus der Transitivität von $\in$.

Schließlich: Wäre $\mathbb{ORD}$ eine Menge, also eine Ordinalzahl, müsste $\mathbb{ORD} \in \mathbb{ORD}$ gelten. □

Die durch $\in$ bzw. $\subseteq$ auf $\mathbb{ORD}$ definierte Wohlordnung wird jetzt auch $<$ geschrieben, die zugehörige nicht-strikte Ordnung $\subseteq$ auch $\leqslant$.

Lemma 6.13 *Wenn α eine Ordinalzahl ist, ist $\alpha + 1 := \alpha \cup \{\alpha\}$ der (unmittelbare)* Nachfolger *von α, d. h. die kleinste Ordinalzahl, die größer als α ist.*

Beweis: Da α transitiv ist, ist $\bigcup(\alpha \cup \{\alpha\}) = (\bigcup \alpha) \cup \alpha \subseteq \alpha \subseteq \alpha \cup \{\alpha\}$. Somit ist auch $\alpha \cup \{\alpha\}$ transitiv. Mit Folgerung 6.11 ist $\alpha \cup \{\alpha\}$ eine Ordinalzahl.

Falls γ eine Ordinalzahl mit $\alpha < \gamma$, also $\alpha \in \gamma$ bzw. $\{\alpha\} \subseteq \gamma$, dann gilt wegen Transitivität von γ auch $\alpha \subseteq \gamma$. Zusammen also $\alpha \cup \{\alpha\} \subseteq \gamma$ bzw. $\alpha \cup \{\alpha\} \leqslant \gamma$. □

Definition 6.14 *Eine Ordinalzahl der Form $\alpha \cup \{\alpha\}$ heißt* Nachfolger(ordinal)zahl*; eine Ordinalzahl $\neq \emptyset$, die nicht dieser Form ist, heißt* Limes(ordinal)zahl.

Nachfolgerordinalzahlen $\alpha + 1$ sind genau die Ordinalzahlen mit einem maximalen Element α. Es gilt $\bigcup(\alpha + 1) = \alpha$.

Lemma 6.15 *Jede Menge a von Ordinalzahlen hat in $\mathbb{ORD}$ die kleinste obere Schranke $\bigcup a$.*

Beweis: Als Vereinigung transitiver Mengen ist $\bigcup a$ auch transitiv. Mit Folgerung 6.11 ist es eine Ordinalzahl. Wenn $\alpha \in a$, ist $\alpha \subseteq \bigcup a$, also ist $\bigcup a$ obere Schranke von a. Ist β eine obere Schranke von a, dann gilt $\alpha \leqslant \beta$ für alle $\alpha \in a$, also $\alpha \subseteq \beta$ und somit $\bigcup a \subseteq \beta$. □

Wenn $a \neq \emptyset$ kein maximales Element hat, ist $\bigcup a$ eine Limesordinalzahl. Für Limesordinalzahlen α ist $\bigcup \alpha = \alpha$.

Lemma 6.16 *Es gibt Limesordinalzahlen.*

Beweis: Man wendet Aussonderung mit $\varphi_{\mathbb{ORD}}$ auf eine durch das Unendlichkeitsaxiom gegebene Menge an und erhält eine (als Teilmenge von $\mathbb{ORD}$ wohlgeordnete) Menge m von Ordinalzahlen. Sie enthält $\emptyset$, ist also nicht leer, und enthält mit jeder Ordinalzahl auch deren Nachfolger, hat also kein maximales Element. $\bigcup m$ ist dann eine transitive Menge von Ordinalzahlen, also nach Folgerung 6.11 selbst eine Ordinalzahl, und hat ebenfalls kein maximales Element, ist also eine Limesordinalzahl. □

Definition 6.17 *Die kleinste Limesordinalzahl wird mit ω oder auch ω_0 bezeichnet. Die Elemente von ω heißen* natürliche Zahlen *(im Sinne von* ZFC*).*

Die $\mathscr{L}_{\mathrm{ML}}$-Formel

$$\varphi_\omega(x) := \big(\varphi_{\mathbb{ORD}}(x) \wedge \neg x \doteq \emptyset \wedge \forall y \in x\,(y \cup \{y\} \in x \wedge (y = \emptyset \vee \exists z \in x\, z \cup \{z\} = y))\big)$$

beschreibt die kleinste Limesordinalzahl. Es gilt $\mathsf{ZFC} \vDash \exists! x\, \varphi_\omega(x)$, also existiert ω in jedem Modell von ZFC. Die $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi_{\mathbb{N}}(x) := \exists y\,(\varphi_\omega(y) \wedge x \in y)$ wird dann in jedem Modell genau von den natürlichen Zahlen erfüllt.

Per Induktion kann man eine Abbildung $\_ : \mathbb{N} \to \mathbb{ORD}$ von den „tatsächlichen“, intuitiv gegebenen natürlichen Zahlen in die Ordinalzahlen definieren durch $n \mapsto \underline{n} = \{\underline{0}, \ldots, \underline{n-1}\}$. Alternativ: $\underline{0} = \emptyset$ und $\underline{n+1} = \underline{n} + 1 = \underline{n} \cup \{\underline{n}\}$. Dann ist $\omega = \{\underline{n} \mid n \in \mathbb{N}\}$, was die Terminologie rechtfertigt. (Da man hier mit einem intuitiven Verständnis von $\mathbb{N}$ arbeitet, ist dies nur ein Plausibilitätsargument und kein Beweis in einer formalen Theorie!) Man kann aber für formale Theorien der natürlichen Zahlen wie die Peano-Arithmetik zeigen, dass ω die geforderten Axiome erfüllt.

In ZFC kann man Folgendes beweisen (hier ohne Beweis):

Satz 6.18 *Für Ordinalzahlen α sind äquivalent:*

(a) *$\alpha \in \omega$, d. h. α ist eine natürliche Zahl.*
(b) *Alle β mit $\emptyset < \beta \leqslant \alpha$ (anders ausgedrückt: $\emptyset \neq \beta \in \alpha \cup \{\alpha\}$) sind Nachfolgerordinalzahlen.*
(c) *Jede nicht-leere Teilmenge von α hat ein größtes Element bzgl. $\in$.*
(d) *Jede Injektion $\alpha \to \alpha$ ist surjektiv.*
(e) *Jede Surjektion $\alpha \to \alpha$ ist injektiv.*

Als Teil eines Beweises kann man benutzen, dass $\alpha \mapsto \alpha + 1 : \omega \to \omega$ eine nicht surjektive Injektion und $\alpha \mapsto \bigcup \alpha : \omega \to \omega$ eine nicht injektive Surjektion ist.

Transfinite Induktion/Rekursion

Satz 6.19 (Beweis per Induktion über Ordinalzahlen)
Eine $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi(x)$ wird genau dann von allen Ordinalzahlen erfüllt, wenn für alle Ordinalzahlen α die Implikation gilt: Wenn φ für alle $\beta < \alpha$ gilt, dann gilt φ auch für α.

Beweis: „$\Rightarrow$“ ist trivial und „$\Leftarrow$“ folgt unmittelbar aus der Wohlgeordnetheit von $\mathbb{ORD}$: Wenn die Konklusion nicht stimmt, gibt es ein kleinstes $\alpha \in \mathbb{ORD}$, das φ nicht erfüllt. Dann gilt φ aber für alle $\beta < \alpha$, also per Voraussetzung auch für α: Widerspruch! □

Variante (A): In Anwendungen nutzt man den Beweis per Induktion oft in der Weise, dass φ für alle Ordinalzahlen gilt, wenn die drei Bedingungen erfüllt sind:

[Induktionsanfang] φ gilt für $\emptyset$;

[Nachfolgerschritt] wenn φ für α gilt, dann auch für $\alpha+1$;

[Limesschritt] wenn φ für alle β kleiner einer Limesordinalzahl λ gilt, dann auch für λ.

Variante (B): Mit der $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi'(x) = (\varphi(x) \vee \gamma \leqslant x)$ ergibt sich folgende Variante des Induktionsbeweises für eine Ordinalzahl γ:

Eine $\mathscr{L}_{\mathrm{ML}}$-Formel $\varphi(x)$ wird genau dann von allen Ordinalzahlen $< \gamma$ erfüllt, wenn für alle Ordinalzahlen $\alpha < \gamma$ die Implikation gilt: Wenn φ für alle $\beta < \alpha$ gilt, dann gilt φ auch für α.

Satz 6.20 (Rekursionssatz: Definition per Induktion über Ordinalzahlen)
Zu jeder klassengroßen Funktion $G : \mathbb{V} \to \mathbb{V}$ gibt es eine eindeutige klassengroße Funktion $F : \mathbb{ORD} \to \mathbb{V}$, so dass die Rekursionsbedingung $F(\alpha) = G(F|_\alpha)$ *für alle Ordinalzahlen α gilt.*

Zur Erinnerung: $F|_\alpha$ ist die Einschränkung von F auf die Menge α, ist also die Menge der Paare $(\beta, F(\beta))$ für $\beta \in \alpha$. Die Bedingung im Satz besagt also, dass jedes $F(\alpha)$ mit Hilfe von G durch die $F(\beta)$ für $\beta < \alpha$ festgelegt ist.

Beweis: Sind $F, F' : \mathbb{ORD} \to \mathbb{V}$ Funktionen, die die Rekursionsbedingung erfüllen, folgt $F(\alpha) = F'(\alpha)$ für alle α per Induktion gemäß Satz 6.19. Die Eindeutigkeit gilt mit Induktion bis α auch für Funktionen $f, f' : \alpha \to \mathbb{V}$, die die Rekursionsbedingung für alle $\beta < \alpha$ erfüllen.

Für die Existenz konstruiert man zunächst für jede Ordinalzahl α diese eindeutige, die Rekursionsbedingung erfüllende Funktion $f_\alpha : \alpha \to \mathbb{V}$ (die am Ende $F|_\alpha$ sein wird).

- Für $\alpha = \emptyset$ ist $f_\emptyset = \emptyset$ und
- für $\alpha = \beta + 1$ ist $f_\alpha = f_\beta \cup \big\{(\beta, G(f_\beta))\big\}$.
- Sei schließlich α Limesordinalzahl: Wegen der aus der Eindeutigkeit folgenden Kompatibilität $f_\gamma = f_\beta|_\gamma$ für $\gamma \in \beta \in \alpha$ ist $f_\alpha := \bigcup\{f_\beta \mid \beta \in \alpha\}$ die gesuchte Funktion. Allerdings muss noch nachweisen, dass dies eine Menge ist. Da sich die Rekursionsbedingung als $\mathscr{L}_{\mathrm{ML}}$-Formel ausdrücken lässt, folgt dies aus dem Ersetzungsaxiom mit der Zuordnung $h_\alpha : \alpha \to \mathbb{V}$, wobei

  $$h_\alpha : \beta \;\mapsto\; (g : \beta \to \mathbb{V} \text{ so, dass } g \text{ die Rekursionsbedingung für alle } \gamma < \beta \text{ erfüllt}),$$

  die wegen der Eindeutigkeit von $g = f_\beta$ funktional ist (und sich außerhalb von α durch $x \mapsto \emptyset$ auf $\mathbb{V}$ fortsetzen lässt).

Per Induktion gilt nun die Eigenschaft „es gibt (eindeutiges) $g : \alpha \to \mathbb{V}$, welches die Rekursionsbedingung für alle $\beta < \alpha$ erfüllt“ für alle Ordinalzahlen α, und die Funktion g ist jeweils f_α. Eine ähnliche Definition wie oben ergibt nun die gesuchte Zuordnung $F : \mathbb{ORD} \to \mathbb{V}$ als

$$F : \alpha \;\mapsto\; g(\alpha) \text{ für } g : \alpha + 1 \to \mathbb{V} \text{ so, dass } g \text{ die Rekursionsbedingung für alle } \beta < \alpha + 1 \text{ erfüllt}$$

Denn zunächst sieht man, dass $F(\beta) = f_{\beta+1}(\beta) = f_\alpha(\beta)$ für alle $\beta < \alpha$, d. h. es ist $f_\alpha = F|_\alpha$. Daraus folgt dann $F(\alpha) = f_{\alpha+1}(\alpha) = G(f_\alpha) = G(F|_\alpha)$. □

Wie zum Beweis per Induktion über Ordinalzahlen kann man auch zum Rekursionssatz diverse Varianten angeben:

- Rekursion bis zu einer Ordinalzahl γ mit Hilfe einer Funtion $g : \gamma \to \mathbb{V}$ (die beliebig zu einer Klassenfunktion $G : \mathbb{V} \to \mathbb{V}$ fortgesetzt wird).

- Die Klassenfunktion G im Rekursionssatz ist auf den Ordinalzahlen per Fallunterscheidung definiert mit den drei Fällen $\emptyset$, Nachfolger- und Limesordinalzahlen (und beliebig fortgesetzt auf Mengen, die keine Ordinalzahlen sind).

Wohlordnungssatz und Zorn'sches Lemma

Satz 6.21 *Jede (strikt) wohlgeordnete Menge ist durch einen eindeutigen Ordnungsisomorphismus zu einer eindeutig bestimmten Ordinalzahl isomorph.*

Beweis: Sei $(m;<)$ wohlgeordnet durch eine strikte Wohlordnung $<$. Definiere mit dem Rekursionssatz $F : \mathbb{ORD} \to m \cup \{m\}$ so, dass

$$F : x \mapsto \begin{cases} \min_< \big(m \setminus \mathrm{im}(F|_x)\big) & \text{falls } x \in \mathbb{ORD} \text{ und } m \neq \mathrm{im}(F|_x) \\ m & \text{sonst} \end{cases}$$

Angenommen es gibt eine Ordinalzahl α mit $F(\alpha) = m$. Für die kleinste solche Ordinalzahl β ist dann $F|_\beta : \beta \to m$ per Konstruktion ordnungstreu (und damit insbesondere injektiv) und nach Wahl von β auch surjektiv, also der gesuchte Ordnungsisomorphimus.

Angenommen es gibt keine solche Ordinalzahl. Dann ist $F : \mathbb{ORD} \to \mathrm{im}(m)$ bijektiv, also auch die Umkehrklassenfunktion $F^{-1} = \{(y,x) \mid (x,y) \in F\} : \mathrm{im}(F) \to \mathbb{ORD}$. Dann wäre aber nach dem Ersetzungsaxiom $\mathbb{ORD}$ eine Menge als Bild der Menge $\mathrm{im}(F)$.

Ein Ordnungsisomorphismus zwischen einer Ordinalzahl und m muss andererseits die Rekursionbedingung erfüllen, ist also nach dem Rekursionssatz eindeutig, ebenso wie β. □

In Sinne des voranstehenden Satzes bilden die Ordinalzahlen also ein (besonders schönes) Repräsentantensystem der Isomorphieklassen von Wohlordnungen.

Satz 6.22 (Wohlordnungssatz von Zermelo)
Aus jeder Menge m gibt es eine Wohlordnung.

Für den Beweis wurde von Zermelo das Auswahlaxiom formuliert, zur Erinnerung:

$$\forall y\big(\neg\emptyset \in y \to \exists w \forall x\big(x \in y \to \exists! x'\big(\,(x,x') \in w \wedge x' \in x\big)\big)\big)$$

Ich habe diese Darstellung gewählt, um das Auswahlaxiom mit möglichst wenigen Vordefinitionen angeben zu können. Im Axiom ist $w \cap (y \times \mathbb{V}) : y \to \mathbb{V}$ eine Funktion. Die übliche (modulo der anderen Axiome von ZFC äquivalente) Formulierung des Auswahlaxioms ist daher:

$$\forall y\big(\neg\emptyset \in y \to \exists f\,(f : y \to \mathbb{V} \wedge \forall x \in y\, f(x) \in x)\big)$$

Beweis: Sei $f : \mathfrak{P}(m) \setminus \{\emptyset\} \to m$ eine Auswahlfunktion. Definiere mit dem Rekursionssatz $F : \mathbb{ORD} \to m \cup \{m\}$ so, dass

$$F : x \mapsto \begin{cases} f\big(m \setminus \mathrm{im}(F|_x)\big) & \text{falls } x \in \mathbb{ORD} \text{ und } m \neq \mathrm{im}(F|_x) \\ m & \text{sonst} \end{cases}$$

Wie im Beweis von Satz 6.21 sieht man, dass es eine Ordinalzahl β gibt, so dass $F|_\beta : \beta \to m$ eine Bijektion ist. Das Bild von $\in|_{\beta\times\beta}$ unter $F|_\beta$ ist dann eine Wohlordnung auf m. □

Der Wohlordnungssatz erlaubt Induktionsbeweise auf beliebigen unendlichen Mengen. Zum Beispiel kann man dmit zeigen, dass jeder Vektorraum V eine Basis besitzt: Man wählt eine

Wohlordnung $<$ auf V. Diese ist zu einer Ordinalzahl ordnungsisomorph, also kann man ohne Einschränkung annehmen, dass $V = \alpha$ eine Ordinalzahl ist. Nun konstruiert man eine maximal linear unabhängige Teilmenge als Bild einer per Rekursion definierten Abbildung $f : \alpha \to \alpha$, indem man $f(\beta) = f[\beta] \cup \{\beta\}$ setzt, wenn β linear unabhängig von $f[\beta]$ ist, und $f(\beta) = f[\beta]$ sonst.

In Anwendungen (so auch hier im Skript bei den Beweisen des Vollständigkeitssatzes und des Darstellungssatzes von Stone) wird diese Induktion häufig durch das folgende Lemma von Zorn abgekürzt:

Satz 6.23 (Zorn'sches Lemma) *Sei $(m; \leqslant_p)$ eine partielle geordnete Menge, in der jede linear geordnete Teilmenge eine obere Schranke besitzt. Dann gibt es in m mindestens ein maximales Element (also ein Element, für das es bzgl. $\leqslant_p$ kein echt größeres gibt).*

Beweis: Sei $<_w$ eine Wohlordnung auf m (die in keinem Zusammenhang mit der gegebenen partiellen Ordnung $\leqslant_p$ steht!). Dann gibt es eine Ordinalzahl μ und einen Ordnungsisomorphismus $h : (\mu; \in) \to (m; <_w)$. Mit dem Rekursionssatz definiert man folgendermaßen eine Funktion $f : \mu + 1 \to m$:

$$\begin{aligned}
\emptyset &\mapsto \text{das } <_w\text{-minimale Element von } m \\
\beta + 1 &\mapsto \begin{cases} h(\beta) & \text{falls } f(\beta) \leqslant_p h(\beta) \\ f(\beta) & \text{sonst} \end{cases} \\
\lambda &\mapsto \begin{cases} \text{die } <_w\text{-minimale obere Schranke} \\ \text{von } \operatorname{im}(f|_\lambda) = \{f(\beta) \mid \beta \in \lambda\} & \text{falls } \operatorname{im}(f|_\lambda) \text{ kein } \leqslant_p\text{-maximales Element hat} \\ f(\beta_0) & \text{falls } f(\beta_0) \leqslant_p\text{-maximal in } \operatorname{im}(f|_\lambda) \text{ ist} \end{cases}
\end{aligned}$$

Per Konstruktion ist jedes $f(\alpha)$ maximales Element in $\{h(\beta) \mid \beta < \alpha\}$ bzgl. $\leqslant_p$. Also ist $h(\mu)$ maximales Element in m. □

Ordinalzahlarithmetik

Neben der Anordnung kann man auch Addition und Multiplikation der natürlichen Zahlen auf Ordinalzahlen ausweiten [14]. Dies geschieht formal am besten per Rekursion:

Definition 6.24 *Addition $+_{\mathbb{ORD}}$ und Multiplikation $\cdot_{\mathbb{ORD}}$ für Ordinalzahlen sind (für jedes α) definiert durch Rekursion über die rechte Seite (λ Limesordinalzahl):*

$$\begin{array}{llll}
\alpha +_{\mathbb{ORD}} \underline{0} & := & \alpha & \alpha \cdot_{\mathbb{ORD}} \underline{0} := \underline{0} \\
\alpha +_{\mathbb{ORD}} (\beta + 1) & := & (\alpha +_{\mathbb{ORD}} \beta) + 1 & \alpha \cdot_{\mathbb{ORD}} (\beta + 1) := (\alpha \cdot_{\mathbb{ORD}} \beta) +_{\mathbb{ORD}} \alpha \\
\alpha +_{\mathbb{ORD}} \lambda & := & \bigcup\{\alpha +_{\mathbb{ORD}} \beta \mid \beta < \lambda\} & \alpha \cdot_{\mathbb{ORD}} \lambda := \bigcup\{\alpha \cdot_{\mathbb{ORD}} \beta \mid \beta < \lambda\}
\end{array}$$

Insbesondere ist $\alpha +_{\mathbb{ORD}} \underline{1} = \alpha + 1$ der unmittelbare Nachfolger von α in $\mathbb{ORD}$. Auf den natürlichen Zahlen stimmen die Ordinalzahladdition und -multiplikation mit der üblichen Addition und Multiplikation überein (da für diese die gleichen Rekursionsgleichungen gelten). Wenn klar ist, dass es sich um Ordinalzahlarithmetik handelt, schreibt man auch kurz $+$ und $\cdot$.

Anschaulich macht man sich die Addition und Multiplikation folgendermaßen:

[14] Und auch die Exponentiation; diese wird aber seltener gebraucht.

Definition 6.25 *Seien $(a, <_a)$ und $(b, <_b)$ total strikt geordnete Mengen. Die* geordnete Summe *von a und b ist die Menge $\big(a \times \{\underline{0}\}\big) \cup \big(b \times \{\underline{1}\}\big)$* [15] *mit der Anordnung*

$$(x,i) < (y,j) :\Longleftrightarrow \begin{cases} i = j = \underline{0} \text{ und } x <_a y \\ i = \underline{0} \text{ und } j = \underline{1} \\ i = j = \underline{1} \text{ und } x <_b y \end{cases}$$

Das geordnete Produkt *von a und b ist die Menge $a \times b$ mit der* antilexikografischen Anordnung

$$(x_1,y_1) < (x_2,y_2) :\Longleftrightarrow \begin{cases} y_1 <_b y_2 \\ y_1 = y_2 \text{ und } x_1 <_a x_2 \end{cases}$$

Illustration der geordneten Summe von a und b :

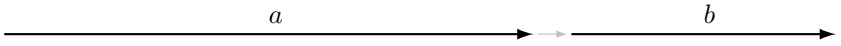
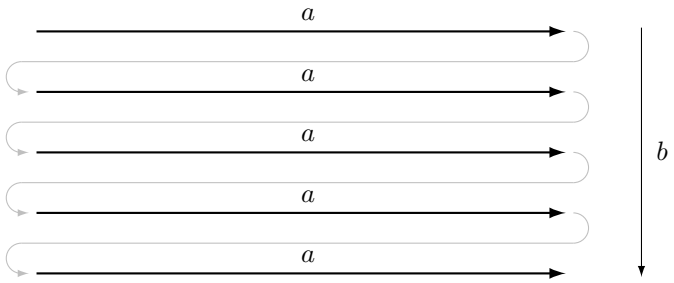


Illustration des geordneten Produkts von a und b :



Satz 6.26 *Die geordnete Summe und das geordnete Produkt von zwei Wohlordnungen sind Wohlordnungen und es gilt, dass $\alpha + \beta$ ordnungsisomorph zu der geordneten Summe von α und β und $\alpha \cdot \beta$ ordnungsisomorph zum geordneten Produkt von α und β ist.*

Beweis: Die erste Behauptung ist eine Übung. Das zweite sieht man dadurch, dass man überprüft, dass die Rekursionsdefinitionen von $+_{\mathbb{ORD}}$ und $\cdot_{\mathbb{ORD}}$ auf die geordnete Summe bzw. das geordnete Produkt zutreffen. □

Ordinalzahladdition und -multiplikation sind nicht kommutativ: Zum Beispiel ist $\underline{1} + \omega = \omega \neq \omega + \underline{1}$ und $\underline{2} \cdot \omega = \omega \neq \omega \cdot \underline{2} = \omega + \omega$. Bei der Multiplikation muss man beachten, dass die Reihenfolge kontraintuitiv ist: Üblicherweise versteht man 2·3 als „zweimal 3“, also 3+3, dagegen $3 \cdot 2$ als „dreimal 2“, also $2 + 2 + 2$. Hier dagegen ist $\omega \cdot \underline{2} = \omega + \omega$. also „zweimal ω“, während $\underline{2} \cdot \omega$ als „ω-mal $\underline{2}$“, d. h. $\underline{2} + \underline{2} + \underline{2} + \ldots$ zu verstehen ist, genauer als $\sup\{\underline{2}, \underline{2} + \underline{2}, \underline{2} + \underline{2} + \underline{2}, \ldots\}$.

Satz 6.27 (Eigenschaften der Ordinalzahladdition und -multiplikation)

- *Addition und Multiplikation sind assoziativ.*
- *$\underline{0}$ ist neutrales Element der Addition und $\underline{1}$ neutrales Element der Multiplikation.*
- *Die Multiplikation ist linksdistributiv über der Addition:*
$$\alpha \cdot (\beta + \gamma) = (\alpha \cdot \beta) + (\alpha \cdot \gamma)$$

[15] Es geht bei dieser Definition nur darum, eine disjunkte Vereinigung zu bekommen.

- *Addition und Multiplikation sind linkskürzbar:*
  $$\text{aus } \alpha+\beta=\alpha+\gamma \text{ folgt } \beta=\gamma; \quad \text{aus } \alpha\cdot\beta=\alpha\cdot\gamma \text{ und } \alpha\neq\underline{0} \text{ folgt } \beta=\gamma$$
- *Genau dann gilt $\alpha\leqslant\gamma$, wenn es ein β mit $\alpha+\beta=\gamma$ gibt.*
- *Addition und Multiplikation sind strikt linksmonoton:*
  $$\text{wenn } \beta<\gamma, \text{ ist } \alpha+\beta<\alpha+\gamma \text{ und, sofern } \alpha\neq\underline{0}, \text{ auch } \alpha\cdot\beta<\alpha\cdot\gamma$$

Beweis: Übung! □

Man bekommt nun eine Fülle unendlicher, immer größer werdender Ordinalzahlen:

$$\begin{aligned}
&\omega\,,\ \omega+1\,,\ \omega+2\,,\ldots\\
&\omega+\omega=\omega\cdot 2\,,\ \omega\cdot 2+1\,,\ \omega\cdot 2+2\,,\ldots,\ \omega\cdot 3\,,\ldots,\ \omega\cdot 4\,,\ldots\\
&\omega\cdot\omega=:\omega^2\,,\ \omega^2+1\,,\ \omega^2+2\,,\ldots,\ \omega^2+\omega\,,\ \omega^2+\omega+1\ \ldots,\ \omega^2\cdot 2\,,\ldots,\ \omega^3\,,\ldots,\ \omega^4\,,\ldots\\
&\sup\{\omega,\omega^2,\omega^3,\omega^4,\ldots\}=:\omega^\omega\,,\ \omega^\omega+1\,,\ldots,\ \omega^{\omega+1}:=\omega^\omega\cdot\omega\,,\ldots\\
&\vdots\\
&\omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}}\,,\ \omega^{\omega^{\cdot^{\cdot^{\cdot^{\omega}}}}}+1\,,\ldots
\end{aligned}$$

Die kleinste Ordinalzahl, die nicht mehr durch mit Addition, Multiplikation und Exponentiation als endlicher Term in ω und natürlichen Zahlen darstellbar ist, wird ε_0 genannt. Alle diese Ordinalzahlen sind aber abzählbar, d. h. stehen in Bijektion zu ω. Die kleinste nicht mehr abzählbare Ordinalzahl heißt ω_1. Dass sie überhaupt existiert, wird im nächsten Abschnitt gezeigt.

6.4 Kardinalzahlen

Definition 6.28

(a) *Zwei Mengen m_1, m_2 heißen* gleichmächtig, *wenn es eine Bijektion $m_1\to m_2$ gibt.*
(b) *Eine* Kardinalzahl *ist eine Ordinalzahl κ, die zu keiner Ordinalzahl $\alpha\in\kappa$ gleichmächtig ist. Die Klasse aller Kardinalzahlen schreibe ich $\mathbb{KARD}$.*
(c) *Für jede Menge m bezeichnet die $|m|$ die eindeutige zu m gleichmächtige Kardinalzahl und wird* Mächtigkeit *oder* Kardinalität von m *genannt.*

Die Existenz einer Bijektion lässt sich durch eine $\mathscr{L}_{\mathrm{ML}}$-Formel ausdrücken. Daher ist zum einen $\mathbb{KARD}$ tatsächlich eine Klasse (und zwar eine echte Klasse, was bereits Cantor wusste), und zum andern bilden alle zu einer Menge m gleichmächtigen Ordinalzahlen eine – nach dem Wohlordnungssatz nicht leere – Teilklasse von $\mathbb{ORD}$, die also ein minimales Element $|m|$ hat, was Teil (c) der Definition rechtfertigt.

Lemma 6.29 *Für je zwei Mengen m_1, m_2 gilt:*

(a) *Genau dann ist $|m_1|=|m_2|$, wenn es eine Bijektion $m_1\to m_2$ gibt.*
(b) *Genau dann ist $|m_1|\leqslant|m_2|$, wenn es eine Injektion $m_1\to m_2$ gibt.*

Beweis: Teil (a) ist klar, da es nach Definition Bijektionen $m_i\to|m_i|=:\kappa_i$ gibt, und „⇒“ von Teil (b) ist ebenso klar, da nach Annahme $\kappa_1\subseteq\kappa_2$.

„⇐“: Mit Hilfe der Bijektionen $m_i\to\kappa_i$ gibt es also eine Injektion $i:\kappa_1\to\kappa_2$. Also steht κ_1 in Bijektion mit einer Teilmenge t von κ_2. Wenn $t=\kappa_2$, ist $\kappa_1=\kappa_2$; wenn $t\subset\kappa_2$, folgt mit dem nächsten Lemma $\kappa_1<\kappa_2$. □

Lemma 6.30 *Jede Teilmenge t einer Ordinalzahl α steht in Bijektion zu einer Ordinalzahl $\beta \leqslant \alpha$. Jede echte Teilmenge t einer natürlichen Zahl $\underline{n}$ steht in Bijektion zu einer natürlichen Zahl $\underline{m} < \underline{n}$.*

Beweis: Für $t = \emptyset$ ist dies trivialerweise richtig; sei daher $t \neq \emptyset$. Mit dem Rekursionssatz definiert man $f : \alpha \cup \{\alpha\} \to t \cup \{t\}$ durch

$$\gamma \mapsto \begin{cases} \min\{x \in t \mid x \notin \mathrm{im}(f|_\gamma)\} & \text{falls } \mathrm{im}(f|_\gamma) \neq t \\ \gamma \mapsto t & \text{sonst} \end{cases}$$

Es ist $\emptyset \leqslant \gamma(\emptyset)$, daraus ergibt sich induktiv $\gamma \leqslant f(\gamma)$ für alle γ mit $f(\gamma) \in t$. Also wird irgendwann t erreicht und es gibt ein minimales $\beta \leqslant \alpha$ mit $f(\beta) = t$. Dann ist $f|_\beta : \beta \to t$ eine Bijektion.

Wenn t echte Teilmenge ist, gibt es ein γ_0 mit $\gamma_0 < f(\gamma_0)$. Daraus folgt dann auch $\gamma_0 + 1 < f(\gamma_0 + 1)$. Da es unterhalb einer natürlichen Zahl keine Limesordinalzahlen gibt, gilt also im Fall $\alpha = \underline{n} = \underline{n'} + 1$, dass $\gamma < f(\gamma)$ für alle $\gamma \geqslant \gamma_0$, so dass $f(\gamma) \in t$. Dann muss $f(\underline{n'}) = t$ gelten und es gilt $\beta \leqslant \underline{n'} < \underline{n}$. □

Folgerung 6.31 (Satz von Cantor, Bernstein, Schröder) *Zwei Mengen m_1, m_2 sind genau dann gleichmächtig, wenn es Injektionen $m_1 \to m_2$ und $m_2 \to m_1$ gibt.*

Den Satz von Cantor-Bernstein-Schröder kann man auch ohne Auswahlaxiom beweisen, dann ist der Beweis etwas anspruchsvoller.

Satz 6.32 *Alle natürlichen Zahlen $\underline{n}$ sowie ω sind Kardinalzahlen.*

Beweis: Für natürliche Zahlen beweist man es per Induktion: Für $\underline{0}$ ist es klar. Angenommen $\underline{n}$ ist eine Kardinalzahl, $\underline{n} + 1$ aber nicht. Dann gibt es eine Bijektion $h : \underline{n} + 1 \to \underline{m} < \underline{n} + 1$. Die Einschränkung von h auf $\underline{n}$ ist dann eine Bijektion $h : \underline{n} \to h[\underline{n}] \subset \underline{m}$. Nach dem vorherigen Lemma gibt es also eine Bijektion $h' : \underline{n} \to \underline{k} < \underline{m} \leqslant \underline{n}$ im Widerspruch zur Annahme, dass $\underline{n}$ eine Kardinalzahl ist.

Wäre ω keine Kardinalzahl, würde es eine Bijektion $h : \omega \to \underline{n} \in \omega$ geben. Wegen der Transitivität der Ordinalzahlen ist $\underline{n} \subseteq \underline{n} + 1 \subseteq \omega$. Die linke Inklusion ergibt eine Injektion $\underline{n} \to \underline{n} + 1$, die rechte Inklusion verkettet mit h eine Injektion $\underline{n} + 1 \to \underline{n}$. Nach dem Satz von Cantor-Schröder-Bernstein widertspricht dies der Eigenschaft von $\underline{n} + 1$, Kardinalzahl zu sein. □

Dass es überhaupt weitere Kardinalzahlen gibt, zeigt:

Satz 6.33 (Satz von Cantor) *Für jede Menge m gilt $|m| < |\mathfrak{P}(m)|$.*

Beweis: Angenommen es gibt eine Bijektion $h : m \to \mathfrak{P}(m)$. Mit Aussonderung gibt es die Menge $r := \{a \in m \mid a \notin h(a)\} \in \mathfrak{P}(m)$ und nach Annahme gibt es somit $a_0 \in m$ mit $h(a_0) = r$. Dann gilt aber $a_0 \in r \iff a_0 \notin h(a_0) = r$: Widerspruch! □

Definition 6.34 *Der unmittelbare Nachfolger einer Kardinalzahl κ in $\mathbb{KARD}$, also die kleinste Kardinalzahl, die größer als κ ist, wird mit κ^+ bezeichnet. Kardinalzahlen dieser Form werden Nachfolgerkardinalzahlen genannt. Kardinalzahlen $\neq \underline{0}$, die keine Nachfolgerkardinalzahlen sind, heißen Limeskardinalzahlen.*

Lemma 6.35 *Das Supremum (in $\mathbb{ORD}$) einer Menge von Kardinalzahlen ist eine Kardinalzahl.*

Beweis: Sei $\{\kappa_i \mid i \in I\}$ eine Menge von Kardinalzahlen. Da alle κ_i Ordinalzahlen sind, existiert nach Lemma 6.15 das Supremum κ. Insbesondere gilt $\kappa_i \subseteq \kappa$ für alle i, und mit Lemma 6.29 folgt $\kappa_i = |\kappa_i| \leqslant |\kappa|$. Also ist $\kappa \leqslant |\kappa|$. Die Ungleichung $|\kappa| \leqslant \kappa$ gilt stets. □

Mit dem Rekursionssatz kann man nun die Klassenfunktion $\aleph : \mathbb{ORD} \to \mathbb{KARD}$ definieren: [16]

$$\aleph_0 := \omega \ , \ \aleph_{\alpha+1} := \aleph_\alpha^+ \ , \ \aleph_\lambda := \sup\{\aleph_\alpha \mid \alpha \in \lambda \text{ Limesordinalzahl}\}$$

Da die $\aleph$-Funktion nach Konstruktion injektiv ist, sieht man, dass $\mathbb{KARD}$ eine echte Klasse ist (sonst wäre $\mathbb{ORD} = \aleph^{-1}[\mathbb{KARD}]$ eine Menge).

Mit $\aleph_\alpha$ bezeichnet man eine Zahl, wenn man sie als Kardinalzahl betrachtet. Betrachtet man sie dagegen als Ordinalzahl, schreibt man ω_α. Es ist also $\aleph_\alpha = \omega_\alpha$, aber die beiden Bezeichnungen drücken eine unterschiedliche Funktionalität aus.

Eine Menge m heißt *endlich*, wenn $|m| \in \omega$. Eine unendliche Menge m heißt *abzählbar*, wenn $|m| = \aleph_0$, sonst *überabzählbar*. Die kleinste überabzählbare Ordinalzahl ist also ω_1.

Kardinalzahlarithmetik

Definition 6.36 *Addition $+_{\mathbb{KARD}}$, Multiplikation $\cdot_{\mathbb{KARD}}$ und Exponentiation für Kardinalzahlen κ, μ sind definiert durch*

$$\begin{array}{rcll} \kappa +_{\mathbb{KARD}} \mu & := & \big|\kappa +_{\mathbb{ORD}} \mu\big| & = & \big|\big(\kappa \times \{\underline{0}\}\big) \cup \big(\mu \times \{\underline{1}\}\big)\big| \\ \kappa \cdot_{\mathbb{KARD}} \mu & := & \big|\kappa \cdot_{\mathbb{ORD}} \mu\big| & = & \big|\kappa \times \mu\big| \\ \kappa^\mu & := & & & \big|\{f \mid f : \mu \to \kappa\}\big| \end{array}$$

Es ist also zum Beispiel $\aleph_0 +_{\mathbb{KARD}} \underline{1} = \aleph_0$, dagegen $\omega_0 +_{\mathbb{ORD}} \underline{1} \neq \omega_0$. Auf den natürlichen Zahlen stimmen Ordinalzahl-, Kardinalzahl- und gewöhnliche Arithmetik überein. Wenn klar ist, dass es sich um Kardinalzahlarithmetik handelt, schreibt man auch kurz $+$ und $\cdot$.

Satz 6.37 (Eigenschaften der Kardinalzahladdition und -multiplikation)

- *Addition und Multiplikation von Kardinalzahlen sind assoziativ und kommutativ und die Multiplikation ist distributiv über der Addition.*
- *$\underline{0}$ ist neutrales Element der Addition und $\underline{1}$ neutrales Element der Multiplikation.*
- *Es gelten für alle Kardinalzahlen $\kappa, \lambda, \mu, \nu$ die Potenzgesetze*

$$\kappa^{\underline{0}} = \underline{1} \quad , \quad \kappa^{\underline{1}} = \kappa \quad , \quad \kappa^{\mu+\nu} = \kappa^\mu \cdot \kappa^\nu \quad , \quad \kappa^{\mu\cdot\nu} = (\kappa^\mu)^\nu \quad , \quad (\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$$

- *Es gelten für alle Kardinalzahlen $\kappa, \lambda, \mu, \nu$ die Monotonieregeln:*

  Wenn $\kappa \leqslant \lambda$, ist $\kappa^\mu \leqslant \lambda^\mu$. Wenn $\mu \leqslant \nu$ und $\max\{\mu, \kappa\} \neq \underline{0}$, ist $\kappa^\mu \leqslant \kappa^\nu$.

Für diese und die folgenden Beweis ist eine in der Mengenlehre übliche Schreibweise nützlich: Für zwei Mengen a, b bezeichent ${}^a b$ die Menge der Funktionen $a \to b$. Es ist dann $|{}^a b| = |a|^{|b|}$.

Beweis: Die ersten beiden Spiegelstriche folgen aus Satz 6.38.

Potenzgesetze: Für jede Menge m gibt genau eine Funktion $\emptyset \to m$, nämlich die leere Funktion $\emptyset$. Für die weiteren Potenzgesetze müssen geeignete Bijektionen angebeben werden:

- $m \xrightarrow{\sim} {}^{\underline{1}}m$: $a \in m$ wird auf die konstante Funktion $\emptyset \mapsto a$ abgebildet.

[16] $\aleph$ („aleph“) ist der erste Buchstabe des hebräischen Alphabets.

- Die Zuordnung $f \in {}^{\mu+\nu}\kappa \mapsto (f|_{\mu\times\{\underline{0}\}}, f|_{\nu\times\{\underline{1}\}})$ ist eine Bijektion ${}^{\mu+\nu}\kappa \to {}^{\mu\times\{\underline{0}\}}\kappa \times {}^{\nu\times\{\underline{1}\}}\kappa$, die leicht zu einer Bijektion ${}^{\mu+\nu}\kappa \to {}^{\mu}\kappa \times {}^{\nu}\kappa$ gemacht werden kann.
- ${}^{\mu\times\nu}\kappa \xrightarrow{\sim} {}^{\nu}({}^{\mu}\kappa)$: $f \mapsto x \mapsto (y \mapsto f((x,y)))$.
- Für $f \in {}^{\mu}(\kappa\times\lambda)$ sei $f(\alpha) = (f_1(\alpha), f_2(\alpha))$. Dann ist $f \mapsto (f_1, f_2)$ die gesuchte Bijektion ${}^{\mu}(\kappa\times\lambda) \to {}^{\mu}\kappa \times {}^{\mu}\lambda$.

Für die Monotonieregeln müssen geeignete Injektionen angegeben werden:

- ${}^{\mu}\kappa \hookrightarrow {}^{\mu}\lambda$, denn die Verknüpfung mit der Inklusionsabbildung $\kappa \hookrightarrow \lambda$ macht aus jeder Funktion $\mu \to \kappa$ eine Funktion $\mu \to \lambda$.
- ${}^{\mu}\kappa \hookrightarrow {}^{\nu}\kappa$: Eine Funktion $\mu \to \kappa$ kann auf $\nu \setminus \mu$ konstant $= \emptyset$ fortgesetzt werden, falls $\kappa \neq \emptyset$. Sonst ist $\mu \neq \emptyset$, also $\emptyset^{\mu} = \emptyset^{\nu} = \underline{1}$. □

Satz 6.38 *Für unendliche Kardinalzahlen $\aleph_\alpha, \aleph_\beta$ und natürliche Zahlen $\underline{n}$ gilt*

$$\aleph_\alpha + \aleph_\beta = \aleph_\alpha \cdot \aleph_\beta = \max\{\aleph_\alpha, \aleph_\beta\} = \aleph_{\max\{\alpha,\beta\}}$$

$$\aleph_\alpha + \underline{n} = \aleph_\alpha \qquad \aleph_\alpha \cdot \underline{n} = \begin{cases} \underline{0} & \text{für } \underline{n} = \underline{0} \\ \aleph_\alpha & \text{sonst} \end{cases} \qquad {\aleph_\alpha}^{\underline{n}} = \begin{cases} \underline{1} & \text{für } \underline{n} = \underline{0} \\ \aleph_\alpha & \text{sonst} \end{cases}$$

Beweis: Es reicht $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$ für alle α zu zeigen, denn dann ist

$$\aleph_\alpha \leqslant \aleph_\alpha + \aleph_\beta \leqslant \aleph_\alpha + \aleph_\alpha = \aleph_\alpha \cdot \underline{2} \leqslant \aleph_\alpha \cdot \aleph_\beta \leqslant \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$$

Für $\underline{n} > \underline{0}$ gilt dann $\aleph_\alpha \leqslant \aleph_\alpha + \underline{n} \leqslant \aleph_\alpha + \aleph_\alpha = \aleph_\alpha$ bzw. $\aleph_\alpha \leqslant \aleph_\alpha \cdot \underline{n} \leqslant \aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. Außerdem gibt es für jede Menge m eine offensichtliche Bijektion zwischen der Menge der Funktionen $\underline{n} \to m$ und der Menge der n-Tupel über m. Die Aussage ${\aleph_\alpha}^{\underline{n}} = \aleph_\alpha$ folgt somit induktiv ebenfalls aus $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. Für $\underline{n} = \underline{0}$ findet man offensichtliche Bijektionen.

Zu zeigen ist also $\aleph_\alpha \cdot \aleph_\alpha = \aleph_\alpha$. Klar ist „$\geqslant$“; es reicht also die umgekehrte Ungleichung zu zeigen. Dazu betrachtet man auf $\mathbb{ORD} \times \mathbb{ORD}$ die folgende klassengroße Wohlordnung:

$$(\alpha_1,\beta_1) < (\alpha_2,\beta_2) :\Longleftrightarrow \begin{cases} \text{entweder} & \max\{\alpha_1,\beta_1\} < \max\{\alpha_2,\beta_2\} \\ \text{oder} & \max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\} \text{ und } \alpha_1 < \alpha_2 \\ \text{oder} & \max\{\alpha_1,\beta_1\} = \max\{\alpha_2,\beta_2\}, \alpha_1 = \alpha_2 \text{ und } \beta_1 < \beta_2 \end{cases}$$

Sei Γ die sich daraus ergebende ordnungserhaltende Bijektion $\mathbb{ORD}\times\mathbb{ORD} \to \mathbb{ORD}$. Ich schreibe kurz $\Gamma(\alpha,\beta)$ für $\Gamma\big((\alpha,\beta)\big)$. Für jede Ordinalzahl α ist $\alpha\times\alpha$ ein Anfangsstück von $\mathbb{ORD}\times\mathbb{ORD}$ bzgl. $<$, also $\Gamma[\alpha\times\alpha]$ eine Ordinalzahl.

Sei nun in Hinblick auf einen Widerspruch κ die kleinste unendliche Kardinalzahl mit $\kappa\cdot\kappa > \kappa$. Dann ist $\Gamma(\kappa,\kappa) > \kappa$, weil $\Gamma[\kappa\times\kappa]$ eine Ordinalzahl der Mächtigkeit κ ist. Es gibt also $\alpha,\beta > \kappa$ mit $\Gamma(\alpha,\beta) = \kappa$. Wähle nun δ mit $\max\{\alpha,\beta\} < \delta < \kappa$ (κ ist Limesordinalzahl!). Dann ist $\kappa = \Gamma(\alpha,\beta =) \in \Gamma[\delta\times\delta] =$ eine Ordinalzahl, und somit $\kappa \subseteq \Gamma[\delta\times\delta]$. Da Γ eine Bijektion ist, folgt daraus nun $|\delta| \leqslant \delta < \kappa \leqslant |\delta\times\delta| = |\delta|\cdot|\delta|$ im Widerspruch zur Wahl von κ. □

Über die charakteristischen Funktionen von Teilmengen sieht man $|\mathfrak{P}(m)| = |{}^{m}\underline{2}| = 2^{|m|}$. Den Satz von Cantor kann man also umformulieren in:

Für beliebige Kardinalzahlen κ gilt $\kappa < 2^\kappa$.

Die von Cantor aufgestellte *Kontinuumshypothese* (CH) besagt $2^{\aleph_0} = \aleph_1$, die *verallgemeinerte Kontinuumshypothese* (GCH) sagt $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ für alle α. Beide sind unabhängig von ZFC, d. h. in ZFC weder beweisbar noch widerlegbar (falls ZFC konsistent ist).

Man kann einige Eigenschaften der Kardinalzahlexponentiation beweisen, zum Beispiel ist

$$\aleph_\alpha{}^{\aleph_\alpha} \leqslant (2^{\aleph_\alpha})^{\aleph_\alpha} = 2^{\aleph_\alpha\cdot\aleph_\alpha} = 2^{\aleph_\alpha} \leqslant \aleph_\alpha{}^{\aleph_\alpha}, \text{ also } 2^{\aleph_\alpha} = \aleph_\alpha{}^{\aleph_\alpha}$$

Konkrete Werte von $\aleph_\alpha{}^{\aleph_\beta}$ kann man dagegen in ZFC typischerweise nicht bestimmen, wenn man nicht mit starken Annahmen wie GCH arbeitet.

Für konkrete Berechnung hilfreich ist das folgende Lemma:

Lemma 6.39 *Sei $m = \bigcup_{i\in I} m_i$ eine Vereinigung von paarweise disjunkten, nicht-leeren Mengen m_i über der unendlichen Indexmegne I.* [17] *Dann ist $|m| = |I| \cdot \sup\big\{|m_i| \;\big|\; i \in I\big\}$.*

Beweis: Sei $\kappa = \sup\big\{|m_i| \;\big|\; i \in I\big\}$. Für jedes i gibt es Injektionen $m_i \to \kappa \times \{i\}$, etwa $a \mapsto (f_i(a), i)$ für eine Bijektion $f_i : m_i \to |m_i|$. Zusammengefasst bekommt man eine Injektion $m \to \kappa \times I$, was „$\leqslant$“ zeigt. Umgekehrt ist $|I| \leqslant |m|$, weil man mit dem Auswahlaxiom eine Funktion $f : I \to m$ mit $f(i) \in m_i$ finden kann. Außerdem gibt es die Inklusionen $m_i \hookrightarrow m$, mithin ist $|m_i| \leqslant |m|$, also $\kappa \leqslant |m|$. Zusammen ergibt sich $|I| \cdot \kappa \leqslant |m| \cdot |m| = |m|$. □

Kurze Warnung: Wenn die Vereinigung nicht disjunkt ist, kann man wenig Allgemeines sagen. Die Kardinalität nicht stetig, d. h. vertauscht nicht mit der Supremumsbildung. Zum Beispiel gilt $\omega_1 = \bigcup\{\alpha \mid \alpha < \omega_1\}$, aber $\aleph_1 = |\omega_1| \neq \bigcup\{|\alpha| \mid \alpha < \omega_1\} = \aleph_0$, da alle Ordinalzahlen unterhalb von ω_1 abzählbar sind.

Auch die Kardinalzahlexponentiation vertauscht im Allgemeinen nicht mit Suprema, d. h. im Allgemeinen ist $k^\lambda \neq \sup\{k^\alpha \mid \alpha < \lambda\}$, Zum Beispiel ist $2^{\aleph_0} > \sup\{2^{\underline{n}} \mid \underline{n} \in \omega\} = \aleph_0$.

Folgerung 6.40 *Für jede unendliche Menge m ist die Menge ${}^{<\omega}m$ der endlichen Folgen von Elementen von m gleichmächtig mit m.*

Beweis: ${}^{<\omega}m$ ist die Vereinigung der ${}^{\underline{n}}m$ für $\underline{n} \in \omega$, also $|{}^{<\omega}m| = \aleph_0 \cdot \sup\big\{|{}^{\underline{n}}m| \;\big|\; \underline{n} \in \omega\big\} = \aleph_0 \cdot |m| = |m|$. □

Satz 6.41 *Es gilt*
$\aleph_0 = |\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\overline{\mathbb{Q}}| = |{}^{<\omega}\mathbb{Q}|$, $2^{\aleph_0} = |\mathfrak{P}(\mathbb{N})| = |{}^{\mathbb{N}}\mathbb{N}| = |\mathbb{R}|$, $2^{2^{\aleph_0}} = |\{f \mid f : \mathbb{R} \to \mathbb{R}\}|$

Beweis: $\aleph_0 = |\mathbb{N}| \leqslant |\mathbb{Z}| \leqslant |\mathbb{Q}| \leqslant |{}^{<\omega}\mathbb{Q}|$ ist klar. Nun kann man $\mathbb{Q}$ über (Vorzeichen, Zähler, Nenner) injektiv in $\underline{2} \times \mathbb{N} \times \mathbb{N}$ abbilden, also $|\mathbb{Q}| \leqslant \underline{2} \cdot \aleph_0 \cdot \aleph_0 = \aleph_0$ Zum andern sind die algebraischen Zahlen über $\mathbb{Q}$ Nullstellen von $\mathbb{Q}$-rationalen Polynomen, die durch die endlichen Folgen ihrer Koeffizienten beschrieben sind und jeweils nur endlich viele Nullstellen haben, also erhält man $|\overline{\mathbb{Q}}| \leqslant |{}^{<\omega}\mathbb{Q}| \cdot \aleph_0 = \aleph_0$, nach den bisherigen Überlegungen und Lemma 6.39.

$|\mathfrak{P}(\aleph_0)| = 2^{\aleph_0} = \aleph_0{}^{\aleph_0}$ ist bereits bewiesen. $|\mathbb{R}| \leqslant |\mathbb{Z}| \cdot |{}^{\mathbb{N}}\mathbb{N}| = |{}^{\mathbb{N}}\mathbb{N}|$ sieht man dadurch, dass man eine reelle Zahl r auf $\lfloor r \rfloor$ und die Folge der Nachkommastellen in der (im Zweifelsfall abbrechenden, durch Nullen ergänzten) Dezimalentwicklung abbildet. $2^{\aleph_0} \leqslant |\mathbb{R}|$ sieht man zum Beispiel, indem man jeder Funktion $f : \mathbb{N} \to \{0, 1\}$ die reelle Zahl $\sum_{i\geqslant 0} f(i) \cdot 10^{-i}$ zuordnet.

Die Menge der reellen Funktionen hat die Mächtigkeit $|{}^{\mathbb{R}}\mathbb{R}| = |\mathbb{R}|^{|\mathbb{R}|} = (2^{\aleph_0})^{2^{\aleph_0}} = 2^{\aleph_0\cdot 2^{\aleph_0}} = 2^{2^{\aleph_0}}$. □

Wie groß die Mächtigkeit von $\mathbb{R}$ (oder gar von ${}^{\mathbb{R}}\mathbb{R}$) ist, lässt sich nicht in ZFC entscheiden. Cantos Kontinuumshypothese behauptet, dass es die nach $\aleph_0$ nächstgrößerer Karidnalität ist,

[17] Insbesondere sind die m_i als *Familie* gegeben, d. h. es gibt eine Funktion $I \to m, i \mapsto m_i$.

also dass jede unendliche Teilemnge von $\mathbb{R}$ entweder abzählbar oder gleichmächtig zu $\mathbb{R}$ ist. Für topologisch schön zu beschreibende Teilmengen von $\mathbb{R}$ kann man dies nachweisen. Dies ist Ziel der sogenannte *deskriptiven Mengenlehre.*

Kurt Gödel hat bewiesen, dass die Kontinuumshypothese mit ZFC konsistent ist, also dass (sofern ZFC konsistent ist) es ein Modell von ZFC gibt (das sogenannte *konstruktible Universum*), in dem die Kontinuumshypothese gilt. Paul Cohen hat später gezeigt, dass auch die Negation der Kontinuumshypothese mit ZFC konsistent ist. In der Folge hat man gesehen, dass $2^{\aleph_0}$ beliebig große Werte annehmen kann und es nur ganz wenige allgemeine Einschränkungen gibt.

Reguläre Kardinalzahlen

Definition 6.42 *Eine Kardinalzahl κ heißt* singulär*, wenn es eine Kardinalzahl $\lambda < \kappa$ und eine Funktion $s : \lambda \to \kappa$ gibt, so dass $\kappa = \sup\{f(\alpha) \mid \alpha < \lambda\}$. Andernfalls heißt κ* regulär*.*

Jede Nachfolgerkardinalzahl ist reguläre. Die erte Limeskardinalzahl $\aleph_\omega$ dagegen ist singulär, denn sie ist Supremum der Abbildung $\omega \to \aleph_\omega$, $\underline{n} \to \aleph_n$.

Es lässt sich in ZFC nicht beweisen, dass es reguläre Limeskardinalzahlen gibt (sogenannte *schwach unerreichbare* Kardinalzahlen). Der „Qualitätssprung“ von unterhalb einer schwach unerreichbare Kardinalzahlen zu dieser hin lässt sich mit dem Qualitätssprung vom endlichen zu $\aleph_0$ hin vergleichen. Untersuchungen dieser Art gehören zur Theorie der *großen Kardinalzahlen.*

Eine schwach unerreichbare Kardinalzahl ist von der Form $\aleph_\alpha = \alpha$, d. h. es ist eine Ordinalzahl α, die zugleich die α-ste unendliche Kardinalzahl ist. Trotz der gigantischen Lücken zwischen zwei Kardinalzahlen haben sich also an solch einer Stelle die Ordinalzahl- und die Kardinalzahlzählung eingeholt. Die erste Kardinalzahl mit dieser Eigenschaft ist $\sup\{\aleph_\omega, \aleph_{\omega_\omega} \aleph_{\omega_{\omega_\omega}}, \dots\}$, die aber qua Konstruktion singulär ist.

7 Rekursivität

7.1 Primitiv rekursive und rekursive Funktionen

In diesem Abschnitt geht es darum, welche Funktionen $\mathbb{N}^k \to \mathbb{N}$ *berechenbar* sind. Es gibt verschiedenste Möglichkeiten, Berechenbarkeit zu definieren: Zum Beispiel durch abstrakte Maschinenmodelle (etwa Turing-Maschinen) oder durch konkrete Programmiersprachen, bei denen man von Ressourcenbeschränkungen abstrahiert. Alle bisherigen Herangehensweisen haben sich als äquivalent erwiesen (oder als schwächer), weshalb man in der sogenannten *Church'schen These* davon ausgeht, dass damit der intuitive Berechenbarkeitsbegriff adäquat umrissen ist und es keinen sinnvollen stärkeren Berechenbarkeitsbegriff gibt.

Äquivalent dazu ist auch der ohne Rückgriff auf Programmiersprachen oder Maschinenmodelle definierte Begriff der *rekursiven Funktion.* Ein historisch früheres, aber nicht ausreichendes Konzept ist das der *primitiven rekursiven Funktionen.*

Definition 7.1
(a) *Eine Funktion $f : \mathbb{N}^k \to \mathbb{N}$ für $k \in \mathbb{N}$ heißt* primitiv rekursiv*, wenn sie eine der* Grund- *oder* Ausgangsfunktionenfunktionen *ist oder sich aus primitiv rekursiven Funktionen durch Anwendung der Regeln* [Komposition] *und* [Primitive Rekursion] *ergibt.*

Grundfunktionen sind dabei

- die konstante Nullfunktion $0 : \mathbb{N}^0 \to \mathbb{N}$
- die Nachfolgerfunktion $S : \mathbb{N} \to \mathbb{N},\ n \mapsto n+1$ („*successor*“)
- die Projektionsfunktionen $\pi_i^k : \mathbb{N}^k \to \mathbb{N},\ (n_1, \ldots, n_k) \mapsto n_i$

und die beiden Regeln funktionieren folgendermaßen:

[**Komposition**] Aus Funktionen $f_1, \ldots, f_l : \mathbb{N}^k \to \mathbb{N}$ und $g : \mathbb{N}^l \to \mathbb{N}$ ergibt sich durch Komposition die Funktion $h : \mathbb{N}^k \to \mathbb{N}$ mit

$$h(n_1, \ldots, n_k) := g\big(f_1(n_1, \ldots, n_k), \ldots, f_l(n_1, \ldots, n_k)\big)$$

[**Primitive Rekursion**] Aus Funktionen $f : \mathbb{N}^k \to \mathbb{N}$ und $g : \mathbb{N}^{k+2} \to \mathbb{N}$ entsteht durch Primitive Rekursion die Funktion $h : \mathbb{N}^{k+1} \to \mathbb{N}$ mit

$$\begin{aligned} h(0, n_1, \ldots, n_k) &:= f(n_1, \ldots, n_k) \\ h(n+1, n_1, \ldots, n_k) &:= g\big(h(n, n_1, \ldots, n_k), n, n_1, \ldots, n_k\big) \end{aligned}$$

(b) *Eine Funktion $f : \mathbb{N}^k \to \mathbb{N}$ heißt* μ-rekursiv *oder kurz* rekursiv*, wenn sie eine der Grundfunktionen ist oder sich aus rekursiven Funktionen durch Anwendung der Regeln* [μ-Rekursion] *im Normalfall,* [Komposition] *und* [Primitive Rekursion] *ergibt.*

Die zusätzlich Regel lautet:

[**μ-Rekursion**] Aus einer Funktion $f : \mathbb{N}^{k+1} \to \mathbb{N}$ entsteht, sofern der *Normalfall* vorliegt, dass für alle $(n_1, \ldots, n_k) \in \mathbb{N}^k$ ein $m \in \mathbb{N}$ mit $f(m, n_1, \ldots, n_k) = 0$ existiert, durch μ-Rekursion die Funktion $h : \mathbb{N}^k \to \mathbb{N}$ mit

$$h(n_1, \ldots, n_k) := \mu m\,[f(m, n_1, \ldots, n_k) = 0]$$

wobei $\mu m\,[E]$ das kleinste $m \in \mathbb{N}$ mit der Eigenschaft E bezeichnet.

(c) *Eine Funktion $f : \mathbb{N}^k \to \mathbb{N}^l$ heißt (primitiv) rekursiv, wenn alle Komponentenfunktionen $f_i = \pi_i^l \circ f : \mathbb{N}^k \to \mathbb{N}$ (primitiv) rekursiv sind.*

Liegt nicht der Normalfall vor, entsteht durch μ-Rekursion eine partielle Funktion. Daher bilden partielle Funktionen den eiegntlich besseren Rahmen für Berechenbarkeitsbetrachtungen.

(d) *Eine* partielle Funktion $f : M \dashrightarrow N$ *ist eine Funktion $f' : D \to N$ für eine Teilmenge $D \subseteq M$, den* Definitionsbereich *von f. Für $m \notin D$ sagt man, dass f in m bzw. $f(m)$* nicht definiert *oder* unbestimmt *ist. Für $m \in D$ ist $f(m) := f'(m)$* definiert */*bestimmt*.*

Eine partielle Funktion $f : \mathbb{N}^k \dashrightarrow \mathbb{N}$ heißt (μ-)rekursiv*, wenn sie eine der Grundfunktionen ist oder sich aus partiellen rekursiven Funktionen durch Anwendung der Regeln* [μ-Rekursion]*,* [Komposition] *und* [Primitive Rekursion] *ergibt, die wie folgt auf partielle Funktionen ausgeweitet werden:*

- Aus partiellen Funktionen $f_1, \ldots, f_l : \mathbb{N}^k \dashrightarrow \mathbb{N}$ und $g : \mathbb{N}^l \dashrightarrow N$ entsteht durch [Komposition] die partielle Funktion $h : \mathbb{N}^k \dashrightarrow \mathbb{N}$ mit

$$h(\bar{a}) := \begin{cases} g\big(f_1(\bar{a}), \ldots, f_l(\bar{a})\big) & \text{falls } f_1(\bar{a}), \ldots, f_l(\bar{a}) \text{ und } g(f_1(\bar{a}), \ldots, f_l(\bar{a})) \text{ definiert} \\ \text{unbestimmt} & \text{sonst} \end{cases}$$

- Aus partiellen Funktionen $f : \mathbb{N}^k \dashrightarrow \mathbb{N}$ und $g : \mathbb{N}^{k+2} \dashrightarrow \mathbb{N}$ entsteht durch [Primitive

Rekursion] die partielle Funktion $h : \mathbb{N}^{k+1} \dashrightarrow \mathbb{N}$ mit

$$h(0, \bar{a}) := \begin{cases} f(\bar{a}) & \text{falls } f(\bar{a}) \text{ definiert} \\ \text{unbestimmt} & \text{sonst} \end{cases}$$

$$h(n+1, \bar{a}) := \begin{cases} g\big(h(n, \bar{a}), n, \bar{a}\big) & \text{falls } h(n, \bar{a}) \text{ und } g(h(n, \bar{a}), n, \bar{a}) \text{ definiert} \\ \text{unbestimmt} & \text{sonst} \end{cases}$$

- Aus einer partiellen Funktion $f : \mathbb{N}^{k+1} \dashrightarrow \mathbb{N}$ entsteht durch [μ-Rekursion] die partielle Funktion $h : \mathbb{N}^k \dashrightarrow \mathbb{N}$ mit

$$h(\bar{a}) := \begin{cases} \mu m\,[f(m, \bar{a}) = 0] & \text{falls ein } m \text{ mit } f(m, \bar{a}) = 0 \text{ existiert} \\ & \text{und } f(n, \bar{a}) \text{ für alle } n \leqslant m \text{ definiert ist} \\ \text{unbestimmt} & \text{sonst} \end{cases}$$

Die Rekursionsschemata sind in der Anwendung nicht so starr, wie es zunächst aussieht. Wenn $f : \mathbb{N}^k \to \mathbb{N}$ (primitiv) rekursiv ist und $\sigma : \{1, \ldots, k\} \to \{1, \ldots, l\}$, dann ist auch die Funktion

$$\mathbb{N}^l \to \mathbb{N},\ (n_1 \ldots, n_l) \mapsto f(n_{\sigma(1)}, \ldots, n_{\sigma(k)}) = f\big(\pi^l_{\sigma(1)}(n_1, \ldots, n_l), \ldots, \pi^l_{\sigma(k)}(n_1, \ldots, n_l)\big)$$

(primitiv) rekursiv. Insbesondere kann man Variable permutieren und „stumme Variable“ hinzufügen. Man muss bei den Rekursionsschemata [Primitive Rekursion] und [μ-Rekursion] die Rekursion daher nicht notwendigerweise über die erste Variable laufen lassen und braucht im Schema der Primitiven Rekursion in der Funktion g nicht unbedingt die zweite Variable n.

Lemma 7.2 *Die folgenden Funktionen sind primitiv rekursiv:*

- *alle konstanten Funktionen* $c^k_m : \mathbb{N}^k \to \mathbb{N}, (n_1 \ldots, , n_k) \mapsto m$
- *Addition, Multiplikation und Exponentiation, jeweils* $\mathbb{N}^2 \to \mathbb{N}$
- *die Fakultätsfunktion* $\mathbb{N} \to \mathbb{N}$, $n \mapsto n!$
- *die* Vorgängerfunktion („predecessor“) $P : \mathbb{N} \to \mathbb{N}$ *mit* $P(0) := 0$ *und* $P(n+1) := n$
- *die* modifizierte Differenz $\dot{-} : \mathbb{N}^2 \to \mathbb{N}$ *mit* $m \dot{-} n := \max\{m - n, 0\}$
- *die Abstandsfunktion* $\mathbb{N}^2 \to \mathbb{N}$, $(m, n) \mapsto |m - n|$
- *die Maximums- und Minimumsfunktionen* $\mathbb{N}^k \to \mathbb{N}$, $(n_1, \ldots, n_k) \mapsto \max\{m_1, \ldots, n_k\}$ *bzw.* $(n_1, \ldots, n_k) \mapsto \min\{m_1, \ldots, n_k\}$

Beweis: (a) Die konstante Funktion c^0_0 ist eine Grundfunktion und die weiteren konstanten Nullfunktionen bekommt man durch iterierte primitive Rekursion:

$$\begin{aligned} c^{k+1}_0(n_1, \ldots, n_k, 0) &= c^k_0(n_1, \ldots, n_k) \\ c^{k+1}_0(n_1, \ldots, n_k, n+1) &= c^{k+1}_0(n_1, \ldots, n_k, n) = \pi^{k+2}_1(c^{k+1}_0(n_1, \ldots, n_k, n), n, n_1, \ldots, n_k) \end{aligned}$$

Für $m \neq 0$ erhält man induktiv $c^k_m = S \circ c^k_{m-1}$ aus der konstanten Nullfunktion durch iterierte Komposition mit der Nachfolgerfunktion S.

(b) Die Addition bekommt man mit primitiver Rekursion über m und Komposition:

$$\begin{aligned} 0 + n &= n = \pi^1_1(n) \\ (m+1) + n &= (m+n) + 1 = S(m+n) = (S \circ \pi^3_1)(m+n, m, n) \end{aligned}$$

ebenso die Multiplikation aus der Addition:

$$0 \cdot n = 0 = c_0^1(n)$$
$$(m+1) \cdot n = (m \cdot n) + n = \pi_1^3(m \cdot n, m, n) + \pi_3^3(m \cdot n, m, n)$$

und die Exponentiation aus der Multiplikation:

$$m \cdot 0 = 1 = c_1^1(m)$$
$$m^{n+1} = m^n \cdot m = \pi_1^3(m^n, m, n) \cdot \pi_2^3(m^n, m, n)$$

(c)–(d) Jeweils mit primitiver Rekursion und Komposition aus bekannten Funktionen:

$$0! = 1 = c_1^0 \quad \text{und} \quad (n+1)! = S(n) \cdot n! = \pi_1^2(n!, n) \cdot (S \circ \pi_2^2)(n!, n)$$
$$P(0) = 0 = c_0^0 \quad \text{und} \quad P(n+1) = n = \pi_2^2(P(n), n)$$
$$m \dot{-} 0 = m = \pi_1^1(m) \quad \text{und} \quad m \dot{-} (n+1) = P(m \dot{-} n) = (P \circ \pi_1^3)(m \dot{-} n, n, m)$$

(f)–(g) Jeweils als Komposition bekannter Funktionen:

$$|m - n| = (m \dot{-} n) + (n \dot{-} m)$$
$$\max\{n_1, n_2\} = n_1 + (n_2 \dot{-} n_1) \quad \text{und} \quad \min\{n_1, n_2\} = n_1 \dot{-} (n_1 \dot{-} n_2)$$

Für höhere Stelligkeiten durch Komposition paarweiser Maxima bzw. Minima. □

Definition 7.3

(a) *Eine Teilmenge $R \subseteq \mathbb{N}^k$ heißt* (primitiv) rekursiv*, wenn ihre charakteristische Funktion $\chi_R : \mathbb{N}^k \to \mathbb{N}$ (primitiv) rekursiv ist.*

(b) *Eine Teilmenge $A \subseteq \mathbb{N}^k$ heißt* rekursiv aufzählbar*, wenn es ein rekursives $R \subseteq \mathbb{N}^{k+1}$ mit*

$$A = \pi[R] = \big\{(n_1, \ldots, n_k) \;\big|\; \text{es gibt } m \in \mathbb{N} \text{ mit } (n_1, \ldots, n_k, m) \in R\big\}$$

gibt, wobei $\pi = (\pi_1^{k+1}, \ldots, \pi_k^{k+1}) : N^{k+1} \to \mathbb{N}^k$ die Projektion auf die ersten k Koordinaten ist.

Beispiel: Die $<$-Relation und die $\leqslant$-Relation sind primitiv rekursiv, denn

$$\chi_{\leqslant}(x, y) = \; 1 \dot{-} (x \dot{-} y) \; = \begin{cases} 1 & x \leqslant y \\ 0 & x > y \end{cases}$$

$$\chi_{<}(x, y) = 1 \dot{-} \chi_{\leqslant}(y, x) = \begin{cases} 1 & x < y \\ 0 & x \geqslant y \end{cases}$$

Die Diagonale $\Delta := \{(n, n) \mid n \in \mathbb{N}\} \subseteq \mathbb{N}^2$ ist primitiv rekursiv, da $\chi_\Delta(x, y) = 1 \dot{-} |x - y|$.

Rekursive Mengen können im folgenden Sinn als Bedingung für die μ-Rekursion verwendet werden: Wenn $R \subseteq \mathbb{N}^{k+1}$ rekursiv ist und $\pi[R] = \mathbb{N}^k$ für die Projektion $\pi = (\pi_1^{k+1}, \ldots, \pi_k^{k+1})$, dann ist die Funktion $f : \mathbb{N}^k \to \mathbb{N}$ mit $f(\bar{n}) = \mu m[(\bar{n}, m) \in R] = \mu m[1 \dot{-} \chi_R(\bar{n}, m) = 0]$ rekursiv.

Lemma 7.4 *Die Menge der (primitiv) rekursiven Teilmengen aller $\mathbb{N}^k$ enthält $\emptyset$ und alle $\mathbb{N}^k$ und ist abgeschlossen bezüglich Komplement, kartesischen Produkten und Urbildern unter rekursiven Funktionen sowie (sofern sinnvoll) Schnitt, Vereinigung und Differenz.*

Beweis: $\mathbb{N}^k$ und $\emptyset \subseteq \mathbb{N}^k$ sind rekursiv, da $\chi_\emptyset = c_0^k$ und $\chi_{\mathbb{N}^k} = c_1^k$. Schnitt und Differenz für $X, Y \subseteq \mathbb{N}^k$ erklären sich durch $\chi_{X \cap Y} = \chi_X \cdot \chi_Y$ und $\chi_{X \setminus Y} = \chi_X \dot{-} \chi_Y$. Damit sind auch Komplemente $\mathbb{N}^k \setminus X$ und Vereinigungen $X \cup Y = \mathbb{N}^k \setminus ((\mathbb{N}^k \setminus X) \cap (\mathbb{N}^k \setminus Y))$ abgedeckt.

Für $A \subseteq \mathbb{N}^k$ und $B \subseteq N^l$ ist $\chi_{A \times B}(n_1, \ldots, n_{k+l}) = \chi_A(n_1, \ldots, n_k) \cdot \chi_B(n_{k+1}, \ldots, n_{k+l})$, und wenn $f : \mathbb{N}^k \to \mathbb{N}^l$ und $R \subseteq \mathbb{N}^l$, ist $\chi_{f^{-1}[R]} = \chi_R \circ f$. □

Mit diesem Lemma sieht man insbesondere:

- Jede rekursive Menge $R \subseteq \mathbb{N}^k$ ist auch rekursiv aufzählbar, weil R die Projektion der rekursiven Menge $R \times \mathbb{N}$ ist.
- Alle endlichen Mengen sind primitiv rekursiv, denn für jedes m ist

  $$\chi_{\{m\}}(n) = 1 \dot{-} |n - m| = 1 \dot{-} |\pi_1^1(n) - c_m^1(n)|$$

  primitiv rekursiv. Mengen, die sich von einer (primitiv) rekursiven Menge nur um eine endliche Mengen unterscheiden, sind daher ebenfalls (primitiv) rekursiv.

Lemma 7.5 *Sei $\beta : \mathbb{N}^k \to \mathbb{N}^l$ eine Bijektion und β, β^{-1} beide (primitiv) rekursiv. Dann ist $X \subseteq \mathbb{N}^k$ genau dann primitiv rekursiv / rekursiv / rekursiv aufzählbar, wenn $\beta[X] \subseteq \mathbb{N}^l$ primitiv rekursiv / rekursiv / rekursiv aufzählbar ist.*

Für die Aussage zu „primitiv rekursiv“ müssen dabei β und β^{-1} primitiv rekursiv sein, für „rekursiv/rekursiv aufzählbar“ reicht es, dass β und β^{-1} rekursiv sind.

Beweis: Sei $X \subseteq \mathbb{N}^k$ (primitiv) rekursiv, dann ist $\chi_{\beta[X]}(n_1, \dots, n_l) = \chi_X(\beta^{-1}(n_1, \dots, n_l))$ (primitiv) rekursiv. Sei $X = \pi[Y] \subseteq \mathbb{N}^k$ rekursiv aufzählbar mit rekursivem $Y \subseteq \mathbb{N}^{k+1}$. Nun ist $\beta \times \mathrm{id}_{\mathbb{N}} : \mathbb{N}^{k+1} \to \mathbb{N}^{l+1}$ bijektiv und rekursiv, ebenso die Umkehrfunktion $\beta^{-1} \times \mathrm{id}_{\mathbb{N}}$. Also ist $(\beta \times \mathrm{id}_{\mathbb{N}})[Y]$ rekursiv und somit $\beta[X] = \pi\big[(\beta \times \mathrm{id}_{\mathbb{N}})[Y]\big]$ rekursiv aufzählbar. □

Lemma 7.6 *Für alle $k, l > 0$ gibt es rekursive Bijektionen $\beta_l^k : \mathbb{N}^k \to \mathbb{N}^l$ mit $\beta_l^k = (\beta_k^l)^{-1}$.*

Beweis: Es reicht aus, den Fall $\{k, l\} = \{1, 2\}$ zu betrachten, den Rest bekommt man durch iterierte Komposition. Für β_1^2 kann man $(m, n) \mapsto \frac{1}{2}(m + n)(m + n + 1) + m$ wählen, was als Komposition rekursiver Funktionen rekursiv ist: Die (abgerundete) Division durch 2 bekommt man etwa durch $\mu m\,[n \leqslant 2m + 1]$. Mit $h(n) = \mu m\,[\frac{1}{2}m(m + 1) > n]$ ist die Umkehrfunktion rekursiv als $\beta_2^1(n) = \big(n \dot{-} \frac{1}{2}h(n)(h(n) \dot{-} 1), \frac{1}{2}h(n)(h(n) + 1) \dot{-} (n + 1)\big)$. □

Satz 7.7
(a) $f : \mathbb{N}^k \to \mathbb{N}$ ist genau dann rekursiv, wenn der Graph $\Gamma_f \subseteq \mathbb{N}^{k+1}$ rekursiv aufzählbar ist, genau dann, wenn Γ_f rekursiv ist. [18]

(b) Eine Menge $X \subseteq \mathbb{N}^l$ ist genau dann rekursiv aufzählbar, wenn sie das Bild einer rekursiven Menge unter einer rekursiven Funktion ist, also wenn es ein rekursives $Y \subseteq \mathbb{N}^k$ und eine rekursive Funktion $f : \mathbb{N}^k \to \mathbb{N}^l$ mit $X = f[Y]$ gibt, und genau dann, wenn X entweder die leere Menge ist oder das Bild $g[\mathbb{N}^k]$ einer rekursiven Funktion $g : \mathbb{N}^k \to \mathbb{N}^l$. In beiden Fällen kann $k = 1$ gewählt werden.

(c) $X \subseteq \mathbb{N}^k$ ist genau dann rekursiv, wenn X und $\mathbb{N}^k \setminus X$ rekursiv aufzählbar sind, und genau dann, wenn es ein rekursives $Y \subseteq \mathbb{N}$ mit $X = f[Y]$ gibt, wobei $f : \mathbb{N} \to \mathbb{N}^k$ eine rekursive Bijektion ist.

Teil (c) erklärt den Begriff „rekursive Aufzählbarkeit“, da jede rekursiv aufzählbare Teilmenge $A \neq \emptyset$ Bild einer rekursiven Funktion $f : \mathbb{N} \to \mathbb{N}^l \supseteq A$ ist, also $A = \{f(0), f(1), f(2), \dots\}$, wobei die Funktionswerte $f(0), f(1), f(2), \dots$ sukzessive berechnet werden können.

Beweis: (a) Es ist $(\bar{x}, y) \in \Gamma_f \iff f(\bar{x}) = y \iff |f(\bar{x}) - y| = 0$, also ist $\chi_{\Gamma_f}(\bar{x}, y) = 1 \dot{-} \big|f(\bar{x}) - y\big|$ mit f ebenfalls rekursiv. Wenn $\Gamma_f = \{(\bar{n}, f(\bar{n}) \mid \bar{n} \in \mathbb{N}^k\}$ rekursiv aufzählbar ist und Projektion der rekursiven Menge $R \subseteq \mathbb{N}^{k+2}$ entlang der ersten Koordinate ist, ist $f(\bar{n}) = \big\langle \mu\beta_1^2(y, z)\big[(z, \bar{n}, y) \in R\big]\big\rangle_0$ rekursiv.

[18] Die zweite Äquivalenz gilt nur für totale Funktionen, die erste auch für partielle.

(b) Eine rekursiv aufzählbare Menge $X \subseteq \mathbb{N}^k$ ist per Definition Bild einer rekursiven Menge $Y \subseteq \mathbb{N}^{k+1}$ unter einer Projektion (die rekursiv ist). Unter Vorschalten von β_k^1 kann man $k = 1$ annehmen.

Ist $\emptyset \neq X = f[Y]$ mit rekursivem $Y \subseteq \mathbb{N}^k$, kann man die Funktion f in die Funktion g mit Bild X abändern, indem man $g(\bar{y}) := f(\bar{y}) \cdot \chi_Y(\bar{y}) + n_0 \cdot \chi_{\mathbb{N}^{k+1} \setminus Y}(\bar{y})$ für ein $n_0 \in X$ setzt. Wenn f rekursiv ist, ist auch g rekursiv.

Wenn $X = g[\mathbb{N}^k]$ für rekursives g, kann man $k = 1$ annehmen (Vorschalten von β_k^1!). Dann ist $\Gamma_g = \{(n, g(n)) \mid n \in \mathbb{N}\}$ und nach Teil (a) rekursiv, und X ist die Projektion von Γ_g (und zwar nach geeigneter Permutation der Koordinaten wie in Definition 7.3).

(c) Mit X ist auch $\mathbb{N}^k \setminus X$ rekursiv, und dann beide rekursiv aufzählbar. Sind umgekehrt $X = \pi[Y]$ und $\mathbb{N}^k \setminus X = \pi[Z]$ Projektionen rekursiver Mengen $Y, Z \subseteq \mathbb{N}^{k+1}$, dann ist $Y \cup Z$ rekursiv nach Lemma 7.4, also auch $\chi_X(\bar{x}) = \chi_Y(\bar{x}, \mu y\,[(\bar{x}, y) \in Y \cup Z])$.

Die zweite Äquivalenz ist Teil von Lemma 7.5, insbesondere ist $X = \beta_k^1[Y]$ mit $Y = \beta_1^k[X]$. □

Aus Teil (a) des Satzes folgt insbesondere, dass $\{\bar{n} \in \mathbb{N}^k \mid f_1(\bar{n}) = \cdots = f_m(\bar{n})\}$ (primitiv) rekursiv ist für (primitiv) rekursive $f_1, \ldots, f_m : \mathbb{N}^k \to \mathbb{N}$, da $= (\mathrm{id}_\mathbb{N}, f_1)^{-1}[\Gamma_{f_1} \cap \cdots \cap \Gamma_{f_m}]$.

Folgerung 7.8 *Die Menge der rekursiv aufzählbaren Teilmengen aller $\mathbb{N}^k$ ist abgeschlossen bezüglich kartesischen Produkten und Bildern und Urbildern unter rekursiven Funktionen sowie (sofern sinnvoll) Schnitt und Vereinigung.*

Beweis: Man benutzt die Eigenschaften rekursiver Mengen aus Lemma 7.4 : Sind $A = \pi[R] \subseteq \mathbb{N}^l$ und $B = \pi[S] \subseteq \mathbb{N}^l$ rekursiv aufzählbar mit rekursiven R, S, dann ist $A \cup B = \pi[R \cup S]$ rekursiv aufzählbar.

Ist A wie oben und $f : \mathbb{N}^k \to \mathbb{N}^l$, dann ist $f^{-1}[A] = \pi'[(f \times \mathrm{id}_\mathbb{N})^{-1}[R]]$ für die Projektion $\pi' : \mathbb{N}^{l+1} \to \mathbb{N}^l$ aus die ersten l Koordinaten, und mit f ist auch $f \times \mathrm{id}_\mathbb{N}$ rekursiv. Also ist $f^{-1}[A]$ rekursiv aufzählbar. Ist zudem $g : \mathbb{N}^l \to \mathbb{N}^m$ rekursive Funktion, dann ist $g[A] = \pi[(g \times \mathrm{id})[R]]$ rekursiv aufzählbar nach Satz 7.7.

Ist A wie oben und nun $B = \pi[S] \subseteq \mathbb{N}^k$ mit rekursivem S, dann ist $A \times B$ die geeignete Projektion von $R \times S \subseteq \mathbb{N}^{l+k+2}$ auf $\mathbb{N}^{l+k}$, also nach Satz 7.7 rekursiv aufzählbar. Im Fall $l = k$ ist die Menge $\Delta = \{(x_1, \ldots, x_{l+k+2}) \mid x_1 = x_{l+2}, \ldots, x_l = x_{2l+2}\}$ rekursiv, also auch $(A \times B) \cap \Delta$. Eine geeignete Projektion davon auf $\mathbb{N}^l$ ist dann $A \cap B$. □

Mengen, die sich von einer rekursiv aufzählbaren Menge nur um eine endliche Mengen unterscheiden, sind ebenfalls rekursiv aufzählbar (da für eine endliche Menge $E \subseteq \mathbb{N}^k$ auch $E \times \mathbb{N} \subseteq \mathbb{N}^{k+1}$ rekursiv ist).

Es folgen nun noch einige weitere Konstruktionselemente für rekursive Funktionen.

Lemma 7.9

[Fallunterscheidung] *Wenn $R_1, \ldots, R_l \subseteq \mathbb{N}^k$ und $f_1, \ldots, f_{l+1} : \mathbb{N}^k \to \mathbb{N}$ (primitiv) rekursiv sind, dann auch die Funktion $f : \mathbb{N}^k \to \mathbb{N}$ mit*

$$f(\bar{n}) := \begin{cases} f_1(\bar{n}) & \bar{n} \in R_1 \\ f_2(\bar{n}) & \bar{n} \in R_2 \setminus R_1 \\ \vdots & \vdots \\ f_l(\bar{n}) & \bar{n} \in R_l \setminus (R_1 \cup \cdots \cup R_{l-1}) \\ f_{l+1}(\bar{n}) & \bar{n} \notin R_1 \cup \cdots \cup R_l \end{cases}$$

[Beschränkte Quantifikation] *Wenn $R \subseteq \mathbb{N}^{k+1}$ (primitiv) rekursiv ist, dann auch*

$$R_\forall := \big\{(\bar{n}, m) \in \mathbb{N}^{k+1} \;\big|\; \forall x < m \;\; (\bar{n}, x) \in R\big\}$$
$$R_\exists := \big\{(\bar{n}, m) \in \mathbb{N}^{k+1} \;\big|\; \exists x < m \;\; (\bar{n}, x) \in R\big\}$$

In beiden Fällen braucht man nicht das Schema der Primitiven Rekursion!

Beweis: (a) Setze $R_{l+1} = \mathbb{N}^k$, dann ist $f(\bar{n}) = \sum_{i=1}^{l+1} f_i(\bar{n}) \cdot \chi_{R_i \setminus (R_1 \cup \cdots \cup R_{i-1})}(\bar{n})$.

(b) Mit primitiver Rekursion:

$$\chi_{R_\forall}(\bar{n}, y) = \begin{cases} 1 & y = 0 \\ \chi_{R_\forall}(\bar{n}, y \dot{-} 1) \cdot \chi_R(\bar{n}, y \dot{-} 1) & y > 0 \end{cases}$$

und ohne primitive Rekursion:

$$\chi_{R_\forall}(\bar{n}, y) = \begin{cases} 1 & \mu z\big[(\bar{n}, z) \notin R \text{ oder } z \geqslant y\big] \geqslant y \\ 0 & \mu z\big[(\bar{n}, z) \notin R \text{ oder } z \geqslant y\big] < y \end{cases}$$

Außerdem gilt $R_\exists = \mathbb{N}^{k+1} \setminus (\mathbb{N}^{k+1} \setminus R)_\forall$. □

Zum Schluss dieses Abschnitts wird gezeigt, dass „primitiv rekursiv beschränkte“ μ-Rekursion primitiv rekursive Funktionen erhält:

Lemma 7.10 *Sei $R \subseteq \mathbb{N}^{k+1}$ primitiv rekursiv. Wenn $\pi[R] \subseteq \mathbb{N}^k$ und es eine primitiv rekursive Funktion $f : \mathbb{N}^k \to \mathbb{N}$ gibt mit $g(\bar{n}) := \mu m\,[\,(\bar{n}, m) \in R\,] \leqslant f(\bar{n})$ für alle $\bar{n} \in \mathbb{N}^k$, dann ist g primitiv rekursiv.*

Für die beschränkte μ-Rekursion wie im Lemma schreibe ich $g(\bar{n}) = \mu m \leqslant f(\bar{n})\,[\,(\bar{n}, m) \in R\,]$.

Beweis: Setze $h(\bar{n}, y) := \mu m[(\bar{n}, m) \in R \text{ oder } m = y]$. Die zusätzliche Variable erlaubt es, h mit Hilfe von Primitiver Rekursion zu definieren über

$$h(\bar{0}) = 0 \quad \text{und} \quad h(\bar{n}, m+1) = \begin{cases} h(\bar{n}, m) & \text{falls } (\bar{n}, h(\bar{n}, m)) \in R \\ m+1 & \text{sonst} \end{cases}$$

Dann ist $g(\bar{n}) = h(\bar{n}, f(\bar{n}))$ ebenfalls primitiv rekursiv. □

Dieses Lemma kann man so verstehen, dass die Berechenbarkeit im Sinne der Primitiven Rekursion „For-Schleifen“ zulässt, während die μ-Rekursion als eine Art „While-Schleife“ angesehen werden kann.

7.2 Kodierung und Gödelisierung

Die Bijektionen $\beta_1^k : \mathbb{N}^k \to \mathbb{N}$ erlauben es, k-Tupel als natürliche Zahlen zu kodieren. Das hat man gebraucht, um zum Beispiel zu zeigen, dass beliebige Projektionen rekursiver Mengen rekursiv aufzählbar sind. In diesem Abschnitt sollen zwei (von Gödel stammende) Kodierungen eingeführt werden, mit denen man endliche Tupel natürlicher Zahlen beliebiger Länge als natürliche Zahlen kodieren kann – einerseits Tupel beliebiger, aber bekannter Länge und andererseits Tupel aller Längen gleichzeitig. Dazu braucht es ein paar Vorbereitungen:

Lemma 7.11 *Primitiv rekursiv sind:*

die zweistellige Teilbarkeitsrelation $x \mid y$

die Menge der Primzahlen $\mathbb{P}$

die „Primzahlaufzählung“, d. h. die Funktion $p : \mathbb{N} \to \mathbb{N}, n \mapsto$ *die $(n+1)$-te Primzahl* p_n

Dabei sind die Teilbarkeitsrelation und die Menge der Primzahlen auch ohne Primitive Rekursion als rekursiv beschreibbar.

Beweis: $x \mid y \iff (x = 0 \text{ und } y = 0)$ oder $(x \neq 0 \text{ und } \mu z\,[x \cdot z \geqslant y] \cdot x \leqslant y)$. Da im zweiten Fall sicher $z \leqslant y$ gilt, ist die Teilbarkeit nach dem voranstehende Lemma auch primitiv rekursiv.

$x \in \mathbb{P} \iff (1 < x \text{ und } \forall y < x\,(y \nmid x \text{ oder } y \leqslant 1))$.

Schließlich ist p rekursiv definiert über $p(0) = 2$ und $p(n+1) = \mu m\,[m \in \mathbb{P} \text{ und } m > p(n)]$. Aus dem Satz von Bertrand, dass für $n > 1$ zwischen n und $2n$ eine Primzahl liegt weiß, bekommt man $p(n) \leqslant 2^{n+1}$, daher ist p nach Lemma 7.10 primitiv rekursiv. □

Satz 7.12 *Die rekursiven Funktionen können alternativ definiert werden als die kleinste Menge von Funktionen, die die* erweiterten Grundfunktionen *enthalten und abgeschlossen sind unter* [Komposition] *und* [μ-Rekursion] *im Normalfall.*

Erweiterte Grundfunktionen sind dabei die konstanten Nullfunktionen c_0^k beliebiger Stelligkeit, alle Projektionen π_i^k, Nachfolger S, Addition $+$, Multiplikation $\cdot$ und die charakteristische Funktion $\chi_<$ der Kleiner- oder $\chi_\leqslant$ der Kleiner-Gleich-Relation.

Beweis: Mit der Nachfolgerfunktion sieht man $\chi_\leqslant(x,y) = \chi_<(x,S(y))$ und $\chi_\leqslant(S(x),y) = \chi_<(x,y)$. Außerdem ist $x \dot{-} y = \mu z\,[x \leqslant y + z] = \mu z[\chi_\leqslant(x, y+z) = 1] = \mu z[\chi_<(x, y+z) = 0]$. Abgesehen von der Primitiven Rekursivität der Exponentiation und der Fakultätsfunktion und von speziellen Ergebnissen über Primitive Rekursivität ist das Schema der Primitiven Rekursion bisher nicht benutzt worden.

Wenn h aus f und g durch Primitive Rekursion entsteht, also

$$h(0,\bar{n}) = f(\bar{n}) \text{ und } h(n+1,\bar{n}) = g(h(n,\bar{n}),n,\bar{n})$$

dann möchte man h im Rekursionsschritt durch die Folge $h(0,\bar{n}), h(1,\bar{n}), \ldots, h(n+1,\bar{n})$ beschreiben. Wenn man diese Folge durch eine Zahl m so kodieren kann, dass man aus m die Länge $n+2$ der Folge und die einzelnen Komponenten $\langle m\rangle_i = f(i,\bar{n})$ rekursiv berechnen kann, dann beschreibt man h im Rekursionsschritt alternativ als

$$\begin{aligned} h(n+1,\bar{n}) = {} & \text{das letzte Folgenglied der durch } m \text{ kodierten Folge der Länge } n+2 \\ & \text{mit } \langle m\rangle_0 = f(\bar{n}) \text{ und } \langle m\rangle_{i+1} = g(\langle m\rangle_i, i, \bar{n}) \text{ für alle } i = 0,\ldots,n \end{aligned}$$

Die Kodierung benutzt den Chinesischen Restsatz, der besagt, dass es für jedes k zu paarweise teilerfremden Zahlen $n_1,\ldots,n_k$ und vorgegebenen „ Resten“ $r_1,\ldots,r_k$ eine Zahl m mit

$$m \equiv r_i \mod n_i \quad \text{d. h.} \quad n_i \mid (m - r_i)$$

gibt. Wenn die n_i gegeben sind und $r_i < n_i$, findet man r_i wieder als Rest von der Division von m durch n_i.

Man sucht nun die kleinste Zahl $\beta_1^2(m,k)$, so dass für alle $i \leqslant n+1$ gilt:

$$m > (n+1)k \qquad i \mid k \qquad r_0 = f(\bar{n}) \qquad r_{i+1} = g(r_i, i, \bar{n})$$

wobei r_i der Rest von m bei Division durch $ik+1$ ist, den man als $\mu x\,[\,ik+1 \mid m \dot{-} x\,]$ rekursiv berechnet (da nach der ersten Bedingung $m \geqslant ik+1$). Die zweite Bedingung stellt sicher, dass die Zahlen $ik+1$ für $i = 0,\ldots,n+1$ paarweise teilerfremd sind. Auf m und k hat man über $\beta_2^1 = (\beta_1^2)^{-1}$ Zugriff, was man mit Hilfe von Lemma 7.10 leicht als primitiv rekursiv erkennt.

Alle Bedingungen lassen sich dann (teils mit Hilfe des beschänkten Quantor $\forall i < n+2$) primitiv rekursiv ausdrücken, und schließlich ist $h(n+1, \bar{n}) = r_{n+1}$. □

Erinnerung: $\mathbb{N}^{<\omega}$ bezeichnet die Menge aller endlichen Folgen natürlicher Zahlen, von beliebiger Länge. Außerhalb der Mengenlehre ist dafür die Schreibweise $\mathbb{N}^*$ üblicher.

Satz 7.13
Es gibt eine Bijektion $\mathbb{N}^{<\omega} \to \mathbb{N}$, $(n_0, \ldots, n_{l-1}) \mapsto \langle n_0, \ldots, n_{l-1}\rangle$, deren Einschränkung auf jedes $\mathbb{N}^k$ primitiv rekursiv ist, und so dass

Längenfunktion	$\lg : \mathbb{N} \to \mathbb{N}$,	$\langle n_0, \ldots, n_{l-1}\rangle \mapsto l$
Komponentenfunktion	$\langle \ldots \rangle_\ldots : \mathbb{N}^2 \to \mathbb{N}$,	$(\langle n_0, \ldots, n_{l-1}\rangle, i) \mapsto \langle n_0, \ldots, n_{l-1}\rangle_i = n_i$

ebenfalls primitiv rekursiv sind. Dafür setzt man $\langle n_0, \ldots, n_{l-1}\rangle_i = 0$ für $i \geqslant l$.

Beweis: Man wählt (Erinnerung: p_i ist die $(i+1)$-te Primzahl):

$$(n_0, \ldots, n_{l-1}) \mapsto \langle n_0, \ldots, n_{l-1}\rangle := p_0^{n_0} \cdot \ldots \cdot p_{l-2}^{n_{l-2}} \cdot p_{l-1}^{n_{l-1}+1} - 1$$

Aufgrund der eindeutigen Primfaktorzerlegung ist dies eine Bijektion (die leere Folge wird auf 0 abgebildet; +1 bei der Multiplizität des größten Primzahl im Produkt erlaubt es, das letzte Folgenglied zu identifizieren, auch wenn es 0 ist).

Die Länge bekommt man als $\lg(s) = \mu k \leqslant s\,[\,p(k) \nmid s+1\,]$, die Komponenten als

$$\langle s\rangle_i = \begin{cases} \mu k \leqslant s\,[\,p(i)^{k+1} \nmid s+1\,] & \text{für } i < \lg(s)-1 \\ \mu k \leqslant s\,[\,p(i)^{k+2} \nmid s+1\,] & \text{für } i = \lg(s)-1 \\ 0 & \text{für } i \geqslant \lg(s) \end{cases}$$

Wegen der Beschränkung in der μ-Rekursion ist nach Lemma 7.10 alles primitiv rekursiv. □

Die Kodierung endlicher Folgen erlaubt folgende Variante der Primitiven Rekursion mit Rückgriff auf alle vorherigen Funktionswerte:

Lemma 7.14 *Wenn $f : \mathbb{N}^k \to \mathbb{N}$ und $g : \mathbb{N}^{k+2} \to \mathbb{N}$ primitiv rekursiv sind, dann auch die Funktion $h : \mathbb{N}^{k+1} \to \mathbb{N}$ mit*

$$\begin{aligned} h(0, \bar{m}) &:= f(\bar{m}) \\ h(n+1, \bar{m}) &:= g\big(\langle h(0, \bar{m}), \ldots, h(n, \bar{m})\rangle, n, \bar{m}\big) \end{aligned}$$

Beweis: Definiere

$$s(n, \bar{m}) = \big\langle h(0, \bar{m}), \ldots, h(n, \bar{m})\big\rangle$$

Dann ist $h(n, \bar{m}) = \langle s(n, \bar{m})\rangle_{n+1}$, weshalb es reicht zu zeigen, dass s primitiv rekursiv ist. Das sieht man mit der Primitiven Rekursion:

$$\begin{aligned} s(0, \bar{m}) &= \langle f(\bar{m})\rangle \\ s(n+1, \bar{m}) &= \big\langle h(0, \bar{m}), \ldots, h(n, \bar{m}), g(\langle h(0, \bar{m}), \ldots, h(n, \bar{m})\rangle, n, \bar{m})\big\rangle \\ &= \tfrac{s(n,\bar{m})+1}{p(n)} \cdot p(n+1)^{g(s(n,\bar{m}),n,\bar{m})+1} - 1 \end{aligned}$$

Die Division durch $p(n)$ bekommt man durch $\frac{1}{p(n)} y = \mu x \leqslant y\,[\,y \leqslant x \cdot p(n)\,]$. □

Gödelisierung

Sei nun $\mathscr{L}$ eine fest gewählte prädikatenlogische Sprache, die aus den endlich vielen (Relations- und Funktions-) Zeichen $\sigma_0, \dots, \sigma_{e-1}$ besteht. Ziel ist es, jeder $\mathscr{L}$-Formel φ so eine natürliche Zahl $\ulcorner \varphi \urcorner$ als Code zuzuordnen, dass alle naheliegenden Eigenschaften von $\mathscr{L}$-Formeln aus ihren Codes primitiv rekursiv berechenbar sind.

Dazu braucht man zunächst eine (willkürliche) Kodierungstafel, die jedem vorkommenden Zeichen einen Code zuweist, die sogenannte *Gödelnummer*. Zum Beispiel

$$\begin{array}{llllllll}
\ulcorner \top \urcorner &= 0 & \ulcorner \leftrightarrow \urcorner &= 6 & \ulcorner \sigma_0 \urcorner &= 12 & \ulcorner v_0 \urcorner &= e+12 \\
\ulcorner \bot \urcorner &= 1 & \ulcorner \forall \urcorner &= 7 & \ulcorner \sigma_1 \urcorner &= 13 & \ulcorner v_1 \urcorner &= e+13 \\
\ulcorner \neg \urcorner &= 2 & \ulcorner \exists \urcorner &= 8 & & \vdots & & \vdots \\
\ulcorner \wedge \urcorner &= 3 & \ulcorner \doteq \urcorner &= 9 & & \vdots & & \vdots \\
\ulcorner \vee \urcorner &= 4 & \ulcorner (\urcorner &= 10 & & \vdots & \ulcorner v_j \urcorner &= e+j+12 \\
\ulcorner \to \urcorner &= 5 & \ulcorner) \urcorner &= 11 & \ulcorner \sigma_{e-1} \urcorner &= e+11 & & \vdots
\end{array}$$

Jeder Zeichenfolge $a_1 \dots a_l$ über diesem Alphabet – insbesondere jedem $\mathscr{L}$-Term und jeder $\mathscr{L}$-Formel – wird nun ihre Gödelnummer $\ulcorner a_1 \dots a_l \urcorner \in \mathbb{N}$ zugewiessen durch

$$\ulcorner a_1 \dots a_l \urcorner = \big\langle \ulcorner a_1 \urcorner, \dots, \ulcorner a_l \urcorner \big\rangle$$

Wenn zum Beispiel $\mathscr{L}$ nur aus dem zweistelligen Relationssymbol R besteht, bekommt die $\mathscr{L}$-Formel $\forall v_0 \exists v_1 R v_0 v_1$ zugewiesen:

$$\begin{aligned}
\ulcorner \forall v_0 \exists v_1 R v_0 v_1 \urcorner &= \big\langle \ulcorner \forall \urcorner, \ulcorner v_0 \urcorner, \ulcorner \exists \urcorner, \ulcorner v_1 \urcorner, \ulcorner R \urcorner, \ulcorner v_0 \urcorner, \ulcorner v_1 \urcorner \big\rangle \\
&= \big\langle 7, 13, 8, 14, 12, 13, 14 \big\rangle \\
&= 2^7 \cdot 3^{13} \cdot 5^8 \cdot 7^{14} \cdot 11^{12} \cdot 13^{13} \cdot 17^{15} - 1
\end{aligned}$$

Aus der Definition heraus ist klar, dass die Gödelnummer einer echten Teilfolge von s kleiner ist als die Gödelnummer von s.

Lemma 7.15 *Die folgenden Mengen (und alle ähnlich definierten) sind primitiv rekursiv:*

$$\begin{aligned}
\mathcal{C}_{\text{Term}} &:= \big\{ \ulcorner \tau \urcorner \mid \tau \textit{ ist } \mathscr{L}\textit{-Term} \big\} \\
\mathcal{C}_{\text{Formel}} &:= \big\{ \ulcorner \varphi \urcorner \mid \varphi \textit{ ist } \mathscr{L}\textit{-Formel} \big\} \\
\mathcal{C}_{\text{Aussage}} &:= \big\{ \ulcorner \varphi \urcorner \mid \varphi \textit{ ist } \mathscr{L}\textit{-Aussage} \big\}
\end{aligned}$$

Beweis: Ich beweise nur das Beispiel $\mathcal{C}_{\text{Term}}$; der Rest geht ähnlich. Ohne Einschränkung seien $\sigma_0, \dots, \sigma_f$ die Funktionszeichen der Stelligkeit $s_i := s(\sigma_i) > 0$ in $\mathscr{L}$. Man drückt nun aus:

$$\chi_{\mathcal{C}_{\text{Term}}}(n) = 1 \iff \begin{cases}
\lg(n) = 1 \text{ und } \langle n \rangle_0 \in \big\{ \underbrace{\ulcorner \sigma_i \urcorner \mid \sigma_i \text{ Konstante}}_{\text{endliche Menge}} \big\} \cup \big\{ \underbrace{\ulcorner v_i \urcorner \mid v_i \text{ Variable}}_{\{k+e+12 \,\mid\, k \in \mathbb{N}\}} \big\} \\
\text{oder } \lg(n) > 1 \text{ und einer der folgenden Fälle für } i = 0, \dots, f: \\
\quad \langle n \rangle_0 = \ulcorner \sigma_i \urcorner \text{ und } \exists n_1 < n \dots \exists n_{s_i} < n \text{ mit} \\
\qquad \chi_{\mathcal{C}_{\text{Term}}}(n_1) = 1 \ \dots \ \chi_{\mathcal{C}_{\text{Term}}}(n_{s_i}) = 1 \text{ und} \\
\qquad \langle n_1 \rangle_0 = \langle n \rangle_1 \quad \dots \ \langle n_1 \rangle_{\lg(n_1)-1} = \langle n \rangle_{\lg(n_1)} \\
\qquad \langle n_2 \rangle_0 = \langle n \rangle_{\lg(n_1)+1} \ \dots \ \langle n_2 \rangle_{\lg(n_2)-1} = \langle n \rangle_{\lg(n_1)+\lg(n_2)} \\
\qquad \qquad \vdots \\
\qquad \dots \ \langle n_{s_i} \rangle_{\lg(n_{s_i}-1)} = \langle n \rangle_{\lg(n)-1}
\end{cases}$$

Das kann man (mit etwas Mühe) primitiv rekursiv ausdrücken; der mehrfache Rückgriff auf $\chi_{\mathcal{C}_{\text{Term}}}(n_1), \ldots, \chi_{\mathcal{C}_{\text{Term}}}(n_{s_i})$ in der Rekursion ist durch Lemma 7.14 gerechtfertigt. Man beachte, dass für jedes der endlich vielen Funktionszeichen σ_i die Stelligkeit s_i fest ist, d. h. darüber wird nicht quantifiziert! □

Für nicht unbedingt endliche, aber *rekursive Sprachen* $\mathscr{L}$ bekommt man analog, dass die Mengen $\mathcal{C}_{\text{Term}}$, $\mathcal{C}_{\text{Formel}}$ etc. rekursiv sind. Dabei heißt eine Sprache rekursiv, wenn die Stelligkeitsfunktion auf den Codes der Relations- und Funktionszeichen, also $\ulcorner\sigma\urcorner \mapsto s(\sigma)$, rekursiv ist.

Definition 7.16 *Eine $\mathscr{L}$-Theorie heißt*

- rekursiv axiomatisierbar*, falls* $\{\ulcorner\varphi\urcorner \mid \varphi \in T\}$ *rekursiv aufzählbar ist,*
- entscheidbar*, falls* $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ *rekursiv ist.*
- semi-entscheidbar [19]*, falls* $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ *rekursiv aufzählbar ist.*

Intuitiv ist eine $\mathscr{L}$-Theorie entscheidbar, wenn es einen Algorithmus gibt, der nach Eingabe einer beliebigen $\mathscr{L}$-Formel φ in endlicher Zeit die Frage beantwortet, ob φ aus T folgt oder nicht. Die Terminologie „rekursiv axiomatisierbar“ kommt daher, dass die Elemente von T als Axiome des *deduktiven Abschlusses* von T, d. h. der Theorie $T^{\vDash} = \{\varphi \mid T \vDash \varphi\}$, aufgefasst werden können. In diesem Sinne ist für eine rekursiv axiomatisierbar Theorie zwar nicht entscheidbar, ob eine $\mathscr{L}$-Formel φ ein Axiom ist oder nicht, aber man kann die Axiome zumindest algorithmisch nach und nach aufzählen und beispielsweise in einen Entscheidungsalgorithmus einspeisen.

Satz 7.17 *Wenn T rekursiv axiomatisierbar ist, dann ist T semi-entscheidbar.*

Beweis: Man beweist den Satz, indem man für den im Beweis des Vollständigkeitssatzes benutzten Kalkül $\mathbb{K}$ zeigt, dass die Menge

$$\left\{\left(\ulcorner\varphi\urcorner, \ulcorner\varphi_0 \ldots, \varphi_n\urcorner\right) \;\middle|\; \varphi_0 \ldots \varphi_n \text{ ist ein } \mathbb{K}\text{-Beweis von } \varphi = \varphi_n\right\} \subseteq \mathbb{N}^2$$

rekursiv aufzählbar ist. Der Vollständigkeitssatz sagt nun aus, dass die (dann ebenfalls rekursiv aufzählbare) Projektion dieser Menge auf die erste Koordinate genau die Menge $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ ist.

⋮

□

Folgerung 7.18 *Wenn T rekursiv axiomatisierbar und vollständig ist, dann ist T entscheidbar.*

Beweis: Es reicht zu zeigen, dass das Komplement von $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ ebenfalls rekursiv aufzählbar ist. Es besteht aus der Vereinigung der rekursiven Menge $\{n \mid n \notin \mathcal{C}_{\text{Aussage}}\}$ mit $A := \{\ulcorner\varphi\urcorner \mid T \nvDash \varphi\}$. Wegen der Vollständigkeit ist $A = \{\ulcorner\varphi\urcorner \mid T \vDash \neg\varphi\}$ und ist das Bild der rekursiv aufzählbaren Menge $\{\ulcorner\psi\urcorner \mid T \vDash \psi \text{ und } \langle\ulcorner\psi\urcorner\rangle_0 = \ulcorner\neg\urcorner\}$ unter der leicht als rekursiv nachzuweisenden Funktion $n \mapsto \begin{cases} \ulcorner\varphi\urcorner & \text{falls } n = \ulcorner\neg\varphi\urcorner \\ 0 & \text{sonst} \end{cases}$. □

Diese Folgerung ist eine wichtige Möglichkeit, um die Entscheidbarkeit von $\mathscr{L}$-Theorien nachzuweisen. Für eine gegebene, rekursiv axiomatisierbare $\mathscr{L}$-Theorie muss man für ihre Anwendung die Vollständigkeit nachweisen: In der Modelltheorie lernt man dafür Techniken kennen.

[19] Dies ist, im Gegensatz zu den beiden anderen Definitionen, für Theorien keine Standard-Terminologie.

Umgekehrt ist die Theorie einer $\mathscr{L}$-Struktur $\mathscr{M}$, das ist $\mathrm{Th}(\mathscr{M}) = \{\varphi \mid \mathscr{M} \vDash \varphi\}$, stets vollständig, und Entscheidbarkeit bekommt man, wenn man eine rekursiv aufzählbare Axiomatisierung findet.

Wichtige Beispiele entscheidbarer Theorien bzw. in diesem Sinne rekursiv axiomatisierbarer Strukturen sind die $\{+,\cdot,-,0,1\}$-Theorien algebraisch abgeschlossen Körper fester Charakteristik 0 oder p, mit Modellen $\mathbb{C}$ bzw. $\widetilde{\mathbb{F}_p}$, und die $\{+,\cdot,-,0,1,<\}$-Theorie der reell abgeschlossenen Körper mit Modell $\mathbb{R}$.

Die Struktur $(\mathbb{N};+,\cdot,<)$ lässt dagegen keine rekursive Axiomatisierung ihrer Theorie zu bzw. jeder rekursiv axiomatisierte Teil der Arithmetik von $+$ und $\cdot$ auf den natürlichen Zahlen ist unvollständig (wie noch gezeigt wird).

Universelle rekursiven Funktionen durch Gödelisierung

Man kann den Aufbau von rekursiven Funktionen f gemäß der Regeln aus Definition 7.1 durch eine endliche Symbolfolge $\hat{f}$ – eine Art verallgemeinerter Term – beschreiben. Ich gebe hier eine von vielen Möglichkeit an, dies zu tun:

Die Symbolfolge beginnt mit einem der Symbole $\alpha, \kappa, \varrho, \mu$, um anzuzeigen, ob die Funktion als Ausgangsfunktion, durch Komposition, Primitive Rekursion oder μ-Rekursion gegeben ist. Es folgt die Stelligkeit der Funktion und dann die Terme der Funktionen, aus denen zusammengesetzt wird:

- Für die Ausgangsfunktionen schreibt man $\alpha\, 0\, c_0^0$, $\alpha\, 1\, S$ bzw. $\alpha\, k\, \pi_i^k$.
- Für die Komposition der Funktion $g : \mathbb{N}^l \to \mathbb{N}$ mit den Funktionen $f_1, \dots, f_l : \mathbb{N}^k \to \mathbb{N}$ schreibt man $\kappa\, k\, \hat{g}\, \hat{f}_1 \dots \hat{f}_l$.
- Für die Primitive Rekursion aus $f : \mathbb{N}^k \to \mathbb{N}$ und $g : \mathbb{N}^{k+2} \to \mathbb{N}$ (wie in Definition 7.1) schreibt man $\varrho\, k+1\, \hat{f}\, \hat{g}$.
- Für die μ-Rekursion aus $f : \mathbb{N}^{k+1} \to \mathbb{N}$ (wie in Definition 7.1) schreibt man $\mu\, k\, \hat{f}$.

Der Nachweis der primitiven Rekursivität der Addition $+ : \mathbb{N}^2 \to \mathbb{N}$ im Beweis von Lemma 7.2 würde zum Beispiel durch die Symbolfolge $\varrho\, 2 (\alpha\, 1\, \pi_1^1)(\kappa\, 3 (\alpha\, 1\, S)(\alpha\, 3\, \pi_1^2))$ beschrieben sein (wobei ich der besseren Lesbarkeit willen Klammern eingefügt habe).

Durch eine geeigneten Gödelisierung der vorkommenden Symbole $\alpha, \kappa, \varrho, \mu, c_0^0, S, \pi_1^k$ und aller möglichen Stelligkeiten [20] kann man so jedem „Aufbaurezept“ einer rekursiven Funktion f eine Gödelnummer zuweisen, die ich der Einfachheit halber $\ulcorner f \urcorner$ schreibe.

Wie in Lemma 7.15 sieht man nun, dass $\mathcal{C}_{\text{prim}} := \{\ulcorner f \urcorner \mid f \text{ primitiv rekursiv}\} \subseteq \mathbb{N}$ primitiv rekursiv ist. Dann ist sicher auch die Einschränkung auf einstellige Funktionen $\mathcal{C}^1_{\text{prim}} := \{\ulcorner f \urcorner \mid f : \mathbb{N} \to \mathbb{N} \text{ primitiv rekursiv}\}$ primitiv rekursiv. Dies allein besagt nicht viel, aber man kann auch zeigen, dass die Auswertungsfunktion

$$\mathrm{eval} : \mathbb{N}^2 \to \mathbb{N},\ (n,m) \mapsto \begin{cases} f(m) & n = \ulcorner f \urcorner \in \mathcal{C}^1_{\text{prim}} \\ 0 & \text{sonst} \end{cases}$$

rekursiv ist. Dazu muss man die Auswertung der Funktion f ihrem Aufbauprozess entlang rekursiv nachbauen. Bei einem Schritt Primitiver Rekursion muss man wieder das Tupel der aufeinander folgenden Funktionswerte $h(0), h(1), \dots, h(m)$ kodieren. Dafür hat man keine Schranke,

[20] Zum Beispiel könnte man für die Stellgikeiten $\ulcorner n \urcorner = 2n$ wählen, dann $\ulcorner \alpha \urcorner = 1$, $\ulcorner \kappa \urcorner = 3$, $\ulcorner \varrho \urcorner = 5$, $\ulcorner c_0^0 \urcorner = 7$, $\ulcorner S \urcorner = 9$ und schließlich $\ulcorner \pi_i^k \urcorner = k(k-1) + 2i + 9$.

da die Kodierung dieser Funktionswerte mit der Kodierung des Aufbauprozesses der Funktion nichts zu tun hat. Daher kann man nicht zeigen, dass eval sogar primitiv rekursiv ist. Im Gegenteil:

Satz 7.19 *Die Funktion* $\delta : n \mapsto \mathrm{eval}(n,n)+1$ *ist rekursiv, aber nicht primitiv rekursiv (was sie wäre, wenn* eval *primitiv rekursiv wäre).*

Beweis: Angenommen δ wäre primitiv rekursiv, dann gäbe es $d = \ulcorner \delta \urcorner \in \mathcal{C}^1_{\mathrm{prim}}$. Einerseits wäre dann $\delta(d) = \mathrm{eval}(d,d)+1$ nach Definition von δ, andererseits $\mathrm{eval}(d,d) = \delta(d)$ nach Definition von eval: Widerspruch! □

Dieses Argument zeigt, dass es keine *universelle primitiv rekursive Funktion* $U_{\mathrm{prim}} : \mathbb{N}^2 \to \mathbb{N}$ gibt. Das wäre eine primitiv rekursive Funktion mit der Eigenschaft, dass die durch U_{prim} definierte Funktionenschar $\{U_{\mathrm{prim}}(\cdot,m) : \mathbb{N} \to \mathbb{N}, n \mapsto U_{\mathrm{prim}}(n,m) \mid m \in \mathbb{N}\}$ genau die Menge der primitiv rekursiven Funktionen $\mathbb{N} \to \mathbb{N}$ ist. (Wie gezeigt, gibt es aber eine rekursive Funktion U mit dieser Eigenschaft.)

Das gleiche Diagonalargument zeigt, dass es keine *universelle (totale) rekursive Funktion* $U_{\mathrm{rek}} : \mathbb{N}^2 \to \mathbb{N}$ gibt. Das wäre eine rekursive Funktion mit der Eigenschaft, dass die durch U_{rek} definierte Funktionenschar $\{U_{\mathrm{rek}}(\cdot,m) : \mathbb{N} \to \mathbb{N}, n \mapsto U_{\mathrm{rek}}(n,m) \mid m \in \mathbb{N}\}$ genau die Menge der rekursiven Funktionen $\mathbb{N} \to \mathbb{N}$ ist. Die Menge der Gödelnummern der „Aufbaurezepte“ rekursiver Funktionen ist zwar ebenfalls (primitiv) rekursiv, allerdings kann man nicht rekursiv prüfen, dass der Normalfall vorliegt. Die Menge $\mathcal{C}_{\mathrm{part}} := \{\ulcorner f \urcorner \mid f \text{ partielle rekursive Funtion}\}$ ist also (primitiv) rekursiv. Die Menge $\mathcal{C}_{\mathrm{rek}} := \{\ulcorner f \urcorner \mid f \text{ totale rekursive Funtion}\}$ kann dagegen nicht rekursiv sein, weil sich sonst über das Diagonalargument ein Widerspruch ergäbe. Für partielle Funktionen entsteht der Widerspruch dagegen nicht, weil im entsprechenden Argument wie oben $\delta(d)$ unbestimmt sein kann (und dann auch sein muss).

Eine konkrete rekursive, aber nicht primitiv rekursive Funktion erhält man aus der Ackermann-Funktion vom Übungblatt 9, nämlich die Funktion $n \mapsto A(n,n)$. Dazu zeigt man, dass jede primitiv rekursive Funktion durch ein A_n beschränkt ist (was $n \mapsto A(n,n)$ nicht ist), und natürlich, dass A rekursiv ist.

7.3 Arithmetik

Mit „Arithmetik“ ist ursprünglich das Rechnen mit natürlichen Zahlen bezeichnet. Im Kontext dieser Vorlesung ist die erststufige Theorie der natürlichen Zahlen $\mathbb{N}$ in einer geeigneten Sprache gemeint. Wir betrachten die (endliche) Sprache $\mathscr{L}_{\mathrm{Ar}} = \{+, \cdot, s, \underline{0}, <\}$ und die $\mathscr{L}_{\mathrm{Ar}}$-Struktur $\mathscr{N} = (\mathbb{N}; +^{\mathbb{N}}, \cdot^{\mathbb{N}}, S, 0, <^{\mathbb{N}})$, wobei die Stelligkeiten und Interpretationen für $+, \cdot, \underline{0}, <$ so sind, wie es die gewählten Symbole nahelegen, und s ein einstelliges Funktionszeichen ist, dessen Interpretation die Nachfolgerfunktion $S : n \mapsto n+1$ ist. Das Konstantenzeichen für 0 schreibe ich nun $\underline{0}$, und induktiv definiert man für jedes $n \in \mathbb{N}$ den $\mathscr{L}_{\mathrm{Ar}}$-Term $\underline{n}$ durch $\underline{n+1} := s\underline{n}$.

Definition 7.20

(a) Robinsons Arithmetik Q *ist folgende* $\mathscr{L}_{\mathrm{Ar}}$*-Theorie:*

$$\begin{array}{ll}
\forall v_0\ v_0 + \underline{0} \doteq v_0 & \forall v_0 \forall v_1 \quad v_0 + sv_1 \doteq s(v_0 + v_1) \\
\forall v_0 \quad v_0 \cdot \underline{0} \doteq \underline{0} & \forall v_0 \forall v_1 \quad v_0 \cdot sv_1 \doteq (v_0 \cdot v_1) + v_0 \\
\forall v_0 \quad \neg v_0 < \underline{0} & \forall v_0 \forall v_1 \ \big(v_0 < sv_1 \leftrightarrow (v_0 < v_1 \vee v_0 \doteq v_1)\big)
\end{array}$$

(b) *Die (erststufige)* Peano-Arithmetik PA *ist die* $\mathscr{L}_{\mathrm{Ar}}$*-Theorie, die sich aus* Q *und dem erststufigen Induktionsschema zusammensetzt, d. h. für jede* $\mathscr{L}_{\mathrm{Ar}}$*-Formel* $\varphi(v_0, v_1, \dots, v_k)$ *ist die folgende* $\mathscr{L}_{\mathrm{Ar}}$*-Aussage ein Axiom von* PA*:* [21]

$$\forall v_1 \dots \forall v_k \left(\left(\varphi[\tfrac{\underline{0}}{v_0}] \wedge \forall v_0 \left(\varphi \to \varphi[\tfrac{sv_0}{v_0}] \right) \right) \to \forall v_0 \, \varphi \right)$$

(c) *Die* volle Arithmetik *ist die* $\mathscr{L}_{\mathrm{Ar}}$*-Theorie* $\mathrm{Th}(\mathscr{N}) = \{\varphi \mid \mathscr{N} \vDash \varphi\}$.

Per Definition ist Q eine endliche und die volle Arithmetik eine vollständige $\mathscr{L}_{\mathrm{Ar}}$-Theorie. Man sieht (ähnlich wie für die $\forall$-Axiome im Beweis von Satz 7.17), dass PA rekursiv axiomatisierbar ist. Klar ist zudem $\mathsf{Q} \subseteq \mathsf{PA} \subseteq \mathrm{Th}(\mathscr{N})$. Insbesondere ist $\mathscr{N} \vDash \mathsf{Q}$ und $\mathscr{N} \vDash \mathsf{PA}$.

Das ursprüngliche von Peano beschriebene Axiomensystem PA_2 für die Arithmetik ist zweitstufig und beinhaltet neben Q das volle Induktionsprinzip

$$\forall V_0^1 \left(\left(V_0^1 \underline{0} \wedge \forall v_0 \left(V_0^1 v_0 \to V_0^1 s v_0 \right) \right) \to \forall v_0 \, V_0^1 v_0 \right)$$

Für jedes Modell $\mathscr{M}$ von PA_2 ist $\iota : n \mapsto \underline{n}^{\mathscr{M}}$ ein Isomorphismus $\mathscr{N} \cong \mathscr{M}$; das Induktionsprinzip angewandt auf $\mathrm{im}(\iota)$ sichert die Surjektivität. Dagegen kann es in Modellen $\mathscr{M}$ von Q, PA oder $\mathrm{Th}(\mathscr{N})$ *Nicht-Standard-Zahlen* geben, also Elemente, die nicht von der Form $\underline{n}^{\mathscr{M}}$ für $n \in \mathbb{N}$ sind. Für solche Elemente m gilt stets $m >^{\mathscr{M}} \underline{n}^{\mathscr{M}}$ für alle $n \in \mathbb{N}$. Daher nennt man sie auch *unendlich große* Zahlen.

Lemma 7.21 *Aus* Q *folgen für alle* $n, m \in \mathbb{N}$ *die* $\mathscr{L}_{\mathrm{Ar}}$*-Aussagen:*

$$\begin{aligned} \underline{n} + \underline{m} &\doteq \underline{n+m} \qquad \text{und folglich} \quad \underline{n} + \underline{m} \doteq \underline{m} + \underline{n} \\ \underline{n} \cdot \underline{m} &\doteq \underline{n \cdot m} \qquad \text{und folglich} \quad \underline{n} \cdot \underline{m} \doteq \underline{m} \cdot \underline{n} \\ \forall v_0 \left(v_0 < \underline{n} \leftrightarrow (v_0 &\doteq \underline{0} \vee \dots \vee v_0 \doteq \underline{n-1}) \right) \end{aligned}$$

Beweis: Beispielhaft beweise ich die erste Aussage, per („Meta-“)Induktion über m:
$\underline{n} + \underline{0} \doteq \underline{n}$ ist in Q und es gilt $\underline{n} = \underline{n+0}$. Dann ist zu zeigen, dass Q $\underline{n} + \underline{m+1} \doteq \underline{(n+m)+1}$ beweist, also $\underline{n} + s\underline{m} \doteq s(\underline{n+m})$. Nach Induktion folgt $s(\underline{n} + \underline{m}) \doteq s(\underline{n+m})$ aus Q, und $\underline{n} + s\underline{m} \doteq s(\underline{n} + \underline{m})$ ist wieder ein Axiom von Q. Der Rest geht ähnlich. □

Die Menge der in dem Lemma aufgeführten Aussagen wird mit Q^* bezeichnet. Q^* ist eine unendliche, aber rekursiv axiomatisierbare $\mathscr{L}_{\mathrm{Ar}}$-Theorie. Im Prinzip besteht $\mathsf{Q}*$ aus der Additions- und Multiplikationstafel von $\mathbb{N}$ und der „Relationstafel“ für die Kleiner-Relation.

Lemma 7.22 *Aus* PA *folgen die* $\mathscr{L}_{\mathrm{Ar}}$*-Aussagen, die besagen:*

- $+$ *ist assoziativ und kommutativ mit neutralem Element* $\underline{0}$.
- $\cdot$ *ist assoziativ und kommutativ mit neutralem Element* $\underline{1}$*, und distributiv über* $+$.
- $<$ *ist eine diskrete* [22] *lineare Ordnung mit kleinstem Element* $\underline{0}$.
- $+$ *ist monoton bzgl.* $<$*, ebenso* $\cdot$ *mit Elementen* $\neq \underline{0}$.
- s *ist injektiv und gibt den unmittelbaren Nachfolger bzgl.* $<$ *an.*

[21] Es reicht, das Induktionsschema für $\mathscr{L}_{\mathrm{Ar}}$-Formeln „ohne Parameter“, d. h. für $\mathscr{L}_{\mathrm{Ar}}$-Formeln $\varphi(v_0)$ mit $k = 0$ zu fordern. Die insgesamt dazu äquivalente Version „mit Parametern“ ist aber für Beweisführungen praktischer.

[22] „Diskret“ heißt, dass jedes Element außer – sofern sie existieren – dem kleinsten und dem größten einen unmittelbaren Vorgänger und einen unmittelbaren Nachfolger hat.

Beweis: Beispielhaft beweise ich die Kommutativität von $+$, im Gegensatz zu Q per „objektsprachlicher“ Induktion, d. h. $\mathsf{PA} \vDash \forall v_0 \forall v_1\, (v_0 + v_1 \doteq v_1 + v_0)$ folgt aus dem Induktionsschema, das zu den Axiomen von PA gehört. Wegen der doppelten Allquantifikation muss man dabei (1) Induktionsanfang und (2) Induktionsschritt ebenfalls mit Hilfe des Induktionsschemas zeigen.

(1) Induktionsanfang für $\mathsf{PA} \vDash \forall v_1\, (\underline{0} + v_1 \doteq v_1 + \underline{0})$ ist das Gleichheitsaxiom $\underline{0} + \underline{0} \doteq \underline{0} + \underline{0}$. Wenn $\mathsf{PA} \vDash \underline{0} + v_1 \doteq v_1 + \underline{0}$, folgt damit (in PA) aus dem Axiom $\underline{0} + sv_1 \doteq s(\underline{0} + v_1)$, zunächst $\underline{0} + sv_1 \doteq s(v_1 + \underline{0})$, und daraus mit zweifacher Axiom-Anwendung $\underline{0} + sv_1 \doteq s(v_1) + \underline{0}$.

(2) Als nächstes zu zeigen ist $\mathsf{PA} \vDash \forall v_0 \big(\forall v_1\, v_0 + v_1 \doteq v_1 + v_0 \to \forall v_1\, sv_0 + v_1 \doteq v_1 + sv_0\big)$. Für beliebiges v_0 ist also in PA unter der Annahme $\forall v_1\, v_0 + v_1 \doteq v_1 + v_0$ zu zeigen, dass $(*)$ $\forall v_1\, sv_0 + v_1 \doteq v_1 + sv_0$. Ich arbeite dazu in einem beliebigen Modell $\mathscr{M}$ von PA mit $0 = \underline{0}^{\mathscr{M}}$ und $S = s^{\mathscr{M}}$ und schreibe der Einfachheit halber v_0, v_1 für beliebige Elemente des Modells.

Induktionsanfang für $(*)$ ist $S(v_0) + 0 \overset{\text{Ax.}}{=} S(v_0) \overset{\text{Ax.}}{=} S(v_0 + 0) \overset{\text{Annahme}}{=} S(0 + v_0) \overset{\text{Ax.}}{=} 0 + S(v_0)$.

Induktionsschritt für $(*)$: Induktionsvoraussetzung ist $S(v_0) + v_1 = v_1 + S(v_0)$. Dann:
$S(v_0) + S(v_1) \overset{\text{Ax.}}{=} S(S(v_0) + v_1) \overset{\text{IV}}{=} S(v_1 + S(v_0)) \overset{\text{Ax.}}{=} S(S(v_1 + v_0)) \overset{\text{Ann.}}{=} S(S(v_0 + v_1)) \overset{\text{Ax.}}{=} S(v_0 + S(v_1)) \overset{\text{Ann.}}{=} S(S(v_1) + v_0) \overset{\text{Ax.}}{=} S(v_1) + S(v_0)$. □

Definition 7.23 *Eine Teilmenge $A \subseteq \mathbb{N}^k$ heißt* arithmetisch*, wenn sie in $\mathscr{N}$ definierbar ist, d. h. wenn es eine $\mathscr{L}_{\text{Ar}}$-Formel $\varphi(v_1, \ldots, v_k)$ gibt, so dass (für beliebige Belegung β)*

$$(a_1, \ldots, a_k) \in A \iff (\mathscr{N}, \beta \tfrac{a_1}{v_1} \ldots \tfrac{a_k}{v_k}) \vDash \varphi$$

Eine Funktion $f : \mathbb{N}^k \to \mathbb{N}$ heißt arithmetisch*, wenn ihr Graph $\Gamma_f \subseteq \mathbb{N}^{k+1}$ arithmetisch ist.*

Statt $(\mathscr{N}, \beta \tfrac{a_1}{v_1} \ldots \tfrac{a_k}{v_k}) \vDash \varphi$ schreibt man auch kurz $\mathscr{N} \vDash \varphi[a_1, \ldots, a_k]$, da β keine Rolle spielt.

Satz 7.24 *Alle rekursiven und rekursiv aufzählbaren Mengen und alle rekursiven Funktionen sind arithmetisch.*

Beweis: Hierfür wurde Satz 7.12 bewiesen: Für Funktionen reicht zu zeigen, dass die erweiterten Grundfunktionen arithmetisch sind und sich [Komposition] und [μ-Rekursion] mit $\mathscr{L}_{\text{Ar}}$-Formel ausdrücken lassen. Für den Graph von Funktionen nehme ich jetzt stets v_0 als Variable für das Bild (etwas kontraintuitiv, dafür aber uniform).

Für die erweiterten Grundfunktionen hat man folgende Definitionen:

c_0^k durch $v_0 \doteq \underline{0} \wedge \bigwedge_{j=1}^k v_j \doteq v_j$ $\quad$ S durch $v_0 \doteq sv_1$ $\quad$ π_i^k durch $v_0 \doteq v_i \wedge \bigwedge_{j=1}^k v_j \doteq v_j$
Addition durch $v_0 \doteq v_1 + v_2$ $\quad$ Multiplikation durch $v_0 \doteq v_1 \cdot v_2$
$\chi_<$ durch $((v_1 < v_2 \to v_0 \doteq s\underline{0}) \wedge (\neg v_1 < v_2 \to v_0 \doteq \underline{0}))$

Wenn $f_1, \ldots, f_l : \mathbb{N}^k \to \mathbb{N}$ durch die $\mathscr{L}_{\text{Ar}}$-Formeln $\varphi_1, \ldots, \varphi_k$ definiert werden und $g : \mathbb{N}^l \to \mathbb{N}$ durch ψ, dann wird die Komposition von g mit $f_1, \ldots, f_l$ definiert durch

$$\exists y_1 \ldots \exists y_l \left(\varphi_1 \left[\tfrac{y_1}{v_0}\right] \wedge \cdots \wedge \varphi_l \left[\tfrac{y_l}{v_0}\right] \wedge \psi \left[\tfrac{y_1}{v_1} \ldots \tfrac{y_l}{v_l}\right]\right)$$

oder auch durch

$$\forall y_1 \ldots \forall y_l \left((\varphi_1 \left[\tfrac{y_1}{v_0}\right] \wedge \cdots \wedge \varphi_l \left[\tfrac{y_l}{v_0}\right]) \to \psi \left[\tfrac{y_1}{v_1}, \ldots, \tfrac{y_l}{v_l}\right]\right)$$

Wenn $f : \mathbb{N}^{k+1} \to \mathbb{N}$ durch $\varphi(v_0, \ldots, v_{k+1})$ definiert ist, dann wird die μ-Rekursion von f („das kleinste v_1, so dass …“) definiert durch

$$\left(\varphi\left[\tfrac{0}{v_0}\right] \wedge \forall y < v_1\, \neg\varphi\left[\tfrac{0}{v_0}, \tfrac{y}{v_1}\right]\right) \left[\tfrac{v_0}{v_1}\right]$$

Eine rekursive Teilmenge $A \subseteq \mathbb{N}^k$ ist nun arithmetisch, weil sie definiert ist durch $\chi_A[\frac{s\underline{0}}{v_0}]$. Schließlich ist eine rekursiv aufzählbare Menge $A \subseteq \mathbb{N}^k$ Projektion einer rekursiven Menge $R \subseteq \mathbb{N}^{k+1}$: Wenn R durch φ definiert ist, wird A durch $\exists v_{k+1}\, \varphi(v_1, \ldots, v_k, v_{k+1})$ definiert. □

Folgerung 7.25 *Es gibt (für jedes $k > 0$) Teilmengen von $\mathbb{N}^k$, die nicht rekursiv aufzählbar sind, und Funktionen $\mathbb{N}^k \to \mathbb{N}$, die nicht rekursiv sind.*

Beweis: Es gibt nur abzählbar viele $\mathscr{L}_{\mathrm{Ar}}$-Formeln, aber überabzählbar viele Teilmengen von $\mathbb{N}^k$ und Funktionen $\mathbb{N}^k \to \mathbb{N}$. □

8 Unvollständigkeit und Unentscheidbarkeit

Satz 8.1 (Unentscheidbarkeit der Arithmetik, Gödel 1931)
Die $\mathscr{L}_{\mathrm{Ar}}$-Theorie $\mathrm{Th}(\mathscr{N})$ ist unentscheidbar.
Jede rekursiv aufzählbare Teiltheorie von $\mathrm{Th}(\mathscr{N})$ ist als $\mathscr{L}_{\mathrm{Ar}}$-Theorie unvollständig.

Die Unentscheidbarkeit wird im Stile der Russell-Zermelo'schen Antinomie gezeigt, wobei man über die Definierbarkeit der rekursiven Funktionen gehen muss, um das Diagonalargument anwenden zu können.

Beweis: Per Induktion über den Aufbau von $\mathscr{L}_{\mathrm{Ar}}$-Formeln zeigt man zunächst, dass die Auswertungsfunktion

$$\mathrm{eval} : \mathbb{N}^2 \to \mathbb{N}, \quad (m,n) \mapsto \begin{cases} \ulcorner \varphi[\frac{n}{v_0}] \urcorner & \text{falls } m = \ulcorner \varphi(v_0) \urcorner \\ 0 & \text{sonst} \end{cases}$$

primitiv rekursiv ist. Analog zu Lemma 7.15 sieht man, dass

$$\mathcal{C}_{\mathrm{Formel}_1} = \left\{ \ulcorner \varphi \urcorner \mid \varphi = \varphi(v_0) \text{ ist } \mathscr{L}\text{-Formel in einer freien Variablen } v_0 \right\}$$

eine (primitiv) rekursive Menge ist. Nach Satz 7.7 gibt es also eine surjektive rekursive Aufzählung $g : \mathbb{N} \to \mathcal{C}_{\mathrm{Formel}_1}$. Sei $\varphi_n(v_0)$ so, dass $g(n) = \ulcorner \varphi_n \urcorner$. Da die (durch 0 fortgesetzte) Abbildung $\ulcorner \neg\psi \urcorner \mapsto \ulcorner \psi \urcorner$ rekursiv ist, ist auch $\{\ulcorner \neg\varphi_n \urcorner \mid n \in \mathbb{N}\}$ als Urbild einer rekursiven Menge rekursiv.

Angenommen $\mathrm{Th}(\mathscr{N})$ ist entscheidbar, also $\{\ulcorner \psi \urcorner \mid \mathscr{N} \vDash \psi\}$ rekursiv. Wegen der Rekursivität von eval ist dann auch $R := \{n \in \mathbb{N} \mid \mathscr{N} \vDash \neg\varphi_n[\frac{n}{v_0}]\}$ rekursiv und nach Satz 7.24 definierbar in $\mathscr{N}$. Es gibt also eine $\mathscr{L}_{\mathrm{Ar}}$-Formel $\varphi(v_0)$, die R in $\mathscr{N}$ definiert.

Sei nun $n_0 := g^{-1}(\ulcorner \varphi \urcorner)$, d. h. $\varphi = \varphi_{n_0}$. Dann

$$\begin{aligned} n_0 \in R &\iff \mathscr{N} \vDash \varphi_{n_0}[n_0] && \text{da } \varphi_{n_0} \text{ die Menge } R \text{ definiert} \\ &\iff \mathscr{N} \vDash \varphi_{n_0}[\tfrac{n_0}{v_0}] && \text{da } \underline{n_0} \text{ in } \mathscr{N} \text{ durch } n_0 \text{ interpretiert wird} \\ &\iff \mathscr{N} \nvDash \neg\varphi_{n_0}[\tfrac{n_0}{v_0}] && \text{da } \mathrm{Th}(\mathscr{N}) \text{ vollständig} \\ &\iff n_0 \notin R && \text{nach Definition von } R \end{aligned}$$

und Widerspruch! Eine vollständige rekursiv aufzählbare Teiltheorie T von $\mathrm{Th}(\mathscr{N})$ wäre entscheidbar, ihr deduktiver Abschluss wäre also verschieden von $\mathrm{Th}(\mathscr{N})$, hätte aber $\mathscr{N}$ als Modell: Widerspruch! □

Die Arithmetik hat aber sehr wohl rekursiv aufzählbare, vollständige Teiltheorien in einer kleineren Sprache, z. B. die *Presburger-Arithmetik*, die die $\{+, 0, 1\}$-Theorie von $\mathbb{N}$ axiomatisiert.

Aus der Unentscheidbarkeit der Arithmetik folgt auch die Unentscheidbarkeit von Theorien, in denen man $\mathbb{N}$ wiederfindet, insbesondere ZFC, aber auch $\mathrm{Th}(\mathbb{Z}; +^{\mathbb{N}}, \cdot^{\mathbb{N}})$ und $\mathrm{Th}(\mathbb{Q}; +^{\mathbb{N}}, \cdot^{\mathbb{N}})$, weil man in beiden Strukturen $\mathbb{N}$ definieren kann: In $\mathbb{Z}$ sieht man das mit dem Vier-Quadrate-Satz von Lagrange, da jede natürliche Zahl als Summe von vier Quadraten geschrieben werden kann. Dass man $\mathbb{N}$ auch in $\mathbb{Q}$ definieren kann, hat Julia Robinson 1948 bewiesen mit Hilfe der

Theorie quadratischer Formen. Dagegen hat Tarski schon 1930 gezeigt, dass $\mathrm{Th}(\mathbb{R}; +^{\mathbb{N}}, \cdot^{\mathbb{N}})$ und $\mathrm{Th}(\mathbb{C}; +^{\mathbb{N}}, \cdot^{\mathbb{N}})$ entscheidbar sind.

Das nächste Ziel ist der Beweis von:

Satz 8.2 (Unentscheidbarkeit von Robinsons Arithmetik, R. Robinson 1950)
Jede konsistente $\mathscr{L}_{\mathrm{Ar}}$-Theorie, die Q^* impliziert – insbesondere also Q – ist unentscheidbar.

Folgerung 8.3 *Für jede rekursiv aufzählbare Teiltheorie T von $\mathrm{Th}(\mathscr{N})$ – wie z. B. Q^*, Q und PA – ist $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ eine rekursiv aufzählbare, nicht rekursive Menge. Insbesondere gibt es also solche Mengen!*

Beweis: Folgt sofort aus Satz 7.17 und Satz 8.2. □

Da Q endlich ist, kann man alle Axiome von Q durch Konjunktion in einer $\mathscr{L}_{\mathrm{Ar}}$-Formel q zusammenfassen. Dies hat eine Reihe von Konsequenzen:

Folgerung 8.4 *Jede Teiltheorie von $\mathrm{Th}(\mathscr{N})$ ist als $\mathscr{L}_{\mathrm{Ar}}$-Theorie unentscheidbar.*

Beweis: Die Hinzunahme einer Formel ändert nichts an der Entscheidbarkeit: Wenn T entscheidbar ist, dann auch $T \cup \{\varphi\}$ (Übung!).

Für $T \subseteq \mathrm{Th}(\mathscr{N})$ ist nun nach Satz 8.2 $T \cup \{q\}$ unentscheidbar, also auch T. □

Folgerung 8.5 (Erster Unvollständigkeitssatz, Gödel 1931) *Jede rekursiv axiomatisierbare, hinreichend ausdrucksstarke $\mathscr{L}$-Theorie T ist entweder widersprüchlich oder unvollständig.*

„Hinreichend ausdrucksstark“ bedeutet dabei, dass T die Definition von Objekten $\underline{n}$ ermöglicht, für die sich aus T die Axiome von Q^* beweisen lassen. Eine solche Theorie ist zum Beispiel ZFC. Dieses Kriterium („Q^* wird in T interpretiert“) kann man technisch sauber definieren; typischerweise wird das in einer Modelltheorie-Vorlesung gemacht.

Satz 8.5 folgt hier für endliche Sprachen $\mathscr{L}$, lässt sich aber auf rekursive Sprachen ausdehnen.

Beweis: Wenn T widerspruchsfrei ist, ist T nach Satz 8.2 unentscheidbar (weil sonst Q^* entscheidbar wäre – das folgt das der sauberen Definition von „hinreichend ausdrucksstark“). Also kann T nach Folgerung 7.18 nicht vollständig sein. □

Folgerung 8.6 (Unentscheidbarkeit der Prädikatenlogik, Church 1936)
Es gibt im Allgemeinen kein Verfahren, mit dem man die Allgemeingültigkeit von $\mathscr{L}$-Formeln entscheiden könnte.
Genauer: Die Menge $\{\ulcorner\varphi\urcorner \mid \varphi$ allgemeingültige $\mathscr{L}$-Formel$\}$ ist im Allgemeinen nicht rekursiv.

Beweis: Wenn $\{\ulcorner\varphi\urcorner \mid \varphi$ allgemeingültige $\mathscr{L}_{\mathrm{Ar}}$-Formel$\}$ rekursiv ist, dann ist auch die Menge $\{\ulcorner q \to \varphi\urcorner \mid \vDash (q \to \varphi)\}$ rekursiv, und damit (wie beim *modus ponens* im Beweis von Satz 7.17) auch $\{\ulcorner\varphi\urcorner \mid \vDash (q \to \varphi)\}$. Das ist aber genau $\{\ulcorner\varphi\urcorner \mid \mathsf{Q} \vDash \varphi\}$, was nicht rekursiv ist, da Q unentscheidbar ist. □

Da man in ZFC mit ω eine Modellisierung der natürlichen Zahlen findet, für die Q gilt, kann man mit dem gleichen Argument die Unentscheidbarkeit der Allgemeingültigkeit von $\mathscr{L}_{\mathrm{ML}}$-Formeln zeigen. Für größere Sprachen gilt die Unentscheidbarkeit erst recht, also hat man die Unentscheidbarkeit für alle Sprachen, die ein zweistelliges Relationszeichen enthalten.

Allgemeiner kann man zeigen, dass die Allgemeingültigkeit von $\mathscr{L}$-Formeln unentscheidbar ist, sobald $\mathscr{L}$ ein mindestens zweistelliges Relationszeichen oder ein mindestens einstelliges Funktionszeichen enthält. Für (endliche oder rekursive) Sprachen, die nur aus einstelligen Funktionszeichen und null- und einstelligen Relationszeichen bestehen, gibt es dagegen Entscheidungsverfahren.

Beweis der Unentscheidbarkeit von Q

Die grundsätzliche Beweisidee für Satz 8.2 ist die gleiche wie für Satz 8.1. Allerdings kann man – da Q^* nicht vollständig ist – nicht von vorneherein mit einem konkreten Modell wie $\mathscr{N}$ arbeiten. Die Definierbarkeit der rekursiven Funktionen in $\mathscr{N}$ wird daher durch eine verwandte Eigenschaft („Repräsentierbarkeit in Q^*“) ersetzt. Um diese Eigenschaft zeigen zu können, braucht man aber zusätzlich, dass rekursive Funktion in der Arithmetik durch besonders einfache $\mathscr{L}_{\mathrm{Ar}}$-Formel definierbar sind, sogenannte Σ_1-Formeln:

Definition 8.7 *Ein* beschränkter Allquantor *ist die Abkürzung $\forall v_i < v_j$ in einer $\mathscr{L}_{\mathrm{Ar}}$-Formel $\forall v_i < v_j\ \varphi$, die für $\forall v_i\,(v_i < v_j \to \varphi)$ steht.* [23]

Eine (arithmetische) Σ_1-Formel *ist eine $\mathscr{L}_{\mathrm{Ar}}$-Formel, die aus quantorenfreien $\mathscr{L}_{\mathrm{Ar}}$-Formeln durch Konjunktion, Disjunktion, Existenzquantifikation und beschränkte Allquantifikation entsteht.*

Insbesondere sind also $\neg, \to, \leftrightarrow$ in einer Σ_1-Formel nur im quantorenfreien Teil erlaubt!

Satz 8.8 *Für jede Σ_1-Formel $\varphi(v_1, \ldots, v_k)$ und alle $n_1, \ldots, n_k \in \mathbb{N}$ gilt*

$$\mathscr{N} \vDash \varphi[n_1, \ldots, n_k] \iff \mathsf{Q}^* \vDash \varphi[\tfrac{n_1}{v_1} \ldots \tfrac{n_k}{v_k}]$$

Insbesondere folgt aus Q^* jede in $\mathscr{N}$ gültige Σ_1-Aussage.

Mit unbeschränkten Allquantoren bekommt man dagegen in $\mathscr{N}$ gültige $\mathscr{L}_{\mathrm{Ar}}$-Aussagen, die sich aus Q^* nicht beweisen lassen, zum Beipsiel die Kommutativität der Addition.

Beweis per Induktion über den Aufbau der Σ_1-Formeln. Die Richtung „$\Leftarrow$“ gilt für alle $\mathscr{L}_{\mathrm{Ar}}$-Formeln, da $\mathscr{N} \vDash \mathsf{Q}$ und $\underline{n}^{\mathscr{N}} = n$.

Um besser Bezug nehmen zu können, bezeichne ich das dritte Axiom von Q^* mit $\mathsf{Q}^*_{<n}$.

(1) Atomare und negiert atomare Formeln: Wenn man in einem $\mathscr{L}_{\mathrm{Ar}}$-Term die Individuenvariable durch einen Term $\underline{n_i}$ substituiert, erhält man einen Term der Form $\underline{m}$. Es reicht daher, an atomaren $\mathscr{L}_{\mathrm{Ar}}$-Formeln die beiden Formeln $\kappa(v_0, v_1) = v_0 < v_1$ und $\gamma(v_0, v_1) = v_0 \doteq v_1$ zu betrachten.

(a) Wenn $\mathscr{N} \vDash \gamma[m, n]$, ist $m = n$ und folglich $\mathsf{Q}^* \vDash \underline{m} \doteq \underline{n}$.

(b) Wenn $\mathscr{N} \vDash \kappa[m, n]$, ist $m < n$. dann zeigt $\mathsf{Q}^*_{<n}$ dass $\mathsf{Q}^* \vDash \underline{m} < \underline{n}$.

(c) Sei nun $\mathscr{N} \vDash \neg\gamma[m, n]$, ohne Einschränkung $m < n$. Zu zeigen ist $\mathsf{Q}^* \vDash \neg\, \underline{m} \doteq \underline{n}$, per Induktion über m:

- $m = 0$: Dann ist $n \neq 0$. Aus $\mathsf{Q}^*_{<n}$ folgt daher $\underline{0} < \underline{n}$, aus $\mathsf{Q}^*_{<0}$ aber $\neg\, \underline{0} < \underline{0}$, also $\mathsf{Q}^* \vDash \neg\, \underline{0} \doteq \underline{n}$.
- $m > 0$: Einerseits gilt nach (b) $\mathsf{Q}^* \vDash \underline{m} < \underline{n}$, andererseits nach Induktion $\mathsf{Q}^* \vDash \neg\, \underline{0} \doteq \underline{m}$ bis $\mathsf{Q}^* \vDash \neg\, \underline{m-1} \doteq \underline{m}$. Mit $\mathsf{Q}^*_{<m}$ ergibt sich daher $\mathsf{Q}^* \vDash \neg\, \underline{m} < \underline{m}$ und folglich $\mathsf{Q}^* \vDash \neg\, \underline{m} \doteq \underline{n}$.

[23] Eine Schreibweise analog zu den relativierten Quantoren in der Mengenlehre.

(d) Sei nun $\mathscr{N} \vDash \neg\kappa[m,n]$, also $n \leqslant m$. Zu zeigen ist $\mathsf{Q}^* \vDash \neg\, \underline{m} < \underline{n}$.

- $n = 0$: Dann besagt $\mathsf{Q}^*_{<0}$, dass $\mathsf{Q}^* \vDash \neg\, \underline{m} < \underline{0}$.
- $n > 0$: mit (c) gilt $\mathsf{Q}^* \vDash \neg\, \underline{m} \doteq \underline{0}$, …, $\mathsf{Q}^* \vDash \neg\, \underline{m} \doteq \underline{n-1}$. Mit $\mathsf{Q}^*_{<n}$ ergibt sich daraus $\mathsf{Q}^* \vDash \neg\, \underline{m} < \underline{n}$.

(2) Konjunktionen und Disjunktionen funktionieren trivialerweise. Damit gilt die Aussage auch für alle quantorenfreien $\mathscr{L}_{\mathrm{Ar}}$-Formeln, da sie (DNF!) logisch äquivalent zu einer Disjunktion von Konjunktionen von atomaren und negiert-atomaren $\mathscr{L}_{\mathrm{Ar}}$-Formeln sind.

(3) Existenzquantifikation: Sei $\psi(v_1,\dots,v_k) = \exists v_0\, \varphi(v_0,v_1,\dots,v_k)$ und $\mathscr{N} \vDash \psi[n_1,\dots,n_k]$. Dann gibt es $n_0 \in \mathbb{N}$ mit $\mathscr{N} \vDash \varphi[n_o,n_1,\dots,n_k]$, nach Induktion also $\mathsf{Q}^* \vDash \varphi[\frac{n_o}{v_0}, \frac{n_1}{v_1}, \dots, \frac{n_k}{v_k}]$. Daraus ergibt sich aber $\mathsf{Q}^* \vDash \exists v_0\, \varphi[\frac{n_1}{v_1}, \dots, \frac{n_k}{v_k}]$ (weil es in jedem $\mathscr{M} \vDash \mathsf{Q}^*$ das Element $n_0^{\mathscr{M}}$ gibt, so dass …).

(4) Beschränkte Allquantifikation: Sei nun $\psi(v_1,\dots,v_k) = \forall v_0 < v_i\, \varphi(v_0,v_1,\dots,v_k)$ und $\mathscr{N} \vDash \psi[n_1,\dots,n_k]$. Für $m = 0,\dots,n_i - 1$ gilt dann $\mathscr{N} \vDash \varphi[m,n_1,\dots,n_k]$, nach Induktion also $\mathsf{Q}^* \vDash \varphi[\frac{m}{v_0}, \frac{n_1}{v_1}, \dots, \frac{n_k}{v_k}]$. Axiom $\mathsf{Q}^*_{<n_i}$ impliziert nun aber $\mathsf{Q}^* \vDash \forall v_0 < \underline{n_i}\, \varphi[\frac{n_1}{v_1}, \dots, \frac{n_k}{v_k}]$, was gerade $\mathsf{Q}^* \vDash \psi[\frac{n_1}{v_1}, \dots, \frac{n_k}{v_k}]$ ist, wenn man die Abkürzung $\forall v_0 < \underline{n_i}$ auflöst. □

Satz 8.9 *Alle rekursiven und rekursiv aufzählbaren Mengen und alle rekursiven Funktionen sind in $\mathscr{N}$ durch Σ_1-Formeln definierbar.*

Beweis: In Satz 7.24 wurde die Definierbarkeit gezeigt, und alle benutzten $\mathscr{L}_{\mathrm{Ar}}$-Formeln sind bereits Σ_1 bis auf den Fall der μ-Rekursion. Dafür gibt es aber eine alternative Σ_1-Formel:

Wenn $f : \mathbb{N}^{k+1}$ wie im Beweis von Satz 7.24 durch φ definiert ist, dann wird die μ-Rekursion von f definiert durch:

$$\psi(v_1,v_2,\dots,v_{k+1}) = \big(\varphi\,[\tfrac{0}{v_0}] \wedge \forall y < v_1\, \exists z\, \big(\varphi\,[\tfrac{z}{v_0}\ \tfrac{y}{v_1}] \wedge \neg z \doteq \underline{0}\big)\big)$$

woebi nun v_1 die Variable für das „Bild-Element“ ist. □

Definition 8.10
(a) *Eine $\mathscr{L}_{\mathrm{Ar}}$-Formel $\varphi(v_0,v_1,\dots,v_k)$* repräsentiert *die Funktion $f : \mathbb{N}^k \to \mathbb{N}$ in Q^*, wenn*

$$\mathsf{Q}^* \vDash \forall v_0\, \big(v_0 \doteq \underline{f(n_1,\dots,n_k)} \leftrightarrow \varphi[\tfrac{n_1}{v_1} \dots \tfrac{n_k}{v_k}]\big)$$

(c) *Eine $\mathscr{L}_{\mathrm{Ar}}$-Formel $\varrho(v_1,\dots,v_k)$* repräsentiert *die Relation $R \subseteq \mathbb{N}^k$ in Q^*, wenn*

$$\begin{aligned}(n_1,\dots,n_k) \in R &\implies \mathsf{Q}^* \vDash \varrho\,[\tfrac{n_1}{v_1} \dots \tfrac{n_k}{v_k}]\\ (n_1,\dots,n_k) \notin R &\implies \mathsf{Q}^* \vDash \neg\varrho\,[\tfrac{n_1}{v_1} \dots \tfrac{n_k}{v_k}]\end{aligned}$$

Wenn φ die Funktion f in Q^* repräsentiert, dann definiert φ auch die Funktion f in $\mathscr{N}$. Umgekehrt sieht man, dass eine f in $\mathscr{N}$ definierende $\mathscr{L}_{\mathrm{Ar}}$-Formel φ die Funktion in Q^* repräsentiert, wenn $\mathsf{Q}^* \vDash \varphi[\frac{f(n_1,\dots,n_k)}{v_o}\, \frac{n_1}{v_1} \dots \frac{n_k}{v_k}]$ für alle $n_1,\dots,n_k$ – was für Σ_1-Formel φ der Fall ist – und wenn Q^* die Funktionalitätseigenschaft von φ beweist, also

$$\mathsf{Q}^* \vDash \forall y \forall y'\, \Big(\big(\varphi[\tfrac{y}{v_0}\, \tfrac{n_1}{v_1} \dots \tfrac{n_k}{v_k}] \wedge \varphi[\tfrac{y'}{v_0}\, \tfrac{n_1}{v_1} \dots \tfrac{n_k}{v_k}]\big) \to y \doteq y'\Big)$$

Lemma 8.11 *Jede rekursive Funktion $f : \mathbb{N}^k \to \mathbb{N}$ und jede rekursive Menge $R \subseteq \mathbb{N}^k$ wird in Q^* durch eine Σ_1-Formel repräsentiert.*

Beweis: Nach der Vorbemerkung sieht man, dass sowohl die im Beweis von Satz 7.24 angegebenen $\mathscr{L}_{\mathrm{Ar}}$-Formeln, die die erweiterten Grundfunktionen in $\mathscr{N}$ definieren, sie auch in Q^* repräsentieren, als auch dass die Konstruktion für die Komposition funktioniert, also aus die Einzelfunktionen repräsentierenden $\mathscr{L}_{\mathrm{Ar}}$-Formeln eine die Komposition repräsentierende $\mathscr{L}_{\mathrm{Ar}}$-Formel macht.

Im Fall der μ-Rekursion ist dagegen die Funktionalitätseigenschaft der im Beweis von Satz 8.9 angegebenen $\mathscr{L}_{\mathrm{Ar}}$-Formel nicht klar. Das Problem hierbei ist, dass Q^* nicht beweist, dass $<$ in jedem Modell eine lineare Ordnung definiert: Zwei die Formel ψ erfüllende Elemente müssen daher bzgl. $<$ nicht kompatibel sein. Das repariert man, indem man die folgende $\mathscr{L}_{\mathrm{Ar}}$-Formel betrachtet:

$$\psi' = \big(\psi \wedge (\underline{0} \doteq v_1 \vee \underline{0} < v_1) \wedge \forall z\,(z < v_1 \to (sz \doteq v_1 \vee sz < v_1))\big)$$

Sei nun $m = \mu z[f(m, n_1, \ldots, n_k) = 0]$. Dann gilt $\mathscr{N} \vDash \psi'[\frac{m}{v_1}\frac{n_1}{v_2}\ldots\frac{n_k}{v_{k+1}}]$, also, da ψ eine Σ_1-Formel, auch $\mathsf{Q}^* \vDash \psi'[\frac{m}{v_1}\frac{n_1}{v_2}\ldots\frac{n_k}{v_{k+1}}]$. Wenn nun $\mathscr{M} \vDash \psi'[m', \underline{n_1}^{\mathscr{M}}, \ldots, \underline{n_k}^{\mathscr{M}}]$ in einem Modell $\mathscr{M} \vDash \mathsf{Q}^*$, enthält die Menge $\{a \in M \mid a <^{\mathscr{M}} m'\}$ das Element $\underline{0}^{\mathscr{M}}$ und ist unter $s^{\mathscr{M}}$ abgeschlossen, enthält also auch $\underline{m}^{\mathscr{M}}$, d. h. $\mathscr{M} \vDash \underline{m}^{\mathscr{M}} < m'$. Dann folgt aber $\underline{m}^{\mathscr{M}} = m'$ wegen der in ψ eingebauten Minimalitätsbedingung (angewandt auf m').

Wenn $R \subseteq \mathbb{N}^k$ und χ_R in Q^* durch $\varphi(v_0, v_1, \ldots, v_k)$ repräsentiert wird, wird R durch $\varphi[\frac{1}{v_0}]$ repräsentiert. □

Beweis von Satz 8.2: Angenommen $T \subseteq \mathrm{Th}(\mathscr{N})$ mit $T \vDash \mathsf{Q}^*$ ist eine entscheidbare $\mathscr{L}_{\mathrm{Ar}}$-Theorie. Sei nun $(\varphi_n)_{n\in\mathbb{N}}$ wie im Beweis von Satz 8.1 eine rekursive Aufzählung aller $\mathscr{L}_{\mathrm{Ar}}$-Formeln in der freien Variable v_0. Die Menge $R := \{n \in \mathbb{N} \mid T \vDash \neg\varphi_n[\frac{n}{v_0}]\}$ ist wegen der angenommenen Entscheidbarkeit von T rekursiv, wird also in Q^* von einer Σ_1-Formel repräsentiert, deren Gödelnummer n_0 sei.

1. Fall: $n_0 \in R$. Dann gilt $\mathsf{Q}^* \vDash \varphi_{n_0}[\frac{n_0}{v_0}]$, weil φ_{n_0} die Relation R in Q^* repräsentiert, also auch $T \vDash \varphi_{n_0}[\frac{n_0}{v_0}]$. Nach Definition von R folgt aus $n_0 \in R$ aber $T \vDash \neg\varphi_n[\frac{n}{v_0}]$, also ist T inkonsistent.

2. Fall: $n_0 \notin R$. Dann gilt $\mathsf{Q}^* \vDash \neg\varphi_{n_0}[\frac{n_0}{v_0}]$, weil φ_{n_0} die Relation R in Q^* repräsentiert, also auch $T \vDash \neg\varphi_{n_0}[\frac{n_0}{v_0}]$. Dies bedeutet nach Definition von R aber, dass $n_0 \in R$: Widerspruch! □

Der Fixpunktsatz und seine Folgen

Satz 8.12 (Fixpunktsatz, Kleene 1938)
Zu jeder $\mathscr{L}_{\mathrm{Ar}}$-Formel $\psi(v_0)$ gibt es eine $\mathscr{L}_{\mathrm{Ar}}$-Aussage φ so, dass

$$\mathsf{Q}^* \vDash \big(\psi[\tfrac{\ulcorner\varphi\urcorner}{v_0}] \leftrightarrow \varphi\big)$$

Wenn ψ Σ_1-Formel ist, kann auch φ als Σ_1-Formel gewählt werden.

Das kann man so interpretieren, dass φ modulo der Gödelisierung eine Aussage über sich selbst macht, und zwar besagt φ: „Die von ψ ausgesagte Eigenschaft gilt von mir“.

Beweis: Die primitiv rekursive Auswertungsfunktion $\mathrm{eval} : \mathbb{N}^2 \to \mathbb{N}$ aus dem Beweis von Satz 8.1 wird in Q^* durch eine Σ_1-Formel $\varepsilon(v_0, v_1, v_2)$ repräsentiert, d. h.

$$\mathsf{Q}^* \vDash \forall y\,\big(\varepsilon[\tfrac{y}{v_0}, \tfrac{\ulcorner\varphi\urcorner}{v_1}, \tfrac{n}{v_2}] \leftrightarrow y \doteq \underline{\ulcorner\varphi[\tfrac{n}{v_0}]\urcorner}\big)$$

Sei nun $\zeta(v_0) := \exists y\,\big(\psi[\frac{y}{v_0}] \wedge \varepsilon[\frac{y}{v_0}, \frac{v_0}{v_1}, \frac{v_0}{v_2}]\big)$. Mit ψ ist auch ζ eine Σ_1-Formel. Es folgt aus der Definition von ζ, dass für jede $\mathscr{L}_{\mathrm{Ar}}$-Formel $\eta(v_0)$:

$$\mathsf{Q}^* \vDash \Big(\zeta[\tfrac{\ulcorner\eta\urcorner}{v_0}] \leftrightarrow \exists y\,\big(\psi[\tfrac{y}{v_0}] \wedge \varepsilon[\tfrac{y}{v_0}, \tfrac{\ulcorner\eta\urcorner}{v_1}, \tfrac{\ulcorner\eta\urcorner}{v_2}]\big)\Big) \sim \Big(\zeta[\tfrac{\ulcorner\eta\urcorner}{v_0}] \leftrightarrow \psi[\tfrac{\eta[\frac{\ulcorner\eta\urcorner}{v_0}]}{v_0}]\Big)$$

Für $\eta = \zeta$ folgt

$$\mathsf{Q}^* \vDash \big(\zeta[\tfrac{\ulcorner\zeta\urcorner}{v_0}] \leftrightarrow \psi[\tfrac{\zeta[\frac{\ulcorner\zeta\urcorner}{v_0}]}{v_0}]\big)$$

Also hat $\varphi = \zeta[\frac{\ulcorner\zeta\urcorner}{v_0}]$ die gewünschte Eigenschaft. Außerdem ist φ Σ_1, wenn ζ es ist. □

Folgerung 8.13 (Nicht-Existenz eines Wahrheitsprädikats, Tarski 1930)
Es gibt keine $\mathscr{L}_{\mathrm{Ar}}$-Formel $\omega(v_0)$ so, dass für alle $\mathscr{L}_{\mathrm{Ar}}$-Aussagen φ gilt:

$$\mathsf{Q}^* \vDash \big(\varphi \leftrightarrow \omega\,[\tfrac{\ulcorner\varphi\urcorner}{v_0}]\big)$$

Beweis: Für jedes $\omega(v_0)$ zeigt ein Fixpunkt von $\neg\omega$, dass ω kein Wahrheitsprädikat ist. □

Wenn ω ein Wahrheitsprädikat wäre, würde ein Fixpunkt von $\neg\omega$ im Sinne der oben angegebene Interpretation von sich sagen: „Ich bin nicht wahr“ oder „Ich bin nicht beweisbar“. Dies wird gerne als ein formales Gegenstück des antiken Lügner-Paradoxons dargestellt.

Folgerung 8.14 *Die $\mathscr{L}_{\mathrm{Ar}}$-Theorie von $\mathscr{N}$ ist nicht arithmetisch, d. h. $\{\ulcorner\varphi\urcorner \mid \mathscr{N} \vDash \varphi\}$ ist in $\mathscr{N}$ nicht definierbar.*

Beweis: Das ist eine Umformulierung der Tatsache, dass es für $\mathscr{N}$ keine Wahrheitsprädikat gibt: Angenommen $\{\ulcorner\varphi\urcorner \mid \mathscr{N} \vDash \varphi\}$ wird durch $\omega(v_0)$ definiert. Für einen Fixpunkt φ von $\neg\omega$ gilt dann auch $\mathscr{N} \vDash \big(\varphi \leftrightarrow \neg\omega\,[\frac{\ulcorner\varphi\urcorner}{v_0}]\big)$. Andererseits gilt nach Wahl von ω gerade $\mathscr{N} \vDash \varphi \iff \mathscr{N} \vDash \omega\,[\frac{\ulcorner\varphi\urcorner}{v_0}]$: Widerspruch! □

Im Beweis von Satz 8.2 ist φ_{n_0} eine Entsprechung von ω; man könnte den Satz auch auf ähnliche Weise mit Hilfe des Fixpunktsatzes beweisen. Hier eine Variante:

Satz 8.15 *Wenn $T \subseteq \mathrm{Th}(\mathscr{N})$ eine rekursiv aufzählbare $\mathscr{L}_{\mathrm{Ar}}$-Theorie ist, oder allgemeiner eine Theorie, für die $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ arithmetisch ist, dann ist T unvollständig.*

Beweis: Sei $\omega(v_0)$ eine $\{\ulcorner\varphi\urcorner \mid T \vDash \varphi\}$ definierende $\mathscr{L}_{\mathrm{Ar}}$-Formel und φ ein Fixpunkt von $\neg\omega$. Dann gilt wie eben

$$\mathscr{N} \vDash \varphi \iff \mathscr{N} \vDash \neg\omega\,[\tfrac{\ulcorner\varphi\urcorner}{v_0}] \iff \mathscr{N} \nvDash \omega\,[\tfrac{\ulcorner\varphi\urcorner}{v_0}] \iff T \nvDash \varphi$$

Da $\mathscr{N} \vDash T$ folgt $T \nvDash \varphi$. □

Insbesondere gibt es also als $\mathscr{L}_{\mathrm{Ar}}$-Aussagen formulierbare Sätze über $\mathbb{N}$, die in der erststufigen Peano-Arithmetik PA nicht beweisbar sind. Aus dem Beweis des vorherigen Lemma könnte man zwar eine konkrete solche Aussage herleiten, die allerdings keine „normale mathematische“ Aussage ist. Der vermutlich bekannteste mathematische Satz über $\mathbb{N}$, von dem man weiß, dass er in PA nicht beweisbar ist, ist der Satz von Goodstein.

Dieser Satz benutzt die iterierte Darstellung $[m]_b$ einer Zahl $m \in \mathbb{N}$ zur Basis b, in der man sukzessive auch alle Exponenten zur Basis b darstellt, so dass keine Zahlen größer als b vorkommen. Beispielweise ist $3^{3^3+2\cdot 3^2+1} + 2 \cdot 3^{2\cdot 3^2} + 2$ eine iterierte Darstellung zur Basis 3. Die *Aufblähungsabbildung* A_k ersetzt in der iterierten Darstellung einer Zahl jedes Vorkommen der Basis b durch k. Zum Beispiel ist $A_4(3^{3^3+2\cdot 3^2+1} + 2 \cdot 3^{2\cdot 3^2} + 2) = 4^{4^4+2\cdot 4^2+1} + 2 \cdot 4^{2\cdot 4^2} + 2$.

Die Goodstein-Folge mit Startwert g_0 wird rekursiv definiert durch $g_{n+1} = A_{n+3}([g_n]_{n+2}) - 1$, falls $g_n \neq 0$, und $g_{n+1} = g_n = 0$ sonst, d. h. in der iterierten Darstellung des n-ten Folgenglieds zur Basis $n+2$ wird die Basis um 1 erhöht und anschließend 1 abgezogen. Wenn beispielsweise

$g_1 = 3^{3^3+2\cdot 3^2+1} + 2\cdot 3^{2\cdot 3^2} + 2$, ist $g_2 = 4^{4^4+2\cdot 4^2+1} + 2\cdot 4^{2\cdot 4^2} + 1$. Wenn es keinen „konstanten Koeffizienten“ gibt, muss die iterierte Darstellung „neu berechnet werden“: Zum Beispiel mit $g_0 = 4 = 2^2$ ergibt sich $g_1 = 3^3 - 1 = 26 = 2\cdot 3^2 + 2\cdot 3 + 2$, also $g_2 = 2\cdot 4^2 + 2\cdot 4 + 1 = 41$.

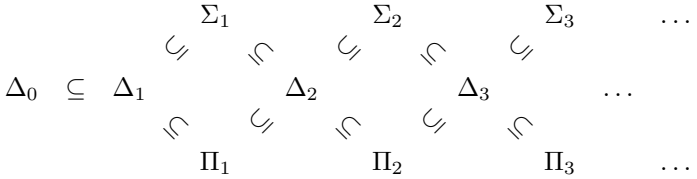
Der Satz von Goodstein besagt nun, dass es für jeden Startwert g_0 ein n mit $g_n = 0$ gibt, dass also jede Goodstein-Folge nach endlich vielen Schritten konstant 0 wird. (Bei $g_0 = 4$ sind dafür allerdings schon über 10^{10^8} Schritte nötig.) Mit etwas Aufwand kann man diesen Satz als $\mathscr{L}_{\mathrm{Ar}}$-Aussage formalisieren: Da Goodstein-Folgen rekursiv definiert sind und damit im intuitiven Sinn berechenbar, muss die Funktion $(g,n) \mapsto$ „n-tes Folgenglied bei Startwert g“ nach Church'scher These im technischen Sinne rekursiv und damit in $\mathscr{N}$ definierbar sein.

Der Beweis ist relativ einfach mit der aus Seite 50 angedeuteten Ordinalzahlexponentiation: $\alpha_n := A_\omega([g_n]_{n+2})$ ist eine Ordinalzahl $< \varepsilon_0$. Offenbar gilt $A_\omega([g_n]_{n+2}) = A_\omega(A_{n+3}([g_n]_{n+2}))$, deshalb ist $\alpha_{n+1} \leqslant \alpha_n - 1$, solange $\alpha_n \neq 0$. Nun hat die Menge $\{\alpha_n \mid n \in \mathbb{N}\}$ wegen der Wohlgeordnetheit von $\mathbb{ORD}$ ein Minimum; dieses kann dann nur 0 sein.

Laurie Kirby und Jeff Paris haben 1982 gezeigt, dass der Satz von Goodstein nicht in PA beweisbar ist, da er für Nicht-Standard-Zahlen nicht gilt. Zuvor hatten Jeff Paris and Leo Harrington 1977 mit einer Verschärfung des Satzes von Ramsey und Harvey Friedman 1981 mit einer Variante des Satzes von Kruskal mathematische Theoreme gefunden, die als $\mathscr{L}_{\mathrm{Ar}}$-Aussagen formulierbar, aber in PA nicht beweisbar sind.

Die arithmetische Hierarchie

Wenn $\mathscr{L}$ eine beliebige prädikatenlogische Sprache ist, klassifiziert man die $\mathscr{L}$-Formeln in folgender Hierarchie: Δ_0 ist die Menge der quantorenfreien $\mathscr{L}$-Formeln. Σ_1 und Π_1 sind die Mengen der $\mathscr{L}$-Formeln, die bis auf logische Äquivalenz von der Form $\exists v_{i_1} \ldots \exists v_{i_k}\, \varphi$ bzw. $\forall v_{i_1} \ldots \forall v_{i_k}\, \varphi$ mit $\varphi \in \Delta_0$ sind. Die Mengen Σ_{n+1} und Π_{n+1} sind analog definiert mit $\varphi \in \Pi_n$ bzw. $\varphi \in \Sigma_n$. Schließlich setzt man $\Delta_n = \Sigma_n \cap \Pi_n$. Häufig sieht man auch die Schreibweisen Σ_i^0, Π_i^0 und Δ_i^0 als Teil einer allgemeineren Hierarchie Σ_i^j, Π_i^j und Δ_i^j für Formeln der $(j+1)$-stufigen Prädikatenlogik.



In Zusammenhang mit der Arithmetik betrachtet man eine Variante davon, die *arithmetische Hierarchie*. Die Definition ist die gleiche, allerdings erlaubt man bei allen Mengen inklusive Δ_0 zusätzlich beschränkte Existenz- und Allquantifikationen.

Lemma 8.16 *Alle Σ_1-definierbaren Teilmengen von $\mathbb{N}^k$ sind rekursiv aufzählbar; alle Σ_1-definierbaren Funktionen $\mathbb{N}^k \to \mathbb{N}$ sind rekursiv.*

Beweis: Per Induktion über den Aufbau der $\mathscr{L}_{\mathrm{Ar}}$-Formeln: Die geschlossenen $\mathscr{L}_{\mathrm{Ar}}$-Terme sind genau die $\underline{n}$ für $n \in \mathbb{N}$. Die Nachfolgerfunktion $s(v_i)$ kann man äquivalent durch $v_i + \underline{1}$ ausdrücken, also beschreiben alle $\mathscr{L}_{\mathrm{Ar}}$-Terme $\tau(v_0, \ldots, v_k)$ Polynome aus $\mathbb{N}[X_0, \ldots, X_k]$ und somit (primitiv) rekursive Funktionen $\tau^{\mathscr{N}}$. Die durch die atomaren $\mathscr{L}_{\mathrm{Ar}}$-Formeln $\tau_1 \doteq \tau_2$ und $\tau_1 < \tau_2$ definierten Mengen sind als Lösungsmengen der Gleichungen $|\tau_1^{\mathscr{N}} - \tau_2^{\mathscr{N}}| = 0$ bzw. $\chi_<(\tau_1^{\mathscr{N}}, \tau_2^{\mathscr{N}}) = 1$ daher ebenfalls (primitiv) rekursiv.

Es folgt, dass alle durch quantorenfreie $\mathscr{L}_{\mathrm{Ar}}$-Formeln definierbare Mengen rekursiv sind, da die rekursiven Teilmengen nach Satz 7.7 unter Schnitt, Vereinigung und Komplementen abgeschlossen sind (und $\wedge, \vee, \neg$ ein vollständiges Junktorensystem ist).

Weiterhin sind die rekursiv aufzählbaren Mengen nach Folgerung 7.8 unter Schnitt, Vereinigung und Projektionen abgeschlossen, also definieren alle aus quantorenfreien $\mathscr{L}_{\mathrm{Ar}}$-Formeln durch Konjunktion, Disjunktion und Existenzquantifikation entstehende $\mathscr{L}_{\mathrm{Ar}}$-Formeln auch rekursiv aufzählbare Mengen. Es bleibt zu zeigen, dass die rekursiv aufzählbar Mangen auch unter beschränkter universeller Quantifikation abgeschlossen sind.

Sei $A \subseteq \mathbb{N}^k$ rekursiv aufzählbar, definiert durch $\exists v_0\, \varphi(v_0, \ldots, v_k)$ mit einer Σ_1-Formel φ, die die rekursive Menge $R \subseteq \mathbb{N}^{k+1}$ definiert, deren Projektion A ist. Dann definiert $\forall v_i < v_j\, \exists v_0\, \varphi$ die gleiche Menge wie $\exists y\, \forall v_i < v_j\, \varphi[\frac{\langle y \rangle_{v_i}}{y}]$, in der die Folge der für $v_i = 0, \ldots, v_j - 1$ nötigen Werte für v_2 in einer Zahl y kodiert ist. Die Teilformel nach „$\exists y$“ definiert nach Lemma 7.9 (b) eine rekursiv aufzählbare Menge: $R_\forall$ wird in der i-ten Koordinate „unterhalb v_j abgeschnitten“ (im Falle $i = 0$ durch den Schnitt mit $[0, v_j) \times \mathbb{N}$, dann Projektion entlang v_i). Also wird insgesamt eine rekursiv aufzählbare Menge definiert.

Schließlich ist eine Funktion genau dann rekursiv, wenn ihr Graph rekursiv aufzählbar ist (Satz 7.7), also genau dann, wenn ihr Graph Σ_1-definierbar ist. □

Satz 8.17 *Es gibt universelle partielle rekursive Funktionen $U^k : \mathbb{N}^{k+1} \dashrightarrow \mathbb{N}$, d. h.*

$$\{U^k(\cdot, m) : \mathbb{N}^k \dashrightarrow \mathbb{N},\ (n_1, \ldots, n_k) \mapsto U^k(n_1, \ldots, n_k, m) \mid m \in \mathbb{N}\}$$

ist die Menge aller partiellen rekursiven Funktionen $\mathbb{N}^k \dashrightarrow \mathbb{N}$.

Beweis: Ähnlich wie in Lemma 7.15 sieht man, dass $\mathcal{C}_{\Sigma_1} = \{\ulcorner \varphi \urcorner \mid \varphi(v_0, \ldots, v_k) \text{ ist } \Sigma_1\text{-Formel}\}$ eine rekursive Menge ist. Sei $g : \mathbb{N} \to \mathcal{C}_{\Sigma_1}$ eine rekursive Aufzählung (also eine rekursive Surjektion). Insbesondere tauchen nach Satz 8.9 die Gödelnummern von allen partiellen rekursiven Abbildungen $\mathbb{N}^k \dashrightarrow \mathbb{N}$ im Bild von g auf.

Aus jeder Σ_1-definierten Menge $R \subseteq \mathbb{N}^{k+1}$ kann man uniform den Graph einer Σ_1-definierbaren partiellen Funktion $\mathbb{N}^k \dashrightarrow \mathbb{N}, (n_1, \ldots, n_k) \mapsto \mu m\, [(n_1, \ldots, n_k, m) \in R]$ machen, die nach Lemma 8.16 rekursiv ist.

Damit ist $U^k(n_1, \ldots, n_k, m) = \mu m\, [(n_1, \ldots, n_k, m) \in R]$ mit $\ulcorner R \urcorner = g(m)$ die gesuchte universeelle Funktion. □

Folgerung 8.18 (Halteproblem, Turing 1937)
Die Menge $H = \{(n_1, \ldots, n_k, m) \in \mathbb{N}^k \mid U^k(n_1, \ldots, n_k, m)$ ist definiert$\}$ ist nicht rekursiv.

In der Sprache der abstrakten Maschinenmodelle wie der Turing-Maschinen entspricht dies dem ursprünglich von Turing beweisenen Ergebnis, dass man nicht entscheiden kann, ob die Maschine mit Programm m bei Eingabe $(n_1, \ldots, n_k)$ nach endlich vielen Schritten anhält.

Beweis: Wäre H rekursiv, könnte man U^k folgendermaßen zu einer universellen totalen rekursiven Funktion abändern:

$$U^k_{\mathrm{tot}}(n_1, \ldots, n_k, m) := \begin{cases} U^k(n_1, \ldots, n_k, m) & \text{falls } (n_1, \ldots, n_k, m) \notin H \\ 0 & \text{falls } (n_1, \ldots, n_k, m) \in H \end{cases}$$

Die modifizierte Diagonalabbildung $n \mapsto U^2_{\mathrm{tot}}(n, n) + 1$ wäre dann eine rekursive Abbildung, die kein $U^k_{\mathrm{tot}}(\cdot, m)$ sein kann. Denn Fall $k > 1$ kann man durch übliche Kodierungen auf $k = 1$ zurückführen. □

Kennen sollte man auch noch den folgenden Satz:

Satz 8.19 (Satz von Rice) *Sei $X \neq \emptyset$ eine echte Teilmenge der Menge aller partiellen rekursiven Funktionen $\mathbb{N}^k \dashrightarrow \mathbb{N}$. Dann ist $I_X = \{m \in \mathbb{N} \mid U^k(\cdot, m) \in X\}$ nicht rekursiv.*

Insbesondere ist in dieser Kodierung der Funktionen definieren $\mathscr{L}_{\mathrm{Ar}}$-Formeln nicht entscheidbar, ob eine rekursive Funktion total ist oder ob zwei $\mathscr{L}_{\mathrm{Ar}}$-Formeln dieselbe Funktion definieren. Es ist auch nicht entscheidbar, ob eine rekursive Funktion primitiv rekursiv. Dies widerspricht nicht den Überlegungen auf Seite 66 f, da dort „Aufbaurezepte“ primitiv rekursiver Funktionen kodiert wurden, hier aber definierende $\mathscr{L}_{\mathrm{Ar}}$-Formeln, also Formeln, die unmittelbar den Graph der Funktion beschreiben.

Auch der Satz von Rice ist ursprünglich ein Satz der theoretischen Informatik und wird typischerweise so formuliert, dass man nicht anhand des Programmcodes eines Programms, das eine Funktion berechnet, allgemein algorithmisch entscheiden kann, ob die berechnete Funktion eine gewisse nicht-trivale Eigenschaft hat. Zum Beispiel kann man kein Programm angeben, dass bei Eingabe eines Programmcodes entscheidet, ob das eingegeben Programm eine gegebene Funktion (z. B. die Addition) berechnet.

Hilberts 10. Problem

Eine *diophantische Gleichung* ist eine Gleichung $P(X_1, \dots, X_n) = 0$ für ein Polynom $P \in \mathbb{Z}[X_1, \dots, X_n]$. Das zehnte in der Liste der berühmten Hilbert'schen Probleme war die Frage, ob es ein Entscheidungsverfahren für die Lösbarkeit diophantischer Gleichungen gibt, also für die Frage, ob es eine Nullstelle in $\mathbb{Z}^n$ gibt. (Äquivalent kann man $\mathbb{Z}$ durch $\mathbb{N}$ ersetzen.) Diese Frage wurde negativ beantwortet durch den

Satz 8.20 (Davis, Putnam, Robinson, Matiyasevich)
Die rekursiv aufzählbaren Mengen $R \subseteq \mathbb{N}^k$ sind genau die diophantischen Mengen

$$R = \big\{(n_1, \dots, n_k) \in \mathbb{N}^k \;\big|\; \text{es gibt } m_1, \dots, m_l \in \mathbb{N} \text{ mit } P(n_1, \dots, n_k, m_1, \dots, m_l) = Q(n_1, \dots, n_k, m_1, \dots, m_l)\big\}$$

für $l \in \mathbb{N}$ und Polynome $P, Q \in \mathbb{N}[X_1, \dots, X_{k+l}]$.

Rekusiv aufzählbare Mengen sind also durch besonders einfache $\mathscr{L}_{\mathrm{Ar}}$-Formeln definierbar, nämlich

$$\exists v_{k+1} \dots \exists v_{k+l}\ \tau_1(v_1, \dots, v_k, v_{k+1}, \dots, v_{l+k}) \doteq \tau_2(v_1, \dots, v_k, v_{k+1}, \dots, v_{l+k})$$

mit $\mathscr{L}_{\mathrm{Ar}}$-Termen τ_1, τ_2. Die Unentscheidbarkeit der Arithmetik liegt also nicht an komplizierten Formeln.

Erstaunlich ist die Folgerung aus dem Satz, dass Komplemente diophantischer Mengen ebenfalls diophantisch sind. So muss zum Beispiel der Menge aller Primzahlen diophantisch sein.

Trotz intensiver Forschung nach wie vor offen ist das 10. Hilbert'sche Problem für $\mathbb{Q}$. Man könnte es negativ lösen, wenn man $\mathbb{Z}$ existenziell in $\mathbb{Q}$ definieren könnte, also durch eine $\{+, \cdot, -, 0, 1\}$-Formel, deren pränexe Normalform keine Allquantoren enthält. Das bis heute beste Ergebnis stammt von Jochen Koenigsmann (2016), der eine universelle Definition von $\mathbb{Z}$ in $\mathbb{Q}$ angeben konnte (woraus folgt, dass die Π_2-Theorie von $(\mathbb{Q}; +^{\mathbb{Q}}, \cdot^{\mathbb{Q}}, -^{\mathbb{Q}}, 0, 1\}$ unentscheidbar ist).

Der zweite Gödel'sche Unvollständigkeitssatz

Hilberts ursprüngliche Motivation für die Entwicklung der Mathematischen Logik in Richtung Beweistheorie war es, die Konsistenz mengentheoretischer Methoden auf der Grundlage unum-

strittener „finitistischer“ Methoden zu zeigen, im Wesentlichen also auf endliche Arithmetik zurückzuführen. Gödels Zweiter Unvollständigkeitssatz besagt, dass rekursiv axiomatisierbare Theorien, die hinreichend ausdrucksstark sind, um die Arithmetik zu interpretieren – also insbeonsdere ZFC und PA – ihre eigene Konsistenz nicht beweisen können.

Sei T die betreffende $\mathscr{L}$-Theorie, also zum Beispiel die $\mathscr{L}_{\mathrm{Ar}}$-Theorie PA oder die $\mathscr{L}_{\mathrm{ML}}$-Theorie ZFC. Man muss nun zunächst eine $\mathscr{L}$-Aussage formulieren, welche die Konsistenz von T ausdrückt. Dazu nutzt man wieder die Gödelisierungs des Kalküls $\mathbb{K}$ – oder eines vergleichbaren Kalküls, der vollständig und korrekt ist – und definiert ähnlich wie im Beweis von Satz 7.17 eine $\mathscr{L}$-Formel – die *Beweisbarkeitsprädikat* gennant wird:

$$\beta(v_0) = \big(v_0 \in \mathcal{C}_{\text{Aussage}} \wedge \exists v_1 \left(\langle v_i\rangle_0 \dots \langle v_i\rangle_{\lg(v_i)-1} \text{ ist ein } \mathbb{K}\text{-Beweis von } v_0 = \langle v_i\rangle_{\lg(v_i)-1} \text{ aus } T\right)\big)$$

Man braucht hierfür, dass die Axiome von T rekursiv aufzählbar sind, und muss die Beweisidee von Satz 7.17 entsprechend abändern, dass neben den Axiomen des Kalküls auch die Axiome von T eingearbeitet sind.

Klar ist: Wenn $T \vDash \varphi$, dann gibt es nach dem Vollständigkeitssatz einen $\mathbb{K}$-Beweis von φ aus T, also gilt $T \vDash \beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}]$, wobei $\underline{n}$ wie in den beiden Beispielen die Interpretation der natürlichen Zahl n in der Theorie T sein soll. Die Umkehrung gilt im Allgemeinen aber nicht, weil es in Modellen von T Nicht-Standard-Zahlen geben könnte und damit „nicht-standard-lange“ Folgen von $\mathscr{L}$-Formeln, von denen „T denkt“, dass sie $\mathbb{K}$-Beweise sind.

In ZFC gelten für das Beweisbarkeitsprädikat die folgenden *Löb-Axiome*, aus denen dann zusammen mit dem Fixpunktsatz (der auch für ZFC gilt) der Zweite Unvollständigkeitssatz folgt. Für die Peano-Arithmetik muss man das Beweisbarkeitsprädikat geschickt modifizieren, damit das dritte Löb-Axiom gilt (Details bei Ziegler, § 20). Andere Theorien muss man sich im Einzelfall anschauen oder auf die Peano-Arithmetik zurückführen.

Die Löb-Axiome für T lauten:

[L1] Wenn $T \vDash \varphi$, dann $T \vDash \beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}]$.

[L2] $T \vDash \big(\big(\beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}] \wedge \beta_T[\frac{\ulcorner(\varphi\to\psi)\urcorner}{v_0}]\big) \to \beta_T[\frac{\ulcorner\psi\urcorner}{v_0}]\big)$ für alle $\mathscr{L}$-Formeln φ und ψ.

[L3] $T \vDash \big(\beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}] \to \beta_T[\frac{\beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}]}{v_0}]\big)$ für alle $\mathscr{L}$-Formeln φ.

[L1] gilt wie oben dargelegt per Definition des Beweisbarkeitsprädikats und [L2] gilt, weil [modus ponens] eine der Kalkülregeln von $\mathbb{K}$ ist. [L3] gilt, grob gesprochen, dann, wenn man den Beweis von [L1] innerhalb von T ausführen kann. Der konkrete Beweis ist typischerweise technisch und subtil (siehe Ziegler, § 20).

[L1] sagt aus, dass T alle in T beweisbaren $\mathscr{L}$-Aussagen auch für beweisbar hält. Umgekehrt könnte es aber $\mathscr{L}$-Aussagen geben, die T für beweisbar hält, ohne dass sie es tatsächlich sind. [L3] sagt aus, dass T denkt, dass aus der Beweisbarkeit einer $\mathscr{L}$-Aussage φ auch die Beweisbarkeit der Beweisbarkeit von φ folgt. Wichtig ist der Unterschied in der Formulierung von [L1] und [L3]: die Implikation in [L1] gilt nur für in T beweisbare $\mathscr{L}$-Aussage φ, die in [L3] für beliebige; $T \vDash \big(\varphi \to \beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}]\big)$ kann nicht allgemein gelten, weil es dem Fixpunktsatz widerspräche! (Zudem wäre [L3] dann ein Spezialfall von [L1].)

Definition 8.21 con_T *ist die* $\mathscr{L}$*-Aussage* $\neg\beta_T[\frac{\ulcorner\bot\urcorner}{v_0}]$.

Satz 8.22 (Zweiter Gödel'scher Unvollständigkeitssatz)
T ist genau dann konsistent, wenn $T \nvDash \mathrm{con}_T$.
Wenn T konsistent ist, kann T also seine eigene Konsistenz nicht beweisen!

Wenn T inkonsistent ist, beweist T alls $\mathscr{L}$-Aussagen, also auch con_T. Umgekehrt ist es möglich, dass T konsistent ist, aber $\neg\mathrm{con}_T$ beweist. Typischerweise erwartet man aber, dass con_T eine von T *unabhängige* $\mathscr{L}$-Aussage ist, also aus T weder beweisbar noch widerlegbar.

Der Beweis ist nahezu kombinatorisch aus den Löb-Axiomen. Der besseren Übersichtlichkeit halber (und im Vorgriff auf den Einblick in die Modallogik) schreibe ich $\Box\varphi$ für $\beta_T[\frac{\ulcorner\varphi\urcorner}{v_0}]$. Die Löb-Axiome lauten dann:

[L1] Wenn $T \vDash \varphi$, dann $T \vDash \Box\varphi$.

[L2] $T \vDash \big((\Box\varphi \wedge \Box(\varphi \to \psi)) \to \Box\psi\big)$

[L3] $T \vDash \big(\Box\varphi \to \Box\Box\varphi\big)$

Beweis von Satz 8.22: Aus den Löb-Axiomen folgen zunächst die beiden Eigenschaften

[L4] Wenn $T \vDash (\varphi \to \psi)$, dann $T \vDash (\Box\varphi \to \Box\psi)$.
[L5] $T \vDash \big((\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)\big)$

Aus $T \vDash (\varphi \to \psi)$ folgt mit [L1] $T \vDash \Box(\varphi \to \psi)$. Mit [L2] folgt daraus [L4].

$(\varphi \to (\psi \to (\varphi \wedge \psi)))$ ist eine $\mathscr{L}$-Tautologie, mit [L4] gilt also $T \vDash (\Box\varphi \to \Box(\psi \to (\varphi \wedge \psi)))$. [L2] liefert $T \vDash ((\Box\psi \wedge \Box(\psi \to (\varphi \wedge \psi))) \to \Box(\varphi \wedge \psi))$; zusammen ergibt sich [L5].

Sei nun φ ein Fixpunkt des negierten Beweisbarkeitsprädikats: $(*)$ $T \vDash (\varphi \leftrightarrow \neg\Box\varphi)$. Daraus folgt mit [L1] $T \vDash \Box(\varphi \to \neg\Box\varphi)$ und daraus wiederum mit [L4] $T \vDash (\Box\varphi \to \Box\neg\Box\varphi)$.
[L3] andererseits zeigt $T \vDash (\Box\varphi \to \Box\Box\varphi)$. Diese beiden Implikationen zusammen mit [L5] ergeben $T \vDash (\Box\varphi \to \Box(\Box\varphi \wedge \neg\Box\varphi)) \sim (\Box\varphi \to \Box\bot)$ bzw. mit Kontraposition $T \vDash (\mathrm{con}_T \to \neg\Box\varphi)$, also mit $(*)$ $T \vDash (\mathrm{con}_T \to \varphi)$.

Falls nun $T \vDash \mathrm{con}_T$, dann gilt auch $T \vDash \varphi$. Daraus folgt einerseits mit [L1] $T \vDash \Box\varphi$ und andererseits mit $(*)$ $T \vDash \neg\Box\varphi$, d. h. T ist inkonsistent. □

Die Peano-Arithmetik wird eigentlich von allen für konsistent gehalten, außer möglicherweise von wenigen, die eine Extremposition vertreten und das Aktual-Unendliche ganz aus der Mathematik verbannen möchten (und die PA dann vermutlich eher als sinnlos denn als widersprüchlich ansehen). Im Standard-Modell $\mathscr{N} \vDash \mathsf{PA}$ gilt übrigens $\mathscr{N} \vDash \mathrm{con}_{\mathsf{PA}}$, weil $\mathscr{N}$ nur Standard-Beweise „kennt“. Insbesondere folgt $\mathsf{PA} \nvdash \neg\mathrm{con}_{\mathsf{PA}}$.

Hinsichtlich ZFC sind die Überzeugungen weniger stark. Man arbeitet seit 100 Jahren mit ZFC, ohne auf einen Widerspruch gestoßen zu sein, was viele als Plausibilitätsargument für die Konsistenz ansehen. Allerdings spielt sich ein Großteil der „tatsächlich betriebenen“ Mathematik in viel schwächerern Teiltheorien ab. Das schwächt zwar einerseits das Plausibilitätsargument ab, lässt aber andererseits darauf hoffen, dass ein in ZFC entdeckter Widerspruch durch zusätzliche oder angepasste Axiome ausgemerzt werden könnte, ohne den Kern der bisher entwickelten Mathematik zu betreffen.

Da ZFC kein Standardmodell hat, kann man $\mathsf{ZFC} \vDash \neg\mathrm{con}_{\mathsf{ZFC}}$ nicht ausschließen. Falls dies nicht der Fall ist, könnte man ZFC „verbessern“ und zu $\mathsf{ZFC}' := \mathsf{ZFC} \cup \{\mathrm{con}_{\mathsf{ZFC}}\}$ erweitern. Der Zweite Gödel'sche Unvollständigkeitssatz gilt aber auch für ZFC', d. h. es gilt dann zwar $\mathsf{ZFC}' \vDash \mathrm{con}_{\mathsf{ZFC}}$, aber dennoch $\mathsf{ZFC}' \nvdash \mathrm{con}_{\mathsf{ZFC}'}$, etc.

Die Merkwürdigkeit der Möglichkeit von $\mathsf{ZFC} \vDash \neg\mathrm{con}_{\mathsf{ZFC}}$ wird in dem folgenden Dialog ausgelotet:

> *It follows that assuming* ZF *is consistent, the "self-hating theory"* $\mathsf{ZF} + \neg\mathrm{con}_{\mathsf{ZF}}$*, or* ZF *plus the assertion of its own inconsistency, must also be consistent. So by the Completeness Theorem,* $\mathsf{ZF} + \neg\mathrm{con}_{\mathsf{ZF}}$ *has a model. What on earth could it be?*

We'll answer this question via a fictional dialogue between you and the axioms of $\mathsf{ZF} + \neg\mathrm{con}_{\mathsf{ZF}}$.

You: Look, you say ZF is inconsistent, from which it follows that there's a proof in ZF that $1 + 1 = 3$. May I see that proof?
Axioms of $\mathsf{ZF}+\neg\mathrm{con}_{\mathsf{ZF}}$**:** I prefer to talk about integers that encode proofs. (Actually *sets* that encode integers that encode proofs. But I'll cut you a break—you're only human, after all.)
You: Then show me the integer.
Axioms: OK, here it is: X.
You: What the hell is X?
Axioms: It's just X, the integer encoded by a set in the universe that I describe.
You: But what is X, as an *ordinary integer*?
Axioms: No, no, no! Talk to the axioms.
You: Alright, let me ask you about X. Is greater or smaller than a billion?
Axioms: Greater.
You: The $10^{10^{1,000,000,000}}$th Ackermann number?
Axioms: Greater than that too.
You: What's $X^2 + 100$?
Axioms: Hmm, let me see... Y.
You: Why can't I just add an axiom to rule out these weird 'nonstandard integers?' Let me try: for all integers X, X belongs to the set obtained by starting from 0 and...
Axioms: Ha ha! This is first-order logic. You're not allowed to talk about sets of objects—even if the objects are *themselves* sets.
You: Argh! I know you're lying about this proof that $1 + 1 = 3$, but I'll never catch you.
Axioms: That right! What Gödel showed is that we can keep playing this game forever. What's more, the infinite sequence of bizarre entities you'd force me to make up—X, Y, and so on—would then constitute a model for the preposterous theory $\mathsf{ZF} + \neg\mathrm{con}_{\mathsf{ZF}}$.
You: But how do you know I'll never trap you in an inconsistency?
Axioms: Because if you did, the Completeness Theorem says that we could convert that into an inconsistency in the original axioms, which contradicts the obvious fact that ZF is consis—no, wait! I'm not supposed to know that! Aaahh! *[The axioms melt in a puddle of inconsistency.]* [24]

Der Zweite Gödel'sche Unvollständigkeitssatz zeigt, dass Fixpunkte des negierten Beweisbarkeitsprädikats nicht beweisbar sind. Fixpunkte des Beweisbarkeitsprädikats sind dagegen beweisbar, wie der Satz von Löb insbesondere zeigt:

Satz 8.23 (Löb, 1955) *Für alle* $\mathscr{L}_{\mathrm{Ar}}$*-Aussagen* φ *gilt* $\mathsf{PA} \vDash (\Box(\Box\varphi \to \varphi) \to \Box\varphi)$.

Beweis: Siehe auch A cartoon proof of Löb's Theorem!

Man startet mit einem Fixpunkt ξ von $(\beta_{\mathsf{PA}}(v_0) \to \varphi)$, so dass also $\mathsf{PA} \vDash (\xi \leftrightarrow (\Box\xi \to \varphi))$. Der Beweis folgt nun zunächst dem Beweis von Satz 8.22: Mit [L1] und [L4] folgt daraus $\mathsf{PA} \vDash (\Box\xi \to \Box(\Box\xi \to \varphi))$. [L3] liefert $\mathsf{PA} \vDash (\Box\xi \to \Box\Box\xi)$. Aus beidem ergibt sich mit [L5] und [L2] die IMplikation $(*)$ $\mathsf{PA} \vDash (\Box\xi \to \Box\varphi)$.

[24] Aus: Scott Aaronson „Is P Versus NP Formally Independent?", Bull. Eur. Assoc. Theor. Comput. Sci. EATCS (2003), no. 81, 109–136. Zitiert nach `https://www.scottaaronson.com/papers/pnp.ps`

Aus $(*)$ und der Implikation $\mathsf{PA} \vDash ((\Box\xi \to \varphi) \to \xi)$ aus der Fixpunktwahl ergibt sich mit dem aussagenlogischen Schluss $(A_0 \to A_1), ((A_0 \to A_2) \to A_3) \vdash ((A_1 \to A_2) \to A_3)$ die Folgerung $\mathsf{PA} \vDash ((\Box\varphi \to \varphi) \to \xi)$. Wieder mit [L1] und [L4] folgt $\mathsf{PA} \vDash (\Box(\Box\varphi \to \varphi) \to \Box\xi)$ und dann erneut mit $(*)$ die Behauptung! □

Merkwürdigerweise folgt der Zweite Gödel'sche Unvollständigkeitssatz aus dem Satz von Löb, und zwar mit $\varphi = \bot$, denn $\mathsf{PA} \vDash (\Box(\Box\bot \to \bot) \to \Box\bot) \sim (\Box\neg\Box\bot \to \Box\bot) = (\Box\,\mathrm{con}_{\mathsf{PA}} \to \Box\bot)$. Wenn nun $\mathsf{PA} \vDash \mathrm{con}_{\mathsf{PA}} = \neg\Box\bot$, folgt $\mathsf{PA} \vDash \Box\,\mathrm{con}_{\mathsf{PA}}$, also $\mathsf{PA} \vDash \Box\bot$. Also ist PA inkonsistent.

Für diejenigen, die tiefer in die Thematik der Gödelschen Unvollständigkeitssätze einsteigen möchten, zwei Literaturhinweise:

- Raymond Smullyan hat sich in vielen Büchern auf spielerische Weise mit den Unvollständigkeitssätzen auseinander gesetzt, etwa in "Forever undecided: A puzzle guide to Gödel.", aber auch mathematisch in "Gödel's Incompleteness Theorems".
- Torkél Franzen: "Gödel's Theorem. An incomplete guide to its use and abuse." (kurze, präzise Darstellung von Gödels erstem Unvollständigkeitssatz und was er nicht aussagt).

9 Ein kurzer Einblick in die Modallogik

Den modallogischen Formalismus kann man für viele Anwendungen nutzen: neben dem ursprünglichen philosophischen Ansatz zum Beispiel für temporale Logiken (Anwendungen in der Informatik), deontische Logiken (Anwendungen in den Rechtswissenschaften), epistemische Logiken (Anwendungen in der Philosophie und der Mathematischen Logik) und zur Modellisierung von nicht-klassischen Logiken wie dem Intuitionismus.

Man kann sowohl die Aussagenlogik als auch die erststufige Prädikatenlogik (und viele andere mehr) zu einer modallogischen Variante erweitern. Der Kürze halber beschränke ich mich auf die aussagenlogische Modallogik; man sieht aber leicht, wie man die Konzepte auf eine prädikatenlogische Modallogik ausdehnen kann.

Die (aussagenlogische) Modallogik erweitert das aussagenlogische Alphabet um zwei Zeichen, die sogenannten *Modaloperatoren* (genauer: Zeichen für die Modaloperatoren):

$\Box$	genannt „*notwendig*", „*Quadrat*" oder engl.„*box*"
$\Diamond$	genannt „*möglich*", „*Raute*" oder engl. „*diamond*"

Die Regeln der aussagenlogischen Grammatik werden zunächst auf modallogische Formeln ausgedehnt. Zusätzlich gibt es die beiden Regeln (in Polnischer wie Infix-Notation):

Wenn φ eine modallogische Formel ist, dann auch $\Box\varphi$ und $\Diamond\varphi$.

Beispiel einer modallogischen Formel ist also etwa $\neg((A_2 \wedge \Diamond A_0) \to \Box(\Diamond\bot \vee \Box\neg\Diamond\Diamond A_0))$.

Ein *modallogisches System* oder kurz *eine Modallogik* definiert für modallogische Formeln eine Folgerungsbegriff, aus dem sich dann auf die aus der Aussagen- und Prädikatenlogik bekannte Weise ein Tautologie- und ein Äquivalenzbegriff ergeben. C. I. Lewis hat 1932 (mit Langford) fünf solche Systeme $\mathsf{S1}$ bis $\mathsf{S5}$ durch Kalkülregeln angegeben. Einen „kontrollierteren" Zugang liefert der semantische Ansatz von Kripke (1959):

Ein *modallogisches Modell* $\mathscr{M} = (W, Z, \beta)$ besteht aus einer nicht-leeren Menge W „*möglicher Welten*", einer *Zugangsrelation* $Z \subseteq W \times W$ und einer *Belegung* $\beta : W \times \mathbb{N} \to \{0, 1\}$, die jeder

Aussagenvariable A_i in jeder möglichen Welt $w \in W$ einen Wahrheitswert $\beta(w,i)$ zuordnet. (W,Z) ist also ein gerichteter Graph (mit Schleifen). Wenn $(w,w') \in Z$ sagt man „*w sieht w'*“.

Per Induktion ordnet man nun jeder modallogischen Formel φ in jeder Welt w eines Modells $\mathscr{M}$ einen Wahrheitswert zu bzw. definiert, wann φ in der Welt des Modells gilt, wofür ich $(\mathscr{M},w) \vDash \varphi$ schreibe:

- $(\mathscr{M},w) \vDash A_i :\iff \beta(w,i) = 1$.
- Wenn $\varphi = *\varphi_1 \dots \varphi_k$ mit einem k-stelligen Junktor $*$ aussagenlogisch zusammengesetzt ist, berechnet sich $(\mathscr{M},w) \vDash \varphi$ wie in der Aussagenlogik gemäß der $*$ zugeordneten Wahrheitswertfunktion aus den Gültigkeiten von $(\mathscr{M},w) \vDash \varphi_1, \dots, (\mathscr{M},w) \vDash \varphi_k$.
- $(\mathscr{M},w) \vDash \Box\varphi :\iff$ für alle w' mit $(w,w') \in Z$ gilt $(\mathscr{M},w') \vDash \varphi$.
- $(\mathscr{M},w) \vDash \Diamond\varphi :\iff$ es gibt ein w' mit $(w,w') \in Z$ und $(\mathscr{M},w') \vDash \varphi$.

Sei $\mathbb{M}$ eine Klasse modallogischer Modelle.

- Eine modallogische Formel φ ist eine $\mathbb{M}$-*Tautologie*, $\mathbb{M} \vDash \varphi$, wenn φ in allen Welten aller Modelle aus $\mathbb{M}$ gilt.
- Eine modallogische Formel φ ist $\mathbb{M}$-*äquivalent* zu einer modallogischen Formel ψ, $\varphi \sim_{\mathbb{M}} \psi$ wenn φ und ψ in allen Welten aller Modelle aus $\mathbb{M}$ den gleichen Wahrheitswert haben.
- Eine modallogische Formel ψ wird von modallogischen Formeln φ_i für $i \in I$ $\mathbb{M}$-*impliziert*, $\{\varphi_i \mid i \in I\} \vdash_{\mathbb{M}} \psi$, wenn in allen Welten aller Modelle aus $\mathbb{M}$, in denen alle φ_i gelten, auch ψ gilt.

Es gelten dann die üblichen Äquivalenzen, etwa sind φ und ψ genau dann $\mathbb{M}$-äquivalent, wenn $(\varphi \leftrightarrow \psi)$ eine $\mathbb{M}$-Tautologie ist.

Man identifiziert in der Regel die auf diese Weise durch $\mathbb{M}$ bestimmte Modallogik mit der Menge ihrer Tautologien. Modallogiken, die auf diese Weise entstehen, indem man bei der Wahl der Modelle Bedingungen an W und Z stellt (nicht aber an β), nennt man *normale modallogische Systeme* oder *normale Modallogiken*. Sie sind durch die folgenden Eigenschaften charakterisiert:

- Es gilt das Prinzip der uniformen und das Prinzip der äquivalenten Substitution (für $\mathbb{M}$-Äquivalenz!).
- $\mathsf{M} \vDash \top$ und $\mathbb{M} \vDash \Box\top$.
- Dualität der Modaloperatoren: $\neg\Box\varphi \sim_{\mathbb{M}} \Diamond\neg\varphi$.
- Axiom K oder *starker modus ponens*: $\mathbb{M} \vDash ((\Box\varphi \wedge \Box(\varphi \to \psi)) \to \Box\psi)$ oder äquivalent: $(\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi))$

Normale Modallogiken erhält man typischerweise auf zwei Weisen:

- Man schränkt die betrachten Modelle durch Anforderungen an die Zugangsrelation ein.
- Man nimmt zusätzlich „Axiome“, also modallogische Formeln, von denen man möchte, dass sie Tautologien werden, und betrachtet den „normalen Abschluss“, d. h. die kleinste normale Modallogik, die diese Axiome enthält.

Erstaunlicherweise ergeben Einschränkungen durch Eigenschaften der Zugangsrelation in vielen Fällen die gleiche Modallogik wie einfache Axiome:

Name	Axiom	duales Axiom	Eigenschaft: Z ist …

D	$(\Box A_0 \to \Diamond A_0)$	selbstdual	linkstotal
T	$(\Box A_0 \to A_0)$	$(A_0 \to \Diamond A_0)$	reflexiv
4	$(\Box A_0 \to \Box\Box A_0)$	$(\Diamond\Diamond A_0 \to \Diamond A_0)$	transitiv
B	$(A_0 \to \Box\Diamond A_0)$	$(\Diamond\Box A_0 \to A_0)$	symmetrisch
5 / E	$(\Diamond A_0 \to \Box\Diamond A_0)$	$(\Diamond\Box A_0 \to \Box A_0)$	„ euklidisch“ [25]
M	$(\Diamond\Box A_0 \to \Box\Diamond A_0)$	selbstdual	[*keine Entsprechung*]
L / G	$(\Box(\Box A_0 \to A_0) \to \Box A_0)$	...	transitiv und anti-wohlgeordnet
tr	$(A_0 \leftrightarrow \Box A_0)$	$(\Diamond A_0 \leftrightarrow A_0)$	die Identität
V	$\Box A_0$	$\neg\Diamond A_0$	leer

Die Entsprechung bedeutet dabei nur, dass z. B. die kleinste normale Modallogik, die 4 enthält, genau die Modallogik ist, die man aus der Betrachtung von Modellen mit transitiver Zugangsrelation erhält. Sie bedeutet nicht, dass die Zugangsrelation in Modellen, in denen überall 4 gilt, transitiv sein muss.

Die kleinste normale Modallogik erhält man, wenn man sämtliche modallogischen Modelle betrachtet. Diese Modallogik heißt ebenfalls K (für „Kripke“). Andere Systeme werden dann nach den charakteristischen Axiomen benannt, häufig mit vorangestelltem K, etwa KTB oder TB.

Zwei der wichtigsten Systeme sind die Systeme S4 und S5 von Lewis, die ihre traditionellen Namen behalten haben. Es gilt $\mathsf{S4} = \mathsf{KT4}$ und $\mathsf{S5} = \mathsf{KT5} = \mathsf{KTB4}$.

Für S5 betrachtet man also Modelle, in denen die Zugangsrelation eine Äquivalenzrelation ist (und es reicht Modelle zu betrachten, in denen jede Welt jede andere sieht). In S5 kollabieren die meisten Modalitäten, jede Aufeinanderfolge von mehreren Modaloperatoren $\triangle_1 \dots \triangle_k \varphi$ mit $k > 0$ und $\triangle_i \in \{\Box, \Diamond\}$ ist S5-äquivalent zu $\Box\varphi$ oder $\Diamond\varphi$.

Für S4 reicht es Modelle zu betrachten, in denen die Zugangsrelation eine partielle Ordnung ist. In S4 kollabieren die Modalitäten „bis auf Doppelungen“, d. h. jede Abfolge von Modaloperatoren wie oben ist S4-äquivalent zu $\Box\varphi$, $\Diamond\varphi$, $\Box\Diamond\varphi$, $\Box\Diamond\varphi$, $\Diamond\Box\Diamond\varphi$ oder $\Box\Diamond\Box\varphi$.

Für S5 und S4 gibt es Entscheidungsverfahren. Andere modallogische Systeme sind aber unentscheidbar.

Die Axiome tr. und V führen beide zu trivialen Modallogiken; beide Axiome stehen nur der Vollständigkeit halber in der Liste. Ihre Namen sind (im Gegensatz zu den andere) kein Standard. Das Axiom tr. (für „trivial“) lässt die Modallogik im Wesentlichen zur Aussagenlogik kollabieren: Jede modallogische Formel φ ist Ktr.-äquivalent zu der Formel, die aus φ entsteht, indem man alle Modaloperatoren entfernt. Das Axiom V (für „Verum“) lässt alle mit $\Box$ beginnenden Formeln KV-äquivalent zu $\top$ sein und alle mit $\Diamond$ beginnenden zu $\bot$.

Der Name des Axioms L steht für „Löb“, die Alternative G für „Gödel“. Die Benennung kommt von der modallogischen Auffassung des Beweisbarkeitsprädikats in der Peano-Arithmetik, also $\Box\varphi$ als Abkürzung für $\beta_{\mathsf{PA}}[\frac{\ulcorner\varphi\urcorner}{v_0}]$. Axiom L entspricht dann genau dem Satz von Löb 8.23. Das zweite Löb-Axiom ist dann Axiom K, das dritte Löb-Axiom ist Axiom 4 und das erste Löb-Axiom gilt auch in jeder normalen Modallogik, denn wenn $\mathbb{M} \vDash \varphi$, ist $\varphi \sim_{\mathbb{M}} \top$, also $\mathbb{M} \vDash \Box\varphi$ nach äquivalenter Substitution. Robert Solovay hat 1976 gezeigt, dass (in einem technisch zu präzisierenden Sinn) $\mathsf{KL} = \mathsf{K4L}$ genau die Formeln beschreibt, die in PA beweisbar sind mit uniformer Ersetzung der Aussagenvariablen durch beliebige $\mathscr{L}_{\mathrm{Ar}}$-Aussagen und Interpretation von $\Box$ durch das Beweisbarkeitsprädikat.

[25] Euklidisch bedeutet: Wenn eine Welt in zwei Welten sieht, sehen diese sich gegenseitig.

Diese Anwendung der Modallogik gehört zu den sogenannten *epistemischen Logiken*. Axiom 4 ist darin ein umstrittenes Prinzip (das „*KK principle*“, da man in epistemischen Logiken statt $\Box$ auch K schreibt für „*know*“).

Ein ausführliches modernes Buch zur Modallogik ist:

- Blackburn, de Rijke, Venema: *Modal Logic*, Cambidge University Press, 2001.