FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 1:
Sprache, Semantik und Syntax der Prädikatenlogik

Autor: Justus Diller

01301-7-01-S 1

Klassische Prädikatenlogik

Kurseinheit 1: Inhalt

Studienhinweise 3

Verzeichnis der definierten Begriffe und der wichtigen Sätze 7

Einleitung 9

1. Sprache, Semantik und Syntax der Prädikatenlogik

§1 Sprache und Theorie 22

1.1 Sprachen der ersten Stufe 22

1.2 Der Begriff der mathematischen Theorie 34

1.3 Zur klammerfreien Schreibweise 38

1.4 Aufgaben 43

§2 Semantik: Strukturen und Gültigkeit 44

2.1 Strukturen und Interpretation 44

2.2 Gültigkeit und Modelle 50

2.3 Einfache semantische Gesetze 55

2.4 Aufgaben 59

§3 Syntax: Herleitungen in mathematischen Theorien 61

3.1 Der Sequenzenkalkül 61

3.2 Korrektheit 65

3.3 Subformel-Eigenschaft und Schnittregel 69

3.4 Aufgaben 71

Lösungen der Übungsaufgaben 73

2. Sätze und Regeln der Prädikatenlogik

3. Vollständigkeit

4. Modelltheorie

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 1: Studienhinweise

1. Zur Funktion der Einleitung im Rahmen des Kurses

Die Einleitung ist zur begleitenden Lektüre vorgesehen. Sie soll Ihnen im Laufe des Kurses immer wieder die Zusammenhänge zwischen den Kurseinheiten verdeutlichen. Manche Passagen geben Hinweise auf Sinn und Zweck eines Lernabschnitts, werden aber erst nach der Lektüre des Abschnitts voll verständlich.

2. Lehrziele

Am Ende der Kurseinheit soll ein klares Verständnis der Grundbegriffe von Semantik und Syntax der klassischen Prädikatenlogik erreicht sein, kurz gesagt der Symbole $\models$ und $\vdash$. Beide fußen auf dem Begriff der Sprache erster Stufe, der in §1 eingeführt wird. Dazu gehören insbesondere die induktiven Definitionen 1.1.4 von Termen und 1.1.8 von Formeln. Die allgemeinen Bemerkungen über induktive Definitionen (1.1.5), induktive Beweise und rekursive Definitionen (1.1.6) sind hier insoweit wichtig, als sie die konkreten induktiv bzw. rekursiv definierten Begriffe wie Term, Formel, Interpretation (2.1.3), Herleitung (3.1.4) verdeutlichen.
Der rein sprachliche Begriff der mathematischen Theorie (1.2.1) steht im Mittelpunkt der folgenden Untersuchungen und wird in 1.2 durch mehrere Beispiele illustriert. Details dieser Beispiele werden später wichtig; zunächst soll deutlich werden, dass Theorien in der Mathematik allgegenwärtig sind und zu sehr verschiedenen Zwecken verwendet werden, wie am Ende von 1.2 ausgeführt wird.
Der semantische Begriff der Interpretation stellt in 2.1 die Verbindung zwischen Sprache und Struktur her. Er ist in 2.1.3 rekursiv auf geschlossenen Termen und Formeln definiert. Damit hängt seine Wohldefiniertheit an der Eindeutigkeit des Aufbaus von Termen und Formeln, die in 1.3.8 bewiesen ist. Das ist auch der Hauptzweck der Betrachtungen zur klammerfreien Schreibweise in 1.3.
Von den Interpretationen kommen wir mit Hilfe der Belegungen zur Gültigkeit und zum Modellbegriff (2.2.3), die beide das Symbol $\models$ verwenden. Zur Illustration führen wir in 2.2 die Modelle der Theorien aus 1.2 ein, die durchweg aus der Mathematik bekannte Strukturen sind. In 2.3 beweisen wir einige fast selbstverständliche, aber doch beweisbedürftige semantische Gesetze. Die-

se Ergebnisse werden später ständig verwendet. Ihre Beweise sind eine erste Übung im Umgang mit semantischen Begriffen.
Als syntaktisches Gegenstück zum Gültigkeitsbegriff wird in 3.1 der Herleitungsbegriff eingeführt. Es soll sofort klar werden und wird aus verschiedenen Blickwinkeln breit diskutiert (2.1.4, Anfang von §3 und von 3.2, auch 3.3), dass das Herleiten eine kombinatorische, finite Tätigkeit ist, die man einem Computer übertragen kann, wohingegen Wahrheit und Gültigkeit i. a. abstrakte, nicht-konstruktive Begriffe sind. Ein zusätzliches Merkmal des Sequenzenkalküls, den wir in 3.1 darstellen, ist seine Schnittfreiheit, die in 3.3 diskutiert wird und zur Subformeleigenschaft 3.3.5 führt. Unser schnittfreier Kalkül liefert in einfachen Fällen einfache Herleitungen, die leicht zu konstruieren bzw. zu finden sind. Der Nutzen der Schnittfreiheit beim Beweisen allgemeiner Ergebnisse wird allerdings erst im 5. Kapitel deutlich.
Für den Aufbau der Prädikatenlogik ist hier der Korrektheitssatz 3.2.1 von Bedeutung. Er stellt einen ersten Zusammenhang zwischen Syntax und Semantik, zwischen Herleitbarkeit und Gültigkeit her:
Alle in einer Theorie herleitbaren Sequenzen gelten in dieser Theorie:

$$T \vdash \Gamma : \Delta \;\Rightarrow\; T \models \Gamma : \Delta.$$

Der Beweis des Korrektheitssatzes ist einfach und naheliegend, wenn er hier auch recht ausführlich dargestellt wird. Aber schon seine Aussage verwendet (mittelbar und unmittelbar) die meisten Begriffe des 1. Kapitels. Ein genaues Studium des Satzes mit seinem Beweis und allen seinen Hintergründen ergibt einen guten Überblick über die Kurseinheit 1.
Der Korrektheitssatz gibt allerdings noch keine weitgehende Information darüber, was der Herleitungsbegriff leistet. Das ist Gegenstand der Kurseinheit 2.
Insgesamt liegt in der Kurseinheit 1 das Schwergewicht auf den Definitionen, die durch zahlreiche Beispiele illustriert werden, und noch nicht auf den Ergebnissen und ihren Beweisen. Das Verhältnis wird sich im Laufe des Kurses ausgleichen.

3. Eingangsvoraussetzungen

§1 setzt keine Vorkenntnisse voraus.

§2 verwendet bei den semantischen Begriffen (etwa: Struktur, Interpretation, Belegung) die übliche mengentheoretische Schreibweise, setzt aber keine

konkreten mathematischen Kenntnisse voraus. Bekanntschaft etwa mit dem Gruppen- und Körperbegriff erleichtert vielleicht die Lektüre der Beispiele in 2.2, ist aber für den systematischen Fortgang ohne Belang.

§3 knüpft in seinen syntaktischen Teilen 3.1 und 3.3 an §1 an und setzt wie dieser keine Vorkenntnisse voraus. In 3.2 wird auf §2 und die dort verwendete mengentheoretische Schreibweise zurückgegriffen; sonst wird nichts vorausgesetzt.

Klassische Prädikatenlogik

Kurseinheit 1: Verzeichnis der definierten Begriffe und der wichtigen Sätze

1.1.1 Grundzeichen einer Sprache L. Logische: freie, gebundene Variablen, Gleichheitszeichen $=$, Junktoren $\bot, \rightarrow$, Quantor $\forall$; nichtlogische: Funktions-, Prädikats-, Aussagezeichen, Konstanten.

1.1.2 Nennformen, Einsetzung von Nennformen

1.1.4 Terme von L

1.1.5 Induktive Definitionen; Ziffern

1.1.6 Induktive Beweise und rekursive Definitionen; vollständige Induktion; Induktionsanfang, -schritt, -voraussetzung IV

1.1.7 Primformeln von L

1.1.8 Formeln von L

1.1.9 Sprache L der ersten Stufe

1.1.12 Abkürzungen: $(A \rightarrow B), \neg, \top, \vee, \wedge, \leftrightarrow, \exists, s = t, \neq$; Klammerersparnis

1.1.13 Sequenzen, Antezedens, Sukzedens

1.1.15 Allabschluss von B, $\forall B$.

1.2.1 Mathematische Theorie T

1.2.2 Gruppentheorie T_G

1.2.3 Theorie T_R der Ringe mit 1, Körpertheorie T_K

1.2.4 Zahlentheorie Z

1.2.5 Theorien LO und DLO der (dichten) linearen Ordnungen

1.2.6 Sprache der Mengenlehre ZF

1.2.7 Algebraische und relationale Sprachen; endlich axiomatisierte, logische, offene Theorien

1.3.1 Q-Terme

1.3.2 Stellenzahl von Grundzeichen

1.3.4 Auftreten

1.3.8 **Lemma:** Eindeutigkeit des Aufbaus von Q-Termen

1.3.10 **Lemma:** Auftreten von Q-Termen

2.1.1 Struktur $\mathcal{A}$ für L; Universum (Individuenbereich); algebraische und relationale Strukturen

2.1.2 Sprache $L(\mathcal{A})$, Name von a

2.1.3 Interpretation $\mathcal{A}$

2.1.5 Wahrheitstafeln

2.2.1 $\mathcal{A}$-Belegung

2.2.3 B bzw. $\Gamma : \Delta$ gilt in $\mathcal{A}$, $\mathcal{A} \models B$, $\mathcal{A} \models \Gamma : \Delta$; $\mathcal{A}$ ist Modell von T, $\mathcal{A} \models T$; B bzw. $\Gamma : \Delta$ gilt in T, $T \models B$, $T \models \Gamma : \Delta$ (Folgerung, allgemeingültig)

2.2.4 Gruppen

2.2.5 Ringe mit 1, Körper

2.2.6 Standardmodell $\mathcal{N}$ der natürlichen Zahlen

2.2.7 (Dicht) geordnete Mengen

2.2.8 Modelle der Mengenlehre

2.2.9 Theorie der Struktur $\mathcal{A}$, $Th(\mathcal{A})$

2.3.1 **Satz:** Homomorphieprinzip

2.3.2 $\Gamma : \Delta \subset_S \Gamma_1 : \Delta_1$

3.1.2 Logische Axiome und Grundschlussregeln des Sequenzenkalküls

3.1.3 Hauptformel, Nebenformeln, Variablenbedingung

3.1.4 Herleitungen in T

3.1.5 $\Gamma : \Delta$ bzw. B ist herleitbar in T, $T \vdash \Gamma : \Delta$, $T \vdash B$; logische Herleitungen, $\vdash \Gamma : \Delta$

3.2.1 **Satz:** Korrektheit

3.3.1 Schnittregel

3.3.2 Direkte Subformel

3.3.3 Subformel

3.3.5 **Lemma:** Subformel-Eigenschaft

Einleitung

1. Adressaten
2. Eingangsvoraussetzungen
3. Einführung: Der Gegenstand der Prädikatenlogik
4. Groblehrziele
5. Historische Bemerkungen
6. Quellen des Kurses
7. Literaturliste
8. Studienhinweise

1. Adressaten

Dieser einführende Kurs „Klassische Prädikatenlogik“ wendet sich an Studierende der Mathematik und der Informatik vom zweiten Studienjahr (3. Fachsemester) an aufwärts. Er kann auch von philosophisch-logisch interessierten Teilnehmern studiert werden.

2. Eingangsvoraussetzungen

Der Kurs setzt in seinen Hauptteilen nur einfache Kenntnisse aus der naiven Mengenlehre voraus. Diese Kenntnisse werden z. B. bei der Einführung in die lineare Algebra vermittelt und auch auf der Schule bei der mathematischen

Begriffsbildung vorausgesetzt. Es ist nicht erforderlich, sich vorher oder parallel Kenntnisse über Mengenlehre zu erarbeiten. Notwendig für ein aktives Verständnis der klassischen Prädikatenlogik ist eher eine gewisse Erfahrung mit mathematischem Argumentieren und im Umgang mit mathematischen Formeln. Nützlich ist ferner die Bekanntschaft mit einigen algebraischen Begriffen wie dem Gruppen- und dem Körperbegriff. Diese Begriffe dienen als Beispiele für allgemeine Begriffsbildungen in der Prädikatenlogik. Auch wenn sie hier genau definiert und erläutert werden, wird der Kursteilnehmer den Sinn der allgemeinen Begriffe leichter und schneller einsehen, wenn er ein typisches Beispiel schon vorher kennt.

3. Einführung: Der Gegenstand der Prädikatenlogik

Die mathematische Logik in ihrer einfachsten Form dient der Präzisierung und Kodifizierung des mathematischen Schließens. Sie liefert eine formale Grundlage der Mathematik in ihrem gesamten klassischen Bestand. In Ergänzung zur landläufigen und richtigen Behauptung, dass die klassische Mathematik in ihrer modernen Formulierung auf der Mengenlehre fußt, halten wir fest, dass die Mengenlehre ebenso wie andere mathematische Theorien im Rahmen der klassischen Prädikatenlogik formuliert ist.

Wir fassen die klassische Prädikatenlogik als Theorie der mathematischen Theorien auf. Eine mathematische Theorie – wie die Gruppentheorie oder die Zahlentheorie – ist durch Axiome gegeben, die in einer bestimmten Sprache formuliert sind. Damit stellt sich eine erste, sehr einfache Aufgabe: Wir müssen mathematische Theorien im Rahmen der Logik zunächst rein sprachlich fixieren. Diese Aufgabe wird in §1 gelöst. Damit ist zwar der Begriff der mathematischen Theorie geklärt, aber sonst noch nicht viel gewonnen. Mathematiker arbeiten mit Theorien, um mathematisch interessante Sätze zu beweisen. Was ist ein Beweis? Charakteristisch für die Mathematik – im Gegensatz zu Erfahrungswissenschaften – ist hierbei, dass mathematische Beweise eine strenge, unbestreitbare Einsicht aus der Sache selbst liefern müssen. Rückgriffe auf Experimente oder Plausibilitätsbetrachtungen helfen oft weiter, lösen aber das Problem nicht. Noch so sorgfältige Messungen an Quadraten können nicht beweisen, dass $\sqrt{2}$ irrational ist, und die Fermatsche Vermutung ist schon Jahrhunderte alt, aber bewiesen wurde sie erst 1993, und zwar mit gewaltigem Aufwand.

Sprache

Ein mathematischer Beweis besteht in einer logisch korrekten Verarbeitung der gegebenen Voraussetzungen, die schließlich die Behauptung ergibt. Wir unterscheiden zwei voneinander scharf getrennte Zugänge zum Beweisbegriff.

Semantik

In der *Semantik* untersucht man die Gültigkeit von Behauptungen in passenden Strukturen. Z. B. gilt in allen Körpern die Nullteilerfreiheit: Ein Produkt ist nur dann 0, wenn einer der Faktoren 0 ist. Der klassischen Semantik zu Grunde liegt der aristotelische Wahrheitsbegriff, nach dem jede Behauptung in jeder passenden Struktur entweder wahr oder falsch ist. Dazu muss die Sprache auf der Struktur interpretiert werden. Allen Sprachteilen wird also Bedeutung beigelegt, und zwar erhalten die jeweils speziellen mathematischen Sprachteile (wie Addition, Null, kleiner, ...) ihre Bedeutung nur relativ zu der passenden Struktur, während die logischen Sprachteile (wie wenn ... dann ...) ihre Bedeutung unabhängig von der Struktur erhalten. Semantische Beweise sind dann Gültigkeitsnachweise, wobei offen bleibt, welche Hilfsmittel für diese Gültigkeitsnachweise herangezogen werden.

Syntax

Beim *syntaktischen* Zugang betrachtet man dagegen formale Herleitungen, endliche baumartige Figuren, die aus Formeln bzw. Sequenzen bestehen. Das syntaktische Beweisen, das formale Herleiten einer Behauptung ist ein endlicher kombinatorischer Prozess, der als einzige Fähigkeit voraussetzt, dass man endliche Zeichenreihen nach einigen wenigen festen Vorschriften – den Grundschlussregeln – zusammensetzen und zerlegen kann.

Der semantische Zugang ist eher verwandt mit den algebraischen Methoden der reinen Mathematik, wogegen der syntaktische Zugang offenbar der Arbeitsweise eines Computers entspricht. Herleitungsverfahren lassen sich programmieren, Gültigkeitsnachweise wegen ihrer methodischen Offenheit dagegen nicht. In jeder Theorie ist einerseits alles syntaktisch Herleitbare auch semantisch gültig (Korrektheitssatz), andererseits sind alle gültigen Formeln und Sequenzen auch herleitbar (Vollständigkeitssatz): Semantik und Syntax liefern dieselben beweisbaren Formeln und Sequenzen. Dieses Ergebnis wirkt nach unseren einleitenden Bemerkungen vielleicht überraschend; es ist auch keineswegs trivial. Es ist das erste wesentliche Ergebnis des Kurses. Mit ihm kommt die Einführung in die Grundbegriffe der Prädikatenlogik zum Abschluss, und es eröffnet den Einstieg in andere Gebiete der mathematischen Logik. Wir führen insbesondere im 4. Kapitel in die Modelltheorie und im 5. Kapitel in die Beweistheorie ein.

Die *Modelltheorie* studiert die Verbindungen zwischen einer formalen Sprache und ihren Modellen. Sie untersucht Klassen von Modellen einer Theorie und entwickelt Methoden zur Konstruktion von Modellen mit speziellen Eigenschaften. Sie ist an sich eine semantische Disziplin, gewinnt aber aus dem Vollständigkeitssatz starke Methoden, die zu überraschenden Ergebnissen, auch in der Algebra führen. In diesem Kurs werden nur elementare Beispiele für die Methoden der Modelltheorie angeführt.

Modelltheorie

Der syntaktische Zugang zum Beweisbegriff wird fortgeführt in der *Beweistheorie*. Während in der Mathematik im Allgemeinen untersucht wird, ob eine Behauptung einen Beweis besitzt, ob sie herleitbar ist, worauf die konkrete Herleitung wieder in Vergessenheit gerät, fragt man in der Beweistheorie nach der Gestalt und den kombinatorischen Eigenschaften der Herleitungen selbst. Eine solche Frage ist die nach der Schnittfreiheit: Kann man Herleitungen so gestalten, dass jeder mathematische Grundbegriff, der in einer Herleitung auftritt, auch in der hergeleiteten Formel oder in den verwendeten mathematischen Axiomen auftritt? Der Begriff der Herleitung wird wesentlich durch diese Frage beeinflusst, obwohl der Begriff der Herleitbarkeit davon nicht berührt wird. Wir definieren den Herleitungsbegriff von vornherein so, dass obige Frage positiv beantwortet wird: Im wesentlichen jeder mathematische Begriff, der in der Prämisse eines logischen Grundschlusses auftritt, tritt auch in dessen Konklusion auf.
In schnittfreien Systemen ist die Gesamtheit der Herleitungen leichter zu überschauen als in anderen. Dies führt gelegentlich zu erstaunlichen Vereinfachungen von Beweisen, auch in Konkurrenz zu semantischen Methoden.

Beweistheorie

Das Vordringen des logischen Programmierens in der Informatik hat das automatische Beweisen in das Blickfeld der Logiker gerückt. Besonders die Resolutionskalküle werden als neueres syntaktisches Gebiet zwischen Logik und theoretischer Informatik studiert. Trotz klar erkennbarer Nachbarschaft liegen die Methoden dieses Gebietes in interessanter Weise „quer“ zu denen der Beweistheorie der Prädikatenlogik. Z. B. der Begriff der Unifikation erweitert den Einsatz syntaktischer Methoden auf früher vernachlässigte Fragen.

Automatisches Beweisen

Am Schluss des Kurses können wir als Gegenstand der Prädikatenlogik ansehen: Die verschiedenen syntaktischen und semantischen Methoden an konkreten logischen Problemen gegeneinander abzuwägen und nach Möglichkeit miteinander zu verbinden.

4. Groblehrziele

Das erste Kapitel führt in die Grundlagen der klassischen Prädikatenlogik ein. In §1 legen wir fest, was wir unter einer Sprache der 1. Stufe und unter einer mathematischen Theorie allgemein verstehen wollen. In §2 wird der klassische semantische Gültigkeitsbegriff, in §3 ein schnittfreier Herleitungsbegriff, ein Gentzen-Kalkül, eingeführt. Am Ende des Kapitels soll der Begriff der mathematischen Theorie auch durch eine Reihe von Beispielen vertraut geworden sein, es soll Sicherheit in der Verwendung der semantischen Begriffe Struktur, Interpretation, Gültigkeit, Modell herrschen, und der Herleitungsbegriff soll soweit klar geworden sein, dass der Korrektheitssatz verstanden wird. Es soll deutlich geworden sein, wie mathematische Begriffe (z. B. der Gruppenbegriff) sich in den allgemeinen logischen Zusammenhang einordnen.

Das zweite Kapitel beschäftigt sich ausschließlich mit dem Herleitungsbegriff. Mit diesem relativ anspruchslosen Kapitel soll eine elementare Fertigkeit im schnittfreien Herleiten erworben werden.

Das dritte Kapitel handelt vom Vollständigkeitssatz, aus dem sich die Äquivalenz von semantischer Gültigkeit und syntaktischer Herleitbarkeit ergibt. Wegen seiner zentralen Rolle beweisen wir ihn in zwei Fassungen verschiedener Allgemeinheit mit grundsätzlich verschiedenen Methoden. Gerade durch den Vollständigkeitssatz soll der Gegensatz zwischen dem abstrakt-realen Ansatz der Semantik und dem konkret-formalen Ansatz der Syntax herausgearbeitet werden.

Im Anschluss daran gibt das vierte Kapitel eine erste Einführung in die Modelltheorie. Im Vordergrund stehen der Kompaktheitssatz und die Löwenheim-Skolem-Satzgruppe. Aus ihnen erhalten wir u. a. die Existenz von Nicht-Standard-Modellen der Zahlentheorie und von abzählbaren Modellen der Mengenlehre – das sog. Löwenheim-Skolem-Paradoxon. Sie führen uns ferner zur Frage nach der Kategorizität mathematischer Theorien.

Im fünften Kapitel treten wieder Fragen der Herleitbarkeit in den Vordergrund. Unser Gentzen-Kalkül erlaubt einfache Beweise des Herbrandschen Satzes, des Interpolations-Satzes von Craig und des Definierbarkeitssatzes von Beth. Definitorische und Skolem-Erweiterungen dagegen behandeln wir aus Gründen der Einfachheit auch mit semantischen Methoden.

Einen alternativen Herleitungsbegriff, den Resolutionskalkül, lernen wir im

sechsten Kapitel kennen. Er ist ein Standard-Hilfsmittel für das logische Programmieren und liefert für genügend einfache Sequenzen effiziente Herleitungen.

Im Ganzen soll der Kurs mit den semantischen und syntaktischen Methoden der klassischen Prädikatenlogik vertraut machen. Dazu werden die einschlägigen klassischen Ergebnisse in diesem Gebiet in einiger Vollständigkeit erarbeitet. Nicht behandelt wird das Entscheidungsproblem der Prädikatenlogik, das Methoden aus der Berechenbarkeitstheorie verwendet und von daher besser in einer Einführung in die theoretische Informatik untergebracht ist. In unserem Rahmen soll die Wichtigkeit der Prädikatenlogik für die Informatik, besonders für das logische Programmieren dadurch verdeutlicht werden, dass wir den Resolutionskalkül in unsere syntaktischen Untersuchungen einbeziehen.

5. Historische Bemerkungen

Die Anfänge der Logik gehen auf die griechische Antike zurück. Es handelt sich dabei um Aristoteles' Syllogistik und aussagenlogische Betrachtungen der megarisch-stoischen Schule.

Eine Anwendung mathematischer Methoden auf die Logik gelang zum ersten Mal George Boole um die Mitte des 19. Jahrhunderts. Gesetze einer Logik, die ausdrucksstark genug ist, um die Mathematik zu erfassen, wurden aber erstmals vor reichlich 100 Jahren formuliert. Die Entdeckung der Prädikatenlogik geht auf den Jenaer Logiker Gottlob Frege und seine 1879 erschienene „Begriffsschrift“ zurück. Allerdings wollte Frege nicht nur den logischen Rahmen für mathematische Theorien schaffen, sondern er wollte die gesamte Mathematik innerhalb der reinen Logik entwickeln. Diese Zielsetzung ist kennzeichnend für die logisch-philosophische Richtung des Logizismus. Sie führte zunächst zur Konzentration auf höhere logische Systeme. Einen Höhepunkt dieser Entwicklung bilden die „Principia Mathematica“ von A. N. Whitehead und B. Russell, die in drei Bänden ab 1910 erschienen. Die in ihnen entwickelte Typentheorie war eine Antwort auf den Widerspruch, den u. a. Russell in Freges typenfreiem System entdeckt hatte – die bekannte Russell-Antinomie von der Menge aller Mengen, die sich nicht selbst als Element enthalten.

Geburt der Prädikatenlogik

Logizismus

Alternativ dazu entwickelte David Hilbert die formale Auffassung der Mathematik, die auch der heutigen Einführung in die klassische Prädikatenlogik als

Formalismus

Theorie der mathematischen Theorien zugrunde liegt. Danach ist mathematisches Schließen nur ein kombinatorisches Spiel mit Zeichenreihen. Die Bedeutung dieser Zeichenreihen, die mit ihnen verbundene Anschauung bleibt unberücksichtigt, natürlich nur soweit die formale Logik betroffen ist: eine radikal syntaktische Einstellung. Um sicherzustellen, dass dieses Spiel mit formalen Systemen seinen Sinn und Inhalt nicht verliert, begründete Hilbert die Beweistheorie. Ihre Aufgabe war es, von mathematisch relevanten Formalismen wie der Zahlentheorie und der Analysis die Widerspruchsfreiheit mit elementaren, sog. finiten Mitteln zu zeigen. Durch das Studium kombinatorischer Eigenschaften der formalen Herleitungen sollte bewiesen werden, dass in den mathematisch relevanten Formalismen keine Widersprüche hergeleitet werden können.

Beweistheorie

Dieser Auffassung widersprach der niederländische Intuitionist L. E. J. Brouwer. Er erklärte es für widersinnig, das Arbeiten in mathematischen Formalismen erst seines Inhalts zu berauben, nur um dann auf der Metastufe, d. h. durch Reflexion über den Formalismus als ganzes, die formale Arbeitsweise durch einen Widerspruchsfreiheitsbeweis zu rechtfertigen. Seine Kritik richtete sich nicht gegen metamathematische Untersuchungen allgemein. Sie richtete sich konkret gegen die nur metamathematische Rechtfertigung von Prinzipien, die inhaltlich nicht zu rechtfertigen sind, wie der Satz vom ausgeschlossenen Dritten oder das klassische Auswahlaxiom. Statt die klassische Mathematik durch die Beweistheorie zu rechtfertigen, verlangte er, die Mathematik insgesamt auf eine neue Basis zu stellen und nur unmittelbar einsichtige, intuitive Mittel zu verwenden. Das subjektive Element, das damit in die Mathematik hineinkommt, nahm er bewusst in Kauf.

Intuitionismus

Die Kontroverse Formalismus gegen Intuitionismus hat die Entwicklung der mathematischen Logik in den zwanziger Jahren sehr vorangebracht. Trotzdem dauerte es bis 1928, bis das erste Lehrbuch zur klassischen Prädikatenlogik erschien, die „Grundzüge der theoretischen Logik“ von Hilbert und Ackermann. Und erst 1930 bewies Kurt Gödel den Vollständigkeitssatz, den wir heute als erstes wesentliches Ergebnis in einer einführenden Vorlesung ansehen. Allerdings wurden tiefliegende Ergebnisse, die sich heute zum Teil als Folgerungen aus dem Vollständigkeitssatz darstellen, bereits früher entdeckt.

erstes Lehrbuch

Vollständigkeitssatz

Durch den Vollständigkeitssatz wurde das methodische Gegeneinander von Semantik und Syntax allgemein bewusst, und beide Zweige konnten sich nun auch getrennt entwickeln und spezifischen Fragestellungen zuwenden.

Das Hilbertsche Programm erhielt einen schweren Stoß durch die Ergebnisse von Gödels Arbeit „Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I“ von 1931. Der zweite Unvollständigkeitssatz aus dieser Epoche machenden Arbeit besagt, grob gesprochen: In einer „vernünftigen“ widerspruchsfreien Theorie, die die Arithmetik umfasst, ist die Widerspruchsfreiheit dieser Theorie nicht herleitbar. Dieses überraschende Ergebnis schien dem Hilbertschen Programm den Boden zu entziehen. Paul Bernays spricht 1939 rückblickend von dem „zeitweiligen Fiasko der Beweistheorie“. Schon 1936 zeigte Gerhard Gentzen, dass eine Erweiterung der finiten Methoden um eine spezielle transfinite Induktion genügt, um jedenfalls die Widerspruchsfreiheit der gewöhnlichen Zahlentheorie zu beweisen. Gentzens Ansatz hat die Beweistheorie seither wesentlich beeinflusst.

Unvollständigkeit

Gödels Arbeit von 1931 wirkte noch in anderen Richtungen. In ihr werden erstmals die primitiv-rekursiven Funktionen eingeführt. Daran schloss sich schon Mitte der dreißiger Jahre die Formulierung des Begriffs der allgemeinrekursiven Funktion durch Church und Kleene, fast gleichzeitig mit Turings berechenbaren Funktionen. Aus diesem Anfang entwickelte sich die Rekursionstheorie, die einerseits theoretische Grundlagen für die Informatik liefert, andererseits klassische Fragen der reinen Mathematik beantwortet hat: Die Unlösbarkeit des 10. Hilbertschen Problems und die Unlösbarkeit des Wortproblems der Gruppentheorie sind wesentlich rekursionstheoretische Ergebnisse.

Rekursionstheorie

Unberührt von der Krise, die die Gödelschen Sätze in der Beweistheorie verursachten, blieben die Untersuchungen zur Semantik. Seit den zwanziger Jahren analysierte Tarski den Wahrheitsbegriff in formalisierten Sprachen und wurde so mit Mostowski und Robinson zum Begründer der Modelltheorie, die von allen Teilgebieten der Logik den stärksten Einfluss auf die reine Mathematik genommen hat. Zu ihr gehört auch die Nicht-Standard-Analysis, die das Rechnen mit unendlich kleinen Größen in der Analysis auf festen mathematischen Boden stellt.

Modelltheorie

Älter sogar als die Prädikatenlogik ist die Mengenlehre. Sie wurde von Georg Cantor 1874 entdeckt und im ersten Drittel unseres Jahrhunderts besonders von Zermelo und Fraenkel, von Neumann und Bernays zu einer axiomatischen Theorie von großer Durchsichtigkeit und Geschlossenheit entwickelt. Seitdem hat sich die Auffassung weitgehend durchgesetzt, dass die Mengenlehre die Grundlage der reinen Mathematik ist. Diese Auffassung ist durch die „Eléments de Mathématique“ von Bourbaki stark beeinflusst. 1938 zeigte Gödel die relati-

Mengenlehre

ve Konsistenz des Auswahlaxioms und der allgemeinen Kontinuumshypothese mit den anderen Axiomen der Mengenlehre, und es dauerte bis 1963, bis P. Cohen mit der Forcing-Methode die Unabhängigkeit beider Hypothesen von der Mengenlehre bewies.

Die mathematische Logik wird heute meistens in folgende Gebiete unterteilt, die alle auf die Prädikatenlogik Bezug nehmen:

Beweistheorie und Intuitionismus

Rekursionstheorie

Modelltheorie

Mengenlehre

Diese Teilgebiete existieren keineswegs getrennt nebeneinander, sondern sie beeinflussen sich gegenseitig und sind in ihren modernen Entwicklungen in vielfältiger Weise verknüpft. In steigendem Maße wirken sie auch in zahlreiche Gebiete der Mathematik und der Informatik hinein.

6. Quellen des Kurses

Der Sequenzenkalkül, den wir für den Herleitungsbegriff zu Grunde legen, geht auf G. Gentzen 1935 zurück. Aufbau und Inhalt des Kurses sind wesentlich von Vorlesungen und Arbeiten von K. Schütte beeinflusst. Diese Gedanken sind zum Teil in den ersten Kapiteln von K. Schütte 1977 zusammengefasst. Die Terminologie zur Semantik folgt Shoenfield 1967, teilweise aber auch Prestel 1986. Die Ausführungen zum Resolutionskalkül sind eine Anpassung von Teilen von Hofbauer/Kutsche 1989 an unseren Kontext.

Der Verfasser ist Professor für mathematische Logik und Grundlagenforschung an der Universität Münster.

7. Literaturliste

Die Liste bevorzugt deutschsprachige Lehrbücher. Einige ältere Werke – Monographien und Zeitschriftenartikel – sind wegen ihrer Bedeutung für die Geschichte der Logik aufgeführt. Ausführliche Literaturverzeichnisse finden Sie in einigen aufgeführten Texten.

Barwise, J. (ed.): Handbook of mathematical logic.
North Holland 1977

Bell, J., Machover, M.: A course in mathematical logic.
North Holland 1977

Ebbinghaus, H.-D., Flum, J., Thomas, W.: Einführung in die mathematische Logik. Wiss. Buchgesellschaft 1978

Frege, G.: Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Nebert, Halle 1879

Gallier, J.H.: Logic for computer science.
Harper and Row 1986

Gentzen, G.: Untersuchungen über das logische Schließen I, II
Math. Z. 39, 176 – 210, 405 – 431, 1935

Gödel, K.: Die Vollständigkeit der Axiome des logischen Funktionenkalküls.
Monatshefte Math.-Phys. 37, 349 – 360, 1930

Gödel, K.: Über formal unentscheidbare Sätze der „Principia Mathematica“ und verwandter Systeme I.
Monatshefte Math.-Phys. 38, 173 – 198, 1931

Heinemann, B., Weihrauch, K.: Logik für Informatiker.
Teubner 1992

Hermes, H.: Einführung in die mathematische Logik, Klassische Prädikatenlogik. Teubner 1963

Hilbert, D., Ackermann, W.: Grundzüge der theoretischen Logik.
Springer 1928

Hilbert, D., Bernays, P.: Grundlagen der Mathematik I, II.
Springer 1934, 1939

Hofbauer, D., Kutsche, R.-D.: Grundlagen des maschinellen Beweisens.
Vieweg 1989

Kreisel, G., Krivine, J.L.: Modelltheorie, eine Einführung in die mathematische Logik und Grundlagentheorie. Springer 1972

Lorenzen, P.: Formale Logik.
de Gruyter 1958

Pohlers, W.: Mathematische Grundlagen der Informatik.
Oldenbourg 1993

Prawitz, D.: Natural deduction, a proof-theoretical study.
Almqvist & Wiksell 1965

Prestel, A.: Einführung in die mathematische Logik und Modelltheorie.
Vieweg 1986

Richter, M. M.: Logik-Kalküle.
Teubner 1978

Schütte, K.: Proof theory.
Springer 1977

Shoenfield, J. R.: Mathematical logic.
Addison-Wesley 1967

Whitehead, A. N., Russell, B.: Principia Mathematica I, II, III.
Cambridge University Press 1910, 1912, 1913

8. Studienhinweise

Mephisto:

Mein teurer Freund, ich rat' Euch drum
Zuerst Collegium Logicum.
Da wird der Geist Euch wohl dressiert,
in spanische Stiefeln eingeschnürt,
...
Wer will was Lebendigs erkennen und beschreiben,
Sucht erst den Geist herauszutreiben,
Dann hat er die Teile in seiner Hand,
Fehlt, leider! nur das geistige Band.

(aus Goethe, Faust, 1. Teil)

Wir würden dieses Collegium Logicum nicht anbieten, wenn sich die Logik heute nicht in einem sehr viel besseren Zustand befände als zu Goethes oder Fausts Zeiten. Trotzdem: Mephisto ist immer unter uns, und seinen Hinweis, dass überm Zergliedern das geistige Band zerrissen werden kann, wollen wir ernst nehmen.

Der Kurs entspricht im Präsenzstudium einer vierstündigen einsemestrigen Vorlesung mit zweistündigen Übungen. Einen ersten wichtigen Erfolg haben Sie erreicht, wenn Sie mit Verständnis bis zum Vollständigkeitssatz und seinen ersten Anwendungen vorgedrungen sind. Die Ergebnisse des 4. und 5. Kapitels sind in der Regel Standardsätze der mathematischen Logik. Sie sollten mit dem methodischen Rüstzeug, das bis zum Vollständigkeitssatz erarbeitet ist, gut zu bewältigen sein. Das letzte Kapitel macht einen neuen Ansatz, der von Bedürfnissen der Informatik inspiriert ist. Es greift nur auf §§ 1 und 2 aus dem 1. Kapitel zurück und kann direkt im Anschluss daran studiert werden.

Bei den vielen einführenden Begriffen sollten Sie jeweils die Beispiele nachrechnen und möglichst durch kleine Abänderungen neue Beispiele konstruieren. Die meisten Übungsaufgaben sind leicht zu lösen. Ihre Lösung verläuft parallel zu einem vorgerechneten Beispiel oder ausgeführten Beweis. Die Aufgaben sollten Sie deshalb unter Rückgriff auf den vorangegangenen Text bearbeiten; sie dienen der Einübung und dem Verständnis der vorangehenden Begriffe und Ergebnisse.

Einige besonders markierte Aufgaben sind anderer Art. Sie führen über den ausgearbeiteten Text hinaus. Gelegentlich setzen sie Kenntnisse aus anderen

Gebieten der Mathematik wie der Algebra oder der Topologie voraus, die dann möglichst genau bezeichnet sind.

Während es in einem ersten Durchgang notwendig ist, den Lehrtext systematisch von vorne nach hinten durchzuarbeiten, empfiehlt sich bei einer Wiederholung das entgegengesetzte Vorgehen: Nehmen Sie sich die zentralen Ergebnisse vor – das sind gar nicht viele – und fragen Sie jedesmal, was dieses Ergebnis denn genau besagt. Damit zwingen Sie sich, die auftretenden Begriffe gewissenhaft bis in ihre Wurzeln zu verfolgen. Dann arbeiten Sie die wesentlichen Beweisideen für das Ergebnis heraus. Damit gewinnen Sie automatisch den Stammbaum der Sätze und Lemmata, die zu diesem zentralen Ergebnis führen. Wenn man einmal verstanden hat, was ein induktiver Beweis ist, laufen die meisten Induktionsbeweise von selbst ab, sobald klar ist, was man beweisen will und wonach die Induktion verläuft. Beim zweiten Durchgang schrumpft der umfangreiche Lehrtext einmal dadurch, dass Sie viele Vorüberlegungen, Beispiele und Aufgaben nicht noch einmal lesen oder nachrechnen müssen; er schrumpft zum andern noch einmal beträchtlich dadurch, dass Sie die tragenden Ideen herausarbeiten. Dann können Sie viele einfache Beweise überspringen.

Wie finden Sie die tragenden Ideen, auf die es ankommt? Wie unterscheiden Sie sie von den unumgänglichen, aber mehr technischen Abschnitten? Die Studienhinweise und Verzeichnisse geben leider nur vage Hinweise darauf. In Diskussionsteilen und Zusammenfassungen versuchen wir, Sie stets darauf hinzuweisen, worauf es ankommt. Allerdings können alle Hinweise erst den gewünschten Erfolg haben, wenn Sie sich wenigstens einmal durch das Kursmaterial hindurchgearbeitet haben. Für die Suche nach dem geistigen Band, von dem Mephistopheles spricht, gibt es kein Patentrezept.

Kapitel 1

Sprache, Semantik und Syntax der Prädikatenlogik

§1 Sprache und Theorie

1.1 Sprachen der ersten Stufe

Zu jeder Sprache gehört zunächst ihr Vokabular. Handelt eine Sprache von einem Teilgebiet der Mathematik wie z. B. der Gruppentheorie, so besteht dieses Vokabular außer aus rein logischen Sprachteilen wie „wenn – dann“ und „für alle“ aus den Begriffen, die für dieses Teilgebiet charakteristisch sind. Diese sind meistens durch logische Verknüpfung anderer, einfacherer Begriffe definiert und können als Abkürzungen aufgefasst werden, mit Ausnahme der Grundbegriffe des Teilgebiets. Die Grundbegriffe sind unanalysierte Bestandteile des Teilgebiets (auch wenn sie in anderen, übergreifenden Teilgebieten einer Analyse zugänglich sein können). Im Fall der Gruppentheorie ist der zentrale Grundbegriff die Gruppenoperation, die eine 2-stellige Funktion ist.
Namen für die Grundbegriffe müssen in die Sprache des Teilgebiets aufgenommen werden. Diese Namen sind die nicht-logischen Grundzeichen der Sprache. Wir haben oben die logischen Sprachteile erwähnt. Zu diesen zählen die *Junktoren* $\top$ (verum, das Wahre) und $\bot$ (falsum, das Falsche), $\neg$ (nicht), $\wedge$ (und), $\vee$ (oder), $\rightarrow$ (wenn – dann) und $\leftrightarrow$ (genau dann – wenn) und die *Quantoren* $\forall$ (für alle) und $\exists$ (es gibt), aber auch das *Gleichheitszeichen* $=$ und die *Variablen*. Die Bezeichnung „Junktor“ (von lat.: iungere - verbinden) rührt daher, dass im allgemeinen ein solches Zeichen mehrere Aussagen miteinander verbindet. Dies ist bei $\wedge, \vee, \rightarrow, \leftrightarrow$ auch der Fall, nicht aber bei $\top, \bot, \neg$.

Das verum $\top$ und das falsum $\bot$ (ein kopfstehendes $\top$) stehen selbst schon für Aussagen. $\top$ steht für die absolut wahre, unbezweifelbare, triviale Aussage, $\bot$ steht für die absolut falsche, absurde Aussage. Sie heißen deshalb auch *Aussagekonstanten*. Wie sich zeigen wird, lassen sich in der klassischen Logik alle Junktoren und Quantoren allein durch $\bot, \to, \forall$ definieren, so dass wir nur diese als Grundzeichen betrachten. Die Gleichheit ist ein 2-stelliges Prädikat wie die Kleiner-Relation. Sie tritt aber in allen hier betrachteten Teilgebieten der Mathematik auf, so dass das Gleichheitszeichen $=$ als logisches Grundzeichen angesehen wird.
Die Rolle der Variablen ist vielleicht am schwersten einzusehen. Deshalb hierzu ein Beispiel aus den rationalen Zahlen $\mathbb{Q}$. In $\mathbb{Q}$ – wie in jedem Körper – gilt das Gesetz

$$(*) \qquad a \neq 0 \to \exists y \quad a \cdot y = 1,$$

wobei $a \neq 0$ abkürzend für $\neg a = 0$ steht. In $(*)$ treten also nacheinander die logischen Symbole $\neg, =, \to, \exists, =$ auf. Der Punkt $\cdot$ steht für die Multiplikation, die 1 für die rationale Zahl 1. Die Buchstaben a, y sind Variablen, und zwar tritt a *frei* und y *gebunden* auf. Für die freie Variable a kann man beliebige rationale Zahlen einsetzen, während man für die gebundene Variable y direkt nichts einsetzen kann; man kann sie nur umbenennen: $\exists y \quad a \cdot y = 1$ und $\exists z \quad a \cdot z = 1$ bedeuten dasselbe. Dass $(*)$ in $\mathbb{Q}$ *gilt*, meint gerade, dass $(*)$ bei jeder Einsetzung einer rationalen Zahl für a wahr wird (in $\mathbb{Q}$). Liest man $(*)$ aber als Aussage über ganze Zahlen, so wird $(*)$ nur bei den Einsetzungen $1, -1, 0$ für a wahr, sonst falsch: Ein y, das etwa die Gleichung $2 \cdot y = 1$ löst, gibt es zwar in $\mathbb{Q}$, aber nicht in $\mathbb{Z}$, weil 2 kein Teiler von 1 ist. Die Wahrheit einer Formel hängt also sowohl von der Struktur ab, in der man sie interpretiert (hier $\mathbb{Q}$ bzw. $\mathbb{Z}$), als auch von den Objekten, die man für die freien Variablen einsetzt, nicht aber von den gebundenen Variablen.

Formal sind freie und gebundene Variablen leicht zu unterscheiden: Gebundene Variablen treten unmittelbar hinter den Quantoren auf, freie nicht.

In unserer formalen Sprache benutzen wir von vornherein verschiedene Zeichen für freie und für gebundene Variablen.

1.1.1 Definition *Grundzeichen einer Sprache L der 1. Stufe sind:*

1. *abzählbar unendlich viele freie (Objekt-)Variablen, mitgeteilt durch a,b,c (auch mit Indizes),*

2. *abzählbar unendlich viele gebundene (Objekt-)Variablen, mitgeteilt durch x,y,z (auch mit Indizes),*

3. *zu jeder natürlichen Zahl $n \geq 0$ die n-stelligen Funktionszeichen, mitgeteilt durch f, g, f^n, g^n (auch mit Indizes), und die n-stelligen Prädikatszeichen, mitgeteilt durch p, q, p^n, q^n (auch mit Indizes), unter den 2-stelligen Prädikatszeichen das Gleichheitszeichen $=$,*

4. *die Junktoren $\perp$ (falsum), $\rightarrow$ (wenn – dann) und der Quantor $\forall$ (für alle).*

Bemerkungen. a. In Definition 1.1.1 wird nicht gesagt, was die Grundzeichen von L sein sollen, weil es darauf in diesem Text nicht ankommt.

b. Die Anzahl der n-stelligen Funktions- und Prädikatszeichen ist beliebig bis auf die Forderung, dass $=$ in jeder Sprache vorkommen soll.

c. Die nullstelligen Funktionszeichen heißen *Konstanten*; die nullstelligen Prädikatszeichen heißen *Aussagezeichen*.

d. Die Grundzeichen $=, \perp, \rightarrow, \forall$ und die Variablen heißen *logische Grundzeichen*, die anderen *nicht-logische Grundzeichen*.

Die Grundzeichen allein ergeben noch keine Sprache. Aus ihnen setzen wir *Terme* (Rechenausdrücke) und *Formeln* zusammmen. Das sind spezielle endliche Zeichenreihen aus Grundzeichen. Um den schon erwähnten Einsetzungsprozess bequem handhaben zu können, führen wir den allgemeinen Begriff der *Nennform* ein.

1.1.2 Definition *Nennformen von L sind endliche Reihen (Verkettungen) von Grundzeichen von L und weiteren Zeichen*

$$*_1, *_2, \ldots, *_n, \ldots,$$

den Nennzeichen. Nennformen werden durch F, G, H, t (auch mit Indizes) mitgeteilt.
Sind $F, G_1, \ldots, G_n$ Nennformen, so bezeichnet

$$F(G_1, \ldots, G_n)$$

(lies: F von G_1 bis G_n) die Zeichenreihe, die aus F hervorgeht, wenn man jedes Vorkommen von $*_i$ in F simultan durch G_i ersetzt für $i = 1, \ldots, n$.
Wir schreiben $F \equiv G$, wenn die durch F, G mitgeteilten Nennformen als Zeichenreihen identisch sind (lies: F ist G, F ist identisch mit G).

Beispiele.

1. $*_1 = *_2(r, s) \equiv r = s$
   $*_2 = *_1(r, s) \equiv s = r$, also
   $*_1 = *_2 \to *_2 = *_1(r, s) \quad \equiv \quad r = s \to s = r$

2. $F(*_1, \ldots, *_n) \equiv F$, denn wenn man jeweils $*_i$ durch $*_i$ ersetzt, ändert man nichts. Speziell:
   $F(*_1) \equiv F$, z.B. $*_2(*_1) \equiv *_2$.

3. $*_i(G_1, \ldots, G_n) \equiv G_i$ für $1 \leq i \leq n$, denn in $*_i$ tritt nur $*_i$ auf, und das wird durch das i-te Argument G_i ersetzt. Speziell:
   $*_1(G) \equiv G$, z.B. $*_1(*_2) \equiv *_2$.

4. Ist $F \equiv F_1F_2$, so ist $F(G) \equiv F_1(G)F_2(G)$, denn die Einsetzung von G für $*_1$ ist überall in F, also sowohl in F_1 als auch in F_2 auszuführen.

1.1.3 Lemma *Das Einsetzen von Nennformen für $*_1$ ist assoziativ:*

$$F(G)(H) \equiv F(G(H))$$

Beweis. Zunächst ist $*_1(G)(H) \equiv G(H) \equiv *_1(G(H))$, und falls $*_1$ nicht in F auftritt, ist $F(G)(H) \equiv F(H) \equiv F \equiv F(G(H))$. Nun hat jede Nennform F eine Gestalt $F \equiv F_0 *_1 F_1 *_1 \ldots *_1 F_n$, wobei $*_1$ in den F_i nicht auftritt und die F_i auch die leere Nennform sein können. Nach Beispiel 4 und dem Gesagten ist dann

$$F(G)(H) \equiv F_0G(H)F_1G(H)\ldots G(H)F_n \equiv F(G(H)).$$

In der Mathematik arbeitet man ständig mit Rechenausdrücken oder *Termen*. Terme, die beim Rechnen mit natürlichen Zahlen vorkommen, sind etwa $2 + a, (a + b) \cdot (a - b), 3^b, 5!$. Die Rechenoperationen werden teils zwischen die zu verknüpfenden Teilterme geschrieben (wie $+$ und $\cdot$), teils dahinter (wie die Fakultät !), teils gar nicht (wie bei der Potenz). Schreibt man die Operationen zwischen die Terme, so muss man Klammern verwenden, um etwa $(a + b) \cdot c$

und $a + (b \cdot c)$ unterscheiden zu können.

Wir wollen den Termbegriff allgemein studieren. Es empfiehlt sich dann, die Zeichen für die Rechenoperationen (Funktionen) einheitlich zu den Argumenten zu stellen, und zwar nicht dazwischen (was soll das bei ein- und dreistelligen Operationen auch heißen?). Wir schreiben das Funktionszeichen einheitlich vor die Terme, die das Funktionszeichen verknüpft. Unsere obigen Ausdrücke gehen dann über in $+2a, \cdot + ab - ab, pot3b, !5$, wenn wir für die Potenz die Silbe *pot* als Funktionszeichen verwenden. Und die Ausdrücke $(a + b) \cdot c$ und $a + (b \cdot c)$, die sich nur durch Klammersetzung unterscheiden, gehen über in $\cdot + abc$ und $+a \cdot bc$: Auf Klammern kann man hier verzichten.

1.1.4 Induktive Definition *der Terme von L*

1. *Jede freie Variable ist ein Term von L.*
2. *Ist f ein n-stelliges Funktionszeichen von L und sind $t_1, \ldots, t_n$ Terme von L, so ist $ft_1 \ldots t_n$ ein Term von L.*

Terme werden mitgeteilt durch r, s, t (auch mit Indizes). Sie sind intendiert als Bezeichnungen für Objekte.

Hier begegnen wir zum ersten Mal einer induktiven Definition. Die Bezeichnung „induktiv“ deutet an, dass der Definition der Terme von L eine *Minimalklausel* in einer der folgenden Fassungen anzufügen ist:

> *M. Alle Terme von L erhält man nach 1. und 2.*
>
> *M'. Die Menge der Terme von L ist die kleinste Menge mit den Eigenschaften 1. und 2.*

Wenn eine Definition ausdrücklich als induktive Definition bezeichnet wird, braucht die Minimalklausel nicht aufgeführt zu werden.

1.1.5 Induktive Definitionen

In der mathematischen Logik ebenso wie in der Informatik werden einige wesentliche Begriffe durch *induktive Definitionen* eingeführt. Eine induktive Definition ist im Wesentlichen eine Vorschrift für ein Konstruktionsverfahren, das nach und nach alle Objekte erzeugt, die unter den durch sie definierten Begriff fallen. Das einfachste Beispiel hierfür ist die folgende Definition der *Ziffern*, die man als Namen für die natürlichen Zahlen verwendet. Ausgehend von den

Grundzeichen 0 (Null) und S (Nachfolger, successor) erhält man die Ziffern wie folgt:

Beispiel. Induktive Definition *der Ziffern*

1. 0 *ist eine Ziffer.*

2. *Wenn n eine Ziffer ist, ist auch Sn eine Ziffer.*

Die Minimalklausel hierzu lautet:

> *M. Alle Ziffern erhält man nach 1. und 2.*

Hiernach ist 0 eine Ziffer nach 1. Damit ist die Prämisse von 2. für $n = 0$ erfüllt; also ist $S0$ eine Ziffer, damit auch $SS0, SSS0$ usw. Bekanntlich schreibt man meist 1 für $S0$, 2 für $SS0$, usw.

Werfen wir noch einen Blick auf die induktive Definition der Terme. Nach 1.1.4,2 sind Konstanten Terme. Ein „Konstruktionsanfang“ steckt also nicht nur in 1., wonach freie Variable Terme sind, sondern auch in 2. für den Fall $n = 0$. Terme, in denen keine freien Variablen auftreten, werden ausschließlich nach 2. erzeugt. Die Ziffern sind übrigens spezielle solche Terme, wenn die Funktionszeichen 0 und S zur Sprache L gehören.

1.1.6 Induktive Beweise und rekursive Definitionen Viele Behauptungen über induktiv definierte Begriffe lassen sich induktiv beweisen. Das allgemeine *Induktionsprinzip* kann man etwa so formulieren:

Wenn sich eine Eigenschaft E über alle Schritte einer induktiven Definition „vererbt“, so trifft E auf alle Objekte zu, die unter den induktiv definierten Begriff fallen.

Das Induktionsprinzip ist eine direkte Folge der Minimalklausel. Den einfachsten Spezialfall des Prinzips liefert wieder die induktive Definition der Ziffern. Wir schreiben die „Vererbung“ einer Eigenschaft E über die Definitionsschritte 1. und 2. explizit hin und erhalten sofort:

Beispiel. *Vollständige Induktion*
Wenn

1. *(Induktionsanfang) eine Eigenschaft E auf* 0 *zutrifft und*

2. *(Induktionsschritt) für jede Ziffer n gilt: Wenn E auf n zutrifft, dann trifft E auch auf Sn zu,*

dann trifft E auf jede Ziffer zu.

Das Prinzip der vollständigen Induktion „folgt“ der Konstruktion der Ziffern aus 1.1.5. Es ist eine direkte Konsequenz der Minimalklausel, dass man mit 1. und 2. aus 1.1.5 nach und nach alle Ziffern erreicht. Ebenso sorgen Induktionsanfang und Induktionsschritt dafür, dass die Eigenschaft E nach und nach auf jede Ziffer zutrifft. Im Induktionsschritt macht man die Voraussetzung: E treffe auf eine beliebige Ziffer n zu. Dies nennt man die *Induktionsvoraussetzung*, abgekürzt IV. Allein aus ihr muss man erschließen, dass E dann auch auf Sn zutrifft. Nur dann hat man den Induktionsschritt bewiesen.

Zu jeder induktiven Definition gehört ein Induktionsprinzip als direkte Konsequenz der jeweiligen Minimalklausel. Das Induktionsprinzip, das der induktiven Definition der Terme „folgt“, ist die *Induktion nach dem Aufbau der Terme*, die man nun leicht selber formuliert.

Zu den induktiven Definitionen, die ihre Objekte in *eindeutiger* Weise erzeugen, gehört aber auch ein Prinzip der *rekursiven Definition*. Es ordnet den Objekten der induktiven Definition nach und nach Objekte aus einem anderen Bereich zu (z. B. Zahlen, Wahrheitswerte). Eine rekursive Definition definiert also eine Funktion auf einem induktiv definierten Bereich. Die eindeutige Erzeugung der Objekte dieses Bereichs garantiert dabei, dass die rekursiv definierte Funktion wohldefiniert ist, also an jeder Stelle auch nur einen einzigen Wert hat.

Die Ziffern werden von ihrer induktiven Definition in eindeutiger Weise erzeugt: Die 0 kommt nur nach 1. zustande, und eine Ziffer Sn entsteht notwendig nach 2. aus ihrem Vorgänger n. Also kann man auf den Ziffern (und auf den natürlichen Zahlen) rekursive Definitionen vornehmen.

Beispiele. 1. Die Fakultät ! ist rekursiv definiert durch

$$0! = 1 \text{ und } (Sn)! = n! \cdot Sn.$$

2. Die endliche Summation $\sum_{i<n} x_i$ ist rekursiv definiert durch

$$\sum_{i<0} x_i = 0 \text{ und } \sum_{i<Sn} x_i = \sum_{i<n} x_i + x_n.$$

Dass die Terme von ihrer induktiven Definition 1.1.4 ebenfalls eindeutig erzeugt werden, ist nicht so einfach zu erkennen. Wir beweisen es in §1.3.

Was den Termen recht ist, ist den Formeln billig.
Ausdrücke wie $a + b = b + a$ und $5 < 3 + a$ sind gut lesbar im Kontext der Arithmetik. In der Prädikatenlogik, in der man es nicht nur mit $=$ und $<$, sondern auch mit anderen Prädikatszeichen beliebiger Stellenzahl zu tun hat, schreibt man zweckmäßig das Gleichheitszeichen an die Spitze von Gleichungen, das Kleinerzeichen an die Spitze von Ungleichungen. Obige Ausdrücke gehen dann über in $= +ab + ba$ und in $< 5 + 3a$. Auch solche, zugegebenermaßen ungewohnten, Ausdrücke kann man durchaus von links nach rechts lesen, etwa: „gleich ist die Summe von a und b zu der Summe von b und a“ und „kleiner ist 5 als die Summe von 3 und a“.

1.1.7 Definition *Primformeln von L sind*

1. *das falsum $\bot$,*
2. *die Zeichenreihen $pt_1 \dots t_n$, wenn p ein n-stelliges Prädikatszeichen von L ist und $t_1, \dots, t_n$ Terme von L sind.*

Primformeln werden mitgeteilt durch P, Q (auch mit Indizes).

1.1.8 Induktive Definition *der Formeln von L*

1. *Jede Primformel von L ist eine Formel von L.*
2. *Sind A, B Formeln von L, so ist auch $\to AB$ eine Formel von L.*
3. *Ist $F(a)$ eine Formel von L, in der die gebundene Variable x nicht auftritt, so ist $\forall x F(x)$ eine Formel von L.*

Formeln werden mitgeteilt durch A, B, C, D, E, F (auch mit Indizes). Sie sind intendiert als Bezeichnungen für Sachverhalte.

Beispiele.

1. $\forall x = xa$ ist eine Formel, weil $= ba$ eine Formel ist, in der x nicht auftritt.
2. $\forall y \forall x = xy$ ist eine Formel, falls $x \not\equiv y$ ist.
3. $\forall x = xy, \forall x \forall x = xx$ sind keine Formeln.

4. $\rightarrow \forall x F(x) \forall x G(x)$ ist eine Formel, falls $F(a), G(a)$ Formeln sind, in denen x nicht auftritt.

1.1.9 Definition *Eine Sprache L der ersten Stufe besteht aus den Grundzeichen, den Termen und den Formeln von L.*

Um verschiedene Sprachen erster Stufe in einfacher Weise vergleichen zu können, treffen wir einige Verabredungen, die in natürlichen Sprachen allerdings nicht immer eingehalten werden.

1.1.10 Konvention

a. Kein Grundzeichen einer Sprache ist zusammengesetzt aus Grundzeichen derselben oder einer anderen Sprache:
Grundzeichen sind atomar.

b. In allen Sprachen werden dieselben Zeichen als freie Variablen und dieselben Zeichen als gebundene Variablen benutzt, und alle Sprachen verwenden die Grundzeichen $\bot, \rightarrow, \forall, =$:
Alle Sprachen verwenden dieselben logischen Grundzeichen in denselben Rollen.

c. Wird in einer Sprache L ein Zeichen als n-stelliges Funktions- bzw. Prädikatszeichen verwendet, so wird es in allen Sprachen L' höchstens so verwendet:
Alle Sprachen verwenden nicht-logische Grundzeichen, wenn überhaupt, dann in derselben Rolle.

1.1.11 Lemma *Haben zwei Sprachen L und L' dieselben nicht-logischen Grundzeichen, so sind sie identisch.*

Beweis. L und L' haben nach Teil b. der Konvention dieselben logischen und nach Voraussetzung dieselben nicht-logischen Grundzeichen. Also haben sie überhaupt dieselben Grundzeichen. Da diese Grundzeichen nach b. und c. in L und L' in denselben Rollen auftreten, stimmen die Definitionen der Terme, Primformeln und Formeln von L und L' buchstäblich überein. Also ist $L = L'$.

In diesem Sinne ist eine Sprache erster Stufe schon allein durch ihre nicht-logischen Grundzeichen eindeutig festgelegt. In vielen Lehrbüchern wird deshalb eine Sprache L direkt mit der Menge der nicht-logischen Grundzeichen von L identifiziert. Allerdings müssen auch dann die Bedingungen der Konvention 1.1.10 erfüllt sein.

Wir haben bei der Definition von Termen und Formeln eine klammerfreie Schreibweise benutzt, bei der das jeweils *regierende* Grundzeichen am Anfang der Zeichenreihe steht. Es steht $= st$ für die Gleichung „s ist gleich t“ und $\rightarrow AB$ für die Implikation „wenn A, dann B“, was unseren Gewohnheiten widerspricht. Ferner verkürzt die Tatsache, dass wir möglichst wenige logische Partikel verwenden (nur zwei Junktoren und einen Quantor) viele Beweise ganz beträchtlich, besonders die Beweise, die durch Induktion nach dem Aufbau der Formeln geführt werden. Beides reduziert allerdings die Lesbarkeit längerer Formeln sehr. Wir führen deshalb folgende Abkürzungen ein, die unseren Lesegewohnheiten entgegenkommen. Solche Abkürzungen brauchen die Mitteilung einer Zeichenreihe nicht unbedingt zu verkürzen; wichtig ist, dass durch sie keine neuen Grundzeichen in die Sprache eingeführt werden.

1.1.12 Abkürzungen

$(A \rightarrow B)$	steht für	$\rightarrow AB$
$\neg A$	steht für	$(A \rightarrow \bot)$
$\top$	steht für	$\neg \bot$
$(A \vee B)$	steht für	$(\neg A \rightarrow B)$
$(A \wedge B)$	steht für	$\neg(A \rightarrow \neg B)$
$(A \leftrightarrow B)$	steht für	$((A \rightarrow B) \wedge (B \rightarrow A))$
$\exists x F(x)$	steht für	$\neg\forall x \neg F(x)$
$s = t$	steht für	$= st$
$s \neq t$	steht für	$\neg s = t$

Zur Klammerersparnis vereinbaren wir außerdem noch:

a. Außenklammern werden meist fortgelassen.

b. $\neg$ bindet am stärksten.

c. $\wedge, \vee$ binden stärker als $\rightarrow, \leftrightarrow$.

d. $\rightarrow$ bindet stärker als $\leftrightarrow$.

e. Von gleichen Junktoren bindet der rechte stärker als der linke (*Rechtsklammerung*).

Beispiele.

1. $\neg A$ steht für $\rightarrow A \bot$;
   $\exists x F(x)$ steht für $\rightarrow \forall x \rightarrow F(x) \bot\bot$.

2. $A \vee B \to B \vee A$ steht für $\to\to\to A \perp B \to\to B \perp A$

3. $r = t \to s = t \to r = s$ steht für $\to= rt \to= st = rs$;

4. $r = t \wedge s = t \to r = s$ steht für $\to (= rt \wedge = st) = rs$,
   also für $\to \neg \to= rt \neg = st = rs$,
   also für $\to\to\to= rt \to= st \perp\perp= rs$.

Die inhaltlich bedeutungsgleichen und abgekürzt gleichlangen Zeichenreihen unter 3. und 4. sind also Abkürzungen für verschieden lange Formeln. Allgemein ist für $n > 1$

$$A_1 \to A_2 \to \ldots \to A_n \to B$$

Abkürzung für eine kürzere Formel als die gleichbedeutende Zeichenreihe

$$A_1 \wedge A_2 \wedge \ldots \wedge A_n \to B.$$

Die Beispiele machen deutlich, dass die Lesbarkeit von Formeln durch die Verwendung von Abkürzungen wesentlich verbessert wird.

Wir werden in §3 nicht Formeln, sondern Sequenzen herleiten.

1.1.13 Definition *Eine Sequenz von L ist ein Paar von endlichen Formelmengen Γ, Δ von L, das wir in der Form $\Gamma : \Delta$ schreiben und „wenn Γ, dann Δ“ lesen. Dabei heißt Γ das Antezedens und Δ das Sukzedens der Sequenz. Im Zusammenhang mit Sequenzen schreibt man abkürzend*

$$\Gamma, \Delta \quad \textit{für} \quad \Gamma \cup \Delta \quad \textit{und} \quad D \quad \textit{für} \quad \{D\}.$$

Ist das Antezedens leer, so kann man es fortlassen und Δ für $\emptyset : \Delta$ schreiben. Die leere Sequenz $\emptyset : \emptyset$ wird oft mit $\Box$ abgekürzt.

Gelegentlich bezeichnen wir mit Γ, Δ auch endliche Mengen von Nennformen. $\Gamma(t)$ steht dann offenbar für die endliche Menge $\{F(t) | F \in \Gamma\}$.
Die intendierte Bedeutung von $\Gamma : \Delta$ ist: Aus allen Formeln aus Γ zusammen folgt mindestens eine Formel aus Δ.

Beispiele. 1. $\Gamma, C : A, B$ steht für $\Gamma \cup \{C\} : \{A, B\}$.
2. $\Gamma, A : \Delta$ und $A, \Gamma : \Delta$ und $\Gamma, A, A : \Delta$ bezeichnen dieselbe Sequenz $\Gamma \cup \{A\} : \Delta$.

3. Ist $\Gamma = \{*_1 = *_1, *_1 = t, t = t\}$, so ist $\Gamma(a) = \{a = a, a = t, t = t\}$ und $\Gamma(t) = \{t = t\}$.
Γ und $\Gamma(t)$ brauchen also nicht gleichmächtig zu sein.

Unter den Termen, Formeln und Sequenzen spielen die, in denen keine freien Variablen auftreten, öfters eine besondere Rolle.

1.1.14 Definition Es sei X eine Nennform oder eine Menge von Nennformen. Dann bezeichnet $FV(X)$ die Menge der freien Variablen, die in X auftreten, und $BV(X)$ die Menge der gebundenen Variablen, die in X auftreten. Tritt in X keine freie Variable auf, ist also $FV(X)$ leer, so heißt X *geschlossen*. Geschlossene Formeln heißen auch *Sätze*. Tritt in X kein Quantor auf, ist also $BV(X)$ leer, so heißt X *quantorenfrei*.

Terme sind stets quantorenfrei. Eine Sequenz ist genau dann geschlossen, wenn sie nur aus Sätzen besteht.

Beispiele: 1. Die Ziffern aus 1.1.5 sind geschlossene Terme.
2. $FV(Sb = b) = FV(\forall x\ x = b) = \{b\}$.
3. $FV(\forall x\ x = 0) = \emptyset$, und $\forall x\ x = 0$ ist ein Satz.

Eine Methode, aus einer Formel einen Satz „ähnlicher Bedeutung“ zu gewinnen, ist der Übergang zu einem *Allabschluss*.

1.1.15 Definition Es sei B eine Formel mit $FV(B) = \{a_1, \ldots, a_n\}$ mit paarweise verschiedenen a_i und $n \geq 0$. Dann hat B eine Gestalt $F(a_1, \ldots, a_n)$ mit $FV(F) = \emptyset$. Dann heißt jede Formel

$$\forall x_1 \ldots \forall x_n F(x_1, \ldots, x_n)$$

ein *Allabschluss* von B und wird mit $\forall B$ bezeichnet.

In einem solchen Allabschluss sind die gebundenen Variablen $x_1, \ldots, x_n$ paarweise verschieden und treten in F und in B nicht auf. Nur ihre Anzahl n ist durch B eindeutig bestimmt als die Anzahl der freien Variablen in B. Die Nennform F ist nur im Fall $n \leq 1$ durch B bestimmt.
Im Fall $n = 0$ ist B ein Satz, und es ist $\forall B \equiv B$. Im Fall $n > 0$ besitzt B unendlich viele Allabschlüsse. $\forall B$ ist stets ein Satz.

Beispiel. $a = b \quad \equiv \quad *_1 = *_2(a, b) \quad \equiv \quad *_2 = *_1(b, a)$. Also sind alle Formeln $\forall x \forall y \quad x = y$ und $\forall x \forall y \quad y = x$ Allabschlüsse von $a = b$.

1.2 Der Begriff der mathematischen Theorie

Um ein Teilgebiet der Mathematik festzulegen, müssen die Grundbegriffe und damit die Sprache dieses Teilgebietes festgelegt werden. Aber das allein genügt nicht. Zusätzlich formuliert man in dieser Sprache die Grundgesetze oder Axiome des Teilgebietes. In den Axiomen werden die grundlegenden Voraussetzungen formuliert, die man über die Grundbegriffe in dem betreffenden Teilgebiet macht. Formal betrachtet, sind die Axiome völlig beliebige Sätze der Sprache.

1.2.1 Definition Eine *Theorie der ersten Stufe* oder eine *mathematische Theorie* T ist ein Paar $(L(T), Ax(T))$, wobei $L(T)$ eine Sprache der ersten Stufe ist, die *Sprache von* T, und $Ax(T)$ eine Menge von Sätzen von $L(T)$ ist, die Menge der *(nicht-logischen) Axiome* von T.

Wir führen als Beispiele einige übliche mathematische Theorien an, auf die wir später öfters zurückkommen werden.

1.2.2 Definition Die Gruppentheorie T_G

Die nicht-logischen Zeichen der *Sprache der Gruppentheorie* sind

ein 0-stelliges Funktionszeichen (Konstante) e,
ein 1-stelliges Funktionszeichen f,
ein 2-stelliges Funktionszeichen g.

Wir schreiben t^{-1} für ft und $(s \circ t)$ für gst und lassen äußere Klammern fort.

Die *Axiome der Gruppentheorie* sind dann

G1. $\forall x \forall y \forall z\ (x \circ y) \circ z = x \circ (y \circ z)$
G2. $\forall x\ x \circ e = x$
G3. $\forall x\ x \circ x^{-1} = e$,

die abkürzend stehen für

G1. $\forall x \forall y \forall z = ggxyzgxgyz$
G2. $\forall x = gxex$
G3. $\forall x = gxfxe$.

Um die Lesbarkeit der Axiome zu erhöhen, lassen wir im folgenden oft die Allquantoren am Formelanfang weg, die nur der Herstellung von Allabschlüssen dienen. Wir erhalten dann etwa

G1. $(a \circ b) \circ c = a \circ (b \circ c)$
G2. $a \circ e = a$
G3. $a \circ a^{-1} = e$,

Diese Formulierung der Gruppentheorie mit drei Funktionszeichen ist unter Logikern allgemein üblich und unter Algebraikern weit verbreitet. Manche Algebraiker bevorzugen jedoch die Formulierung mit dem einzigen Funktionszeichen $\circ$, wodurch dann das zweite und dritte Axiom zusammen folgende Fassung erhalten:

$$\exists u \forall x (x \circ u = x \quad \wedge \quad \exists y \; x \circ y = u).$$

Wegen der Schachtelung wechselnder Quantoren $\exists \forall \ldots \exists$ ist dieses Axiom formal komplizierter als die oben aufgeführten quantorenfreien Axiome.

1.2.3 Definition Die Theorie T_R der Ringe mit Eins
Die nicht-logischen Grundzeichen der *Sprache* $L(T_R)$ sind
zwei Konstanten 0 und 1,
ein 1-stelliges Funktionszeichen $-$,
zwei 2-stellige Funktionszeichen $+$ und $\cdot$.

Wie üblich schreiben wir $(s+t)$ für $+st$ und $(s \cdot t)$ für $\cdot st$ und sparen Klammern in der üblichen Weise.

Die *Axiome von* T_R sind dann Allabschlüsse von
R1. $(a + b) + c = a + (b + c)$
R2. $a + 0 = a$
R3. $a + (-a) = 0$
R4. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
R5. $a \cdot 1 = a \quad \wedge \quad 1 \cdot a = a$ l
R6. $\neg \quad 1 = 0$
R7. $a \cdot (b + c) = a \cdot b + a \cdot c$
R8. $(a + b) \cdot c = a \cdot c + b \cdot c$

Die Theorie $T_R = (L(T_R), Ax(T_R))$ ist die *Theorie der Ringe mit 1.*
Erweitern wir $Ax(T_R)$ um das Axiom
R9. $a \cdot b = b \cdot a$,
so erhalten wir die Theorie $(L(T_R), Ax(T_R) \cup \{\text{R9}\})$, die *Theorie der kommutativen Ringe mit 1.*

Erweitern wir $Ax(T_R)$ um das Axiom

R10. $\quad a \neq 0 \to \exists y \quad a \cdot y = 1,$

so erhalten wir die *Theorie der Schiefkörper*. Schließlich ist

$T_K = (L(T_R), Ax(T_R) \cup \{\text{R9,R10}\})$ die *Theorie der Körper*.

Wir vermeiden es, in T_K analog zur Gruppentheorie eine Inversenfunktion $^{-1}$ einzuführen. Dadurch kommen wir gar nicht erst in die Verlegenheit, den Term 0^{-1} bilden zu müssen.

1.2.4 Definition Die gewöhnliche Zahlentheorie Z.

Die nicht-logischen Grundzeichen von $L(Z)$ sind
eine Konstante 0 (*Null*),
ein 1-stelliges Funktionszeichen S (*Nachfolger*, *successor*)
zwei 2-stellige Funktionszeichen $+$ (*Summe*) und $\cdot$ (*Produkt*).

Wie üblich schreiben wir $(s + t)$ für $+st$ und $(s \cdot t)$ für $\cdot st$ und lassen äußere Klammern fort.

Die *Axiome von Z* sind dann Allabschlüsse von

$$\begin{array}{lll} \neg Sa = 0, & Sa = Sb & \to \quad a = b \\ a + 0 = a, & a + Sb = S(a + b) & \\ a \cdot 0 = 0, & a \cdot Sb = a \cdot b + a & \end{array}$$

und für jede Formel $F(a)$, in der x nicht auftritt, ein Allabschluss der Formel

$$F(0) \to \forall x(F(x) \to F(Sx)) \to \forall x F(x).$$

Dieses Formelschema ist das *Schema der vollständigen Induktion*. Es ist im Rahmen der Sprache von Z die Formalisierung des Prinzips der vollständigen Induktion aus 1.1.6.

Außer den Induktionsaxiomen enthalten die Axiome von Z zu jedem Funktionszeichen $S, +, \cdot$ genau zwei Axiome, den Null-Fall und den Nachfolgerfall. Die Symmetrie wird noch deutlicher, wenn wir das erste Axiom in der Form $Sa = 0 \to \bot$ schreiben.

1.2.5 Definition Die Theorie LO der linearen Ordnung.

Einziges nicht-logisches Grundzeichen von $L(LO)$ ist ein 2-stelliges Prädikatszeichen $<$. Wir schreiben $a < b$ statt $< ab$.

$Ax(LO)$ besteht aus Allabschlüssen von

$LO1.$	$a < b \to b < c \to a < c$	(Transitivität)
$LO2.$	$\neg\, a < a$	(Antireflexivität)
$LO3.$	$a < b \quad \vee \quad a = b \quad \vee \quad b < a$	(Linearität)

Aus LO entsteht die *Theorie DLO der dichten linearen Ordnungen ohne erstes und letztes Element*, indem man zu $Ax(LO)$ Allabschlüsse folgender Formeln hinzufügt:

$LO4.$	$a < b \to \exists y(a < y \wedge y < b)$
$LO5.$	$\exists y \quad y < a$
$LO6.$	$\exists y \quad a < y$

Diese Theorien der Kleiner-Relation, die keine Funktionszeichen verwenden, haben kurze, einfach formulierte Axiome mit höchstens einem Existenzquantor. Die Theorie DLO hat eine interessante modelltheoretische Eigenschaft, auf die wir im 4. Kapitel zurückkommen.

1.2.6 Definition Die Sprache der Zermelo-Fraenkelschen Mengenlehre ZF.

Die Sprache $L(ZF)$ enthält als einziges nicht-logisches Grundzeichen das 2-stellige Prädikatszeichen $\in$. Man schreibt $a \in b$ statt $\in ab$.

ZF hat wie Z unendlich viele Axiome; $Ax(ZF)$ ist also eine unendliche Menge von Sätzen. Wir brauchen hier die Axiome von ZF nicht aufzuführen. Wir werden später nur verwenden, dass einige geläufige mengentheoretische Definitionen und Schlüsse sich in ZF ausführen lassen. Sie lassen sich also in der extrem einfachen Sprache von ZF hinschreiben. Um einen Eindruck von der gewaltigen Ausdruckskraft der Theorie ZF zu erhalten, muss man ohnehin ausführlich in die Mengenlehre – und auch in die Mathematik – eindringen.

Einzelne dieser Theorien werden uns im Laufe des Textes immer wieder als Beispiele dienen, um die Bedeutung allgemeiner Ergebnisse zu erläutern. Einige werden wir mit logischen, insbesondere modell-theoretischen Methoden untersuchen, um Aussagen über diese Theorien zu gewinnen.

Wir führen einige Bezeichnungen ein, mit denen wir Sprachen und Theorien grob klassifizieren können.

1.2.7 Definitionen Eine Sprache L nennen wir *algebraisch*, wenn alle nicht-logischen Grundzeichen von L Funktionszeichen sind; wir nennen sie *relational*,

wenn alle nicht-logischen Grundzeichen von L Prädikatszeichen oder Konstanten sind.
Eine Theorie T ist *endlich axiomatisiert*, wenn $Ax(T)$ endlich ist. Wir nennen T eine *logische Theorie*, wenn die Menge $Ax(T)$ leer ist, wenn T also keine theorie-eigenen Axiome besitzt.
T heißt *offen*, wenn $Ax(T)$ nur aus Allabschlüssen quantorenfreier Formeln besteht.

Beispiele. Die Sprachen von T_G, T_R (und T_K), Z sind algebraisch; solche Theorien werden in der Algebra bevorzugt. Die Sprachen von LO (und DLO) und ZF sind relational. Die Theorie der geordneten Körper, die wir nicht aufgeführt haben, verwendet die Grundzeichen von T_K und von LO; ihre Sprache ist also weder algebraisch noch relational. Die Zahlentheorie Z und die Mengenlehre ZF haben unendlich viele Axiome, darunter Axiome mit beliebig vielen Quantoren. Die übrigen aufgeführten Theorien, von T_G bis DLO, sind endlich axiomatisiert. Von ihnen sind T_G, T_R und LO offene Theorien, während die Erweiterungen T_K von T_R und DLO von LO nicht mehr offen sind, weil die Zusatzaxiome $R10$ bzw. $LO4$ bis $LO6$ Existenzquantoren enthalten.

Der Begriff der mathematischen Theorie ist von zentraler Bedeutung. Zunächst ist dieser Begriff rein sprachlicher Natur. Wir haben noch nicht festgelegt, wie Formeln in einer Theorie als *gültig* nachzuweisen sind und was eine *Herleitung* in einer Theorie ist. Dieses wird in den folgenden Paragraphen geschehen. In diesem Paragraphen untersuchen wir noch einfache Eigenschaften von Zeichenreihen, insbesondere von Termen und Formeln.

1.3 Zur klammerfreien Schreibweise

Wir haben bei der Definition der Terme und Formeln die klammerfreie Schreibweise benutzt. Ist dies berechtigt? Kann man Terme und Formeln eindeutig in ihre Bestandteile zerlegen? Ist z. B. bei einer Formel $\to AB$ eindeutig bestimmt, wo die Teilformel A aufhört und B anfängt? Dasselbe Problem ergibt sich übrigens auch bei Verwendung von Klammern: Ist bei $(A \to B)$ eindeutig klar, welches Auftreten des Zeichens $\to$ das regierende Zeichen ist? Für eine einheitliche Behandlung dieser Frage fassen wir Terme und Formeln unter einem Begriff zusammen.

1.3.1 Definition Ist $F(a_1, \ldots, a_n)$ ein Term oder eine Formel, so heißt $F(x_1, \ldots, x_n)$ ein *Q-Term (Quasi-Term).*

Q-Terme gehen also aus Termen und Formeln hervor, indem man einzelne (Auftreten von) freien Variablen durch gebundene Variablen ersetzt.

Beispiele

a. Terme und Formeln sind Q-Terme.

b. Ist $\forall x F(x)$ ein Q-Term (z. B. eine Formel), so sind x und $F(x)$ Q-Terme.

Der Quantor $\forall$ fasst also zwei spezielle Q-Terme x und $F(x)$ wieder zu einem Q-Term $\forall x F(x)$ zusammen, ebenso wie ein zweistelliges Funktionszeichen f zwei Terme s und t wieder zu einem Term fst zusammenfasst. Es liegt danach nahe, auch den logischen Grundzeichen eine Stellenzahl zuzuordnen.

1.3.2 Definition *Stellenzahl von Grundzeichen*

1. *Variablen und $\bot$ sind 0-stellig;*

2. *n-stellige Funktions- und Prädikatszeichen sind n-stellig;*

3. *$\rightarrow$ und $\forall$ sind 2-stellig.*

Hiermit ergibt sich sofort folgende einfache, aber grundlegende Aussage über die Gestalt von Q-Termen:

1.3.3 Lemma Jeder Q-Term hat eine Gestalt $ft_1 \ldots t_n$, wobei f ein n-stelliges Grundzeichen ist und $t_1, \ldots, t_n$ Q-Terme sind.

Beweis. Wir bezeichnen mit $\bar{u}$ die Q-Terme, die aus einem Q-Term u hervorgehen, wenn man in u freie durch gebundene Variablen ersetzt. Nun gehen wir für u die Fälle der Definitionen 1.1.4 und 1.1.8 der Terme und Formeln durch.

1. Ist u eine freie Variable oder $\bot$, so ist $\bar{u}$ eine freie oder gebundene Variable oder $\bot$, also 0-stellig.

2. Ist $u \equiv fs_1 \ldots s_n$ mit einem n-stelligen Funktions- oder Prädikatszeichen f und n Termen $s_1, \ldots, s_n$, so ist $\bar{u} \equiv f\bar{s_1} \ldots \bar{s_n}$ mit n Q-Termen $\bar{s_1}, \ldots, \bar{s_n}$ hinter dem n-stelligen Grundzeichen f.

3. Ist $u \equiv \ \rightarrow AB$ mit Formeln A, B, so ist $\bar{u} \equiv \ \rightarrow \bar{A}\bar{B}$ mit zwei Q-Termen $\bar{A}, \bar{B}$ hinter dem 2-stelligen Grundzeichen $\rightarrow$.

4. Ist $u \equiv \forall x F(x)$ mit einer Formel $F(a)$, so ist $\bar{u} \equiv \forall x \bar{F}(x)$, wobei $\bar{F}(x)$ zugleich ein Q-Term $\bar{F}(\bar{a})$ ist. Also besteht $\bar{u}$ aus dem 2-stelligen Grundzeichen $\forall$, gefolgt von zwei Q-Termen x und $\bar{F}(\bar{a})$.

Damit sind alle Q-Terme $\bar{u}$ erfasst.

1.3.4 Definition *Ein Auftreten einer Zeichenreihe u in einer Zeichenreihe v ist eine Nennform F, in der $*_1$ genau einmal vorkommt, so dass $F(u) \equiv v$ ist. Dieses Auftreten steht am Anfang von v, wenn F die Gestalt $\ *_1 G\ $ hat. u tritt in v auf, wenn es ein Auftreten von u in v gibt.*

Beispiele.

a. v tritt in v auf wegen $*_1(v) \equiv v$.

b. Das „*Auftreten in*“ ist transitiv: Ist F ein Auftreten von u in v und G ein Auftreten von v in w, so ist

$$G(F)(u) \equiv G(F(u)) \equiv G(v) \equiv w,$$

so dass $G(F)$ ein Auftreten von u in w ist.

c. B tritt in $A :\equiv B \rightarrow C \vee B$ zweimal auf. Denn es ist $*_1 \rightarrow C \vee B(B) \equiv B \rightarrow C \vee *_1(B) \equiv A$. Diese beiden Auftreten sind offenbar verschieden.

Ein Auftreten von u in v fixiert also auch die Stelle, an der u in v auftritt, neben der Tatsache, dass u ein zusammenhängendes Stück der Zeichenreihe v ist.

Wir überlegen jetzt, dass für jeden Q-Term die in Lemma 1.3.3 angegebene Gestalt eindeutig ist.

1.3.5 Lemma *Sind $s_1, \ldots, s_m, \quad t_1, \ldots, t_n$ Q-Terme und tritt die Zeichenreihe $s_1 \ldots s_m$ am Anfang der Zeichenreihe $t_1 \ldots t_n$ auf, so ist $m \leq n$ und $s_i \equiv t_i$ für $1 \leq i \leq m$.*

Beweis durch Induktion nach der Länge von $s_1 \dots s_m$, also nach der Summe der Längen der s_i:

1. Ist diese Länge 0, so ist $m = 0$, und es ist nichts zu zeigen.

2. Ist diese Länge positiv, so ist s_1 ein Q-Term, der nach Lemma 1.3.3 eine Gestalt $fr_1 \dots r_k$ hat mit einem k-stelligen Grundzeichen f und Q-Termen $r_1, \dots, r_k$. Dann beginnt auch t_1 mit f, und nach 1.3.3 ist $t_1 \equiv fr'_1 \dots r'_k$ für geeignete Q-Terme $r'_1, \dots, r'_k$. Dann tritt $r_1 \dots r_k s_2 \dots s_m$ am Anfang von $r'_1 \dots r'_k t_2 \dots t_n$ auf und ist (um das Zeichen f) kürzer als $s_1 \dots s_m$. Nach Induktionsvoraussetzung (IV) ist dann

$$\begin{gathered} k + m - 1 \leq k + n - 1, \text{ also } m \leq n, \\ r_i \equiv r'_i \text{ für } 1 \leq i \leq k, \text{ also } s_1 \equiv t_1, \text{ und} \\ s_i \equiv t_i \text{ für } 2 \leq i \leq m. \end{gathered}$$

Wir notieren zwei Spezialfälle dieses Lemmas:

1.3.6 Korollar *Tritt ein Q-Term s am Anfang eines Q-Terms t auf, so ist $s \equiv t$.*
Dies ist der Fall $m = n = 1$ des Lemmas.

1.3.7 Korollar *Ist für Q-Terme $s_1, \dots, s_m, t_1, \dots, t_n$*

$$s_1 \dots s_m \equiv t_1 \dots t_n \quad ,$$

so ist $m = n$ und $s_i \equiv t_i$ für $1 \leq i \leq m = n$.

Dies ist der symmetrische Fall des Lemmas, das dann sowohl $m \leq n$ als auch $n \leq m$ und damit die Behauptung ergibt.

1.3.8 Lemma *Eindeutigkeit des Aufbaus von Q-Termen*
Jeder Q-Term besitzt genau eine Darstellung $ft_1 \dots t_n$ mit einem Grundzeichen f und Q-Termen $t_1, \dots, t_n$. Die Anzahl n dieser Q-Terme ist dann die Stellenzahl von f.

Beweis. Die Existenz einer solchen Darstellung ist Lemma 1.3.3.
Zur Eindeutigkeit: Sei $gs_1 \dots s_m \equiv ft_1 \dots t_n$. Dann ist $f \equiv g$, so dass auch g n-stellig ist, und $s_1 \dots s_m \equiv t_1 \dots t_n$. Nach 1.3.7 ist dann $m = n$ und $s_i \equiv t_i$ für $1 \leq i \leq n$.

Bemerkung. Mit 1.3.8 ist die klammerfreie Schreibweise von Q-Termen, Termen und Formeln gerechtfertigt. Hierdurch werden auch erst die oben eingeführten Abkürzungen wohldefiniert. Denn eine Formel $\to AB$ ist notwendig die Implikation von A und B, wenn dies Formeln sind; es gibt keine anderen Formeln A', B' mit $A'B' \equiv AB$.

Kann ein Q-Term, der in $\to AB$ auftritt, in der Schreibweise $(A \to B)$ „zerrissen“ werden, weil sein Anfang in A und sein Ende in B auftritt? Dann wären nicht alle Teil-Q-Terme von $\to AB$ in der Schreibweise $(A \to B)$ zu erkennen. Wir zeigen, dass auch dies nicht möglich ist.

1.3.9 Lemma *Das Grundzeichen g trete in dem Q-Term t auf. Dann steht g am Anfang eines Q-Terms s, der in t auftritt.*

Beweis durch Induktion nach der Länge von t:

1. Fall: g steht am Anfang von t. Dann ist t der gesuchte Q-Term s.

2. Fall: g steht nicht am Anfang von t. Dann ist $t \equiv ft_1 \dots t_n$ für gewisse Q-Terme $t_1, \dots, t_n$ und ein Grundzeichen f nach 1.3.3, und g tritt in einem t_i auf $(1 \leq i \leq n)$. Da die Länge von t_i kleiner ist als die Länge von t, gilt nach IV: g steht am Anfang eines Q-Terms s, der in t_i, also auch in t auftritt.

1.3.10 Lemma *Auftreten von Q-Termen*

Sei f ein n-stelliges Grundzeichen und seien $t_1, \dots, t_n$ Q-Terme. Wenn ein Q-Term s in $ft_1 \dots t_n$ auftritt, dann gilt

$$s \equiv ft_1 \dots t_n \text{ oder } s \text{ tritt in einem } t_i \text{ auf } (1 \leq i \leq n) :$$

Teil-Q-Terme treten nicht „übergreifend“ auf.

Beweis. 1. Fall: s tritt am Anfang von $ft_1 \dots t_n$ auf. Dann ist $s \equiv ft_1 \dots t_n$ nach 1.3.6.
2. Fall: Das erste Zeichen g von s tritt in einem t_i $(1 \leq i \leq n)$ auf. Dann steht g nach Lemma 1.3.9 am Anfang eines Q-Terms s', der in t_i auftritt. Nach 1.3.6 ist $s \equiv s'$.

Damit beenden wir die Untersuchung allgemeiner Eigenschaften von Q-Termen. Wir können im Folgenden ohne Bedenken die klammerfreie Schreibweise und die gewohnte „abkürzende“ Schreibweise nebeneinander verwenden.

1.4 Aufgaben

1.4.1 a. Zeigen Sie für beliebige Nennformen F, G_i, H_j:

$$F(G_1, G_2)(H_1, H_2) \equiv F(G_1(H_1, H_2), G_2(H_1, H_2))$$

b. Geben Sie Nennformen F, G, H_1, H_2 an, für die $F(G)(H_1, H_2)$ und $F(G(H_1, H_2))$ nicht übereinstimmen.

1.4.2 **(Induktive Definition)** der *Nf von L*.
1. Die leere Zeichenreihe ist eine Nf von L.
2. Ist F eine Nf von L und z ein Grundzeichen von L, so ist Fz eine Nf von L.
3. Ist F eine Nf von L, so ist $F*_i$ eine Nf von L (für $i > 0$).
a. Zeigen Sie: Die Nf von L sind genau die Nennformen von L gemäß 1.1.2.
b. Geben Sie eine rekursive Definition der Substitution $F(G_1, \ldots, G_n)$ an.

1.4.3 Zeigen Sie durch vollständige Induktion:

$$\sum_{i<n+1} i \quad = \quad \frac{1}{2} \cdot n \cdot (n+1)$$

1.4.4 a. $+$ sei ein 2-stelliges Funktionszeichen von L. Ist $= + + a + abb + +ab + ab$ eine Formel (Gleichung) von L? Wenn ja, schreiben Sie sie in der „üblichen“ Weise mit $=$ und $+$ zwischen den Argumenten.
b. Schreiben Sie die Formel $A \wedge B \to B \wedge A$ ohne Abkürzungen gemäß 1.1.8 und geben Sie alle Auftreten von B in dieser Formel an für den Fall, dass B in A nicht auftritt.
c. Schreiben Sie $a \neq 0 \to \exists y\ a \cdot y = 1$ als Formel von $L(T_R)$ gemäß 1.1.8 und geben Sie alle Auftreten von $=$ in dieser Formel an.

§2 Semantik: Strukturen und Gültigkeit

2.1 Strukturen und Interpretation

Im vorigen Paragraphen haben wir Sprachen der ersten Stufe nur *semiotisch* betrachtet, d.h. wir haben Zeichenreihen auf ihren Aufbau und auf das Auftreten von Teilausdrücken untersucht. Wir haben bisher nicht berücksichtigt, dass Terme und Formeln auch etwas bedeuten können.

Terme und Formeln erhalten Bedeutung, indem wir sie in passenden Strukturen als Objekte bzw. als Sachverhalte interpretieren. Gruppen, Körper, geordnete Mengen sind Beispiele für Strukturen. Jedes Funktionszeichen der Sprache muss durch eine Funktion auf der Trägermenge, dem Individuenbereich der Struktur interpretiert werden, jedes Prädikatszeichen durch eine Relation. Dann ist die Bedeutung der Terme und Formeln kanonisch festgelegt (vgl. 2.1.3 und 2.2.3). Wir formulieren als erstes den allgemeinen Strukturbegriff.

2.1.1 Definition *Sei L eine Sprache der ersten Stufe. Eine Struktur $\mathcal{A}$ zu L ist ein Tripel (A, F, R) mit:*

(i) *einer nichtleeren Menge $\mid \mathcal{A} \mid := A$, genannt das Universum, die Trägermenge oder der Individuenbereich von $\mathcal{A}$.*

(ii) *einer Familie $F = (f_{\mathcal{A}})_{f \in L}$, die zu jedem n-stelligen Funktionszeichen f von L genau eine n-stellige Funktion auf A*

$$f_{\mathcal{A}} : A^n \to A$$

enthält. Wir identifizieren 0-stellige Funktionen auf A mit ihrem einzigen Wert $\in A$.

(iii) *einer Familie $R = (p_{\mathcal{A}})_{p \in L}$, die zu jedem n-stelligen nicht-logischen Prädikatszeichen p von L genau eine n-stellige Relation*

$$p_{\mathcal{A}} \subseteq A^n$$

enthält.

Ist $\mathcal{A}$ eine Struktur zu L, so sagen wir auch, $\mathcal{A}$ passt zu L. Ist L eine algebraische Sprache, so ist R leer, und A heißt algebraische Struktur. Ist L eine relationale Sprache, so enthält F höchstens Elemente von A, und A heißt relationale Struktur.

Bemerkung. Eine n-stellige Funktion φ auf A bestimmt zu je n nacheinander gewählten Elementen $a_1, \ldots, a_n$ von A genau ein Element c von A, den Funktionswert von φ für $a_1, \ldots a_n$. Im Fall $n = 0$ werden gar keine Elemente gewählt, und die Funktion φ bestimmt unmittelbar ein $c \in A$ als den Funktionswert von φ schlechthin. Ebenso wie 0-stellige Funktionszeichen Terme sind, schreiben wir hier $\varphi = c \in A$: 0-stellige Funktionen auf A sind Elemente von A.

Beispiele. Gruppen, Ringe, Körper sind algebraische Strukturen; geordnete Mengen sind relationale Strukturen (vgl. 2.2.4 bis 7). Gruppen passen zur Sprache der Gruppentheorie T_G, sie sind Strukturen zur Sprache von T_G. Ebenso passen geordnete Mengen zur Sprache von LO.

2.1.2 Definition *Sei $\mathcal{A}$ eine Struktur zu L. Man erhält die Sprache $L(\mathcal{A})$ aus L, wenn man zu L für jedes Individuum $a \in \mid \mathcal{A} \mid$ eine Konstante, den Namen von a hinzufügt, und zwar verschiedene Namen für verschiedene Individuen. Den Namen von a bezeichnen wir wieder mit a. Individuen aus $\mid \mathcal{A} \mid$ und ihre Namen bezeichnen wir also gleich.*

Beispiel. Sei L die Sprache der Körpertheorie (das ist die Sprache der Ringe mit 1 aus 1.2.3) und $\mathbb{R}$ der Körper der reellen Zahlen. Dann enthält $L(\mathbb{R})$ einen Namen für jede reelle Zahl.

In der Sprache $L(\mathcal{A})$ lassen sich Sachverhalte *in der Struktur* $\mathcal{A}$ gut formulieren, weil für jedes Individuum ein Name zur Verfügung steht, z.B. e und π in $L(\mathbb{R})$ für die entsprechenden transzendenten Zahlen aus $\mathbb{R}$. In $L(\mathcal{A})$ kann man aber immer noch nicht *über die Struktur* $\mathcal{A}$ sprechen; z. B. lässt sich folgende Aussage sicher nicht in $L(\mathbb{R})$ formulieren: $\mathbb{R}$ ist bis auf Isomorphie der einzige archimedisch und vollständig angeordnete Körper. Solche Aussagen höherer Stufe sind erst in der Mengenlehre oder in einer höheren Logik formulierbar.

Die Kernfrage dieses Paragraphen ist: Was bedeuten die Terme und Formeln einer Sprache? Terme sollen Objekte bezeichnen, Formeln sollen Sachverhalte ausdrücken, die zutreffen oder nicht zutreffen. Treten freie Variablen auf, so hängen diese Objekte und Sachverhalte offenbar noch von den Objekten ab, die

man den freien Variablen als Werte zugeordnet hat. Als erstes Ziel präzisieren wir deshalb die Bedeutung der geschlossenen Terme und Formeln der Sprache $L(\mathcal{A})$ durch den Begriff der *Interpretation*. Zutreffende Sachverhalte werden dabei als *wahr* bezeichnet und erhalten den *Wahrheitswert* w; nicht zutreffende Sachverhalte werden als *falsch* bezeichnet und erhalten den *Wahrheitswert* f.

2.1.3 Rekursive Definition *der Interpretation $\mathcal{A}$ der geschlossenen Terme, Formeln und Sequenzen von $L(\mathcal{A})$ in der Struktur $\mathcal{A}$ zur Sprache L.*

1. *Die Interpretation $\mathcal{A}$ ordnet jedem geschlossenen Term von $L(\mathcal{A})$ ein Element von $\mid \mathcal{A} \mid$ wie folgt zu:*

   1.1. *Für $c \in \mid \mathcal{A} \mid$ sei* $\mathcal{A}(c) := c$

   1.2. $\mathcal{A}(ft_1 \ldots t_n) := f_{\mathcal{A}}(\mathcal{A}(t_1), \ldots, \mathcal{A}(t_n))$

2. *Die Interpretation $\mathcal{A}$ ordnet jedem Satz von $L(\mathcal{A})$ einen Wahrheitswert w bzw. f wie folgt zu:*

   2.1. $\mathcal{A}(s = t) = w \quad :\Leftrightarrow \quad \mathcal{A}(s) = \mathcal{A}(t)$

   2.2. $\mathcal{A}(pt_1 \ldots t_n) = w \quad :\Leftrightarrow \quad (\mathcal{A}(t_1), \ldots, \mathcal{A}(t_n)) \in p_{\mathcal{A}}$

   2.3. $\mathcal{A}(\bot) = f$

   2.4. $\mathcal{A}(B \to C) = w \quad :\Leftrightarrow \quad$ *aus* $\mathcal{A}(B) = w$ *folgt* $\mathcal{A}(C) = w$

   2.5. $\mathcal{A}(\forall x F(x)) = w \quad :\Leftrightarrow \quad \mathcal{A}(F(c)) = w$ *für alle* $c \in \mid \mathcal{A} \mid$

3. *Ist $\Gamma : \Delta$ eine geschlossene Sequenz von $L(\mathcal{A})$, so ist*
   $\mathcal{A}(\Gamma : \Delta) = w : \quad \Leftrightarrow \quad$ *Wenn* $\mathcal{A}(C) = w$ *ist für alle* $C \in \Gamma$,
   so ist $\mathcal{A}(D) = w$ *für mindestens ein* $D \in \Delta$.

Erläuterungen. 1. Wir verwenden als Bezeichnung für die Interpretation in der Struktur $\mathcal{A}$ wieder den Buchstaben $\mathcal{A}$. Eine Verwechslung der beiden Verwendungen von $\mathcal{A}$ ist so gut wie ausgeschlossen. Denn die Interpretation $\mathcal{A}$ ist eine Abbildung, die auf der Menge aller geschlossenen Terme, Formeln und Sequenzen von $L(\mathcal{A})$ definiert ist, während die Struktur $\mathcal{A}$ keine Abbildung ist.

2. Wegen der Zweiwertigkeit der Interpretation $\mathcal{A}$, also weil für $\mathcal{A}(B)$ nur die beiden Wahrheitswerte w (wahr) und f (falsch) zur Verfügung stehen, wird in den anderen Fällen unter 2. und 3. notwendig der Wert f angenommen. Wegen

der Zweiwertigkeit ist $\mathcal{A}(B \to C) = w$ auch gleichwertig mit: $\mathcal{A}(B) = f$ oder $\mathcal{A}(C) = w$. Denn sowohl für $\mathcal{A}(B) = w$ als auch für $\mathcal{A}(B) = f$ ist diese Festsetzung äquivalent zu der in Definitionsschritt 2.4.

3. Wie der Definitionsschritt 2.5 zeigt, muss man die Interpretation $\mathcal{A}$ für alle geschlossenen Formeln aus $L(\mathcal{A})$ definieren und nicht nur für die aus L, weil $F(c)$ nicht aus L zu sein braucht, auch wenn $\forall x F(x)$ aus L ist.

4. Ist p ein 0-stelliges Prädikatszeichen, so gibt es für $p_{\mathcal{A}}$ nur die beiden Möglichkeiten $p_{\mathcal{A}} = \{\emptyset\}$ und $p_{\mathcal{A}} = \{\}$, wobei $\emptyset$ das leere Tupel bezeichnet. Dann ergibt Definitionsschritt 2.2

$$\begin{array}{ccccc} \mathcal{A}(p) = w & \Leftrightarrow & \emptyset \in p_{\mathcal{A}} & \Leftrightarrow & p_{\mathcal{A}} = \{\emptyset\}, \\ \mathcal{A}(p) = f & \Leftrightarrow & \emptyset \notin p_{\mathcal{A}} & \Leftrightarrow & p_{\mathcal{A}} = \{\}. \end{array}$$

Man kann danach die nicht-leere 0-stellige Relation $\{\emptyset\}$ mit dem Wahrheitswert w und die leere 0-stellige Relation $\{\}$ mit dem Wahrheitswert f identifizieren.

5. Für die leere Sequenz $\square = \emptyset : \emptyset$ ist $\mathcal{A}(\square) = f$, weil sowohl für alle $C \in \emptyset$ als auch für kein $C \in \emptyset \quad \mathcal{A}(C) = w$ ist.

2.1.4 Eine grundlagenkritische Bemerkung Grob gesprochen ist die Interpretation $\mathcal{A}$ eine Abbildung von einer Sprache in eine „fertige Wirklichkeit“ oder „reale Welt“, nämlich von $L(\mathcal{A})$ in $\mathcal{A}$. Bei der Interpretation werden die logischen Grundzeichen $=, \perp, \to, \forall$ mit ihrer üblichen inhaltlichen Bedeutung versehen. Die Zweiwertigkeit der Interpretation $\mathcal{A}$ beruht auf der Vorstellung, dass jeder in der Sprache $L(\mathcal{A})$ formulierbare Sachverhalt in dieser realen Welt so oder so entschieden ist, unabhängig davon, ob man die wahren Sachverhalte als solche erkennen kann oder nicht.
Andere theoriebildende Disziplinen, allen voran die theoretische Physik, stehen oft vor dem Problem, eine *angemessene* Interpretation ihrer Sprache zu finden, um überhaupt verständlich zu werden und eben um Bedeutung zu erlangen. Dieses Problem stellt sich uns hier nur insofern, als die Strukturen, in denen wir unsere Sprachen interpretieren, zu diesen Sprachen passen müssen.

Es ist üblich, Formeln als abstrakt und Strukturen als konkret anzusehen, weil Formeln zunächst (vgl. § 1) bedeutungslos sind und viele Interpretationen zulassen, und weil einige Strukturen wie die natürlichen Zahlen und der euklidische Raum als Realität vorlagen, bevor man Theorien über sie machte. Diese

Einstellung erschwert den Zugang zur Logik ebenso wie zur theoretischen Informatik. Denn offenbar sind Terme und Formeln endliche Zeichenreihen, die man hinschreiben und im Computer verarbeiten kann, und somit konkrete Gebilde. Dagegen sind die meisten in der Mathematik wichtigen Strukturen wie Hilbert-Räume, komplexe Räume, Lie-Gruppen, ja schon reelle Vektorräume höherer Dimension abstrakte Gebilde, die sich nicht – weder im Labor, noch auf dem Papier – herstellen lassen, ganz zu schweigen von Begriffsumfängen wie der Klasse aller Gruppen. Eine Interpretation $\mathcal{A}$ ordnet also konkreten sprachlichen Gebilden Objekte und Wahrheitswerte in der i.a. abstrakten Struktur $\mathcal{A}$ zu. Wir fassen diese Überlegung in einem Diagramm zusammen:

$$\begin{array}{ccc} & \text{Interpretation} & \\ \text{Sprache} & \longrightarrow & \text{Struktur} \\ \textit{formal} & & \textit{real} \\ \textit{konkret} & & \textit{abstrakt} \end{array}$$

2.1.5 Wahrheitstafeln Sind B und C Sätze von $L(\mathcal{A})$, so hängt der Wahrheitswert $\mathcal{A}(B \to C)$ nur von den Wahrheitswerten $\mathcal{A}(B)$ und $\mathcal{A}(C)$ ab und nicht mehr von der Gestalt von B und C. Dasselbe gilt für alle Junktoren, die wir mit Hilfe von $\bot$ und $\to$ in 1.1.12 eingeführt haben. Man kann dann für jeden Junktor eine Tabelle anlegen, in der der Wahrheitswert eines mit dem Junktor gebildeten Satzes neben den Wahrheitswerten der Teilsätze aufgeführt ist. Solche Tabellen heißen *Wahrheitstafeln*. Die Wahrheitstafel eines n-stelligen Junktors hat 2^n Zeilen, weil man aus w und f genau 2^n Folgen der Länge n bilden kann. 0-stellige Junktoren haben dann 1-zeilige Wahrheitstafeln, ihre Wahrheitswerte hängen von keinen Teilformeln ab. Die Wahrheitstafeln der Junktoren aus 1.1.12 sind:

$\mathcal{A}(\bot)$	$\mathcal{A}(\top)$	$\mathcal{A}(B)$	$\mathcal{A}(\neg B)$
f	w	w	f
		f	w

$\mathcal{A}(B)$	$\mathcal{A}(C)$	$\mathcal{A}(B \to C)$	$\mathcal{A}(B \vee C)$	$\mathcal{A}(B \wedge C)$	$\mathcal{A}(B \leftrightarrow C)$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	w	w	f	f
f	f	w	f	f	w

Die Junktoren werden also alle so interpretiert, wie man es von der Umgangssprache her erwartet. Für $\bot$ und $\to$ ist dies Bestandteil von 2.1.3. Für die in 1.1.12 eingeführten Junktoren lesen wir aus den Wahrheitstafeln ab:

$\mathcal{A}(\top)$ ist stets wahr.
$\mathcal{A}(\neg B) = w \quad \Leftrightarrow \quad \mathcal{A}(B)$ ist nicht wahr,
$\mathcal{A}(B \vee C) = w \quad \Leftrightarrow \quad \mathcal{A}(B) = w$ oder $\mathcal{A}(C) = w$,

wobei wir das nicht-ausschließende *oder* (im Gegensatz zum *entweder – oder*) verwenden.

$\mathcal{A}(B \wedge C) = w \quad \Leftrightarrow \quad \mathcal{A}(B) = w$ und $\mathcal{A}(C) = w$.
$\mathcal{A}(B \leftrightarrow C) = w \quad \Leftrightarrow \quad \mathcal{A}(B) = w$ ist äquivalent zu $\mathcal{A}(C) = w$.

Ferner erhält man durch Induktion nach n für Sätze $B_1, \ldots, B_n, C$ aus $L(\mathcal{A})$:

a. $\mathcal{A}(B_1 \to \ldots \to B_n \to C) = w \quad \Leftrightarrow$ wenn $\mathcal{A}(B_i) = w$ ist für alle i mit $1 \leq i \leq n$, so ist $\mathcal{A}(C) = w$.

b. $\mathcal{A}(B_1 \vee \ldots \vee B_n) = w \quad \Leftrightarrow$ $\mathcal{A}(B_i) = w$ für mindestens ein i mit $1 \leq i \leq n$.

c. $\mathcal{A}(B_1 \wedge \ldots \wedge B_n) = w \quad \Leftrightarrow$ $\mathcal{A}(B_i) = w$ für alle i mit $1 \leq i \leq n$.

b. und c. ergeben für geschlossene Sequenzen $\Gamma : \Delta$ aus $L(\mathcal{A})$:

d. Ist $\Gamma = \{C_1, \ldots, C_m\}$ und $\Delta = \{D_1, \ldots, D_n\}$, so ist

$$\mathcal{A}(\Gamma : \Delta) = w \quad \Leftrightarrow \quad \mathcal{A}(C_1 \wedge \ldots \wedge C_m \to D_1 \vee \ldots \vee D_n) = w.$$

Bei a. benutzen wir wesentlich unsere Konvention 1.1.12 zur Klammerersparnis (Rechtsklammerung).

2.1.6 Lemma *Sei $\forall x F(x)$ ein Satz aus $L(\mathcal{A})$. Dann gilt:*

a. $\mathcal{A}(\exists x F(x)) = w \quad \Leftrightarrow \quad \mathcal{A}(F(c)) = w$ *für mindestens ein* $c \in |\, \mathcal{A} \,|$.

b. $\mathcal{A}(\forall x \forall y (F(x) \to F(y) \to x = y)) = w$
$\Leftrightarrow \quad \mathcal{A}(F(c)) = w$ *für höchstens ein* $c \in |\, \mathcal{A} \,|$.

Beweis. a. $\exists x F(x)$ ist $\neg\forall x\neg F(x)$. Also ergeben 2.3 bis 2.5 aus 2.1.3

$\mathcal{A}(\neg\forall x\neg F(x)) = w \quad \Leftrightarrow \quad \mathcal{A}(\forall x\neg F(x)) = f$
$\Leftrightarrow \quad \mathcal{A}(\neg F(c)) = w$ gilt nicht für alle $c \in |\mathcal{A}|$
$\Leftrightarrow \quad \mathcal{A}(\neg F(c)) = f$ gilt für mindestens ein $c \in |\mathcal{A}|$
$\Leftrightarrow \quad \mathcal{A}(F(c)) = w$ gilt für mindestens ein $c \in |\mathcal{A}|$.

b. $\mathcal{A}(\forall x\forall y(F(x) \to F(y) \to x = y)) = w$
$\Leftrightarrow \quad$ für alle $c, d \in |\mathcal{A}|$ folgt $c = d$ aus $\mathcal{A}(F(c)) = \mathcal{A}(F(d)) = w$
$\Leftrightarrow \quad$ es gibt höchstens ein $c \in |\mathcal{A}|$ mit $\mathcal{A}(F(c)) = w$.

„Es gibt (mindestens) ein ...“ ist also eine Existenzaussage, *„es gibt höchstens ein ...“* dagegen eine Allaussage.

2.2 Gültigkeit und Modelle

Um auch Termen und Formeln, in denen freie Variable auftreten, eine Bedeutung zuzuordnen, führen wir den Begriff der Belegung ein:

2.2.1 Definition *Eine $\mathcal{A}$-Belegung ist eine Abbildung*
$' : \{a \mid a \text{ freie Variable}\} \to \{c \mid c \text{ ist Name eines Elements von } |\mathcal{A}|\}$

$$a \quad \longmapsto \quad a'$$

Ist $'$ eine $\mathcal{A}$-Belegung, so geht die Nennform F' aus der Nennform F hervor, indem man jede freie Variable a in F durch a' ersetzt. Ferner ist $\Gamma' := \{C' \mid C \in \Gamma\}$ für Formelmengen Γ und $(\Gamma : \Delta)' := \Gamma' : \Delta'$.

Man erkennt unmittelbar:

a. Tritt in F keine freie Variable auf, so ist $F' \equiv F$.

b. Es ist $(FG)' \equiv F'G'$

c. Es ist $(F(G))' \equiv F'(G')$

2.2.2 Lemma *Es sei X ein Term, eine Formel oder eine Sequenz von $L(\mathcal{A})$ und $'$ eine $\mathcal{A}$-Belegung.*
1. Dann ist X' ein geschlossener Term, eine geschlossene Formel bzw. eine geschlossene Sequenz von $L(\mathcal{A})$.
2. Ist $^\circ$ eine weitere $\mathcal{A}$-Belegung, die mit $'$ auf $FV(X)$ übereinstimmt, so ist $X' \equiv X^\circ$.

Der **Beweis** von 1. durch Induktion nach dem Aufbau von Term und Formel X ist trivial. 2. besagt, dass X' von der Belegung $'$ nur an den Stellen a abhängt, die in X auftreten. Das ist aber klar.

Der Begriff der Belegung führt vom Begriff der Interpretation zu dem zentralen Begriff der Gültigkeit in Strukturen, Modellen und Theorien.

2.2.3 Definition *Es sei L eine Sprache der 1. Stufe und T eine Theorie mit der Sprache L.*

1. *Es sei $\mathcal{A}$ eine Struktur zu L. Eine Formel B von L gilt in $\mathcal{A}$ oder ist $\mathcal{A}$-gültig, $\mathcal{A} \models B$, wenn $\mathcal{A}(B') = w$ ist für jede $\mathcal{A}$-Belegung $'$. Ebenso gilt eine Sequenz $\Gamma : \Delta$ von L in $\mathcal{A}, \mathcal{A} \models \Gamma : \Delta$, wenn $\mathcal{A}(\Gamma' : \Delta') = w$ ist für jede $\mathcal{A}$-Belegung $'$.*

2. *Eine Struktur $\mathcal{A}$ zu L heißt Modell von T, $\mathcal{A} \models T$, wenn jedes Axiom von T in $\mathcal{A}$ wahr ist.*

3. *Eine Formel B von L gilt in T, $T \models B$, wenn B in jedem Modell von T gilt. Ebenso gilt $\Gamma : \Delta$ in T, $T \models \Gamma : \Delta$, wenn $\Gamma : \Delta$ in jedem Modell von T gilt.*

Bemerkungen. a. Sätze B aus L gelten in $\mathcal{A}$ genau dann, wenn sie wahr sind: $\mathcal{A} \models B \Leftrightarrow \mathcal{A}(B) = w$. Denn Sätze werden von Belegungen nicht berührt.

b. Ein Satz B aus L gilt in einer Struktur $\mathcal{A}$ zu L genau dann, wenn $\mathcal{A}$ ein Modell der Theorie $(L, \{B\})$ ist. Dies ergibt sich aus 2.2.3, 2.

c. Viele Lehrbücher benutzen die Begriffe „Folgerung“ und „allgemeingültig“. Eine Formel B aus L *folgt* oder *ist eine Folgerung aus* einer Menge M von Sätzen aus L, wenn B in der Theorie (L, M) gilt. Eine Formel B aus L ist *allgemeingültig*, wenn B in der logischen Theorie $(L, \emptyset)$ gilt.

Eine Formel B aus L ist also genau dann allgemeingültig, wenn B in jeder Struktur zu L gilt.

Parallel zur Formulierung mathematischer Theorien in §1.2 können wir nun über den Modellbegriff übliche mathematische Begriffe einführen.

2.2.4 Gruppen Die *Gruppen* sind genau die Modelle der Gruppentheorie. Eine Gruppe ist also eine algebraische Struktur $G = (| G |; e_G, {}^{-1}_G, \circ_G)$, in der die Gruppenaxiome G1 bis G3 gelten.

Das kommutative Gesetz $a \circ b = b \circ a$ gilt in einer Gruppe G genau dann, wenn für alle Elemente $k, l \in G$ (genauer: $\in|\ G\ |$) $k \circ_G l = l \circ_G k$ ist, wenn also G kommutativ ist. Eine Formel aus der Sprache der Gruppentheorie gilt in der Gruppentheorie genau dann, wenn sie in allen Gruppen gilt. Das kommutative Gesetz gilt also nicht in der Gruppentheorie; denn es gibt nicht-kommutative Gruppen. Dagegen gilt $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$ in der Gruppentheorie; denn dies gilt in jeder Gruppe.

Beispiele. $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ bezeichne die Menge der ganzen, rationalen, reellen bzw. komplexen Zahlen. Dann sind

$$(\mathbb{Z}; 0, -, +), (\mathbb{Q}; 0, -, +), (\mathbb{R}; 0, -, +), (\mathbb{C}; 0, -, +)$$

kommutative Gruppen. Dabei bezeichnet $-$ jeweils den (1-stelligen) Übergang zum Negativen,

$$- : k \mapsto -k,$$

und nicht etwa die (2-stellige) Differenz. Ebenso sind

$$(\mathbb{Q} \setminus \{0\}; 1, ^{-1}, \cdot), (\mathbb{R} \setminus \{0\}; 1, ^{-1}, \cdot), (\mathbb{C} \setminus \{0\}; 1, ^{-1}, \cdot)$$

kommutative Gruppen. Dagegen ist $(\mathbb{Z}; 1, -, \cdot)$ keine Gruppe. Eine nicht-kommutative Gruppe ist etwa die symmetrische Gruppe S_3 (Gruppe der Permutationen von drei Elementen).

2.2.5 Ringe mit 1, Körper Die *Ringe mit* 1 sind genau die Modelle der Theorie T_R der Ringe mit 1. Ebenso sind die *kommutativen Ringe mit* 1, die *Schiefkörper* und die *Körper* genau die Modelle der entsprechenden Theorien aus 1.2.3. Eine Formel, die in allen Schiefkörpern, also auch in allen Körpern, nicht aber in allen (kommutativen) Ringen mit 1 gilt, ist die *Nullteilerfreiheit*:

$$a \cdot b = 0 \to a = 0 \vee b = 0.$$

Diese Formel gilt also in der Körpertheorie T_K, nicht aber in der Theorie T_R der Ringe mit 1.

Beispiele. $(\mathbb{Z}; 0, 1, -, +, \cdot)$ ist ein kommutativer, nullteilerfreier Ring mit 1, aber kein Körper. Man bezeichnet ihn oft kurz mit $\mathbb{Z}$.

$$(\mathbb{Q}; 0, 1, -, +, \cdot), (\mathbb{R}; 0, 1, -, +, \cdot) \text{ und } (\mathbb{C}; 0, 1, -, +, \cdot)$$

sind Körper. Man bezeichnet diese Körper meistens nur mit $\mathbb{Q}, \mathbb{R}$ bzw. $\mathbb{C}$.

Sei K ein Körper, $n > 0$ und $M(n \times n, K)$ die Menge der $n \times n$-Matrizen über K. Dann ist $(M(n \times n, K); 0, E_n, -, +, \cdot)$ ein Ring mit 1, der im Fall $n > 1$ weder kommutativ noch nullteilerfrei ist. Dabei bezeichnet 0 die Nullmatrix, E_n die $n \times n$-Einheitsmatrix und $\cdot$ die Matrizenmultiplikation.

2.2.6 Die natürlichen Zahlen Die gewöhnliche Zahlentheorie Z haben wir in 1.2.4 eingeführt. Ein Modell von Z ist offenbar die Struktur der natürlichen Zahlen

$$\mathcal{N} := (\mathbb{N}; 0, +1, +, \cdot).$$

Die Struktur hat die Menge $\mathbb{N}$ der natürlichen Zahlen als Individuenbereich, und in ihr werden die Funktionszeichen $0, S, +, \cdot$ durch die Zahl 0, die Addition von 1 (Nachfolgerfunktion), die Addition und die Multiplikation der natürlichen Zahlen interpretiert. Diese Struktur $\mathcal{N}$ heißt das *Standardmodell* von Z.

Das Induktionsschema gilt in $\mathcal{N}$ sogar für beliebige Eigenschaften und nicht nur für die Eigenschaften, die durch Formeln $F(a)$ aus $L(Z)$ ausgedrückt werden. Wie wir im 4. Kapitel sehen werden, ist diese Tatsache dafür verantwortlich, dass die Theorie Z noch andere Modelle als $\mathcal{N}$ besitzt.

2.2.7 Geordnete Mengen Die linear geordneten Mengen oder linearen Ordnungen sind genau die Modelle der Theorie LO. Sie sind Paare $(M; <)$, wobei $<$ eine lineare Ordnungsrelation auf M ist.
$(\mathbb{Z}; <), (\{0, 1\}; <)$ sind Modelle von LO, aber nicht von DLO, weil $LO4$ in diesen Strukturen nicht gilt.
$(\mathbb{Q}; <), (\mathbb{R}; <)$ und $(\mathbb{R} \setminus \{0\}; <)$ sind Modelle von DLO.

2.2.8 Modelle der Mengenlehre Jedem Modell $(V; \in)$ der Mengenlehre ZF haftet etwas Paradoxes an, weil die Trägermenge V eines solchen Modells nach 2.1.1 eine Menge ist, wogegen das Universum aller Mengen keine Menge ist, etwa wegen der Russellschen Antinomie: V ist zwar eine Menge, aber V tritt in dem Modell $(V; \in)$ nicht als Menge auf. Ein Standardmodell von ZF in dem Sinne, wie $\mathcal{N}$ das Standardmodell von Z ist, kann es deshalb nicht geben.

Wichtig an diesen Beispielen sind uns die folgenden beiden Punkte.

(1) Wichtige mathematische Begriffe wie Gruppe, Körper, natürliche Zahlen treten hier unter dem einheitlichen Blickwinkel des Modellbegriffs auf. Es liegt nahe, dass sich dann weitere Begriffe wie Homomorphismus, Untergruppe, Unterring auch im allgemeinen Rahmen entwickeln lassen.

(2) Mathematische Theorien werden unter gegensätzlichen Gesichtswinkeln studiert. Die meisten Theorien werden aufgestellt, um möglichst viele Strukturen gemeinsam behandeln zu können. Da die Ergebnisse dieser Theorien – z. B. der Gruppentheorie – an vielen Stellen in der Mathematik angewendet werden können, haben diese Theorien große Bedeutung für einen rationellen, sparsamen Aufbau der Mathematik insgesamt.
Die Zahlentheorie Z dagegen ist intendiert für die Beschreibung einer einzigen Struktur, nämlich ihres Standardmodells $\mathcal{N}$. Es ist nicht trivial, dass diese Intention ihr Ziel nicht erreichen kann. Die Zermelo-Fraenkelsche Mengenlehre ZF beschreibt möglichst explizit unseren naiven Mengenbegriff, der unmittelbar von keinem Modell im Sinne von 2.2.3 erfasst wird. In diesen beiden Fällen tritt das Ökonomieprinzip offenbar zurück hinter dem Ziel, eine formale Grundlage für Teilgebiete der Mathematik zu schaffen.

Bisher haben wir in diesem Abschnitt Modelle vorher gegebener Theorien betrachtet. Man kann aber auch umgekehrt in kanonischer Weise eine Theorie zu einer vorher gegebenen Struktur einführen.

2.2.9 Definition *Es sei $\mathcal{A}$ eine Struktur zu einer Sprache L. Die Theorie von $\mathcal{A}$, $Th(\mathcal{A})$, ist die Theorie mit Sprache L, deren Axiome die sämtlichen in $\mathcal{A}$ wahren Sätze sind:*

$$Ax(Th(\mathcal{A})) := \{C \text{ ist Satz aus } L \mid \mathcal{A}(C) = w\}.$$

Dieser Begriff ist dem der logischen Theorie entgegengesetzt. Während eine logische Theorie keine Axiome, daher viele Modelle hat und dementsprechend schwach ist, hat $Th(\mathcal{A})$ unendlich viele Axiome, die im Rahmen ihrer Sprache die Struktur $\mathcal{A}$ möglichst vollständig beschreiben.

Der Begriff der Theorie einer Struktur legt weitere Begriffsbildungen nahe. Wir begnügen uns hier mit einem einfachen Lemma.

2.2.10 Lemma *Es sei $\mathcal{A}$ eine Struktur zu L.*
1. $\mathcal{A}$ ist ein Modell von $Th(\mathcal{A})$.
2. Für alle Sätze C aus L ist

$$\mathcal{A} \models C \Leftrightarrow C \in Ax(Th(\mathcal{A})) \Leftrightarrow Th(\mathcal{A}) \models C$$

Beweis. 1. Die Axiome von $Th(\mathcal{A})$ gelten per Definition in $\mathcal{A}$. Also ist (vgl. 2.2.3, 2) $\mathcal{A}$ Modell von $Th(\mathcal{A})$.
2. Die erste Äquivalenz steckt in obiger Definition. Zur zweiten Äquivalenz: Ist C ein Axiom einer Theorie, so gilt C in jedem Modell der Theorie, also auch in der Theorie selbst. Gilt umgekehrt ein Satz C in $Th(\mathcal{A})$, so gilt C nach 1. in $\mathcal{A}$.

2.3 Einfache semantische Gesetze

Wir wenden uns nun einfachen Gesetzen der Semantik zu, die man in der Mathematik auch ohne Beweis ständig verwendet. Das *Homomorphieprinzip* haben wir schon bei der Einführung der Wahrheitstafeln in 2.1.5 in einem Spezialfall angesprochen:
Für die Interpretation eines Terms oder einer Formel kommt es nicht auf die Gestalt der Teilausdrücke an, sondern nur auf deren Interpretation.

2.3.1 Satz *Homomorphieprinzip Gegeben sei eine Struktur $\mathcal{A}$ zu einer Sprache L und eine $\mathcal{A}$-Belegung $'$. Es sei s ein Term aus $L(\mathcal{A})$.*

1. *Ist $t(a)$ ein Term, so ist auch $t(s)$ ein Term, und es ist*

$$\mathcal{A}(t(s)') = \mathcal{A}(t'(\mathcal{A}(s'))).$$

2. *Ist $F(a)$ eine Formel, so ist auch $F(s)$ eine Formel, und es ist*

$$\mathcal{A}(F(s)') = \mathcal{A}(F'(\mathcal{A}(s'))).$$

3. *Sind B und $F(B)$ Formeln, so sind $F(\top), F(\bot)$ Formeln, und:*
   Aus $\mathcal{A}(B') = w$ folgt $\mathcal{A}(F(B)') = A(F'(\top))$,
   Aus $\mathcal{A}(B') = f$ folgt $\mathcal{A}(F(B)') = A(F'(\bot))$.

Die Interpretation ist ein mit allen Funktions- und Prädikatszeichen und allen logischen Partikeln verträglicher Homomorphismus.

Beweis. Sei $k := \mathcal{A}(s') \in |\ \mathcal{A}\ |$.
1. Wir machen Induktion nach dem Aufbau von $t(a)$.
1.1. t ist $*_1$. Dann ist $t(s) \equiv s$ ein Term, und es ist

$$\mathcal{A}(t(s)') = \mathcal{A}(s') = \mathcal{A}(k) = \mathcal{A}(t'(k)).$$

1.2. t ist eine Variable b. Dann ist $t(s) \equiv b$ ein Term, und es ist

$$\mathcal{A}(t(s)') = \mathcal{A}(b') = \mathcal{A}(t'(k)).$$

1.3. Sonst ist $t \equiv ft_1 \dots t_n$ mit einem n-stelligen Funktionszeichen f und Termen $t_i(a)$, die nach 1.3.8 eindeutig bestimmt sind. Nach IV sind die $t_i(s)$ Terme, also ist $t(s) \equiv ft_1(s) \dots t_n(s)$ ein Term, und es ist

$$\begin{array}{rcll} \mathcal{A}(t(s)') & = & \mathcal{A}(ft_1(s)' \dots t_n(s)') & \\ & = & f_{\mathcal{A}}(\mathcal{A}(t_1(s)'), \dots, \mathcal{A}(t_n(s)')) & \text{nach 2.1.3} \\ & = & f_{\mathcal{A}}(\mathcal{A}(t_1'(k)), \dots, \mathcal{A}(t_n'(k))) & \text{nach IV} \\ & = & \mathcal{A}(ft_1'(k) \dots t_n'(k)) & \text{nach 2.1.3} \\ & = & \mathcal{A}(t'(k)). & \end{array}$$

Mit Induktion folgt, dass 1. auf alle Terme $t(a)$ zutrifft.

Beweis von 2. durch Induktion nach dem Aufbau von $F(a)$:

2.1. F ist $t_1 = t_2$ mit Termen $t_i(a)$. Nach 1. sind die $t_i(s)$ Terme, also ist $F(s) \equiv t_1(s) = t_2(s)$ eine Formel, und es ist

$$\begin{array}{rcll} \mathcal{A}(F(s)') = w & \Leftrightarrow & \mathcal{A}(t_1(s)') = \mathcal{A}(t_2(s)') & \text{nach 2.1.3} \\ & \Leftrightarrow & \mathcal{A}(t_1'(k)) = \mathcal{A}(t_2'(k)) & \text{nach 1.} \\ & \Leftrightarrow & \mathcal{A}(t_1'(k) = t_2'(k)) = w & \text{nach 2.1.3} \end{array}$$

und das ist die Behauptung.

2.2. F ist $pt_1 \dots t_n$ mit Termen $t_i(a)$. Nach 1. sind die $t_i(s)$ Terme, also ist $F(s) \equiv pt_1(s) \dots t_n(s)$ eine (Prim-)Formel, und es ist

$$\begin{array}{rcll} \mathcal{A}(F(s)') = w & \Leftrightarrow & \mathcal{A}(pt_1(s)' \dots t_n(s)') = w & \\ & \Leftrightarrow & (\mathcal{A}(t_1(s)'), \dots, \mathcal{A}(t_n(s)')) \in p_{\mathcal{A}} & \text{nach 2.1.3} \\ & \Leftrightarrow & (\mathcal{A}(t_1'(k)), \dots, \mathcal{A}(t_n'(k))) \in p_{\mathcal{A}} & \text{nach 1.} \\ & \Leftrightarrow & \mathcal{A}(pt_1'(k) \dots t_n'(k)) = w & \text{nach 2.1.3} \end{array}$$

und das ist die Behauptung.

2.3. Für $F \equiv \bot$ ist nichts zu beweisen.

2.4. F ist $(F_1 \to F_2)$, genauer $\to F_1F_2$, mit Formeln $F_i(a)$. Nach IV sind die $F_i(s)$ Formeln, also ist $F(s) \equiv (F_1(s) \to F_2(s))$ eine Formel, und es gilt:

$$\begin{array}{lcll} \mathcal{A}(F(s)') = w & \Leftrightarrow & \text{aus } \mathcal{A}(F_1(s)') = w \text{ folgt } \mathcal{A}(F_2(s)') = w & \text{nach 2.1.3} \\ & \Leftrightarrow & \text{aus } \mathcal{A}(F_1'(k)) = w \text{ folgt } \mathcal{A}(F_2'(k)) = w & \text{nach IV} \\ & \Leftrightarrow & \mathcal{A}(F'(k)) = w & \text{nach 2.1.3} \end{array}$$

2.5. F ist $\forall x G(*_1, x)$ mit einer Formel $G(a, b)$, in der x nicht auftritt. Nach IV ist $G(s, b)$ eine Formel. In dieser tritt x nicht auf, weil x in s nicht auftreten kann. Also ist $\forall x G(s, x)$ eine Formel, und es gilt:

$$\begin{array}{lcll} \mathcal{A}(F(s)') = w & \Leftrightarrow & \mathcal{A}(G(s, c)') = w \text{ für alle } c \in\mid \mathcal{A} \mid & \text{nach 2.1.3} \\ & \Leftrightarrow & \mathcal{A}(G'(k, c)) = w \text{ für alle } c \in\mid \mathcal{A} \mid & \text{nach IV} \\ & \Leftrightarrow & \mathcal{A}(F'(k)) = w & \text{nach 2.1.3} \end{array}$$

Mit Induktion folgt, dass 2. auf alle Formeln $F(a)$ zutrifft.

Beweis von 3. durch Induktion nach dem Aufbau von $F(\bot)$:
3.1. $F(\bot)$ ist $\bot$ wegen $F \equiv *_1$. Dann ist stets $F(B) \equiv B$, und wie im Fall 1.1 ist (fast) nichts zu zeigen.

3.2. $F(\bot)$ ist eine Primformel, und F ist nicht $*_1$. Dann tritt $*_1$ nicht in F auf, es ist $F(B) \equiv F$, und wie im Fall 1.2 ist nichts zu zeigen.

Die Fälle 3.3 $F \equiv (F_1 \to F_2)$ und 3.4 $F \equiv \forall x G(*_1, x)$ entsprechen den Fällen 2.4 und 2.5, wenn man dort s durch B und k durch $\top$ bzw. $\bot$ ersetzt. Wenn gebundene Variablen in B auftreten, lässt sich allerdings der Schluss von 2.5 nur so auf 3.4 übertragen, dass $F(\top), F(\bot)$ nach IV Formeln sind, nicht aber $F(B)$ für beliebiges B. Deshalb ist dies in die Voraussetzung von 3. aufgenommen.

Damit ist der Satz bewiesen.

Die Wahrheit von Sequenzen bleibt offenbar beim Übergang zu „größeren“ Sequenzen erhalten.

2.3.2 Definition *Aus einer Sequenz* $\Gamma : \Delta$ *folgt strukturell eine Sequenz* $\Gamma_1 : \Delta_1$, *man schreibt* $\Gamma : \Delta \subset_S \Gamma_1 : \Delta_1$, *wenn*

$$\Gamma \subseteq \Gamma_1 \quad \textit{und} \quad \Delta \subseteq \Delta_1 \cup \{\bot\}.$$

$\subset_S$ bedeutet also für die *Ante*zedenten die übliche mengentheoretische Inklusion; für die *Suk*zedenten spielt es dagegen keine Rolle, ob und wo die Formel $\bot$ in ihnen auftritt. Diese Sonderrolle von $\bot$ wird wie folgt gerechtfertigt.

2.3.3 Lemma *Sei* $\Gamma : \Delta \subset_S \Gamma_1 : \Delta_1$ *in* L, *sei* $\mathcal{A}$ *eine Struktur zu* L *und* $'$ *eine* $\mathcal{A}$*-Belegung.*

$$\textit{Wenn } \mathcal{A}(\Gamma' : \Delta') = w \textit{ ist, ist auch } \mathcal{A}(\Gamma_1' : \Delta_1') = w.$$

Beweis. Sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma_1$. Wegen $\Gamma \subseteq \Gamma_1$ gilt dies auch für alle $C \in \Gamma$. Ist nun $\mathcal{A}(\Gamma' : \Delta') = w$, so ist dann $\mathcal{A}(D') = w$ für ein $D \in \Delta$. Wegen $\Delta \subseteq \Delta_1, \bot$ ist dieses $D \in \Delta_1$, weil $\mathcal{A}(\bot) = w$ nicht möglich ist. Also ist $\mathcal{A}(\Gamma_1' : \Delta_1') = w$.

2.3.4 Korollar *Sei* $\Gamma : \Delta \subset_S \Gamma_1 : \Delta_1$ *in* L *und sei* $\mathcal{A}$ *eine Struktur zu* L.

$$\textit{Aus } \mathcal{A} \models \Gamma : \Delta \textit{ folgt } \mathcal{A} \models \Gamma_1 : \Delta_1.$$

Beweis. Sei $'$ eine $\mathcal{A}$-Belegung. Dann ist $\mathcal{A}(\Gamma' : \Delta') = w$ nach Voraussetzung, also $\mathcal{A}(\Gamma_1' : \Delta_1') = w$ nach 2.3.3. Mithin gilt $\Gamma_1 : \Delta_1$ in $\mathcal{A}$.

Die Gültigkeit von Sequenzen lässt sich ohne weiteres auf die Gültigkeit von Formeln zurückspielen.

2.3.5 Lemma *Es sei* $\Gamma : \Delta$ *eine Sequenz von* L, $\Gamma = \{C_1, \ldots, C_m\}$, $\Delta = \{D_1, \ldots, D_n\}$ *und* $\mathcal{A}$ *eine Struktur zu* L. *Dann gilt:*

$$\mathcal{A} \models \Gamma : \Delta \quad \Leftrightarrow \quad \mathcal{A} \models C_1 \wedge \ldots \wedge C_m \to D_1 \vee \ldots \vee D_n$$

Beweis. Für jede $\mathcal{A}$-Belegung $'$ ist $\Gamma' : \Delta'$ eine geschlossene Sequenz aus $L(\mathcal{A})$. Dann ist nach 2.1.5, d.

$$\mathcal{A}(\Gamma' : \Delta') = w \Leftrightarrow \mathcal{A} \models C_1' \wedge \ldots \wedge C_m' \to D_1' \vee \ldots \vee D_n'$$

für jede $\mathcal{A}$-Belegung $'$, und daraus folgt die Behauptung.

Wir notieren noch eine grundlegende Eigenschaft des Allquantors.

2.3.6 Lemma *Es sei* $F(a)$ *eine Formel einer Sprache* L, *so dass* a *in der Nennform* F *nicht auftritt, und* $\mathcal{A}$ *sei eine Struktur zu* L. *Dann ist*

$$\mathcal{A} \models F(a) \quad \Leftrightarrow \quad \mathcal{A} \models \forall x F(x).$$

Beweis. $\Rightarrow$: Wir setzen die linke Seite voraus. Sei $'$ eine $\mathcal{A}$-Belegung und $k \in |\, \mathcal{A} \,|$. Dann sei * die $\mathcal{A}$-Belegung mit

$$a^* := k \quad \text{und} \quad b^* := b' \text{ für } b \not\equiv a.$$

Weil a in F nicht auftritt, ist $F^* \equiv F'$, also

$$\mathcal{A}(F'(k)) = \mathcal{A}(F^*(a^*)) = w$$

wegen der linken Seite. Da $k \in |\, \mathcal{A} \,|$ und $'$ beliebig sind, folgt $\mathcal{A}(\forall x F(x)') = w$ und schließlich $\mathcal{A} \models \forall x F(x)$.

$\Leftarrow$: Die rechte Seite besagt: $\mathcal{A}(F'(k)) = w$ für alle $k \in |\, \mathcal{A} \,|$, speziell auch für $k = a'$, also $\mathcal{A}(F'(a')) = w$, und dies für alle $\mathcal{A}$-Belegungen $'$. Daraus folgt $\mathcal{A} \models F(a)$.

2.3.7 Korollar *Es sei B eine Formel aus L, $\forall B$ ein Allabschluss von B und $\mathcal{A}$ eine Struktur für L.*

$$\mathcal{A} \models B \quad \Leftrightarrow \quad \mathcal{A}(\forall B) = w$$

Beweis. Sei $B \equiv F(a_1, \ldots, a_n)$ wie in 1.1.15. Wir induzieren nach n.

1. $n = 0$. Dann ist $B \equiv \forall B$ ein Satz, und es ist nichts zu beweisen.

2. $n \rightsquigarrow n+1$. Da a_{n+1} von den übrigen a_i verschieden ist, ist nach 2.3.6

$$\mathcal{A} \models B \quad \Leftrightarrow \quad \mathcal{A} \models C \text{ für } C :\equiv \forall x_{n+1} F(a_1, \ldots, a_n, x_{n+1}).$$

Nach IV ist die rechte Seite äquivalent zu $\mathcal{A}(\forall C) = w$, und einer von diesen Allabschlüssen von C ist der gegebene $\forall B$.

2.4 Aufgaben

2.4.1 Prüfen Sie die Wahrheitstafeln aus 2.1.5 nach.

2.4.2 (Definition) *Der Junktor $\sqcup$ (lies: entweder ... oder ...) hat die Interpretation*

$$\mathcal{A}(B \sqcup C) = w \quad \Leftrightarrow \quad \textit{entweder } \mathcal{A}(B) = w \textit{ oder } \mathcal{A}(C) = w$$

a. Stellen Sie die Wahrheitstafel für $\sqcup$ auf.

b. Geben Sie eine Definition von $\sqcup$ mittels $\neg, \wedge, \vee, \rightarrow$ an, d.h. geben Sie eine Formel $F(B, C)$ an, so dass

   i. stets $\mathcal{A}(B \sqcup C) = \mathcal{A}(F(B, C))$ ist und

   ii. F allein mit $\neg, \wedge, \vee, \rightarrow$ aus Nennzeichen aufgebaut ist.

c. Beweisen bzw. widerlegen Sie:

   i. Es ist stets $\mathcal{A}(B \sqcup C) = w \Leftrightarrow \mathcal{A}(B) \neq \mathcal{A}(C)$

   ii. Es ist stets $\mathcal{A}(B \sqcup C) = \mathcal{A}(\neg B \sqcup \neg C)$

   iii. Es ist stets $\mathcal{A}(B \sqcup B) = f$

   iv. $\mathcal{A}(B \sqcup (C \sqcup D)) = w \Leftrightarrow$ genau einer der drei Werte $\mathcal{A}(B), \mathcal{A}(C), \mathcal{A}(D)$ ist w

   v. Es ist stets $\mathcal{A}(B \sqcup (C \sqcup D)) = \mathcal{A}((B \sqcup C) \sqcup D)$

2.4.3 Beweisen Sie die Behauptung a. aus 2.1.5 für $n > 0$ durch Induktion nach n.

2.4.4 Sei $F(a)$ eine Formel von L, $FV(F) = \emptyset$. Geben Sie Sätze $\exists_{\geq 2} x F(x)$ und $\exists_{\leq 2} x F(x)$ an, die für beliebige Strukturen $\mathcal{A}$ für L folgendes besagen:

a. $\mathcal{A}(\exists_{\geq 2} x F(x)) = w \quad \Leftrightarrow$ es gibt mindestens 2 Elemente $c \in \mid \mathcal{A} \mid$ mit $\mathcal{A}(F(c)) = w$

b. $\mathcal{A}(\exists_{\leq 2} x F(x)) = w \quad \Leftrightarrow$ es gibt höchstens 2 Elemente $c \in \mid \mathcal{A} \mid$ mit $\mathcal{A}(F(c)) = w$

2.4.5 Man gebe eine Nennform F an, so dass $F(a) \equiv a = a$ ist und $\forall x F(x)$ nicht in allen Strukturen gilt. Man charakterisiere die Strukturen, in denen $\forall x F(x)$ gilt.

2.4.6 Zeigen Sie:

$$\text{Aus } T \models \Gamma : B, \Delta \text{ und } T \models \Gamma, B : \Delta \text{ folgt } T \models \Gamma : \Delta$$

2.4.7 Beweisen Sie die Gültigkeit folgender Formeln in T_G:

a. $a \circ c = b \circ c \quad \rightarrow \quad a = b$

b. $\exists x \; a \circ x = x \quad \rightarrow \quad a = e$

§3 Syntax: Herleitungen in mathematischen Theorien

Wie kann man sich davon überzeugen, dass eine Formel B (allgemeiner: eine Sequenz $\Gamma : \Delta$) in einer Theorie T gültig ist? Wollte man nach 2.2.3 vorgehen, müsste man sich eine Übersicht über *sämtliche* Modelle von T verschaffen und in jedem Modell $\mathcal{A}$ von T einzeln nachprüfen, ob B in $\mathcal{A}$ gilt oder nicht. Dies ist in der Regel nicht möglich, weil

a. eine Theorie unendlich viele Modelle besitzen kann (so hat die Gruppentheorie etwa alle Gruppen als Modelle) und

b. ein „einfaches Kriterium“ fehlt, mit dem man entscheiden kann, ob B in einem Modell $\mathcal{A}$ von T gilt oder nicht.

In diesem Paragraphen verschaffen wir uns einen konstruktiven Zugang zu dem Begriff der Gültigkeit in T in dem Sinne, dass wir eine Art „Maschine“ angeben, die genau die in T gültigen Sequenzen ausdruckt. Damit liefert sie auch genau die in T gültigen Formeln. Die „Maschine“ ist ein Kalkül, der Beweise syntaktisch erzeugt. Er ist gegeben durch Axiome und durch Schlussregeln, wie man aus bereits hergeleiteten Sequenzen durch rein syntaktisches Umformen (also ohne inhaltliche Einsicht) neue Sequenzen herleitet.

3.1 Der Sequenzenkalkül

Unser Sequenzenkalkül geht auf Gentzen 1935 zurück. Wir beziehen uns jeweils auf eine feste Sprache L.

3.1.1 Definition *Schlussregeln teilen wir in der Form*

$$\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n \vdash \Gamma : \Delta$$

mit, worin $\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n; \Gamma : \Delta$ *Sequenzenschemata aus der Sprache* L *mitteilen. Ein Einzelfall*

$$\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n \vdash \Gamma : \Delta$$

einer solchen Schlussregel mit konkreten Sequenzen $\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n; \Gamma : \Delta$ *ist ein Schluss, und zwar der Schluss von* $\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n$ *auf* $\Gamma : \Delta$.

Dabei heißen $\Gamma_1 : \Delta_1; \ldots; \Gamma_n : \Delta_n$ *die Prämissen,* $\Gamma : \Delta$ *die Konklusion des Schlusses.*

3.1.2 Definition *Logische Axiome sind alle Sequenzen der Gestalt*

$$\Gamma, P : P, \Delta,$$

worin P *eine Primformel von* L *ist, und*

$$\Gamma, \bot : \Delta.$$

Logische Grundschlussregeln sind

$(\to S)$ $\Gamma, A : B, \Delta \vdash \Gamma : A \to B, \Delta$

$(\to A)$ $\Gamma : A, \Delta; \Gamma, B : \Delta \vdash \Gamma : A \to B : \Delta$

$(\forall S)$ $\Gamma : F(a), \Delta \vdash \Gamma : \forall x F(x), \Delta$,
wenn a *weder in* Γ, Δ *noch in* F *auftritt.*

$(\forall A)$ $\Gamma, F(t) : \Delta \vdash \Gamma, \forall x F(x) : \Delta$

$(= \mathrm{I})$ $\Gamma, t = t : \Delta \vdash \Gamma : \Delta$

$(= \mathrm{F})$ $\Gamma, fs_1 \ldots s_n = ft_1 \ldots t_n : \Delta \vdash \Gamma, s_1 = t_1, \ldots, s_n = t_n : \Delta$, *falls* $n > 0$ *ist.*

$(= \mathrm{P})$ $\Gamma, pt_1 \ldots t_n : \Delta \vdash \Gamma, ps_1 \ldots s_n,\ s_1 = t_1, \ldots, s_n = t_n : \Delta$, *falls* $n > 0$ *ist.*

Sei $T = (L, Ax(T))$ eine Theorie mit Sprache L. Die *T-Grundschlussregel* ist

(T) $\Gamma, B : \Delta \vdash \Gamma : \Delta$, *falls* B *ein Axiom der Theorie* T *ist.*

Bemerkung. Zu den logischen Partikeln $\to$ und $\forall$ gibt es jeweils zwei Grundschlussregeln, die diese Partikel ins Sukzedens bzw. Antezedens der Konklusion einführen. Entsprechend liest man

$(\to S)$	als $\to$-Einführung im Sukzedens,
$(\to A)$	als $\to$-Einführung im Antezedens,
$(\forall S)$	als $\forall$-Einführung im Sukzedens,
$(\forall A)$	als $\forall$-Einführung im Antezedens.

$\bot$ wird nur durch ein Axiomschema eingeführt.

Die drei Gleichheitsregeln

(= I)	Identitätsregel
(= F)	Gleichheitsregel für Funktionszeichen
(= P)	Gleichheitsregel für Prädikatszeichen

erlauben es, Gleichungen oder andere Primformeln unter den angegebenen Bedingungen fortzulassen. Dabei ist die Forderung $n > 0$ keine Einschränkung. Denn im Falle $n = 0$ wird aus (= F)

$$\Gamma, f = f : \Delta \vdash \Gamma : \Delta,$$

was schon unter (= I) fällt, und bei (= P) fallen Prämisse und Konklusion zusammen, so dass der Schluss überflüssig ist.

3.1.3 Definition *Bei den Grundschlüssen nach den Regeln* $(\rightarrow S), (\rightarrow A)$, $(\forall S), (\forall A)$ *heißt die in der Konklusion des Schlusses besonders bezeichnete Formel* $A \rightarrow B$ *bzw.* $\forall x F(x)$ *die Hauptformel des Schlusses. Entsprechend heißen die Formeln aus den Formelmengen* Γ, Δ *Nebenformeln des Schlusses. Schlüsse nach den anderen Grundschlussregeln haben keine Hauptformel.*

Die bei $(\forall S)$ *aufgeführte Bedingung des Inhalts, dass die freie Variable* a *nicht in der Konklusion des* $(\forall S)$*-Schlusses auftreten darf, nennt man die Variablenbedingung.*

Zur Beachtung. Die Hauptformel eines Schlusses kann auch (aber muss nicht) als Nebenformel desselben Schlusses auftreten, so dass durch einen solchen Schluss nicht unbedingt eine neue Formel in die Konklusion hineinkommen muss. Ebenso können auch die direkten Subformeln $A, B, F(a), F(t)$ der Hauptformeln $A \rightarrow B$ bzw. $\forall x F(x)$ dieser Schlüsse auch als Nebenformeln auftreten; sie müssen also bei diesen Schlüssen nicht notwendig verlorengehen. In extremen Fällen können Prämisse und Konklusion eines Schlusses sogar übereinstimmen.

Formale Beweise in einem solchen Formalismus sind baumartige Figuren, an deren unterem Ende die bewiesene Sequenz steht, an deren Spitzen die logischen Axiome stehen und in denen die Prämissen eines Grundschlusses jeweils nebeneinander über der Konklusion dieses Grundschlusses stehen. Diese Vorstellung wird durch die folgende Definition wiedergegeben.

3.1.4 Induktive Definition *der Herleitungen von Sequenzen in einer Theorie T.*

1. *Jedes logische Axiom $\Gamma : \Delta$ ist eine Herleitung von $\Gamma : \Delta$ in T.*
2. *Ist für jedes $i = 1, \ldots, n (n > 0) H_i$ eine Herleitung von $\Gamma_i : \Delta_i$ in T und ist*

   $$\Gamma_1 : \Delta_1, \ldots, \Gamma_n : \Delta_n \vdash \Gamma : \Delta$$

   ein (logischer oder T-) Grundschluss, so ist

   $$\frac{H_1 \ldots H_n}{\Gamma : \Delta}$$

   eine Herleitung von $\Gamma : \Delta$ in T.

In dieser Definition kann man 1. als den Fall $n = 0$ von 2. ansehen. Der Fall $n = 2$ von 2. kommt in unserem Sequenzenkalkül bei der Regel $(\rightarrow A)$ vor, sonst nur der Fall $n = 1$. Man kann den Sequenzenkalkül aber auch so formulieren, dass die Gleichheitsregeln für n-stellige Funktions- und Prädikatszeichen n bzw. $n+1$ Prämissen haben. Auch für solche anderen Kalküle legt die Definition 3.1.4 offenbar einen Herleitungsbegriff fest. Wir gehen darauf hier nicht weiter ein.

3.1.5 Definition *Eine Sequenz $\Gamma : \Delta$ ist herleitbar in T, wir schreiben*

$$T \vdash \Gamma : \Delta,$$

wenn es eine Herleitung von $\Gamma : \Delta$ in T gibt. Die Herleitbarkeit einer Formel B identifizieren wir mit der Herleitbarkeit der Sequenz $\emptyset : \{B\}$.

Eine Herleitung H heißt logisch, wenn H eine Herleitung in einer logischen Theorie ist, wenn also in H keine T-Grundschlüsse vorkommen. Wir schreiben $\vdash \Gamma : \Delta$, wenn die Theorie, in der $\Gamma : \Delta$ herleitbar ist, aus dem Kontext klar ist oder wenn es auf sie nicht ankommt, insbesondere wenn $\Gamma : \Delta$ logisch herleitbar ist.

3.1.6 Beispiele logischer Herleitungen P, Q seien Primformeln. Dann ist

$$\dfrac{\dfrac{\dfrac{P, Q : Q, \bot}{P, Q : P, \bot} \qquad \dfrac{P, Q, \bot : \bot}{P, Q, \neg Q : \bot}}{P, Q, P \rightarrow \neg Q : \bot}}{P, Q : \neg(P \rightarrow \neg Q)}$$

eine Herleitung von $P, Q : P \wedge Q$. Denn die drei Sequenzen an den Spitzen sind logische Axiome, die beiden oberen Schlüsse sind $(\rightarrow A)$-Schlüsse, und der unterste Schluss ist ein $(\rightarrow S)$-Schluss.

$F(t)$ sei eine Primformel. Dann ist

$$\dfrac{\dfrac{F(t) : \quad F(t), \bot \qquad\qquad F(t), \bot : \bot}{F(t), \neg F(t) : \bot}}{\dfrac{F(t), \forall x \neg F(x) : \bot}{F(t) : \neg \forall x \neg F(x)}}$$

eine Herleitung von $F(t) : \exists x F(x)$. Denn die beiden obersten Sequenzen sind logische Axiome, und die drei Schlüsse sind nacheinander ein $(\rightarrow A)$-, ein $(\forall A)$- und ein $(\rightarrow S)$-Schluss.

t sei ein Term. Dann ist

$$\dfrac{t = t \quad : t = t}{: t = t}$$

eine Herleitung der Sequenz $\quad : t = t$, also der Formel $t = t$. Denn die obere Sequenz ist ein logisches Axiom, und der einzige Schluss ist ein $(= I)$-Schluss.

Der systematischen Untersuchung von Herleitungen und herleitbaren Sequenzen (und Formeln) ist das ganze nächste Kapitel gewidmet.

Wie wir nach dem Aufbau der Terme und Formeln einer Sprache induzieren können, so können wir auch nach dem Aufbau der Herleitungen in einer Theorie Induktion machen, einfach weil die Herleitungen induktiv definiert sind. Diese Induktion bezeichnet man als *Herleitungsinduktion*.

3.2 Korrektheit

Der Herleitungsbegriff und der Gültigkeitsbegriff entstammen begrifflich weit voneinander getrennten Bereichen. Trotzdem besteht zwischen der Herleitbarkeit und der Gültigkeit in einer Theorie T ein offenkundiger Zusammenhang: Wie wir gleich genau ausführen, gelten die logischen Axiome in jeder Struktur, und die Grundschlussregeln sind *korrekt* in dem Sinne, dass sich die Gültigkeit (in einer Theorie T) von den Prämissen eines Grundschlusses stets auf seine Konklusion vererbt. Daher können auch nur gültige Sequenzen herleitbar sein.

3.2.1 Korrektheitssatz *Jede in T herleitbare Sequenz ist gültig in T:*

$$T \vdash \Gamma : \Delta \quad \Rightarrow \quad T \models \Gamma : \Delta$$

Beweis. Gegeben sei eine Herleitung H einer Sequenz in T und ein Modell $\mathcal{A}$ von T. Nach 2.2.3 haben wir zu zeigen, dass diese Sequenz in $\mathcal{A}$ gilt. Das zeigen wir durch Herleitungsinduktion, also durch Induktion nach dem Aufbau von H.

1.1. H ist ein logisches Axiom $\Gamma, P : P, \Delta$. Ist $'$ eine $\mathcal{A}$-Belegung mit $\mathcal{A}(C') = w$ für alle C aus dem Antezedens, so ist insbesondere $\mathcal{A}(P') = w$, so dass $\mathcal{A}(D') = w$ ist für ein D aus dem Sukzedens, nämlich für P. Also ist $\mathcal{A}(\Gamma', P' : P', \Delta') = w$. Da $'$ eine beliebige $\mathcal{A}$-Belegung war, gilt $\Gamma, P : P, \Delta$ in $\mathcal{A}$.

1.2. H ist ein logisches Axiom $\Gamma, \bot\colon \Delta$. Da bei jeder $\mathcal{A}$-Belegung $'$ $\mathcal{A}(\bot') = \mathcal{A}(\bot) = f$ ist, ist niemals $\mathcal{A}(C') = w$ für alle C aus dem Antezedens. Also ist $\mathcal{A}(\Gamma', \bot'\colon \Delta') = w$, und es folgt $\mathcal{A} \models \Gamma, \bot\colon \Delta$.

2. H endet mit einem Grundschluss

$$\Gamma_1 : \Delta_1 \vdash \Gamma_0 : \Delta_0 \text{ oder } \Gamma_1 : \Delta_1; \Gamma_2 : \Delta_2 \vdash \Gamma_0 : \Delta_0.$$

Nach IV gelten dann $\Gamma_i : \Delta_i$ für $i = 1$ oder $i = 1, 2$ in $\mathcal{A}$. Wir unterscheiden nach der Regel, zu der dieser Grundschluss gehört:

$(\to S)$ $\Gamma, A : B, \Delta \vdash \Gamma : A \to B, \Delta$.

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$. Ist $\mathcal{A}(A' \to B') = w$, so ist $\mathcal{A}(\Gamma : A \to B, \Delta)' = w$. Sei nun $\mathcal{A}(A' \to B') = f$. Dann ist $\mathcal{A}(A') = w$ und $\mathcal{A}(B') = f$. Da nach IV $\Gamma, A : B, \Delta$ in $\mathcal{A}$ gilt, ist dann $\mathcal{A}(D') = w$ für ein $D \in \Delta$.
Also ist in jedem Fall $\mathcal{A}(\Gamma : A \to B, \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma : A \to B, \Delta$.

$(\to A)$ $\Gamma : A, \Delta$ und $\Gamma, B : \Delta \vdash \Gamma, A \to B : \Delta$.

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma, A \to B$. Dann ist insbesondere $\mathcal{A}(A' \to B') = w$, also $\mathcal{A}(A') = f$ oder $\mathcal{A}(B') = w$. In beiden Fällen ist dann $\mathcal{A}(D') = w$ für ein $D \in \Delta$, und zwar im Fall $\mathcal{A}(A') = f$ wegen der IV für die 1. Prämisse $\Gamma : A, \Delta$, und im Fall $\mathcal{A}(B') = w$ wegen der IV für die 2. Prämisse $\Gamma, B : \Delta$.
Also ist in jedem Fall $\mathcal{A}(\Gamma, A \to B : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma, A \to B : \Delta$.

$(\forall S)$ $\Gamma : F(a), \Delta \vdash \Gamma : \forall x F(x), \Delta$, wobei a nicht in Γ, Δ, F auftritt.

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$. Ist $\mathcal{A}(\forall x F(x)') = w$, so ist $\mathcal{A}(\Gamma : \forall x F(x), \Delta)' = w$. Sei nun $\mathcal{A}(\forall x F'(x)) = f$. Dann gibt es ein $k \in |\mathcal{A}|$ mit $\mathcal{A}(F'(k)) = f$. Wir definieren eine $\mathcal{A}$-Belegung * durch

$$a^* := k \quad \text{und} \quad b^* := b' \text{ für } b \not\equiv a.$$

Wegen der Variablenbedingung ist dann $C^* \equiv C'$ für $C \in \Gamma, \Delta$ und $F^* \equiv F'$. Also ist $\mathcal{A}(C^*) = w$ für alle $C \in \Gamma$ und $\mathcal{A}(F(a)^*) = \mathcal{A}(F^*(k)) = \mathcal{A}(F'(k)) = f$. Dann gibt es nach IV ein $D \in \Delta$, so dass $\mathcal{A}(D') = \mathcal{A}(D^*) = w$ ist.
Also ist in jedem Fall $\mathcal{A}(\Gamma : \forall x F(x), \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma : \forall x F(x), \Delta$.

$(\forall A) \quad \Gamma, F(t) : \Delta \vdash \Gamma, \forall x F(x) : \Delta.$

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma, \forall x F(x)$. Dann ist $\mathcal{A}(\forall x F'(x)) = w$, also nach 2.3.1 auch $\mathcal{A}(F(t)') = \mathcal{A}(F'(\mathcal{A}(t')) = w$. Dann gibt es nach IV ein $D \in \Delta$ mit $\mathcal{A}(D') = w$. Also ist $\mathcal{A}(\Gamma, \forall x F(x) : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma, \forall x F(x) : \Delta$.

$(= \text{I}) \quad \Gamma, t = t : \Delta \vdash \Gamma : \Delta.$

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$. Da offenbar $\mathcal{A}(t') = \mathcal{A}(t')$, also $\mathcal{A}(t' = t') = w$ ist, gibt es dann nach IV ein $D \in \Delta$ mit $\mathcal{A}(D') = w$. Also ist $\mathcal{A}(\Gamma : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma : \Delta$.

$(=\text{F}) \quad \Gamma, f s_1 \dots s_n = f t_1 \dots t_n : \Delta \vdash \Gamma,\ s_1 = t_1, \dots, s_n = t_n : \Delta.$

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$ und $\mathcal{A}(s_i' = t_i') = w$, also $\mathcal{A}(s_i') = \mathcal{A}(t_i')$ für $i = 1, \dots, n$. Dann ist

$$\mathcal{A}(f s_1 \dots s_n)' = f_{\mathcal{A}}(\mathcal{A}(s_1'), \dots, \mathcal{A}(s_n')) = f_{\mathcal{A}}(\mathcal{A}(t_1'), \dots, \mathcal{A}(t_n')) = \mathcal{A}(f t_1 \dots t_n)'$$

und daher $\mathcal{A}(f s_1 \dots s_n = f t_1 \dots t_n)' = w$. Dann gibt es nach IV ein $D \in \Delta$ mit $\mathcal{A}(D') = w$. Also ist $\mathcal{A}(\Gamma, s_1 = t_1, \dots, s_n = t_n : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma, s_1 = t_1, \dots, s_n = t_n : \Delta$.

$(= \text{P}) \quad \Gamma, p t_1 \dots t_n : \Delta \vdash \Gamma, p s_1 \dots s_n, s_1 = t_1, \dots, s_n = t_n : \Delta.$

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$ und $\mathcal{A}(s_i' = t_i') = w$, also $\mathcal{A}(s_i') = \mathcal{A}(t_i')$ für $i = 1, \dots, n$, und $\mathcal{A}(p s_1 \dots s_n)' = w$, also

$$(\mathcal{A}(t_1'), \dots, \mathcal{A}(t_n')) = (\mathcal{A}(s_1'), \dots, (\mathcal{A}(s_n')) \in p_{\mathcal{A}}$$

und damit $\mathcal{A}(p t_1 \dots t_n)' = w$. Dann gibt es nach IV ein $D \in \Delta$ mit $\mathcal{A}(D') = w$. Also ist $\mathcal{A}(\Gamma, p s_1 \dots s_n, s_1 = t_1, \dots, s_n = t_n : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma, p s_1 \dots s_n, s_1 = t_1, \dots s_n = t_n : \Delta$.

(T) $\quad \Gamma, B : \Delta \vdash \Gamma : \Delta$, wobei $B \in Ax(T)$ ist.

Sei $'$ eine $\mathcal{A}$-Belegung, und sei $\mathcal{A}(C') = w$ für alle $C \in \Gamma$. Da $\mathcal{A}$ ein Modell von T und B ein Axiom von T ist, ist $\mathcal{A}(B') = \mathcal{A}(B) = w$. Dann gibt es nach IV ein $D \in \Delta$ mit $\mathcal{A}(D') = w$. Also ist
$\mathcal{A}(\Gamma : \Delta)' = w$, und es folgt $\mathcal{A} \models \Gamma : \Delta$.

Mit Herleitungsinduktion folgt, dass jede in T herleitbare Sequenz in $\mathcal{A}$ gilt. Da $\mathcal{A}$ ein beliebiges Modell von T ist, ist damit der Korrektheitssatz bewiesen.

Der Beweis des Korrektheitssatzes funktioniert für jedes Modell von T einzeln: Als Induktionsvoraussetzung verwenden wir nicht, dass die Prämissen des letzten Schlusses in allen Modellen von T gelten, sondern nur, dass sie in einem Modell $\mathcal{A}$ von T gelten. Dann gilt auch die Konklusion in diesem Modell $\mathcal{A}$. Der Beweis funktioniert aber nicht für jede $\mathcal{A}$-Belegung einzeln, wie man am Induktionsschritt für $(\forall S)$ sieht. Dort muss die Belegung $'$ eventuell für die Variable abgeändert werden, die der Variablenbedingung unterworfen ist. Die Grundschlussregel $(\forall S)$ mit ihrer Variablenbedingung nimmt insofern eine Sonderstellung ein.

Im Spezialfall, dass $\Gamma : \Delta$ eine einzige Formel B ist, dass also $\Gamma = \emptyset$ und $\Delta = \{B\}$ ist, erhält man aus 3.2.1 unmittelbar die „Hilbert-artige“ Fassung des Korrektheitssatzes:

3.2.2 Korollar *Jede in einer Theorie T herleitbare Formel gilt in T:*

$$T \vdash B \quad \Rightarrow \quad T \models B$$

Der Korrektheitssatz stellt sicher, dass jedenfalls nicht zu viele Sequenzen oder Formeln in T herleitbar sind. Dass es auch nicht zu wenige sind, sagt der *Vollständigkeitssatz*:

Jede in T gültige Formel ist herleitbar in T.

Er wird in Kapitel 3 bewiesen. Korrektheit und Vollständigkeit zusammen besagen, dass der (formale, syntaktische) Herleitbarkeitsbegriff gleichwertig ist zum (inhaltlichen, semantischen) Gültigkeitsbegriff.

3.2.3 Anwendung auf die Gruppentheorie Für die Gruppentheorie und ihre Modelle, die Gruppen, sagt der Korrektheitssatz: Jede in der Gruppentheorie herleitbare Formel gilt in jeder Gruppe.

Der Vollständigkeitssatz sagt: Jede Formel in der Sprache der Gruppentheorie, die in jeder Gruppe gilt, ist in der Gruppentheorie herleitbar: Zu jedem Beweis einer solchen Formel mit höheren Mitteln (etwa Homomorphismen oder Kategorien) gibt es einen *elementaren* Beweis derselben Formel.

3.2.4 Anwendung auf die Zahlentheorie Während es ein Ziel der Gruppentheorie ist, möglichst viele Strukturen unter einem einheitlichen Gesichtspunkt zu betrachten, wollte man mit der Zahlentheorie genau eine Struktur beschreiben, nämlich das Standardmodell $\mathcal{N}$. Dies ist jedoch nicht möglich: Selbst die Theorie $Th(\mathcal{N})$, deren Axiome *alle* in $\mathcal{N}$ gültigen Sätze von $L(Z)$ sind, hat Modelle, die nicht isomorph zu $\mathcal{N}$ sind. Dies folgern wir in Kapitel 4 aus dem Vollständigkeitssatz.

Der Korrektheitssatz sagt: Jede in Z herleitbare Formel gilt in jedem Modell von Z, speziell also in $\mathcal{N}$.

Der Vollständigkeitssatz sagt: Jede in allen Modellen von Z gültige Formel von $L(Z)$ ist herleitbar in Z.

Davon zu unterscheiden ist die *Unvollständigkeit* von Z bzgl. $\mathcal{N}$:

Es gibt eine in $\mathcal{N}$ gültige Formel von $L(Z)$, die nicht in Z herleitbar ist.

In $\mathcal{N}$ gelten also Formeln, die nicht in allen Modellen von Z gelten; in $Th(\mathcal{N})$ sind mehr Formeln herleitbar als in Z.

Diese Aussage ist ein Spezialfall des ersten Gödelschen Satzes (Gödel 1931). Zu seinem Beweis braucht man Methoden der Rekursionstheorie. Wir beweisen diesen Satz hier nicht.

3.3 Subformel-Eigenschaft und Schnittfreiheit

Wir diskutieren noch eine besondere Eigenschaft unseres Herleitungsbegriffs, nämlich seine Schnittfreiheit: Jede Formel aus der Prämisse eines logischen Grundschlusses tritt selbst in der Konklusion dieses Grundschlusses auf, oder sie ist direkte Subformel dieser Konklusion. Einzige Ausnahme von diesem Prinzip bilden die (trivialen) Gleichungen $t = t$, die bei (= I)-Schlüssen verlorengehen oder (aus einer Herleitung) *herausgeschnitten* werden. Prädikatszeichen außer = ebenso wie der logische Aufbau der Formeln (mit $\bot, \to, \forall$) können dagegen niemals durch logische Grundschlüsse, höchstens durch T-Schlüsse weggeschnitten werden. Unser Herleitungsbegriff heißt deshalb (in seinem logischen Teil) *schnittfrei*.

3.3.1 Schnittregel Bei der *Schnittregel*

$$(\mathit{Schnitt}) \quad \Gamma : B, \Delta; \quad \Gamma, B : \Delta \vdash \Gamma : \Delta$$

wird eine ganze Formel, die sog. *Schnittformel* B, weggeschnitten: Weder ihr logischer Aufbau noch die in B auftretenden Prädikatszeichen lassen sich i.a. aus der Konklusion $\Gamma : \Delta$ rekonstruieren. Die Schnittregel zerstört die Schnittfreiheit. Sie lässt sich deshalb nicht aus logischen Grundschlüssen zusammensetzen, obwohl sie eine korrekte Regel ist.

Unser Sequenzenkalkül geht mit Termen offenbar weniger sorgfältig um als mit Prädikatszeichen außer $=$ und dem Aufbau der Formeln: Bei $(\forall A)$-Schlüssen geht u.U. der Term t aus der Antezedensformel $F(t)$ verloren, während der logische Aufbau von $F(t)$ sich, angereichert um einen Allquantor, in der Hauptformel $\forall x F(x)$ wiederfindet.
Bei $(=\mathrm{F})$-Schlüssen gehen u.U. die Terme $fs_1 \ldots s_n,\ ft_1 \ldots t_n$ und damit das Funktionszeichen f verloren. Bei $(=\mathrm{P})$-Schlüssen geht zwar u.U. eine Primformel $pt_1 \ldots t_n$ verloren, aber eine Primformel $ps_1 \ldots s_n$ mit demselben Prädikatszeichen p bleibt erhalten, auch wenn p das Gleichheitszeichen ist.

Alles dies wird in der folgenden Definition berücksichtigt.

3.3.2 Definition *der direkten Subformeln einer Formel*

1. *Jede Formel $pt_1 \ldots t_n$ ist direkte Subformel von $ps_1 \ldots s_n$.*
2. *$\bot$ hat keine direkten Subformeln.*
3. *A und B sind direkte Subformeln von $(A \to B)$.*
4. *Jede Formel $F(t)$ ist direkte Subformel von $\forall x F(x)$.*

Durch Iteration des Übergangs zu direkten Subformeln erhält man daraus den Begriff der Subformel:

3.3.3 Induktive Definition *der Subformeln einer Formel C.*

1. C ist Subformel von C.

2. Ist A Subformel von B und B direkte Subformel von C, so ist A Subformel von C.

3.3.4 Bemerkung A ist also genau dann Subformel von C, wenn es Formeln $B_0, \ldots, B_k$ gibt, so dass $B_0 \equiv A, B_k \equiv C$ und für alle $i < k$ B_i direkte Subformel von B_{i+1} ist.

3.3.5 Lemma *Subformel-Eigenschaft. In einer logischen Herleitung bestehen alle Sequenzen nur aus Subformeln von Formeln der Endsequenz und aus Gleichungen.*

Beweis durch Herleitungsinduktion. Sei H eine logische Herleitung.

1. H sei ein logisches Axiom. Dann ist H mit seiner Endsequenz identisch, und es ist nichts zu beweisen.

2. H ende mit einem logischen Grundschluss außer (= I). Inspektion der Grundschlussregeln zeigt unmittelbar, dass die Prämissen außer Formeln der Endsequenz nur direkte Subformeln einer Formel der Endsequenz enthalten. Dies gilt auch für (= F), weil wegen der Beschränkung auf positiv-stellige Funktionszeichen auch die Endsequenz noch wenigstens eine Gleichung enthält, von der nach 3.3.2 jede Gleichung direkte Subformel ist. Mit IV folgt die Behauptung des Lemmas.

3. H ende mit einem (= I)-Schluss. Dann stimmt die Prämisse mit der Endsequenz bis auf eine zusätzliche Gleichung $t = t$ überein, so dass die IV unmittelbar die Behauptung ergibt.

Induktion nach dem Aufbau von H ergibt nun das Lemma.

Die Subformeleigenschaft erlaubt eine wirksame Kontrolle über die möglichen Herleitungen einer gegebenen Sequenz. Sie liefert sogar ein Standardverfahren zum Auffinden einer Herleitung einer Sequenz, sofern diese Sequenz überhaupt eine Herleitung besitzt. Wir gehen darauf in Kapitel 3 ein. Schnittfreie Kalküle allgemein sind in der Beweistheorie von entscheidender Bedeutung. Wir werden die Schnittfreiheit unseres Kalküls in Teilen des Kapitels 5 ausnutzen.

3.4 Aufgaben

3.4.1 Bilden Sie logische Herleitungen von

a. $P \vee Q : P, Q$

b. $P \to \neg Q : Q \to \neg P$

c. $\forall x(px \rightarrow qx), \forall ypy : \forall zqz$

d. $: \exists x \quad t = x$

3.4.2 Zeigen Sie: Das kommutative Gesetz $a \circ b = b \circ a$ ist in der Gruppentheorie nicht herleitbar.

3.4.3 Beweisen Sie die Bemerkung 3.3.4.

3.4.4 Folgern Sie aus dem Korrektheitssatz: Es gibt Sequenzen $F(a) : \forall x F(x)$, die logisch nicht herleitbar sind. Geben Sie eine solche Sequenz an, wobei a in F nicht auftritt.

Klassische Prädikatenlogik

Kurseinheit 1: Lösungen zu den Übungsaufgaben

1.4.1 a. Wir induzieren nach der Länge der Nennform F.

1. F ist die leere Zeichenreihe $\emptyset$. Dann ändert die Substitution nichts, und man erhält $\emptyset \equiv \emptyset$.

2. F hat positive Länge. Wir unterscheiden nach dem letzten Zeichen von F.

2.1. F ist $F_0 z$ mit einem Grundzeichen z oder einem Nennzeichen $z \equiv *_i$ mit $i > 2$. Dann ist stets $z(G_1, G_2) \equiv z$, und es folgt

$$\begin{aligned} F(G_1, G_2)(H_1, H_2) &\equiv F_0(G_1, G_2)(H_1, H_2)z \\ &\equiv F_0(G_1(H_1, H_2), G_2(H_1, H_2))z \quad \text{nach IV} \\ &\equiv F(G_1(H_1, H_2), G_2(H_1, H_2)). \end{aligned}$$

2.2. F ist $F_0 *_i$ mit $i = 1$ oder $i = 2$. Dann ist $*_i(G_1, G_2) \equiv G_i$, also

$$\begin{aligned} F(G_1, G_2)(H_1, H_2) &\equiv F_0(G_1, G_2)(H_1, H_2)G_i(H_1, H_2) \\ &\equiv F_0(G_1(H_1, H_2), G_2(H_1, H_2))G_i(H_1, H_2) \quad \text{nach IV} \\ &\equiv F(G_1(H_1, H_2), G_2(H_1, H_2)). \end{aligned}$$

Mit Induktion folgt a.

b. Wir wählen $F \equiv *_1 = *_2$, $G \equiv a, H_1 \equiv H_2 \equiv b$. Dann ist

$$\begin{aligned} &F(G)(H_1, H_2) \equiv a = *_2(b, b) \equiv a = b \text{ und} \\ &F(G(H_1, H_2)) \equiv F(a) \equiv a = *_2. \end{aligned}$$

Bemerkung: Aufgabe a. zeigt, wie die Gleichung richtig lautet.
Wegen $F(G) \equiv F(G, *_2)$ und $*_2(H_1, H_2) \equiv H_2$ ergibt a.

$$\begin{aligned} F(G)(H_1, H_2) &\equiv F(G, *_2)(H_1, H_2) \\ &\equiv F(G(H_1, H_2), *_2(H_1, H_2)) \quad \text{nach a.} \\ &\equiv F(G(H_1, H_2), H_2) \end{aligned}$$

1.4.2 a. $\subseteq$: Wir zeigen durch Induktion nach dem Aufbau der Nf: Jede Nf von L ist eine Nennform von L.

1. $\emptyset$ ist die Nennform der Länge 0 von L

2. und 3. Ist F nach IV eine Nennform von L (etwa der Länge n), so sind Fz und $F*_i$ (für $i > 0$) Nennformen von L (der Länge $n+1$).

$\supseteq$: Wir zeigen durch Induktion nach der Länge der Nennformen: Jede Nennform von L ist eine Nf von L.

1. Hat F die Länge 0, so ist $F \equiv \emptyset$ eine Nf nach 1.

2. Hat F positive Länge, so ist $F \equiv F_0 z$, und F_0 hat kleinere Länge. Nach IV ist F_0 eine Nf von L. Ist z ein Grundzeichen von L, so ist F eine Nf von L nach 2., und ist z ein Nennzeichen, so ist F eine Nf von L nach 3.

Mit $\subseteq$ und $\supseteq$ ist a. bewiesen.

b. Rekursive Definition von $F(G_1, \ldots, G_n)$.

1. $\emptyset(G_1, \ldots, G_n) \equiv \emptyset$ die leere Nf.

2. $Fz(G_1, \ldots, G_n) \equiv F(G_1, \ldots, G_n)z$ für Grundzeichen z von L und für Nennzeichen $z \equiv *_i$ mit $i > n$.

3. $F *_i (G_1, \ldots, G_n) \equiv F(G_1, \ldots, G_n)G_i$ für $1 \leq i \leq n$.

1.4.3

1. $n = 0$. $\sum_{i<1} i = 0 = \frac{1}{2} \cdot 0 \cdot 1$

2. $n \rightsquigarrow n+1$. $\sum_{i<n+2} i = \sum_{i<n+1} i + (n+1)$
$= \frac{1}{2} \cdot n \cdot (n+1) + (n+1)$ nach IV
$= (\frac{1}{2} \cdot n + 1) \cdot (n+1) = \frac{1}{2} \cdot (n+1) \cdot (n+2)$.

Aus 1. und 2. folgt mit vollständiger Induktion die Behauptung.

1.4.4 a. Ja, es handelt sich um die Gleichung

$$(a + (a + b)) + b = (a + b) + (a + b).$$

b. $A \wedge B \to B \wedge A$ steht für $\neg(A \to \neg B) \to \neg(B \to \neg A)$, und das ist unter Verwendung von $\neg$:

$$\to \neg \to A \neg B \neg \to B \neg A,$$

also schließlich $\to\to\to A \to B \bot\bot \to\to B \to A \bot\bot$.

Die Auftreten von B sind zunächst

$$\to\to\to A \to *_1 \bot\bot \to\to B \to A \bot\bot \text{ und}$$
$$\to\to\to A \to B \bot\bot \to\to *_1 \to A \bot\bot.$$

Für den Fall $B \equiv \neg A \equiv\to A\bot$ gibt es das weitere Auftreten

$$\to\to\to A \to B \bot\bot \to\to B *_1 \bot.$$

c. $\neg a = 0 \to \neg\forall y \neg a \cdot y = 1$ ist unter Verwendung von $\neg$:

$$\to \neg = a0 \neg\forall y \neg = \cdot ay1$$

also schließlich $\to\to= a0\bot \to \forall y \to= \cdot ay1\bot\bot$.

Die Auftreten von $=$ sind

$$\to\to *_1 a0\bot \to \forall y \to= \cdot ay1\bot\bot \text{ und}$$
$$\to\to= a0\bot \to \forall y \to *_1 \cdot ay1\bot\bot.$$

2.4.1

1. $\mathcal{A}(\bot) = f$ nach 2.1.4, 2.3.
2. Definition 2.1.4, 2.4 ergibt wegen der Zweiwertigkeit

   $$\mathcal{A}(B \to C) = f \Leftrightarrow \mathcal{A}(B \to C) \neq w \Leftrightarrow \mathcal{A}(B) = w \text{ und } \mathcal{A}(C) = f.$$

   Also ist $\mathcal{A}(B \to C) = w$ in den anderen 3 Fällen.
3. Hiernach ist

   $$\mathcal{A}(\neg B) = \mathcal{A}(B \to \bot) = f \Leftrightarrow (\mathcal{A}(B) = w \text{ und } \mathcal{A}(\bot) = f) \Leftrightarrow \mathcal{A}(B) = w$$

   nach 2.1.4, 2.3. Wegen der Zweiwertigkeit ist dann umgekehrt

   $$\mathcal{A}(\neg B) = w \Leftrightarrow \mathcal{A}(B) = f.$$

4. $\mathcal{A}(B \vee C) = \mathcal{A}(\neg B \to C) = f \Leftrightarrow \mathcal{A}(\neg B) = w$ und $\mathcal{A}(C) = f$
   $\Leftrightarrow \mathcal{A}(B) = f$ und $\mathcal{A}(C) = f$.
   Also ist $\mathcal{A}(B \vee C) = w$ in den anderen 3 Fällen.

5. $\mathcal{A}(B \wedge C) = \mathcal{A}(\neg(B \to \neg C)) = w \Leftrightarrow \mathcal{A}(B \to \neg C) = f$ nach 3.
   $\Leftrightarrow \mathcal{A}(B) = w$ und $\mathcal{A}(\neg C) = f$ nach 2.
   $\Leftrightarrow \mathcal{A}(B) = w$ und $\mathcal{A}(C) = w$ nach 3.
   Also ist $\mathcal{A}(B \wedge C) = f$ in den anderen 3 Fällen.

6. $\mathcal{A}(B \leftrightarrow C) = \mathcal{A}((B \to C) \wedge (C \to B)) = w$
   $\Leftrightarrow \mathcal{A}(B \to C) = w$ und $\mathcal{A}(C \to B) = w$ nach 5.
   $\Leftrightarrow$ aus $\mathcal{A}(B) = w$ folgt $\mathcal{A}(C) = w$, und
   aus $\mathcal{A}(C) = w$ folgt $\mathcal{A}(B) = w$
   $\Leftrightarrow \mathcal{A}(B) = \mathcal{A}(C) = w$ oder $\mathcal{A}(B) = \mathcal{A}(C) = f$.
   Also ist $\mathcal{A}(B \leftrightarrow C) = f$ in den anderen 2 Fällen.

7. $\mathcal{A}(\top) = w \Leftrightarrow$ aus $\mathcal{A}(\bot) = w$ folgt $\mathcal{A}(\bot) = w$.
   Also ist $\mathcal{A}(\top) = w$.

2.4.2 a.

$\mathcal{A}(B)$	$\mathcal{A}(C)$	$\mathcal{A}(B \sqcup C)$
w	w	f
w	f	w
f	w	w
f	f	f

b. Es gibt mehrere Möglichkeiten.

$\mathcal{A}(B \sqcup C) = w \Leftrightarrow \mathcal{A}(B \to \neg C) = w$ und $\mathcal{A}(\neg B \to C) = w$, also

$$\mathcal{A}(B \sqcup C) = \mathcal{A}((B \to \neg C) \wedge (\neg B \to C)).$$

Die Wahrheitstafel aus a. ergibt unmittelbar

$$\mathcal{A}(B \sqcup C) = \mathcal{A}((B \wedge \neg C) \vee (\neg B \wedge C)).$$

Damit sind (in abgekürzter Schreibweise)

$$(*_1 \to \neg *_2) \wedge (\neg *_1 \to *_2) \text{ und } (*_1 \wedge \neg *_2) \vee (\neg *_1 \wedge *_2)$$

Nennformen F, die i. und ii. erfüllen.

c. i. liest man unmittelbar aus der Wahrheitstafel in a. ab.

ii. Wegen der Zweiwertigkeit und $\mathcal{A}(\neg B) \neq \mathcal{A}(B)$ ist

$$\mathcal{A}(B \sqcup C) = w \Leftrightarrow \mathcal{A}(B) \neq \mathcal{A}(C) \Leftrightarrow \mathcal{A}(\neg B) \neq \mathcal{A}(\neg C)$$
$$\Leftrightarrow \mathcal{A}(\neg B \sqcup \neg C) = w$$

und daraus folgt ii.

iii. folgt aus i., weil $\mathcal{A}(B) = \mathcal{A}(B)$ ist.

iv. ist falsch! Für $\mathcal{A}(B) = \mathcal{A}(C) = \mathcal{A}(D) = w$ ist $\mathcal{A}(C \sqcup D) = f$, also wieder $\mathcal{A}(B \sqcup (C \sqcup D)) = w$, im Widerspruch zur $\Rightarrow$-Richtung.

v. ist richtig. Denn für $\mathcal{A}(B) = w$ ist
$\mathcal{A}(B \sqcup (C \sqcup D)) = w \Leftrightarrow \mathcal{A}(C \sqcup D) = f \Leftrightarrow \mathcal{A}(C) = \mathcal{A}(D)$,
$\mathcal{A}(B \sqcup C) \neq \mathcal{A}(C)$, also
$\mathcal{A}((B \sqcup C) \sqcup D) = w \Leftrightarrow \mathcal{A}(B \sqcup C) \neq \mathcal{A}(D) \Leftrightarrow \mathcal{A}(C) = \mathcal{A}(D)$.
Ähnlich ist für $\mathcal{A}(B) = f$
$\mathcal{A}(B \sqcup (C \sqcup D)) = w \Leftrightarrow \mathcal{A}(C \sqcup D) = w \Leftrightarrow \mathcal{A}(C) \neq \mathcal{A}(D)$,
$\mathcal{A}(B \sqcup C) = \mathcal{A}(C)$, also
$\mathcal{A}((B \sqcup C) \sqcup D) = w \Leftrightarrow \mathcal{A}(C) = \mathcal{A}(B \sqcup C) \neq \mathcal{A}(D)$.
Damit ist in jedem Fall

$$\mathcal{A}(B \sqcup (C \sqcup D)) = w \Leftrightarrow \mathcal{A}((B \sqcup C) \sqcup D) = w,$$

und das ergibt wegen der Zweiwertigkeit die Behauptung.

2.4.3

1. Für $n = 0$ reduzieren sich beide Seiten der Äquivalenz auf $\mathcal{A}(C) = w$.

2. $n \rightsquigarrow n + 1$. $\mathcal{A}(B_1 \to \ldots \to B_{n+1} \to C) = w$
$\Leftrightarrow$ aus $\mathcal{A}(B_1) = w$ folgt $\mathcal{A}(B_2 \to \ldots \to B_{n+1} \to C) = w$
$\Leftrightarrow$ aus $\mathcal{A}(B_1) = w$ folgt: wenn $\mathcal{A}(B_i) = w$ ist für $i = 2, \ldots, n + 1$,
so ist $\mathcal{A}(C) = w$ nach IV
$\Leftrightarrow$ aus $\mathcal{A}(B_1) = w$ und $\mathcal{A}(B_i) = w$ für $i = 2, \ldots, n + 1$ folgt $\mathcal{A}(C) = w$
$\Leftrightarrow$ aus $\mathcal{A}(B_i) = w$ für $i = 1, \ldots, n + 1$ folgt $\mathcal{A}(C) = w$.

Induktion nach n ergibt die Behauptung.

2.4.4 **a.** Die rechte Seite ist gleichwertig mit:

$$\text{Es gibt 2 verschiedene } c, d \in |\mathcal{A}| \text{ mit } \mathcal{A}(F(c)) = \mathcal{A}(F(d)) = w$$

$\Leftrightarrow$ es gibt $c, d \in |\mathcal{A}|$ mit $c \neq d$ und $\mathcal{A}(F(c)) = \mathcal{A}(F(d)) = w$
$\Leftrightarrow$ $\mathcal{A}(\exists x \exists y(x \neq y \wedge F(x) \wedge F(y))) = w$.
Also ist $\exists x \exists y(x \neq y \wedge F(x) \wedge F(y))$ eine Formel $\exists_{\geq 2} x F(x)$.

b. Die rechte Seite ist gleichwertig mit:

$$\text{Von 3 Elementen } c \in |\mathcal{A}| \text{ mit } \mathcal{A}(F(c)) = w \text{ sind mindestens 2 gleich}$$

$\Leftrightarrow$ für $c, d, e \in |\mathcal{A}|$ mit $\mathcal{A}(F(c)) = \mathcal{A}(F(d)) = \mathcal{A}(F(e)) = w$
ist stets $c = d$ oder $c = e$ oder $d = e$
$\Leftrightarrow$ $\mathcal{A}(\forall x \forall y \forall z(F(x) \wedge F(y) \wedge F(z) \rightarrow x = y \vee x = z \vee y = z)) = w$.
Also ist $\forall x \forall y \forall z(F(x) \wedge F(y) \wedge F(z) \rightarrow x = y \vee x = z \vee y = z)$ eine Formel $\exists_{\leq 2} x F(x)$.

2.4.5 Setzt man $F :\equiv a = *_1$, so ist $F(a) \equiv a = a$.
Die Formel $\forall x F(x) \equiv \forall x a = x$ gilt nicht in allen Strukturen. Genauer:

$$\begin{aligned} \mathcal{A} \models \forall x a = x \quad &\Leftrightarrow \text{ für jedes } c \in |\mathcal{A}| \text{ ist } \mathcal{A}(\forall x c = x) = w \\ &\Leftrightarrow \text{ für jedes } c \in |\mathcal{A}| \text{ ist } c = d \text{ für alle } d \in |\mathcal{A}| \\ &\Leftrightarrow \text{ alle } c \in |\mathcal{A}| \text{ sind gleich} \\ &\Leftrightarrow |\mathcal{A}| \text{ enthält höchstens, also genau ein Element.} \end{aligned}$$

Die Formel $\forall x F(x)$ gilt also nur in den einelementigen Strukturen.

2.4.6 Sei $\mathcal{A}$ ein Modell von T und $'$ eine $\mathcal{A}$-Belegung. Dann ist

$$(1)\ \mathcal{A}(\Gamma' : B', \Delta') = w \text{ und } (2)\ \mathcal{A}(\Gamma', B' : \Delta') = w$$

nach den beiden Voraussetzungen. Sei nun $\mathcal{A}(C') = w$ für alle $C \in \Gamma$.

1. Ist $\mathcal{A}(B') = f$, so ist nach (1) $\mathcal{A}(D') = w$ für ein $D \in \Delta$,
   also $\mathcal{A}(\Gamma' : \Delta') = w$.

2. Ist $\mathcal{A}(B') = w$, so ist nach (2) $\mathcal{A}(D') = w$ für ein $D \in \Delta$,
   also $\mathcal{A}(\Gamma' : \Delta') = w$.

Damit ist stets $\mathcal{A}(\Gamma' : \Delta') = w$, $\Gamma : \Delta$ gilt in $\mathcal{A}$ und, weil $\mathcal{A}$ ein beliebiges Modell von T war, auch in T.

2.4.7 a. Sei G eine Gruppe und $'$ eine G-Belegung.
Wir setzen $i := G(a')$, $j := G(b'), k := G(c')$ und schreiben $\circ$ für $\circ_G$, $^{-1}$ für $^{-1}_G$. Aus der Voraussetzung $i \circ k = j \circ k$ folgt dann

$$\begin{aligned} &i = i \circ e_G = i \circ (k \circ k^{-1}) = (i \circ k) \circ k^{-1} = (j \circ k) \circ k^{-1} \\ &= j \circ (k \circ k^{-1}) = j \circ e_G = j, \end{aligned}$$

wobei wir nacheinander $G2$, $G3$, $G1$, die Voraussetzung, $G1$, $G3$, $G2$ verwandt haben. Weil dies für jede G-Belegung $'$ und jede Gruppe G so ist, gilt

$$a \circ c = b \circ c \to a = b$$

in G und schließlich in T_G.

b. Mit den Festlegungen wie unter a. folgt für ein beliebiges $l \in |G|$ aus $i \circ l = l$ stets

$$i = i \circ e_G = i \circ (l \circ l^{-1}) = (i \circ l) \circ l^{-1} = l \circ l^{-1} = e_G,$$

wobei wir nacheinander $G2$, $G3$, $G1$, die Voraussetzung und $G3$ verwandt haben. Wenn es also ein $l \in |G|$ mit $i \circ l = l$ gibt, ist $i = e_G$. Also folgt wie oben, dass

$$\exists x\, a \circ x = x \to a = e$$

in G und schließlich in T_G gilt.

3.4.1 a. Wir empfehlen, die Herleitungen jeweils von der Endsequenz aus aufzubauen.

$$\dfrac{\dfrac{P : \bot, P, Q}{: P \to \bot, P, Q} \qquad Q : P, Q}{\neg P \to Q : P, Q}$$

ist eine logische Herleitung von $P \vee Q : P, Q$. Denn die beiden oberen Sequenzen sind logische Axiome, der obere Schluss ist ein $(\to S)$-, der untere ein $(\to A)$-Schluss.

b.

$$\dfrac{\dfrac{\dfrac{}{P,Q:P,\bot}\qquad\dfrac{\dfrac{}{P,Q:Q,\bot}\qquad\dfrac{}{P,Q,\bot:\bot}}{P,Q,Q\to\bot:\bot}}{P,Q,P\to\neg Q:\bot}}{\dfrac{Q,P\to\neg Q:P\to\bot}{P\to\neg Q:Q\to\neg P}}$$

ist eine logische Herleitung. Denn an den Spitzen stehen drei logische Axiome, die oberen beiden Schlüsse sind zwei ($\to A$)-, die unteren beiden zwei ($\to S$)-Schlüsse.

c.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{pa:pa,qa}\qquad\dfrac{}{qa,pa:qa}}{pa\to qa,pa:qa}}{\forall x(px\to qx),pa:qa}}{\dfrac{\forall x(px\to qx),\forall ypy:qa}{\forall x(px\to qx),\forall ypy:\forall zqz}}$$

ist eine logische Herleitung. Denn oben stehen zwei logische Axiome, auf die ein ($\to A$)-Schluss angewandt wird, und es folgen zwei ($\forall A$)-Schlüsse und schließlich ein ($\forall S$)-Schluss, bei dem a die Variablenbedingung erfüllt.

d.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{t=t:t=t,\bot}}{:t=t,\bot}\qquad\dfrac{}{\bot:\bot}}{t=t\to\bot:\bot}}{\dfrac{\forall x\neg t=x:\bot}{:\forall x\neg t=x\to\bot}}$$

ist eine logische Herleitung von $:\exists x\ t=x$. Denn oben stehen zwei logische Axiome, es folgen ein ($=I$)-, ein ($\to A$)-, ein ($\forall A$)- und ein ($\to S$)-Schluss.

3.4.2 Wäre $T_G \vdash a\circ b=b\circ a$, so wäre nach dem Korrektheitssatz $T_G \models a\circ b=b\circ a$. Das ist aber nicht der Fall, weil es nicht-kommutative Gruppen gibt (vgl. 2.2.4). Also ist $T_G \not\vdash a\circ b=b\circ a$.

3.4.3 Eine Folge $B_0,\dots,B_k$ von Formeln, so dass für $i<k$ B_i direkte Subformel von B_{i+1} ist, nennen wir eine Subformelkette von B_0 nach B_k. Wir zeigen zunächst: Ist A Subformel von C nach 3.3.3, so gibt es eine Subformelkette von A nach C. Wir induzieren nach der induktiven Definition der Subformeln.

1. $A \equiv C$. Dann ist $k = 0$, wir setzen $A \equiv B_0 \equiv C$, und B_0 ist Subformelkette von A nach C.

2. A ist Subformel von B und B ist direkte Subformel von C. Dann gibt es nach IV eine Subformelkette von $A \equiv B_0$ nach $B \equiv B_k$. Setzen wir $B_{k+1} \equiv C$, so ist $B_0, \ldots, B_{k+1}$ eine Subformelkette von A nach C.

Induktion ergibt nun obige Behauptung.

Wir zeigen umgekehrt: Gibt es eine Subformelkette von A nach C, so ist A Subformel von C. Wir induzieren nach der Länge k der Subformelkette.

1. Ist $k = 0$, so ist $A \equiv B_0 \equiv C$, und A ist Subformel von C nach 3.3.3, 1.

2. $k \rightsquigarrow k + 1$. Dann ist A nach IV Subformel von B_k und B_k direkte Subformel von $B_{k+1} \equiv C$. Also ist A Subformel von C nach 3.3.3, 2.

Vollständige Induktion nach k ergibt nun die Umkehrung. Damit ist die Äquivalenz vollständig bewiesen.

3.4.4 Seien a, b verschiedene freie Variablen. Wir setzen $F :\equiv *_1 = b$, also $F(a) \equiv a = b$. Dann tritt a in F nicht auf. Ist $\mathcal{A}$ eine Struktur mit mindestens zwei Elementen, so gibt es eine $\mathcal{A}$-Belegung $'$ mit $a' = b'$, also $\mathcal{A}(a' = b') = w$, aber $\mathcal{A}(\forall x F'(x)) = \mathcal{A}(\forall x\ x = b') = f$, weil es eben $k \in |\mathcal{A}|$ mit $k \neq \mathcal{A}(b')$ gibt. Dann ist $\mathcal{A}((a = b : \forall x\ x = b)') = f$, und $a = b : \forall x\ x = b$ ist logisch nicht gültig. Dann kann diese Sequenz nach dem Korrektheitssatz auch nicht logisch herleitbar sein.

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 2:
Syntaktische Sätze und Regeln der Prädikatenlogik

Autor: Justus Diller

01301-7-02-S 1

Klassische Prädikatenlogik

Kurseinheit 2: Inhalt

Studienhinweise 85

Verzeichnis der definierten Begriffe und der wichtigen Sätze 86

1. Sprache, Semantik und Syntax der Prädikatenlogik

2. Syntaktische Sätze und Regeln der Prädikatenlogik

§4 Aussagenlogik 89

4.1 Der Satz von der Identität 90
4.2 Tautologien 92
4.3 Wahrheitsfunktionen 98
4.4 Disjunktive und konjunktive Normalform 101
4.5 Aufgaben 106

§5 Zulässige Regeln der Prädikatenlogik 108

5.1 Schwache Schlussregeln 108
5.2 Junktoren-Regeln 115
5.3 Quantoren-Regeln 118
5.4 Das Deduktionstheorem 121
5.5 Aufgaben 125

§6 Gleichheit und Äquivalenz 127

6.1 Gleichheit 127
6.2 Logische Äquivalenz 132
6.3 Exkurs: Eine Hilbert-artige Formalisierung der Prädikatenlogik 136
6.4 Aufgaben 143

3. Vollständigkeit

4. Modelltheorie

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 2: Studienhinweise

1. Lehrziele

In dieser Kurseinheit wird der Herleitungsbegriff aus §3 eingehend studiert. Einerseits werden viele Gesetze der Logik syntaktisch hergeleitet, andererseits wird der Herleitungsbegriff auch global untersucht, und besonders in Abschnitt 5.1 werden Abgeschlossenheitseigenschaften des Herleitbarkeitsbegriffs bewiesen. Einerseits sollen Sie demnach das Herleiten üben, Sicherheit im Konstruieren konkreter Herleitungen erwerben und herleitbare Formeln und Sequenzen, speziell auch in §4 Tautologien als solche erkennen lernen. Andererseits sollen Ihnen Herleitungs– und Herleitbarkeitsbegriff als mathematische Begriffe vertraut werden, so dass Sie auch über diese Begriffe global zu reflektieren lernen. Diese Arbeitsweise wird hier in Abschnitt 5.1 geübt. Sie wird in Kapitel 5 weiter vertieft werden.
Zwar bauen wir den Begriffsapparat aus Kapitel 1, §§1 und 3, kräftig aus, aber es wird hier kein grundlegend neuer Apparat aufgebaut. Das Schwergewicht verschiebt sich gegenüber Kapitel 1 deutlich in Richtung Ergebnisse und Beweise. Allerdings sind die Beweise recht leicht, die Ergebnisse naheliegend. Während im Kapitel 1 die Beispiele wesentliche Begriffe wie Theorie und Modell illustrieren, behandeln sie hier meistens typische Spezialfälle von allgemeinen Ergebnissen. Die Abschnitte 5.2 und 5.3 kann man insgesamt als Sammlung von Anwendungsbeispielen der Ergebnisse aus 5.1 lesen.
Wie die Gliederung der Kurseinheit begründet ist, erläutern wir in der Einleitung zum Kapitel 2.

2. Eingangsvoraussetzungen

In jedem Abschnitt der Kurseinheit wird auf den *Herleitungsbegriff* aus §3 zurück gegriffen. Das Prinzip der Induktion nach einem induktiv definierten Begriff sollte verstanden sein. Induktionen über $\mathbb{N}$ (das sind die vollständigen Induktionen), Induktionen nach dem Aufbau von Formeln, dem Aufbau oder der Ordnung von Herleitungen werden ständig verwendet und bilden das Rückgrat der Beweisführung in diesem Kapitel.

Klassische Prädikatenlogik

Kurseinheit 2: Verzeichnis der definierten Begriffe und der wichtigen Sätze

4.1.2 **Satz von der Identität** $\vdash \Gamma, C : C, \Delta$

4.1.4 **Tertium non datur**

4.2.1 Aussagenelemente

4.2.2 Wertungen $V : AE \to \{w, f\}$

4.2.3 Wert der Formeln (und Sequenzen) von L unter V

4.2.5 Tautologien, aussagenlogisch gültig

4.2.7 Entscheidungsverfahren für Tautologien

4.2.9 Einfache Sequenzen, aussagenlogische Axiome

4.2.15 **Tautologiesatz**

4.2.17 2. Entscheidungsverfahren für Tautologien

4.3.1 n-stellige Wahrheitsfunktionen

4.3.2 Darstellung von Wahrheitsfunktionen

4.3.4 Vollständige Junktorensysteme

4.3.5 **Satz** $\{\top, \bot, \neg, \to, \wedge\}$ ist vollständig

4.4.2 Iterierte Disjunktionen und Konjunktionen

4.4.5 Disjunktive und konjunktive Normalformen

4.4.6 **Satz** Jede Formel besitzt eine disjunktive Normalform

4.4.7 **Satz** Jede Formel besitzt eine konjunktive Normalform

5.1.1 Zulässige Schlussregeln

5.1.2 Ordnung einer Herleitung

5.1.3 $T \overset{n}{\vdash} \Gamma : \Delta$

5.1.4 Schwache Schlussregeln

5.1.5 **Substitutionssatz** $(Subst)$

5.1.6 Strukturelle Folgerung

5.1.7 **Satz** Strukturschlussregel (Str)

5.1.8 – 10 **Satz** Inversionsregeln ($\rightarrow SInv$), ($\rightarrow AInv$) ($\forall SInv$)

5.3.1 **Lemma** Existenz-Regeln ($\exists A$), ($\exists S$), ($\exists AInv$)

5.3.5 **Lemma** Regeln für den Allabschluss

5.4.1 Theorie $T + \Sigma$

5.4.2 **Deduktionstheorem**

5.4.3 **Endlichkeitslemma**

5.4.4 **Verallgemeinertes Deduktionstheorem**

6.1.4 **Gleichheitsregel**

6.1.6 **Gleichheitssatz**

6.2.3 **Äquivalenzsatz**

6.3.1 Axiome und Regeln des Hilbert-Kalküls

6.3.2 Herleitungen im Hilbert-Kalkül, $T \vdash_H C$

6.3.8 **Satz** Äquivalenz von Hilbert- und Sequenzenkalkül

Kapitel 2

Syntaktische Sätze und Regeln der Prädikatenlogik

Wir studieren in diesem Kapitel einfache Eigenschaften des in § 3 eingeführten Herleitungsbegriffs. Semantische Überlegungen spielen dabei nur in § 4 eine Rolle, und auch dort nur in eingeschränkter Form. Nach dem Korrektheitssatz sind in jeder mathematischen Theorie nur gültige Sequenzen und Formeln herleitbar. Wir untersuchen, welche üblichen, grundlegenden Sequenzen bzw. Sequenzenschemata tatsächlich herleitbar und damit „Gesetze der Logik“ sind, und wie dabei der Herleitungsbegriff eingesetzt wird. Die Wirkungsweise einzelner Grundschlussregeln wird schwerpunktmäßig in eigenen Paragraphen untersucht: Die logischen Axiome und die Implikationsregeln $(\to S)$ und $(\to A)$ stehen in § 4 im Vordergrund, die Quantorenregeln $(\forall S)$ und $(\forall A)$ in § 5, zusammen mit der Theorie-Schlussregel in 5.4, und die Gleichheitsregeln schließlich in § 6.

§4 Aussagenlogik

4.1 Der Satz von der Identität

4.2 Tautologien

4.3 Wahrheitsfunktionen

4.4 Disjunktive und konjunktive Normalform

4.5 Aufgaben

4.1 Der Satz von der Identität

Wir betrachten die Aussagenlogik als den Teil der Prädikatenlogik, in dem es auf die Eigenschaften der Terme und der Grundzeichen $=$ und $\forall$ nicht ankommt. In der Aussagenlogik werden also Primformeln (außer dem Falsum $\bot$) und Allformeln behandelt wie Aussagezeichen. Formeln und Sequenzen, die auch bei dieser groben Sichtweise, „aus aussagenlogischen Gründen“, gültig sind, werden wir *Tautologien* nennen. Z. B. sind Sequenzen $C : C$ und Formeln $C \to C$ und $A \wedge B \to A$ aus aussagenlogischen Gründen gültig und damit Tautologien. Sequenzen $: a = a$ und $\forall x F(x) : F(t)$ gelten zwar auch in jeder Struktur, aber nicht aussagenlogisch, sondern wegen bekannter Eigenschaften von $=$ und $\forall$, und somit sind sie keine Tautologien.
Wir haben unsere logischen Axiome so sparsam gewählt, dass nicht einmal Sequenzen $C : C$ stets Axiome sind. Wir beginnen mit der Herleitung von Sequenzen, die „Wenn C, dann C“ verallgemeinern und oft als *Satz von der Identität* behandelt werden.

4.1.1 Schreibweise Wir schreiben im folgenden $\vdash \Gamma : \Delta$, wenn $\Gamma : \Delta$ herleitbar ist in jeder logischen Theorie T und damit in jeder Theorie T mit $\Gamma : \Delta$ aus $L(T)$.

4.1.2 Satz von der Identität

$$\vdash \Gamma, C : C, \Delta$$

Beweis durch Induktion nach dem Aufbau der Formel C.

1. C ist eine Primformel. Dann ist $\Gamma, C : C, \Delta$ ein logisches Axiom, also herleitbar.

2. C ist $A \to B$. Dann ist nach Induktionsvoraussetzung

   $$\vdash \Gamma, A : A, B, \Delta \text{ und } \vdash \Gamma, A, B : B, \Delta.$$

   Mit $(\to A)$ folgt $\vdash \Gamma, A, A \to B : B, \Delta$, woraus sich mit $(\to S)$
   $\vdash \Gamma, A \to B : A \to B, \Delta$ ergibt.

3. C ist $\forall x \mathcal{F}(x)$. Sei a eine freie Variable, die nicht in $\Gamma, \Delta, \mathcal{F}$ auftritt. Dann ist nach Induktionsvoraussetzung

   $$\vdash \Gamma, \mathcal{F}(a) : \mathcal{F}(a), \Delta.$$

Mit $(\forall A)$ folgt $\vdash \Gamma, \forall x\, \mathcal{F}(x) : \mathcal{F}(a), \Delta$, woraus sich mit $(\forall S)$ $\vdash \Gamma, \forall x\, \mathcal{F}(x) : \forall x\, \mathcal{F}(x), \Delta$ ergibt, weil die Variablenbedingung nach Wahl von a erfüllt ist.

Mit Induktion nach C folgt die Behauptung.

Den Sequenzen $C : C$ entsprechen inhaltlich die Formeln $C \to C$; deren Herleitbarkeit ergibt sich offenbar mit einem $(\to S)$-Schluss aus 4.1.2:

4.1.3 Korollar $\vdash \Gamma : C \to C, \Delta$.

Als *Tertium non datur (Satz vom ausgeschlossenen Dritten)* bezeichnet man das Formelschema $C \vee \neg C$, dem die Sequenzen $: C, \neg C$ inhaltlich entsprechen. Als *Satz vom Widerspruch* bezeichnet man das Schema $\neg(\neg C \wedge C)$, dem die Sequenzen $C, \neg C :$ entsprechen. Die Herleitbarkeit dieser Formeln und Sequenzen ergibt sich unmittelbar aus 4.1.2 und 4.1.3.

4.1.4 Korollar (Tertium non datur)

$$\vdash \Gamma : C, \neg C, \Delta \quad \text{und} \quad \vdash \Gamma : C \vee \neg C, \Delta.$$

Beweis: Nach 4.1.2 ist $\vdash \Gamma, C : C, \bot, \Delta$. Daraus folgt mit einem $(\to S)$-Schluss die erste Behauptung. Nach 4.1.3 ist $\vdash \Gamma : \neg C \to \neg C, \Delta$, und das ist die zweite Behauptung.

4.1.5 Korollar (Satz vom Widerspruch)

$$\vdash \Gamma, C, \neg C : \Delta \quad \text{und} \quad \vdash \Gamma : \neg(\neg C \wedge C), \Delta.$$

Beweis: Nach 4.1.2 ist $\vdash \Gamma, C : C, \Delta$, und es ist $\vdash \Gamma, C, \bot : \Delta$ als logisches Axiom. Mit einem $(\to A)$-Schluss folgt die erste Behauptung.
Nach 4.1.3 ist $\vdash \Gamma : \neg C \to \neg C, \bot, \Delta$, und es ist $\vdash \Gamma, \bot : \bot, \Delta$ als logisches Axiom. Mit $(\to A)$ folgt

$$\vdash \Gamma, \neg(\neg C \to \neg C) : \bot, \Delta,$$

und da $\neg(\neg C \to \neg C) \equiv \neg C \wedge C$ ist, folgt mit $(\to S)$ die zweite Behauptung.

4.2 Tautologien

Tautologien sind „aus aussagenlogischen“ Gründen gültig. Um diese vage Vorstellung zu fixieren, wollen wir aus dem prädikatenlogischen Begriff der Interpretation den aussagenlogischen Anteil isolieren. Das führt uns zu dem Begriff der aussagenlogischen *Wertung*, der eine Vergröberung des Begriffs der Interpretation darstellt. Die aussagenlogischen Ergebnisse, wie wir sie hier darstellen, sind einfacher, aber auch schwächer als die prädikatenlogischen. In kleinem Maßstab geben sie einen Eindruck von den Problemen der Prädikatenlogik.

4.2.1 Definition Alle Formeln, die keine Implikationen sind, also Primformeln und Allformeln, nennen wir *Aussagenelemente*.

4.2.2 Definition Sei AE die Menge der von $\perp$ verschiedenen Aussagenelemente einer Sprache L. Jede Abbildung

$$V : AE \longrightarrow \{w, f\}$$

heißt *(aussagenlogische) Wertung* der Sprache L.

Diese Abbildungen werden analog zu den Interpretationen auf alle Formeln und Sequenzen fortgesetzt.

4.2.3 Rekursive Definition des *Wertes* $V(C)$ für jede Formel C von L bei der Wertung V.

1. Ist $C \in AE$, so ist $V(C)$ nach 4.2.2 schon gegeben.
2. $V(\perp) = f$.
3. $V(A \to B) = w \Leftrightarrow$ Aus $V(A) = w$ folgt $V(B) = w$.

4.2.4 Definition des *Wertes* $V(\Gamma : \Delta)$ für Sequenzen $\Gamma : \Delta$ von L bei der Wertung V.

$$V(\Gamma : \Delta) = w \Longleftrightarrow \begin{array}{l} \text{Ist } V(C) = w \text{ für alle } C \in \Gamma, \text{ so ist} \\ V(D) = w \text{ für (mindestens) ein } D \in \Delta. \end{array}$$

4.2.5 Definition Eine Formel C bzw. eine Sequenz $\Gamma : \Delta$ von L ist eine *Tautologie*, sie ist *aussagenlogisch gültig*, wenn

$$V(C) = w \text{ bzw. } V(\Gamma : \Delta) = w$$

ist für jede Wertung V von L.

Wenn man die Aussagenelemente $\not\equiv \bot$ als Aussagezeichen auffasst, wird eine Wertung zu einer Interpretation aussagenlogischer Formeln. Wertungen beachten nicht die innere Struktur der Aussagenelemente, im Gegensatz zu den Interpretationen. Nur wenn zwei Aussagenelemente als Zeichenreihen übereinstimmen, erhalten sie bei jeder Wertung denselben Wahrheitswert.

Beispiele: Es gibt Wertungen V mit $V(a = a) = f$, $V(\forall x(px \to px)) = f$ und $V(\forall x \bot) = w$. Dagegen ist stets $V(\top) = V(\bot \to \bot) = w$, allgemeiner $V(A \to A) = w$. Danach sind Formeln $C \to C$ Tautologien.
$V(A \leftrightarrow B) = w$ ist äquivalent mit $V(A) = V(B)$, und es ist stets $V(C) \neq V(\neg C) \neq V(\neg\neg C)$. Weil es nur zwei Wahrheitswerte gibt, ist daher stets $V(C) = V(\neg\neg C)$, so dass auch Formeln $C \leftrightarrow \neg\neg C$ Tautologien sind.

Der Wert einer Formel bzw. einer Sequenz unter einer Wertung V hängt offenbar nur von den Werten der (endlich vielen) Aussagenelemente ab, die in ihr tatsächlich auftreten.

4.2.6 Lemma Ist $V(A) = V'(A)$ für alle $A \in AE$, die in einer Formel C bzw. einer Sequenz $\Gamma : \Delta$ auftreten, so ist

$$V(C) = V'(C) \text{ bzw. } V(\Gamma : \Delta) = V'(\Gamma : \Delta).$$

Der **Beweis** verwendet für C eine triviale Induktion entlang der Definition 4.2.3.

1. Für Aussagenelemente C (auch $C \equiv \bot$) ist die Behauptung nach 4.2.3 trivial.

2. Ist $C \equiv C_1 \to C_2$, so ist $V(C_i) = V'(C_i)$ nach Induktionsvoraussetzung, weil genau die Aussagenelemente, die in C auftreten, in C_1 oder in C_2 auftreten, und mit 4.2.3, 3 folgt $V(C) = V'(C)$,

Mit Induktion folgt nun das Lemma für Formeln C. Unter Rückgriff auf 4.2.4 folgt es dann auch für Sequenzen $\Gamma : \Delta$.

4.2.7 Bemerkung So selbstverständlich dieses Ergebnis auch ist, so beinhaltet es doch ein *Entscheidungsverfahren* für Tautologien. Ist eine Formel C aus n verschiedenen Aussagenelementen $A_1, \ldots, A_n$ und eventuell $\bot$ mit $\to$ aufgebaut, so ist für den Wert $V(C)$ nach 4.2.6 nur das Wahrheitswert-n-tupel $(V(A_1), \ldots, V(A_n))$ relevant, und davon gibt es nur 2^n verschiedene. Offenbar ist die Berechnung von $V(C)$ aus dem n-tupel $(V(A_1), \ldots, V(A_n))$ ein endlicher Prozess, und C ist genau dann eine Tautologie, wenn man $V(C) = w$ berechnet für jedes der 2^n „relevanten“ Wahrheitswert-n-tupel.

Beispiel. $C \equiv (A \to B) \vee (B \to A) \equiv \neg(A \to B) \to (B \to A)$ ist eine Tautologie wegen

$V(A)$	$V(B)$	$V(A \to B)$	$V(\neg(A \to B))$	$V(B \to A)$	$V(C)$
w	w	w	f	w	w
w	f	f	w	w	w
f	w	w	f	f	w
f	f	w	f	w	w

$C \equiv (A \to B) \to (B \to A)$ ist dagegen keine Tautologie, falls A, B verschiedene Aussagenelemente $\not\equiv \bot$ sind. Denn für $V(A) = f$ und $V(B) = w$ ist $V(A \to B) = w$ und $V(B \to A) = f$, also $V(C) = f$.

4.2.8 Lemma Alle Sequenzen $\Gamma, C : C, \Delta$ und $\Gamma, \bot : \Delta$ sind Tautologien.

Beweis. Sei V eine Wertung.

1. Ist $V(A) = w$ für alle $A \in \Gamma \cup \{C\}$, so ist insbesondere $V(C) = w$, und damit ist $V(D) = w$ für ein $D \in \{C\} \cup \Delta$. Also ist $V(\Gamma, C : C, \Delta) = w$.

2. Wegen $V(\bot) = f$ ist in keinem Fall $V(A) = w$ für alle $A \in \Gamma \cup \{\bot\}$. Also ist $V(\Gamma, \bot : \Delta) = w$.

Da 1. und 2. für jede Wertung V gelten, sind diese Sequenzen Tautologien.

4.2.9 Definition Eine Sequenz $\Gamma : \Delta$ heißt *einfach*, wenn alle Formeln aus $\Gamma \cup \Delta$ Aussagenelemente sind. Einfache Sequenzen der Gestalten $\Gamma, A : A, \Delta$ und $\Gamma, \bot : \Delta$ nennen wir *aussagenlogische Axiome*.

Unter den einfachen Sequenzen lassen sich die Tautologien sehr leicht charakterisieren.

4.2.10 Lemma Eine einfache Sequenz ist genau dann eine Tautologie, wenn sie eine Gestalt

$$\Gamma, A : A, \Delta \text{ oder } \Gamma, \bot : \Delta$$

hat, wenn sie also ein aussagenlogisches Axiom ist.

Beweis: Die Richtung $\Leftarrow$ ist ein Spezialfall von Lemma 4.2.8.
Zur Richtung $\Rightarrow$. Sei $\Gamma : \Delta$ eine Tautologie. Wir definieren eine Wertung V durch

$$V(A) = w \text{ für alle } A \in \Gamma \cap AE$$
$$V(A) = f \text{ für alle } A \in AE - \Gamma.$$

1. Fall. Es gibt ein $A \in \Delta$ mit $V(A) = w$. Weil $\Gamma : \Delta$ einfach (und $V(\bot) = f$) ist, ist dann dieses $A \in AE$. Nach der Wahl von V ist dann $A \in \Gamma \cap \Delta$, und $\Gamma : \Delta$ hat eine Gestalt $\Gamma, A : A, \Delta$.
2. Fall. Für alle $A \in \Delta$ ist $V(A) = f$. Weil $\Gamma : \Delta$ eine Tautologie ist, gibt es dann ein $C \in \Gamma$ mit $V(C) = f$. Weil $\Gamma : \Delta$ einfach ist, ist dieses C ein Aussagenelement, aber nach Wahl von V ist dieses $C \notin AE$. Dann ist notwendig $C \equiv \bot$, und $\Gamma : \Delta$ hat eine Gestalt $\Gamma, \bot : \Delta$.

Man beachte, dass aussagenlogische Axiome der Gestalt $\Gamma, A : A, \Delta$ im allgemeinen keine logischen Axiome sind, wenn A keine Primformel ist. Nach dem Satz von der Identität sind sie aber jedenfalls herleitbar, und wie alle Tautologien sind sie allgemeingültig.
Wir betrachten jetzt die nicht-einfachen Tautologien.

4.2.11 Lemma Für Wertungen V gilt stets:

$$V(\Gamma, A : B, \Delta) = w \Leftrightarrow V(\Gamma : A \to B, \Delta) = w.$$

Beweis. Sei $V(C) = w$ für alle $C \in \Gamma$ und $V(D) = f$ für alle $D \in \Delta$. Dann besagt die linke Seite: Ist $V(A) = w$, so ist $V(B) = w$. Das ist aber dasselbe wie $V(A \to B) = w$, und das besagt gerade die rechte Seite.

4.2.12 Korollar $\Gamma, A : B, \Delta$ ist eine Tautologie genau dann, wenn $\Gamma : A \to B, \Delta$ eine Tautologie ist.

Dies folgt unmittelbar aus 4.2.11.

4.2.13 Lemma Für Wertungen V gilt stets:

$$V(\Gamma : A, \Delta) = V(\Gamma, B : \Delta) = w \Leftrightarrow V(\Gamma, A \to B : \Delta) = w.$$

Beweis. Sei $V(C) = w$ für alle $C \in \Gamma$ und $V(D) = f$ für alle $D \in \Delta$. Dann besagt die linke Seite: $V(A) = w$ und $V(B) = f$. Das ist aber äquivalent zu $V(A \to B) = f$, und das besagt gerade die rechte Seite.

4.2.14 Korollar $\Gamma : A, \Delta$ und $\Gamma, B : \Delta$ sind Tautologien genau dann, wenn $\Gamma, A \to B : \Delta$ eine Tautologie ist.

Dies folgt unmittelbar aus 4.2.13.

Nach 4.2.12 und 4.2.14 führen $(\to S)$- und $(\to A)$-Schlüsse von Tautologien wieder zu Tautologien, und umgekehrt: Ist die Konklusion eines solchen Schlusses eine Tautologie, so sind auch seine Prämissen Tautologien. Damit können wir die Tautologien durch ihre Herleitungen charakterisieren.

4.2.15 Tautologiesatz Eine Sequenz ist genau dann eine Tautologie, wenn sie aus aussagenlogischen Axiomen allein mit $(\to S)$- und $(\to A)$-Schlüssen herleitbar ist.

Beweis. Gegeben sei eine Tautologie $\Gamma : \Delta$. Wir induzieren nach der Länge, also nach der Anzahl der Auftreten von Grundzeichen in $\Gamma : \Delta$ und zeigen, dass $\Gamma : \Delta$ eine Herleitung der geforderten Gestalt besitzt.

1. Ist $\Gamma : \Delta$ eine einfache Tautologie, so ist sie nach Lemma 4.2.10 ein aussagenlogisches Axiom, ist also trivial aus solchen Sequenzen herleitbar.

2. Sei $\Gamma : \Delta$ eine Tautologie einer Gestalt $\Gamma : A \to B, \Delta_1$, wobei $A \to B \notin \Delta_1$ ist. Nach 4.2.12 ist dann auch $\Gamma, A : B, \Delta_1$ eine Tautologie, die kürzer ist, weil in ihr ein $\to$ weniger auftritt. Nach Induktionsvoraussetzung ist daher $\Gamma, A : B, \Delta_1$ aus aussagenlogischen Axiomen allein mit $(\to S)$- und $(\to A)$-Schlüssen herleitbar. Dann ist auch $\Gamma : \Delta$ so herleitbar, und zwar mit einem $(\to S)$-Schluss mehr.

3. Sei $\Gamma : \Delta$ eine Tautologie einer Gestalt $\Gamma_1, A \to B : \Delta$, wobei $A \to B \notin \Gamma_1$ ist. Nach 4.2.14 sind dann $\Gamma_1 : A, \Delta$ und $\Gamma_1, B : \Delta$ kürzere Tautologien. Nach Induktionsvoraussetzung sind sie daher aus aussagenlogischen Axiomen allein mit $(\to S)$- und $(\to A)$-Schlüssen herleitbar. Dann ist auch $\Gamma : \Delta$ so herleitbar, und zwar mit einem zusätzlichen $(\to A)$-Schluss.

Mit Induktion folgt nun die Richtung von links nach rechts.

Gegeben sei umgekehrt eine Herleitung einer Sequenz $\Gamma : \Delta$ aus aussagenlogischen Axiomen allein mit $(\to S)$- und $(\to A)$-Schlüssen. Wir induzieren nach der Länge, also nach der Anzahl der $(\to S)$- und $(\to A)$-Schlüsse in dieser Herleitung und zeigen, dass $\Gamma : \Delta$ eine Tautologie ist.

1. $\Gamma : \Delta$ ist ein aussagenlogisches Axiom. Dann ist $\Gamma : \Delta$ eine Tautologie nach 4.2.10.

2. $\Gamma : \Delta$ hat eine Gestalt $\Gamma : A \to B, \Delta_1$ und geht mit einem $(\to S)$-Schluss aus $\Gamma, A : B, \Delta_1$ hervor. Nach Induktionsvoraussetzung ist dann $\Gamma, A : B, \Delta_1$ eine Tautologie, so dass nach 4.2.12 auch $\Gamma : \Delta$ eine Tautologie ist.

3. $\Gamma : \Delta$ hat eine Gestalt $\Gamma_1, A \to B : \Delta$ und geht mit einem $(\to A)$-Schluss aus $\Gamma_1 : A, \Delta$ und $\Gamma_1, B : \Delta$ hervor. Nach Induktionsvoraussetzung sind dies dann Tautologien, so dass nach 4.2.14 auch $\Gamma : \Delta$ eine Tautologie ist.

Mit Induktion folgt die Richtung von rechts nach links. Damit ist der Satz bewiesen.

4.2.16 Korollar Tautologien sind (in logischen Theorien) herleitbar.

Dies folgt unmittelbar aus 4.1.2 und 4.2.15.

Mit dem Tautologiesatz sind die Tautologien syntaktisch charakterisiert. Die Richtung $\Leftarrow$ ist ein aussagenlogischer Korrektheitssatz, die Richtung $\Rightarrow$ ein aussagenlogischer Vollständigkeitssatz. Im Fall der Aussagenlogik laufen die Induktionen über die Länge der Sequenzen für die Richtung $\Rightarrow$ und über die Länge der Herleitungen für die Richtung $\Leftarrow$ Schritt für Schritt parallel: Den Induktionsanfang liefert jeweils 4.2.10, und die Induktionsschritte liefern in beiden Fällen 4.2.12 und 4.2.14. Dies gelingt, weil $(\to S)$- und $(\to A)$-Schlüsse in *dieser* Situation von kürzeren zu längeren Sequenzen führen.
Die Herleitbarkeit *aller* logisch gültigen Sequenzen, also die Vollständigkeit der Prädikatenlogik, kann nicht so direkt und einfach bewiesen werden.

4.2.17 Bemerkung Neben dem semantischen Entscheidungsverfahren 4.2.7 liefert der Tautologiesatz ein zweites, syntaktisches Entscheidungsverfahren

für Tautologien: Eine gegebene Sequenz $\Gamma : \Delta$ führt man, über $(\to S)$- und $(\to A)$-Schlüsse rückwärts gehend, auf einfache Sequenzen zurück. $\Gamma : \Delta$ ist nach 4.2.12 und 4.2.14 genau dann eine Tautologie, wenn alle diese einfachen Sequenzen Tautologien sind, und das ist nach 4.2.10 leicht zu entscheiden. Da man hierbei oft deutlich weniger als 2^n einfache Sequenzen erhält, ist dieses Verfahren oft vorzuziehen.

4.2.18 Beispiel Die PEIRCEsche Formel

$$F :\equiv ((A \to B) \to A) \to A$$

ist eine Tautologie, wie man nach 4.2.7 wie folgt sieht:

$V(A)$	$V(B)$	$V(A \to B)$	$V((A \to B) \to A)$	$V(F)$
w	w	w	w	w
w	f	f	w	w
f	w	w	f	w
f	f	w	f	w

F besitzt aber auch die Herleitung gemäß 4.2.15

$$\dfrac{\dfrac{\dfrac{A : B, A}{: A \to B, A} \qquad A : A}{(A \to B) \to A : A}}{F}$$

und $A : B, A$ und $A : A$ sind Tautologien nach 4.2.8. Um eine Sequenz als Tautologie nachzuweisen, braucht man also nicht immer über $(\to S)$- und $(\to A)$-Schlüsse rückwärts bis zu einfachen Sequenzen vorzudringen; es genügt, wenn man überhaupt Sequenzen $\Gamma, C : C, \Delta$ oder $\Gamma, \bot : \Delta$ erreicht.

4.3 Wahrheitsfunktionen

4.3.1 Definition Eine *n-stellige Wahrheitsfunktion* $(n \geq 0)$ ist eine Abbildung

$$\varphi : \{w, f\}^n \longrightarrow \{w, f\}.$$

Einzelne Wahrheitsfunktionen, insbesondere zweistellige, sind aus § 2.1 bekannt. Die Wahrheitstafeln aus 2.1.5 sind Tabellen für Wahrheitsfunktionen, die von den betreffenden Junktoren syntaktisch *dargestellt* werden.

4.3.2 Definition Eine Formel C *stellt* die n-stellige Wahrheitsfunktion φ in den n verschiedenen Aussagenelementen $P_1, \ldots, P_n \not\equiv \bot$ *dar*, wenn

$$\varphi(V(P_1), \ldots, V(P_n)) = V(C)$$

ist für jede Wertung V.

Hierbei wird nicht verlangt, dass C genau die Aussagenelemente $P_1, \ldots, P_n$ enthält. Welche Wahrheitsfunktion C darstellt, hängt offenbar nicht nur von der Menge $\{P_1, \ldots, P_n\} \subseteq AE$ ab, sondern auch von der Reihenfolge der P_i.

Beispiele:

a) Die 0-stellige Wahrheitsfunktion $\varphi = w$ wird dargestellt durch $\bot \to \bot$, aber auch durch $P \to P$, $\neg P \vee P$ und überhaupt durch jede Tautologie.

b) Die Wahrheitstafel zum Junktor j definiert eine Wahrheitsfunktion φ_j. Dann gilt stets

$$\varphi_j(V(P), V(Q)) = V(PjQ)$$

für die Junktoren $j = \to, \wedge, \vee, \leftrightarrow$. Also stellt PjQ die Wahrheitsfunktion φ_j in P, Q dar.

Die Übereinstimmung von Wahrheitsfunktionen ist leicht zu charakterisieren:

4.3.3 Lemma Wenn in den Formeln C, D höchstens $P_1, \ldots, P_n \in AE$ auftreten, so sind äquivalent:

(i) C und D stellen dieselbe Wahrheitsfunktion in $P_1, \ldots, P_n$ dar.

(ii) Für jede Wertung V ist $V(C) = V(D)$.

(iii) $C \leftrightarrow D$ ist eine Tautologie.

Beweis. C stelle φ, D stelle ψ in $P_1, \ldots, P_n$ dar. Dann ist:
(i) $\Leftrightarrow$ $\varphi = \psi$ $\Leftrightarrow$ für jede Wertung V ist

$$V(C) = \varphi(V(P_1), \ldots, V(P_n)) = \psi(V(P_1), \ldots, V(P_n)) = V(D)$$

$\Leftrightarrow$ (ii) $\Leftrightarrow$ stets ist $V(C \leftrightarrow D) = w$ $\Leftrightarrow$ (iii).

Es gibt 2^n n-tupel aus w und f. Also hat der Definitionsbereich einer n-stelligen Wahrheitsfunktion 2^n Elemente. Demnach gibt es 2^{2^n} n-stellige Wahrheitsfunktionen. Für $n = 2$ sind das 16, für $n = 3$ sind das 256 Funktionen.
Sind alle diese Wahrheitsfunktionen durch geeignete Formeln darstellbar? Welche Junktoren muss man verwenden, um alle darstellen zu können?

4.3.4 Definition Eine Menge S von Junktoren heißt ein *vollständiges Junktorensystem*, wenn sich jede Wahrheitsfunktion durch eine Formel darstellen lässt, die aus Aussagenelementen $\not\equiv \bot$ allein mit Junktoren aus S aufgebaut ist.

4.3.5 Satz $\{\top, \bot, \neg, \rightarrow, \wedge\}$ ist ein vollständiges Junktorensystem.

Beweis: Wir zeigen durch Induktion nach n: Jede n-stellige Wahrheitsfunktion besitzt eine Darstellung, die nur diese fünf Junktoren verwendet.

1. φ ist 0-stellig. Dann ist

   $$\varphi = w = V(\top) \quad \text{oder} \quad \varphi = f = V(\bot).$$

   $\top$ und $\bot$ stellen also w und f dar.

2. φ ist $(n+1)$-stellig. Hält man das letzte Argument von φ fest ($= w$ oder $= f$), so erhält man zwei n-stellige Wahrheitsfunktionen. Nach Induktionsvoraussetzung gibt es daher Formeln A und B in den angegebenen Junktoren, so dass

   $$\begin{aligned} &\varphi(V(P_1), \ldots, V(P_n), w) = V(A) \quad \text{und} \\ &\varphi(V(P_1), \ldots, V(P_n), f) = V(B) \end{aligned}$$

   für jede Wertung V gilt. Dann ist

   $$\varphi(V(P_1), \ldots, V(P_n), V(P)) = w$$

   äquivalent zu:

   aus $V(P) = w$ *folgt* $V(A) = w$ *und aus* $V(P) = f$ *folgt* $V(B) = w$,

   also äquivalent zu

   $$V((P \rightarrow A) \wedge (\neg P \rightarrow B)) = w.$$

   Demnach stellt $(P \rightarrow A) \wedge (\neg P \rightarrow B)$ die Funktion φ dar. Wie A und B, so ist auch diese Formel allein mit den angegebenen Junktoren gebildet.

Mit vollständiger Induktion folgt die Behauptung.

4.3.6 Korollar $\{\bot, \to\}$ ist ein vollständiges Junktorensystem.

Beweis: In 1.1.12 sind die Junktoren $\top, \neg, \wedge$ mit Hilfe von $\bot, \to$ definiert; ihre Wahrheitstafeln werden in 2.1.5 gerade an Hand dieser Definitionen errechnet. Wir können also $\top, \neg, \wedge$ in den darstellenden Formeln als Abkürzungen für Ausdrücke lesen, die nur mit $\bot, \to$ gebildet sind.

4.3.7 Korollar $\{\neg, \wedge\}$ ist ein vollständiges Junktorensystem.

Beweis: Wie man mit 2.1.5 nachrechnet, stellen

$$\begin{array}{lll} \top & \text{und} & \neg(P \wedge \neg P), \\ \bot & \text{und} & P \wedge \neg P, \\ A \to B & \text{und} & \neg(A \wedge \neg B) \end{array}$$

jeweils die gleichen Wahrheitsfunktionen dar. Wenn man in einer φ darstellenden Formel C überall $\top, \bot$ und $A \to B$ durch die rechts stehenden Formeln ersetzt, erhält man also eine Formel, die immer noch φ darstellt und aus Elementen von AE allein mit $\neg$ und $\wedge$ gebildet ist.

4.4 Disjunktive und konjunktive Normalform

$\{\neg, \wedge, \vee\}$ ist ein vollständiges Junktorensystem, sogar ein redundantes. Lassen sich Formeln in $\neg, \wedge, \vee$ in einer speziellen Form darstellen, etwa derart, dass die Negation nur ganz innen bei den Aussagenelementen auftritt, dann die Konjunktion, und nur außen die Disjunktion? Folgendes Lemma ist ein Schritt zu einer positiven Lösung.

4.4.1 Lemma (de Morgan'sche Regeln)

$$\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B) \text{ und } \neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$$

sind Tautologien.

Beweis. Sei V eine Wertung.

$$\begin{aligned} V(\neg(A \vee B)) = w &\iff V(A \vee B) = f \\ &\iff V(A) = f \text{ und } V(B) = f \\ &\iff V(\neg A) = w \text{ und } V(\neg B) = w \\ &\iff V(\neg A \wedge \neg B) = w. \end{aligned}$$

Also stellen die beiden Seiten der ersten Äquivalenz dieselbe Wahrheitsfunktion dar, und damit ist die erste Äquivalenz eine Tautologie.
Die zweite Behauptung beweist man analog.

4.4.2 Definition Formeln der Gestalt

$$\bigvee_{i \leq n} B_i :\equiv B_0 \vee B_1 \vee \ldots \vee B_n$$

heißen *iterierte Disjunktionen*, und Formeln der Gestalt

$$\bigwedge_{i \leq n} B_i :\equiv B_0 \wedge B_1 \wedge \ldots \wedge B_n$$

heißen *iterierte Konjunktionen*. Wie stets, ist auch hier Rechtsklammerung fortgelassen.

Da in dieser Definition $n = 0$ sein kann, ist jede Formel sowohl iterierte Disjunktion als auch iterierte Konjunktion. Die Begriffe werden erst nützlich, wenn man zusätzliche Forderungen an die Teilformeln B_i stellt.

4.4.3 Lemma Verallgemeinertes Assoziativgesetz.

$$(A_0 \vee \ldots \vee A_m) \vee (B_0 \vee \ldots \vee B_n) \leftrightarrow (A_0 \vee \ldots \vee A_m \vee B_0 \vee \ldots \vee B_n)$$

und

$$(A_0 \wedge \ldots \wedge A_m) \wedge (B_0 \wedge \ldots \wedge B_n) \leftrightarrow (A_0 \wedge \ldots \wedge A_m \wedge B_0 \wedge \ldots \wedge B_n)$$

sind für alle m, n Tautologien.

Beweis. Sei V eine Wertung. Es ist

$$\begin{aligned} & V((A_0 \vee \ldots \vee A_m) \vee (B_0 \vee \ldots \vee B_n)) = w \\ \iff\ & V(A_0 \vee \ldots \vee A_m) = w \text{ oder } V(B_0 \vee \ldots \vee B_n) = w \\ \iff\ & \text{für ein } i \leq m \text{ ist } V(A_i) = w \text{ oder für ein } j \leq n \text{ ist } V(B_j) = w \\ & \iff V(A_0 \vee \ldots \vee A_m \vee B_0 \vee \ldots \vee B_n) = w \end{aligned}$$

Die Behauptung für iterierte Konjunktionen beweist man analog.

4.4.4 Lemma Verallgemeinertes Distributivgesetz.

$$(\bigvee_{i\leq m} A_i) \wedge (\bigvee_{j\leq n} B_j) \leftrightarrow \bigvee_{i\leq m}\bigvee_{j\leq n} (A_i \wedge B_j)$$

und

$$(\bigwedge_{i\leq m} A_i) \vee (\bigwedge_{j\leq n} B_j) \leftrightarrow \bigwedge_{i\leq m}\bigwedge_{j\leq n} (A_i \vee B_j)$$

sind für alle m, n Tautologien.

Beweis. Sei V eine Wertung. Es ist

$$\begin{array}{ll} & V((\bigvee_{i\leq m} A_i) \wedge (\bigvee_{j\leq n} B_j)) = w \\ \Longleftrightarrow & V(\bigvee_{i\leq m} A_i) = w \text{ und } V(\bigvee_{j\leq n} B_j) = w \\ \Longleftrightarrow & \text{für ein } i \leq m \text{ ist } V(A_i) = w \text{ und für ein } j \leq n \text{ ist } V(B_j) = w \\ \Longleftrightarrow & \text{für ein } i \leq m \text{ und ein } j \leq n \text{ ist } V(A_i \wedge B_j) = w \\ \Longleftrightarrow & \text{für ein } i \leq m \text{ ist } V(\bigvee_{j\leq n} (A_i \wedge B_j)) = w \\ \Longleftrightarrow & V(\bigvee_{i\leq m}\bigvee_{j\leq n} (A_i \wedge B_j)) = w. \end{array}$$

Das andere distributive Gesetz beweist man analog.

4.4.5 Definition Eine Formel D ist *in disjunktiver Normalform (in DNF)*, wenn D eine iterierte Disjunktion $C_0 \vee \ldots \vee C_n$ von iterierten Konjunktionen $C_i \equiv B_{i0} \wedge \ldots \wedge B_{ik_i} (i = 0, \ldots, n)$ von Aussagenelementen $B_{ij} \equiv A_{ij} (\not\equiv \bot)$ und negierten Aussagenelementen $B_{ij} \equiv \neg A_{ij}$ ist. D ist *disjunktive Normalform (DNF) von* einer Formel F, wenn D in DNF ist und $F \leftrightarrow D$ eine Tautologie ist.
Analog ist eine Formel C *in konjunktiver Normalform (KNF)*, wenn C eine iterierte Konjunktion $D_0 \wedge \ldots \wedge D_n$ von iterierten Disjunktionen $D_i \equiv B_{i0} \vee \ldots \vee B_{ik_i} (i = 0, \ldots, n)$ von Aussagenelementen und negierten Aussagenelementen ist. C ist *konjunktive Normalform (KNF) von* einer Formel F, wenn C in KNF ist und $F \leftrightarrow C$ eine Tautologie ist.

Beispiele. A_1, A_2, A_3 seien Aussagenelemente.

1. $D \equiv A_1 \vee (\neg A_2 \wedge A_3)$ ist in DNF, und $C \equiv (A_1 \vee \neg A_2) \wedge (A_1 \vee A_3)$ ist in KNF, und C ist KNF von D und D ist DNF von C.

2. $A_1 \vee \neg A_2 \vee A_3$ und $A_1 \wedge \neg A_2 \wedge A_3$ sind sowohl in DNF als auch in KNF.

3. Jede Formel in DNF ist DNF von sich selbst, jede Formel in KNF ist KNF von sich selbst.

4. Ist $F \leftrightarrow G$ eine Tautologie, so haben F und G dieselben DNF und dieselben KNF.

4.4.6 Satz Jede Formel besitzt eine disjunktive Normalform.

Beweis. Weil $\{\neg, \wedge\}$ ein vollständiges Junktorensystem ist, können wir uns die gegebene Formel F schon aus Aussagenelementen mit $\neg, \wedge$ aufgebaut denken. Wir zeigen durch Induktion nach einem solchen Aufbau der Formel F: Sowohl F als auch $\neg F$ besitzen DNF F^D bzw. $(\neg F)^D$.

1. F ist ein Aussagenelement. Dann sind F und $\neg F$ in DNF, es ist $F^D \equiv F$ und $(\neg F)^D \equiv \neg F$.

2. F ist $\neg G$. Nach Induktionsvoraussetzung haben G und $F \equiv \neg G$ DNF G^D und $(\neg G)^D$. Ferner ist $\neg F \leftrightarrow G$ eine Tautologie, so dass mit $G \leftrightarrow G^D$ auch $\neg F \leftrightarrow G^D$ eine Tautologie ist. Also ist G^D auch DNF von $\neg F$.

3. F ist $G \wedge H$. Nach Induktionsvoraussetzung haben G und H DNF

$$G^D \equiv \bigvee_{i \leq m} G_i \text{ und } H^D \equiv \bigvee_{j \leq n} H_j.$$

Nach Lemma 4.4.4 ist

$$G^D \wedge H^D \leftrightarrow \bigvee_{i \leq m} \bigvee_{j \leq n} (G_i \wedge H_j)$$

eine Tautologie. Aus der rechten Seite dieser Äquivalenz erhält man durch systematisches Rechtsklammern eine $m \cdot n$-gliedrige Disjunktion F^D in DNF, und nach Lemma 4.4.3 ist

$$\bigvee_{i \leq m} \bigvee_{j \leq n} (G_i \wedge H_j) \leftrightarrow F^D$$

eine Tautologie. Da nach Induktionsvoraussetzung $G \leftrightarrow G^D$ und $H \leftrightarrow H^D$ Tautologien sind, ist zunächst $F \leftrightarrow G^D \wedge H^D$ und schließlich $F \leftrightarrow F^D$ eine Tautologie, so dass F^D tatsächlich eine DNF von F ist.
Nach Lemma 4.4.1 ist

$$\neg F \leftrightarrow \neg G \vee \neg H$$

eine Tautologie. Nach Induktionsvoraussetzung haben $\neg G$ und $\neg H$ DNF $(\neg G)^D$ und $(\neg H)^D$. Aus $(\neg G)^D \vee (\neg H)^D$ erhält man durch Rechtsklammern eine Formel $(\neg F)^D$ in DNF, und nach Lemma 4.4.3 ist

$$(\neg G)^D \vee (\neg H)^D \leftrightarrow (\neg F)^D$$

eine Tautologie. Mit Induktionsvoraussetzung folgt $\neg F \leftrightarrow (\neg F)^D$, so dass $(\neg F)^D$ eine DNF von $\neg F$ ist.

Mit Induktion nach dem $\{\neg, \wedge\}$-Aufbau von F folgt der Satz.

Es gibt einen anderen Beweis dieses Satzes, der mittelbar auf den Begriff der Wahrheitsfunktion zurückgreift und auch i. a. zu anderen, meist längeren DNF führt: Unter $A_0, \ldots, A_n$ seien alle Aussagenelemente $\not\equiv \bot$, die in F auftreten. Zu jeder Wertung V setzen wir

$$\begin{aligned} B_i^V :\equiv A_i, \quad &\text{falls } V(A_i) = w, \text{ und } B_i^V :\equiv \neg A_i, \text{ falls } V(A_i) = f \\ C^V :\equiv &\bigwedge_{i \leq n} B_i^V. \end{aligned}$$

Für jede Wertung V' ist

$$\begin{aligned} V'(C^V) = w \quad &\Longleftrightarrow \; V'(B_i^V) = w = V(B_i^V) \text{ für } i = 0, \ldots, n \\ &\Longleftrightarrow \; V'(A_i) = V(A_i) \text{ für } i = 0, \ldots, n. \end{aligned}$$

Wertungen V, die auf $\{A_0, \ldots, A_n\}$ übereinstimmen, definieren also dieselbe Formel C^V (und umgekehrt), so dass wir solche Wertungen identifizieren können. F^D sei nun eine Disjunktion aller der Formeln C^V, für die $V(F) = w$ ist. (Ist stets $V(F) = f$, so sei $F^D :\equiv A_0 \wedge \neg A_0$.) Offenbar ist F^D in DNF.
Man sieht nun leicht, dass F und F^D dieselbe Wahrheitsfunktion in $A_0, \ldots, A_n$ definieren, dass also $F \leftrightarrow F^D$ eine Tautologie ist. Für jede Wertung V' ist nämlich

$$\begin{aligned} V'(F^D) = w \quad &\Longleftrightarrow \; V'(C^V) = w \text{ für ein Disjunktionsglied } C^V \text{ von } F^D \\ &\Longleftrightarrow \; V' = V \text{ auf } \{A_0, \ldots, A_n\} \text{ für ein } V \text{ mit } V(F) = w \\ &\Longleftrightarrow \; V'(F) = w. \end{aligned}$$

Also ist $F^D \leftrightarrow F$ eine Tautologie.
In diesem Beweis kodiert F^D die Wahrheitsfunktion, die F in $A_0, \ldots, A_n$ definiert.

4.4.7 Satz Jede Formel besitzt eine konjunktive Normalform.

Den Beweis kann man in Analogie insbesondere zum ersten Beweis von Satz 4.4.6 selbst führen. Man kann diesen Satz aber auch als Korollar aus Satz 4.4.6 gewinnen.

4.5 Aufgaben

4.5.1 Folgern Sie aus 4.3.3, dass folgende Formeln stets Tautologien sind:

$(\vee Ass)$ $\quad (A \vee B) \vee C \leftrightarrow A \vee (B \vee C)$

$(\vee Komm)$ $\quad A \vee B \leftrightarrow B \vee A$

$(\vee Id)$ $\quad A \vee A \leftrightarrow A$

Welcher Spezialfall von 4.4.3 ist $(\vee Ass)$?

4.5.2 Folgern Sie aus 4.3.3: Sind $A \leftrightarrow B$ und $B \leftrightarrow C$ Tautologien, so ist auch $A \leftrightarrow C$ eine Tautologie.

4.5.3 Welche Wahrheitsfunktion (in A, B) wird dargestellt durch $(A \to B) \leftrightarrow (B \to A)$?

4.5.4 Zeigen Sie, dass folgende Distributivgesetze Tautologien sind:

$(\wedge \vee Distr)$ $\quad A \wedge (B \vee C) \leftrightarrow (A \wedge B) \vee (A \wedge C)$

$(\vee \wedge Distr)$ $\quad A \vee (B \wedge C) \leftrightarrow (A \vee B) \wedge (A \vee C)$

Welche Spezialfälle von 4.4.4 sind diese Distributivgesetze?

4.5.5 Skizzieren Sie für den Fall, dass A, B, C Primformeln sind, eine Herleitung von $A \wedge (B \vee C) : (A \wedge B) \vee (A \wedge C)$.

4.5.6 Zeigen Sie:

a) $\{\neg, \vee\}$ ist ein vollständiges Junktorensystem.

b) $\{\top, \bot, \neg\}$ ist kein vollständiges Junktorensystem.

c) (etwas schwerer) $\{\top, \bot, \wedge, \vee\}$ ist kein vollständiges Junktorensystem.

4.5.7 Der SHEFFERsche Strich | ist definiert durch

$$A|B :\equiv A \to \neg B.$$

a) Geben Sie die Wahrheitstafel von | an. Welche Wahrheitsfunktion $\varphi_|$ definiert der Junktor |?

b) Zeigen Sie: $\{|\}$ ist ein vollständiges Junktorensystem.

4.5.8 Geben Sie für $A, B, C \in AE$ eine DNF von $(A \leftrightarrow B) \leftrightarrow C$ an.

4.5.9 Beweisen Sie Satz 4.4.7: Jede Formel besitzt eine KNF.

§5 Zulässige Regeln der Prädikatenlogik

5.1 Schwache Schlussregeln

5.2 Junktoren-Regeln

5.3 Quantoren-Regeln

5.4 Das Deduktionstheorem

5.5 Aufgaben

5.1 Schwache Schlussregeln

Für die genauere Kenntnis unseres Herleitungsbegriffs ist es wichtig, nicht nur Herleitungen von Sequenzen und Sequenzenschemata zu bilden, sondern auch Transformationen zu untersuchen, die aus gegebenen Herleitungen neue Herleitungen mit mehr oder weniger ähnlichen Endsequenzen machen. *Zulässige Regeln* sind das Ergebnis solcher Transformationen.

5.1.1 Definition Eine Schlussregel von $\Gamma_1 : \Delta_1;\ \Gamma_2 : \Delta_2; \ldots ; \Gamma_k : \Delta_k$ auf $\Gamma : \Delta$ heißt *zulässig* (in T), wenn aus der Herleitbarkeit der $\Gamma_i : \Delta_i$ (in T) $(i = 1, \ldots, k)$ die Herleitbarkeit von $\Gamma : \Delta$ (in T) folgt:

$$T \vdash \Gamma_1 : \Delta_1, \ldots, T \vdash \Gamma_k : \Delta_k \Rightarrow T \vdash \Gamma : \Delta.$$

Die $\Gamma_i : \Delta_i$ und $\Gamma : \Delta$ bezeichnen hier *Sequenzenschemata*, in denen Mitteilungszeichen für Sequenzen, Formeln, Terme etc. auftreten können.

Ist die Regel in jeder Theorie zulässig, so schreiben wir:

$$\vdash \Gamma_1 : \Delta_1, \ldots, \vdash \Gamma_k : \Delta_k \Rightarrow \vdash \Gamma : \Delta.$$

Beispiele:

a. Die Grundschlussregeln sind selbstverständlich zulässige Regeln. Bei ihnen besteht die oben erwähnte Transformation im Anhängen eines Grundschlusses an die ein oder zwei gegebenen Herleitungen.

b. Herleitbare Sequenzenschemata können auch als zulässige Regeln aufgefasst werden, und zwar als Regeln mit $k = 0$ Prämissen.

c. $\vdash\colon A \to B \Rightarrow \vdash A : B$ und $\vdash\colon \forall x\, \mathcal{F}(x) \Rightarrow \vdash\colon \mathcal{F}(t)$ sind Spezialfälle von Inversions- Schlussregeln, die wir unten behandeln.

d. $\vdash A : B;\ \vdash B : C \Rightarrow \vdash A : C$,
eine Form des Kettenschlusses, ist ein Spezialfall der Mix-Regel

$$\vdash \Gamma_1, B : \Delta_1;\ \vdash \Gamma_2 : B, \Delta_2 \Rightarrow \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2,$$

die ihrerseits als Spezialfall die Schnittregel

$$\vdash \Gamma, B : \Delta;\ \vdash \Gamma : B, \Delta \Rightarrow \vdash \Gamma : \Delta$$

enthält. Die Zulässigkeit dieser Regeln ist nicht trivial in unserem schnittfreien Sequenzenkalkül. Sie ergibt sich einmal semantisch aus dem Korrektheits- und Vollständigkeitssatz, zum anderen direkt mit dem Hauptsatz von Gentzen, den wir erst in Kapitel 5 beweisen.

Bemerkung. In den Beispielen a. und b. entsteht die gesuchte Herleitung durch Anhängen von Grundschlüssen an die gegebenen Herleitungen. Solche Regeln nennt man auch *direkt herleitbar*. Wegen der Schnittfreiheit des Sequenzenkalküls sind die Regeln in den Beispielen c. und d. nicht direkt herleitbar. Bei den Regeln in Beispiel c. ist eine Herleitung der Konklusion im wesentlichen in der gegebenen Herleitung schon enthalten; insbesondere ist die gesuchte Herleitung nicht länger als die gegebene. Solche Regeln nennen wir *schwache* Schlussregeln. Um diesen Begriff zu präzisieren, definieren wir ein Maß für die Länge von Herleitungen.

5.1.2 Rekursive Definition der *Ordnung* einer Herleitung in einer Theorie T.

1. Jedes logische Axiom hat die Ordnung 0.
2. Hat die Herleitung H in T die Ordnung k und ist $\frac{H}{\Gamma:\Delta}$ eine Herleitung in T, so hat $\frac{H}{\Gamma:\Delta}$ die Ordnung $k+1$.
3. Haben die Herleitungen H_1 und H_2 in T die Ordnungen k bzw. l und ist $\frac{H_1 \quad H_2}{\Gamma:\Delta}$ eine Herleitung in T, so hat diese die Ordnung $\max(k,l)+1$.

5.1.3 Definition $\Gamma : \Delta$ ist in T *mit einer Ordnung* $\leq n$ *herleitbar*, und wir schreiben $T \overset{n}{\vdash} \Gamma : \Delta$, wenn es in T eine Herleitung von $\Gamma : \Delta$ mit einer Ordnung $\leq n$ gibt.

Die Ordnung einer Herleitung H ist die Länge (Anzahl der Grundschlüsse) eines längsten Fadens in dem Baum H. Durch die Einführung der Herleitungsordnung geht die Herleitungsinduktion in eine vollständige Induktion über.

5.1.4 Definition Eine Schlussregel von $\Gamma_1 : \Delta_1$ auf $\Gamma : \Delta$ heißt *schwache Schlussregel*, wenn für jedes $n \in \mathbb{N}$ gilt:

$$\overset{n}{\vdash} \Gamma_1 : \Delta_1 \Rightarrow \overset{n}{\vdash} \Gamma : \Delta.$$

Schwache Schlussregeln beruhen auf einfachen Verfahren, die eine gegebene Herleitung H_1 von $\Gamma_1 : \Delta_1$ in eine Herleitung H von $\Gamma : \Delta$ transformieren:
Logische Axiome von H_1 werden in logische Axiome von H überführt (für $n = 0$). Gewöhnlich gehen (für $n > 0$) Grundschlüsse in H_1 auch wieder in Grundschlüsse nach derselben Regel über. Wenn dies stets der Fall ist, sind H_1 und H völlig strukturgleich: H_1 und H unterscheiden sich wohl in den Sequenzen, nicht aber in der Aufeinanderfolge der Anwendung der Grundschlussregeln.
Es können auch einzelne Grundschlüsse aus H_1 bei der Transformation wegfallen und dann in H fehlen. In diesen Fällen kann die Ordnung von H sogar kleiner sein als die von H_1, was nach 5.1.3 auch erlaubt ist.
Wir geben die wichtigsten schwachen Schlussregeln an.

5.1.5 Substitutionssatz Die *Substitutionsregel* ist eine schwache Regel:
$(Subst)$ $\overset{n}{\vdash} \Gamma(b) : \Delta(b) \Rightarrow \overset{n}{\vdash} \Gamma(s) : \Delta(s)$, falls b nicht in $\Gamma : \Delta$ auftritt.

Beweis durch Induktion nach n:

1. $\overset{0}{\vdash} \Gamma(b) : \Delta(b)$. Dann hat $\Gamma(b) : \Delta(b)$ eine Gestalt $\Gamma(b),\ P(b) : P(b), \Delta(b)$ oder $\Gamma(b), \bot : \Delta(b)$ mit einer Primformel $P(b)$, wobei b in Γ, Δ, P nicht auftritt. Dann ist auch $P(s)$ eine Primformel, und $\Gamma(s) : \Delta(s)$ ist wieder ein logisches Axiom.

2. $\overset{n+1}{\vdash} \Gamma(b) : \Delta(b)$, und der letzte Grundschluss ist kein $(\forall S)$-Schluss und hat die Prämissen $\overset{n}{\vdash} \Gamma_i(b) : \Delta_i(b)$ ($i = 1, 2$ oder $i = 1$), b nicht in $\Gamma_i : \Delta_i$. Nach Induktionsvoraussetzung ist $\overset{n}{\vdash} \Gamma_i(s) : \Delta_i(s)$, und daraus schließt man mit derselben Grundschlussregel auf $\overset{n+1}{\vdash} \Gamma(s) : \Delta(s)$.

3. $\overset{n+1}{\vdash} \Gamma(b) : \Delta(b)$, und der letzte Schluss ist ein $(\forall S)$-Schluss

$$\overset{n}{\vdash} \Gamma(b) : \mathcal{F}(a, b), \Delta_1(b) \vdash \Gamma(b) : \forall x\, \mathcal{F}(x, b), \Delta_1(b),$$

wobei a nicht in Γ, $\mathcal{F}(*_1, b), \Delta_1$ und b nicht in Γ, $\mathcal{F}, \Delta_1$ auftritt. Sei nun c eine freie Variable, die nicht in Γ, $\mathcal{F}, \Delta_1, b, s$ auftritt. (Es gibt unendlich viele solche c.)
Nach Induktionsvoraussetzung ist

$$\vdash^n \Gamma(b) : \mathcal{F}(c, b), \Delta_1(b)$$

und wieder nach Induktionsvoraussetzung

$$\vdash^n \Gamma(s) : \mathcal{F}(c, s), \Delta_1(s).$$

Hieraus kann man mit einem $(\forall S)$-Schluss auf

$$\vdash^{n+1} \Gamma(s) : \forall x\, \mathcal{F}(x, s), \Delta_1(s) \equiv \Gamma(s) : \Delta(s)$$

schließen, weil die Variablenbedingung für c erfüllt ist.

Mit vollständiger Induktion nach n folgt nun $(Subst)$.

Hier wird, wie der Beweis zeigt, die gegebene Herleitung $H(b)$ von $\Gamma(b) : \Delta(b)$ in eine völlig strukturgleiche Herleitung von $\Gamma(s) : \Delta(s)$ überführt, allerdings nicht buchstäblich in $H(s)$. Wenn in $H(b)$ freie Variablen aus s einer Variablenbedingung unterliegen, müssen diese erst in $H(b)$ umbenannt werden. Ähnlich sieht es bei der Strukturschlussregel aus, deren Korrektheit wir bereits in 2.3.4 gezeigt haben. Wir wiederholen Definition 2.3.2:

5.1.6 Definition *Aus* einer Sequenz $\Gamma : \Delta$ *folgt* eine Sequenz $\Gamma^+ : \Delta^+$ *strukturell*, wir schreiben $\Gamma : \Delta \subset_S \Gamma^+ : \Delta^+$, wenn $\Gamma \subseteq \Gamma^+$ und $\Delta \subseteq \Delta^+ \cup \{\bot\}$ ist.

Im Wesentlichen sind also Ante- und Sukzedens der ersten Sequenz im Ante- bzw. Sukzedens der zweiten Sequenz enthalten. Allerdings kann ein $\bot$ im Sukzedens verloren gehen, ohne die strulturelle Folgerung zu stören.

5.1.7 Satz Die *Strukturschlussregel* ist eine schwache Regel:
(Str) Wenn $\Gamma : \Delta \subset_S \Gamma^+ : \Delta^+$, dann $\vdash^n \Gamma : \Delta \Rightarrow \vdash^n \Gamma^+ : \Delta^+$.

Beweis durch Induktion nach n.

1. $n = 0$. Dann ist $\Gamma : \Delta$ ein logisches Axiom, und es gibt eine Primformel $P \in \Gamma \cap \Delta$, oder es ist $\bot \in \Gamma$. Wegen $\Gamma : \Delta \subset_S \Gamma^+ : \Delta^+$ ist dann auch $P \in \Gamma^+ \cap \Delta^+$, oder es ist jedenfalls $\bot \in \Gamma^+$. Also ist auch $\Gamma^+ : \Delta^+$ ein logisches Axiom.

2. $\vdash^{n+1} \Gamma : \Delta$, und der letzte Schluss ist kein $(\forall S)$-Schluss und hat die Prämissen $\vdash^{n} \Gamma_i : \Delta_i$ $(i = 1, 2$ oder $i = 1)$.
   Setzt man

$$\Gamma_i^+ = \Gamma_i \cup \Gamma^+ \text{ und } \Delta_i^+ = \Delta_i \cup \Delta^+ \text{ oder } \Delta_i^+ = (\Delta_i - \{\bot\}) \cup \Delta^+,$$

   so ist $\Gamma_i : \Delta_i \subset_S \Gamma_i^+ : \Delta_i^+$, so dass nach Induktionsvoraussetzung $\vdash^{n} \Gamma_i^+ : \Delta_i^+$ (für $i = 1, 2$ oder $i = 1$) ist. Ein Grundschluss nach derselben Regel ergibt dann $\vdash^{n+1} \Gamma^+ : \Delta^+$ bei Wahl der richtigen Alternative für Δ_i^+.

3. $\vdash^{n+1} \Gamma : \Delta$, und der letzte Schluss ist ein $(\forall S)$-Schluss mit der Prämisse $\vdash^{n} \Gamma : \mathcal{F}(a), \Delta_1$. Ist c eine freie Variable, die in Γ^+, Δ^+ nicht auftritt, so folgt aus dem Substitutionssatz 5.1.5 $\vdash^{n} \Gamma : \mathcal{F}(c), \Delta_1$. Nach Induktionsvoraussetzung ist dann auch $\vdash^{n} \Gamma^+ : \mathcal{F}(c), \Delta^+$, und wegen $\forall x\, \mathcal{F}(x) \in \Delta^+$ ergibt ein $(\forall S)$-Schluss $\vdash^{n+1} \Gamma^+ : \Delta^+$.

Mit Induktion nach n folgt (Str).

Beispiel. $\vdash^{n} \Gamma : \bot, \Delta \Rightarrow \vdash^{n} \Gamma : A, \Delta$ für jede Formel A.

Auf der Basis dieser beiden schwachen Schlussregeln können wir auch einige Grundschlussregeln umkehren.

5.1.8 Satz Die $(\to S)$-*Inversion* ist eine schwache Schlussregel:

$$(\to SInv) \quad \vdash^{n} \Gamma : A \to B,\ \Delta \Rightarrow \vdash^{n} \Gamma, A : B, \Delta.$$

Beweis. Ist in der Voraussetzung $A \to B \in \Delta$, so geht die Behauptung aus ihr durch einen Strukturschluss, nämlich eine Abschwächung, hervor. Also können wir uns auf den Fall $A \to B \notin \Delta$ beschränken.
Wir beweisen den Satz in diesem Fall durch Induktion nach n.

1. $n = 0$, die Voraussetzung ist ein logisches Axiom. Da die logische Axiom-Eigenschaft am geeigneten Auftreten einer Primformel hängt, ist sogar $\Gamma : \Delta$, also auch $\Gamma, A : B, \Delta$ ein logisches Axiom.

2. $\vdash^{n+1} \Gamma : A \to B, \Delta$, und der letzte Schluss ist *kein* $(\to S)$-Schluss mit Hauptformel $A \to B$, hat also Prämissen $\vdash^{n} \Gamma_i : A \to B, \Delta_i$ $(i = 1$ oder $i = 1, 2)$. Nach Induktionsvoraussetzung ist dann $\vdash^{n} \Gamma_i, A : B, \Delta_i$, und ein Schluss nach derselben Grundschlussregel ergibt $\vdash^{n+1} \Gamma, A : B, \Delta$.

3. $\vdash^{n+1} \Gamma : A \to B, \Delta$, und der letzte Schluss hat die Hauptformel $A \to B$, ist also ein $(\to S)$-Schluss mit einer Prämisse $\vdash^{n} \Gamma, A : B, \Delta_1$ mit $A \to B, \Delta_1 = A \to B, \Delta$.

   3.1 Ist $\Delta_1 = \Delta$, so folgt die Behauptung, und zwar sogar unter Fortfall des letzten $(\to S)$-Schlusses.

   3.2 Weil $A \to B \notin \Delta$ ist, ist sonst $\Delta_1 = A \to B, \Delta$, also

$$\vdash^{n} \Gamma, A : B, A \to B, \Delta,$$

   und mit Induktionsvoraussetzung folgt wieder $\vdash^{n} \Gamma, A : B, \Delta$.

Mit Induktion folgt $(\to S\ Inv)$.

Die anderen beiden Inversionsregeln werden weitgehend analog bewiesen.

5.1.9 Satz Die $(\to A)$-*Inversionen* sind schwache Schlussregeln:

$$(\to AInv) \quad \vdash^{n} \Gamma, A \to B : \Delta \Rightarrow \vdash^{n} \Gamma : A,\ \Delta \text{ und } \vdash^{n} \Gamma,\ B : \Delta.$$

Beweis. Analog zum vorigen Beweis können wir uns auf den Fall $A \to B \notin \Gamma$ beschränken und nach n induzieren.

1. $n = 0$, die Voraussetzung ist ein logisches Axiom. Dann ist wie eben $\Gamma : \Delta$, also auch $\Gamma : A, \Delta$ und $\Gamma,\ B : \Delta$ ein logisches Axiom.

2. $\vdash^{n+1} \Gamma, A \to B : \Delta$, und der letzte Schluss ist *kein* $(\to A)$-Schluss mit Hauptformel $A \to B$. Mit Induktionsvoraussetzung folgt dann die Behauptung wie eben.

3. $\vdash^{n+1} \Gamma, A \to B : \Delta$, und der letzte Schluss hat die Hauptformel $A \to B$, ist also ein $(\to A)$-Schluss mit Prämissen $\vdash^{n} \Gamma_1 : A, \Delta$ und $\vdash^{n} \Gamma_1, B : \Delta$ mit $\Gamma_1, A \to B = \Gamma, A \to B$.

   3.1 Ist $\Gamma_1 = \Gamma$, so folgen die Behauptungen unter Fortfall des letzten $(\to A)$-Schlusses.

   3.2 Sonst ist wie eben $\Gamma_1 = \Gamma, A \to B$, also

$$\vdash^{n} \Gamma, A \to B : A, \Delta \text{ und } \vdash^{n} \Gamma, A \to B, B : \Delta,$$

   und mit Induktionsvoraussetzung folgen wieder

$$\vdash^{n} \Gamma : A, \Delta \text{ und } \vdash^{n} \Gamma, B : \Delta.$$

Mit Induktion folgt ($\rightarrow AInv$).

5.1.10 Satz Die ($\forall S$)-*Inversion* ist eine schwache Schlussregel:

$$(\forall SInv) \quad \vdash^n \Gamma : \forall x\, \mathcal{F}(x), \Delta \Rightarrow \vdash^n \Gamma :\ \mathcal{F}(s), \Delta.$$

Beweis. Wie eben können wir uns auf den Fall $\forall x\, \mathcal{F}(x) \notin \Delta$ beschränken und nach n induzieren.

1. $n = 0$, die Voraussetzung ist ein logisches Axiom. Dann ist wie eben $\Gamma : \Delta$, also auch $\Gamma :\ \mathcal{F}(s), \Delta$ ein logisches Axiom.

2.1 $\vdash^{n+1} \Gamma\ :\ \forall x\, \mathcal{F}(x), \Delta$, und der letzte Schluss ist *kein* ($\forall S$)-Schluss. Mit Induktionsvoraussetzung folgt dann die Behauptung wie eben.

2.2 $\vdash^{n+1} \Gamma : \forall x\, \mathcal{F}(x), \Delta$, und der letzte Schluss ist ein ($\forall S$)-Schluss, aber *nicht* mit Hauptformel $\forall x\, \mathcal{F}(x)$ mit einer Prämisse $\vdash^n \Gamma\ :\ \forall x\, \mathcal{F}(x), G(b), \Delta_1$ mit $\forall y G(y), \Delta_1 \equiv \Delta$ und b nicht in $\Gamma, \mathcal{F}, G, \Delta_1$. Sei c eine freie Variable, die außerdem auch nicht in s auftritt. Mit der Substitutionsregel folgt $\vdash^n \Gamma : \forall x\, \mathcal{F}(x), G(c), \Delta_1$, und mit Induktionsvoraussetzung folgt die Behauptung wie eben.

3. $\vdash^{n+1} \Gamma : \forall x\, \mathcal{F}(x), \Delta$, und der letzte Schluss hat die Hauptformel $\forall x\, \mathcal{F}(x)$, ist also ein ($\forall S$)-Schluss mit einer Prämisse $\vdash^n \Gamma :\ \mathcal{F}(a), \Delta_1$ mit $\forall x\, \mathcal{F}(x), \Delta_1 = \forall x\, \mathcal{F}(x), \Delta$ und a nicht in $\Gamma, \mathcal{F}, \Delta_1$.

3.1 Ist $\Delta_1 = \Delta$, so folgt mit der Substitutionsregel die Behauptung.

3.2 Sonst ist wie eben $\Delta_1 = \forall x\, \mathcal{F}(x), \Delta$, also nach ($Subst$)

$$\vdash^n \Gamma :\ \mathcal{F}(s),\ \forall x\, \mathcal{F}(x), \Delta$$

und mit Induktionsvoraussetzung folgt wieder $\vdash^n \Gamma :\ \mathcal{F}(s), \Delta$.

Mit Induktion folgt ($\forall SInv$).

5.2 Junktoren-Regeln

Wir übertragen die Grundschluss- und die Inversionsregeln für die Implikation auf die definierten Junktoren $\neg$, $\vee$, $\wedge$.

In 1.1.12 ist die Negation $\neg A$ definiert als $A \to \bot$.

5.2.1 Lemma Folgende Negationsregeln sind zulässig:

$$(\neg S) \qquad \vdash^n \Gamma, A : \Delta \Rightarrow \vdash^{n+1} \Gamma : \neg A, \Delta$$

$$(\neg A) \qquad \vdash^n \Gamma : A, \Delta \Rightarrow \vdash^{n+1} \Gamma, \neg A : \Delta$$

$$(\neg SInv) \qquad \vdash^n \Gamma : \neg A, \Delta \Rightarrow \vdash^n \Gamma, A : \Delta$$

$$(\neg AInv) \qquad \vdash^n \Gamma, \neg A : \Delta \Rightarrow \vdash^n \Gamma : A, \Delta$$

Beweis. Zu $(\neg S)$: Aus $\vdash^n \Gamma, A : \Delta$ folgt mit einem Strukturschluss $\vdash^n \Gamma, A : \bot, \Delta$, woraus mit $(\to S)$ die Behauptung folgt.
Zu $(\neg A)$: Aus $\vdash^n \Gamma : A, \Delta$ und dem logischen Axiom $\vdash^0 \Gamma, \bot : \Delta$ folgt mit $(\to A)$ die Behauptung.
Zu $(\neg SInv)$: Aus $\vdash^n \Gamma : A \to \bot, \Delta$ folgt mit $(\to SInv)$ $\vdash^n \Gamma, A : \bot, \Delta$ und daraus mit einem Strukturschluss die Behauptung.
$(\neg AInv)$ ist der Spezialfall $B \equiv \bot$ der ersten $(\to A)$-Inversion.

Bemerkung. Von diesen Regeln ist nur $(\neg A)$ direkt herleitbar, wie der Beweis zeigt, die anderen drei nicht. Für $(\neg S)$ kann das hinzutretende $\bot$ im Sukzedens i. a. nur durch einen Strukturschluss gewonnen werden, und die Inversionsregeln verletzen das Subformelprinzip und können deshalb nicht direkt herleitbar sein.
Durch Koppelung von jeweils zwei Negations-Schlüssen erhält man, wenn man auf die Kontrolle der Herleitungsordnung verzichtet:

5.2.2 Korollar Folgende Stabilitätsregeln sind zulässig:

$$(\neg\neg S) \qquad \vdash \Gamma : A, \Delta \Leftrightarrow \vdash \Gamma : \neg\neg A, \Delta$$

$$(\neg\neg A) \qquad \vdash \Gamma, A : \Delta \Leftrightarrow \vdash \Gamma, \neg\neg A : \Delta$$

5.2.3 Korollar Folgende Kontrapositionsregeln sind zulässig:

$$\begin{aligned}
&\vdash \Gamma, A : B, \Delta \Leftrightarrow \quad \vdash \Gamma, \neg B : \neg A, \Delta \\
&\vdash \Gamma, A : \neg B, \Delta \Rightarrow \quad \vdash \Gamma, B : \neg A, \Delta \\
&\vdash \Gamma, \neg A : B, \Delta \Rightarrow \quad \vdash \Gamma, \neg B : A, \Delta
\end{aligned}$$

Die Disjunktion $A \vee B$ ist in 1.1.12 definiert als $\neg A \to B$.

5.2.4 Lemma Folgende Disjunktionsregeln sind zulässig:

$(\vee S)$ $\quad \overset{n}{\vdash} \Gamma : A, B, \Delta \Rightarrow \overset{n+2}{\vdash\!\!\!-} \Gamma : A \vee B, \Delta$

$(\vee A)$ $\quad \overset{n}{\vdash} \Gamma, A : \Delta$ und $\overset{n}{\vdash} \Gamma, B : \Delta \Rightarrow \overset{n+2}{\vdash\!\!\!-} \Gamma, A \vee B : \Delta$

$(\vee SInv)$ $\quad \overset{n}{\vdash} \Gamma : A \vee B, \Delta \Rightarrow \overset{n}{\vdash} \Gamma : A, B, \Delta$

$(\vee AInv)$ $\quad \overset{n}{\vdash} \Gamma, A \vee B : \Delta \Rightarrow \overset{n}{\vdash} \Gamma, A : \Delta$ und $\overset{n}{\vdash} \Gamma, B : \Delta$

Beweis: Zu $(\vee S)$: Die Voraussetzung ergibt mit $(\neg A)$ $\overset{n+1}{\vdash\!\!\!-} \Gamma, \neg A : B, \Delta$, woraus mit $(\to S)$ folgt

$$\overset{n+2}{\vdash\!\!\!-} \Gamma : \neg A \to B, \Delta,$$

und das ist die Behauptung.
Zu $(\vee A)$: Die erste Voraussetzung ergibt mit $(\neg S)$ $\overset{n+1}{\vdash\!\!\!-} \Gamma : \neg A, \Delta$, und hieraus und der zweiten Voraussetzung folgt mit $(\to A)$ $\overset{n+2}{\vdash\!\!\!-} \Gamma, \neg A \to B : \Delta$, und das ist die Behauptung.
Zu $(\vee SInv)$: Mit $(\to SInv)$ ergibt die Voraussetzung $\overset{n}{\vdash} \Gamma, \neg A : B, \Delta$, und hieraus folgt mit $(\neg AInv)$ die Behauptung.
Zu $(\vee AInv)$: Mit $(\to AInv)$ ergibt die Voraussetzung $\overset{n}{\vdash} \Gamma : \neg A, \Delta$ und $\overset{n}{\vdash} \Gamma, B : \Delta$. Damit haben wir schon die zweite Behauptung, und ein $(\neg SInv)$-Schluss liefert auch die erste Behauptung.

Durch Koppelung von Disjunktions-Schlüssen erhält man wieder leicht:

5.2.5 Korollar Die Regeln für die Kommutativität und Assoziativität der Disjunktion sind zulässig:

$(\vee SKomm)$ $\quad \vdash \Gamma : A \vee B, \Delta \Rightarrow \vdash \Gamma : B \vee A, \Delta$

$(\vee AKomm)$ $\quad \vdash \Gamma, A \vee B : \Delta \Rightarrow \vdash \Gamma, B \vee A : \Delta$

$(\vee SAss)$ $\quad \vdash \Gamma : (A \vee B) \vee C, \Delta \iff \vdash \Gamma : A \vee (B \vee C), \Delta$

$(\vee AAss)$ $\quad \vdash \Gamma, (A \vee B) \vee C : \Delta \iff \vdash \Gamma, A \vee (B \vee C) : \Delta$

Die Konjunktion $A \wedge B$ ist in 1.1.12 definiert als $\neg(A \to \neg B)$.

5.2.6 Lemma Folgende Konjunktionsregeln sind zulässig:

$(\wedge S)$ $\quad \overset{n}{\vdash} \Gamma : A, \Delta$ und $\overset{n}{\vdash} \Gamma : B, \Delta \Rightarrow \overset{n+3}{\vdash\!\!\!-} \Gamma : A \wedge B, \Delta$

$(\wedge A)$ $\quad \overset{n}{\vdash} \Gamma, A, B : \Delta \Rightarrow \overset{n+3}{\vdash\!\!\!-} \Gamma, A \wedge B : \Delta$

$(\wedge SInv)$ $\quad \overset{n}{\vdash} \Gamma : A \wedge B, \Delta \Rightarrow \overset{n}{\vdash} \Gamma : A, \Delta$ und $\overset{n}{\vdash} \Gamma : B, \Delta$

$(\wedge AInv)$ $\quad \overset{n}{\vdash} \Gamma, A \wedge B : \Delta \Rightarrow \overset{n}{\vdash} \Gamma, A, B : \Delta$

Die Beweise überlegt man sich in Analogie zu 5.2.4 leicht selbst.

Durch Koppelung von Konjunktions-Schlüssen erhält man analog zu 5.2.5:

5.2.7 Korollar Die Regeln für die Kommutativität und Assoziativität der Konjunktion sind zulässig:

$(\wedge SKomm)$ $\quad \vdash \Gamma : A \wedge B, \Delta \Rightarrow \vdash \Gamma : B \wedge A, \Delta$

$(\wedge AKomm)$ $\quad \vdash \Gamma, A \wedge B : \Delta \Rightarrow \vdash \Gamma, B \wedge A : \Delta$

$(\wedge SAss)$ $\quad \vdash \Gamma : (A \wedge B) \wedge C, \Delta \iff \vdash \Gamma : A \wedge (B \wedge C), \Delta$

$(\wedge AAss)$ $\quad \vdash \Gamma, (A \wedge B) \wedge C : \Delta \iff \vdash \Gamma, A \wedge (B \wedge C) : \Delta$

Wenn man Disjunktions- und Konjunktions-Schlüsse geeignet mischt, erhält man ferner:

5.2.8 Korollar Die Regeln für die Distributivität von $\wedge$ und $\vee$ sind zulässig:

$(\wedge \vee Distr)$ $\quad \vdash \Gamma : A \wedge (B \vee C), \Delta \iff \vdash \Gamma : (A \wedge B) \vee (A \wedge C), \Delta$

$(\vee \wedge Distr)$ $\quad \vdash \Gamma : A \vee (B \wedge C), \Delta \iff \vdash \Gamma : (A \vee B) \wedge (A \vee C), \Delta$

(und analog auch im Antezedens.)

Bemerkung. Aus dem Satz 4.1.2 von der Identität erhält man mit 5.2.5

$$\vdash A \vee B : B \vee A \text{ und } \vdash (A \vee B) \vee C : A \vee (B \vee C),$$

mit 5.2.7

$$\vdash A \wedge B : B \wedge A \text{ und } \vdash (A \wedge B) \wedge C : A \wedge (B \wedge C),$$

und mit 5.2.8

$$\vdash A \wedge (B \vee C) : (A \wedge B) \vee (A \wedge C) \text{ und } \vdash A \vee (B \wedge C) : (A \vee B) \wedge (A \vee C).$$

Alle diese Sequenzen sind aber Tautologien und deshalb auch schon nach dem Korollar 4.2.16 des Tautologiesatzes herleitbar. Die Bedeutung der Junktoren-Regeln liegt in ihrer Anwendung außerhalb des rein aussagenlogischen Zusammenhangs.

5.3 Quantoren-Regeln

Wir verwenden den Allquantor $\forall$ als Grundzeichen und den Existenzquantor $\exists$ als definiertes Zeichen. Die Existenzformel $\exists x\, \mathcal{F}(x)$ ist in 1.1.12 definiert als

$$\neg\forall x \neg\, \mathcal{F}(x) \equiv \forall x(\, \mathcal{F}(x) \to \bot) \to \bot.$$

Existenzformeln sind also negierte Allformeln. Angesichts der Negationsregeln 5.2.1 besteht eine starke Symmetrie in dem Verhältnis der All- und der Existenzformeln zum Ante- und zum Sukzedens von Sequenzen: Existenzformeln verhalten sich im Sukzedens so wie Allformeln im Antezedens, und umgekehrt. Dieser Symmetrie wollen wir nachgehen. Zunächst stellen wir für den Existenzquantor die Regeln auf, die den uns bekannten Regeln $(\forall S)$, $(\forall A)$ und $(\forall SInv)$ entsprechen. Wir setzen dabei stets voraus, dass Zeichenreihen $\forall x\, \mathcal{F}(x)$, die wir betrachten, tatsächlich Formeln sind und dass dabei x in der Nennform $\mathcal{F}$ nicht auftritt.

5.3.1 Lemma Folgende Existenz-Regeln sind zulässig:

$(\exists A)$	$\vdash^n \Gamma, \mathcal{F}(a) : \Delta \Rightarrow \vdash^{n+3} \Gamma, \exists x\, \mathcal{F}(x) : \Delta,$	falls a nicht in $\Gamma, \mathcal{F}, \Delta$ auftritt.
$(\exists S)$	$\vdash^n \Gamma : \mathcal{F}(t), \Delta \Rightarrow \vdash^{n+3} \Gamma : \exists x\, \mathcal{F}(x), \Delta$	
$(\exists AInv)$	$\vdash^n \Gamma, \exists x\, \mathcal{F}(x) : \Delta \Rightarrow \vdash^n \Gamma, \mathcal{F}(t) : \Delta$	

Beweis. Zu $(\exists A)$: Die Voraussetzung ergibt mit $(\neg S)$ $\overset{n+1}{\vdash\!\!\!\!-}\, \Gamma : \neg\, \mathcal{F}(a), \Delta$. Daraus macht ein $(\forall S)$-Grundschluss $\overset{n+2}{\vdash\!\!\!\!-}\, \Gamma : \forall x \neg\, \mathcal{F}(x), \Delta$, und das ergibt mit $(\neg A)$

$$\overset{n+3}{\vdash\!\!\!\!-}\, \Gamma, \neg \forall x \neg\, \mathcal{F}(x) : \Delta,$$

und das ist die Behauptung.
Zu $(\exists S)$: Aus der Voraussetzung folgt mit $(\neg A)$ $\overset{n+1}{\vdash\!\!\!\!-}\, \Gamma, \neg\, \mathcal{F}(t) : \Delta$. Daraus macht ein $(\forall A)$-Grundschluss $\overset{n+2}{\vdash\!\!\!\!-}\, \Gamma, \forall x \neg\, \mathcal{F}(x) : \Delta$, und das ergibt mit $(\neg S)$ die Behauptung.
Zu $(\exists AInv)$: Aus der Voraussetzung $\overset{n}{\vdash} \Gamma, \neg \forall x \neg\, \mathcal{F}(x) : \Delta$ folgt mit $(\neg AInv)$ $\overset{n}{\vdash} \Gamma : \forall x \neg\, \mathcal{F}(x), \Delta$. Hieraus ergibt sich mit $(\forall SInv)$ $\overset{n}{\vdash} \Gamma : \neg\, \mathcal{F}(t), \Delta$, woraus mit $(\neg SInv)$ die Behauptung folgt.

Bemerkung. Alle $\exists$-Regeln sind wie $(\neg S)$ nur zulässig und nicht wie die Grundschlussregeln $(\forall S)$ und $(\forall A)$ direkt herleitbar. Das liegt bei $(\exists A)$ und $(\exists S)$ an der Verwendung von $(\neg S)$, bei $(\exists AInv)$ an der Verletzung des Subformelprinzips.

Wir stellen weitere naheliegende Gesetze für den All- und den Existenzquantor paarweise auf. Die Gesetze für den Existenzquantor lassen sich wegen der oben besprochenen Symmetrie parallel zu den entsprechenden Gesetzen für den Allquantor beweisen.

5.3.2 Lemma $\vdash \Gamma, \forall x\, \mathcal{F}(x) : \mathcal{F}(t), \Delta$ und $\vdash \Gamma, \mathcal{F}(t) : \exists x\, \mathcal{F}(x), \Delta$.

Beweis. Nach dem Satz von der Identität ist $\vdash \Gamma, \mathcal{F}(t) : \mathcal{F}(t), \Delta$, woraus mit einem $(\forall A)$-Schluss die erste und mit einem $(\exists S)$-Schluss die zweite Behauptung folgt.

Beispiel. Mit $(\rightarrow S)$ erhält man hieraus im Fall $\Gamma = \Delta = \emptyset$

$$\vdash \forall x\, \mathcal{F}(x) \rightarrow \mathcal{F}(t) \text{ und } \vdash \mathcal{F}(t) \rightarrow \exists x\, \mathcal{F}(x).$$

5.3.3 Lemma (Verteilungsregeln)

a. $\vdash \Gamma, \mathcal{F}(a) : G(a), \Delta \Rightarrow \quad \vdash \Gamma, \forall x\, \mathcal{F}(x) : \forall y G(y), \Delta$

b. $\vdash \Gamma, \mathcal{F}(a) : G(a), \Delta \Rightarrow \quad \vdash \Gamma, \exists x\, \mathcal{F}(x) : \exists y G(y), \Delta$

falls a nicht in $\Gamma, \mathcal{F}, G, \Delta$ auftritt.

Beweis. Aus der (gemeinsamen) Voraussetzung folgt mit $(\forall A)$ bzw. $(\exists S)$

$$\vdash \Gamma, \forall x\, \mathcal{F}(x) : G(a), \Delta \text{ und } \vdash \Gamma, \mathcal{F}(a) : \exists y G(y), \Delta$$

und daraus, da nun die Variablenbedingung erfüllt ist, mit $(\forall S)$ bzw. $(\exists A)$ die jeweilige Behauptung.

5.3.4 Korollar (gebundene Umbenennung)

$$\vdash \Gamma, \forall x\, \mathcal{F}(x) : \forall y\, \mathcal{F}(y), \Delta \text{ und } \vdash \Gamma, \exists x\, \mathcal{F}(x) : \exists y\, \mathcal{F}(y), \Delta.$$

Beweis. Nach dem Satz von der Identität ist $\vdash \Gamma, \mathcal{F}(a) : \mathcal{F}(a), \Delta$, wobei a so gewählt werden kann, dass a nicht in $\Gamma, \mathcal{F}, \Delta$ auftritt. Mit den Verteilungsregeln ergeben sich daraus die Behauptungen.

Die Regeln und Schemata dieses Abschnitts gehen bisher mit einfachsten Schlüssen aus dem Satz von der Identität und den Negationsregeln aus 5.2 hervor. Sie sollen eine gewisse Sicherheit im Umgang mit quantorenlogischen Gesetzen vermitteln.
Achten Sie allerdings sorgfältig darauf, dass bei den entsprechenden Schlüssen die Variablenbedingung erfüllt ist. Das sind die Schlüsse nach den Regeln $(\forall S)$, $(\exists A)$ und der Substitutionsregel. Oft erreicht man eine gewünschte Variablenbedingung, indem man erst einige Auftreten einer freien Variablen mit $(\forall A)$- oder $(\exists S)$-Schlüssen bindet, bis sie nur noch in den gewünschten Positionen auftritt. Der Beweis der Verteilungsregeln 5.3.3 illustriert dieses Vorgehen.

Wir schließen mit einigen Bemerkungen zum Allabschluss.

5.3.5 Lemma Für den Allabschluss sind folgende Regeln zulässig:

$$\begin{array}{ll}
(\forall^* S) & \vdash \Gamma : B, \Delta \Rightarrow \ \vdash \Gamma : \forall B, \Delta, \\
 & \quad \text{falls keine freie Variable aus } B \text{ in } \Gamma, \Delta \text{ auftritt.} \\
(\forall^* A) & \vdash \Gamma, B : \Delta \Rightarrow \ \vdash \Gamma, \forall B : \Delta \\
(\forall^* SInv) & \vdash \Gamma : \forall B, \Delta \Rightarrow \ \vdash \Gamma : B, \Delta.
\end{array}$$

Beweis. Es sei $B \equiv \mathcal{F}(a_1, \ldots, a_n)$ und $\forall B \equiv \forall x_1 \ldots \forall x_n\, \mathcal{F}(x_1, \ldots, x_n)$.
Wir induzieren nach n.

1. $n = 0$. Dann ist $B \equiv \forall B$, und die Voraussetzung ist schon jeweils die Behauptung.

2. $n > 0$. Dann ist a_n von $a_1, \ldots, a_{n-1}$ verschieden.

($\forall^* S$) a_n tritt in $\Gamma, \Delta, \mathcal{F}$ nach Voraussetzung nicht auf. Dann ist die Variablenbedingung erfüllt, und mit ($\forall S$) folgt

$$\vdash \Gamma : \forall x_n \mathcal{F}(a_1, \ldots, a_{n-1}, x_n), \Delta.$$

Mit Induktionsvoraussetzung folgt hieraus die Behauptung, weil $\forall B$ auch ein Allabschluss von $\forall x_n \mathcal{F}(a_1, \ldots, a_{n-1}, x_n)$ ist.

($\forall^* A$) Aus der Voraussetzung folgt mit ($\forall A$)

$$\vdash \Gamma, \forall x_n \mathcal{F}(a_1, \ldots, a_{n-1}, x_n) : \Delta,$$

und mit Induktionsvoraussetzung folgt hieraus wie eben die Behauptung.

($\forall^* SInv$) Aus der Voraussetzung folgt mit Induktionsvoraussetzung

$$\vdash \Gamma : \forall x_n \mathcal{F}(a_1, \ldots, a_{n-1}, x_n), \Delta,$$

und ein ($\forall SInv$)-Schluss ergibt die Behauptung.

5.3.6 Korollar Sind $\forall B$ und $\forall' B$ zwei Allabschlüsse von B, so gilt

a. $\vdash B \iff \vdash \forall B \iff \vdash \forall' B$

b. $\vdash \forall B : \forall' B$.

Beweis. a. ergibt sich unmittelbar aus ($\forall^* S$) und ($\forall^* SInv$) im Fall $\Gamma = \Delta = \emptyset$. b. folgt aus $\vdash B : B$ mit ($\forall^* A$) und ($\forall^* S$).

5.4 Das Deduktionstheorem

In der Mathematik beweist man eine Implikation $B \to C$ häufig, indem man B annimmt und dann auf C schließt. Dieser Gedanke liegt dem sogenannten *natürlichen Schließen* zu Grunde. Im Sequenzenkalkül sind die Antezedens-Formeln als Annahmen aufzufassen. Um $B \to C$ zu beweisen, kann man also B ins Antezedens schreiben und die Sequenz $B : C$ herleiten.

Sei nun B ein Satz, und unter Verwendung von B habe man C bewiesen. Für

diesen Beweis macht es kaum einen Unterschied, ob man B angenommen hat und der Beweis von C deshalb von der Annahme des Satzes B abhängt, oder ob B ein Axiom der betrachteten Theorie ist und C in dieser Theorie bewiesen ist. Nur das Ergebnis des Beweises ist ein anderes: Im ersten Fall hat man die Sequenz $B : C$ bewiesen, im zweiten Fall hat man C selbst bewiesen, gestützt auf das Axiom B.

Wir können also in einem inhaltlichen Beweis Axiome in geschlossene Antezedensformeln oder Implikationsvorderglieder umwandeln, und umgekehrt. Diese Möglichkeit besteht auch für unseren Sequenzenkalkül.

5.4.1 Definition Es sei T eine Theorie und Σ eine Menge von Sätzen aus $L(T)$. Dann bezeichnet $T + \Sigma$ die Theorie $(L(T), Ax(T) \cup \Sigma)$.

Bemerkung. Jede Theorie T lässt sich darstellen in der Form $T = T_0 + Ax(T)$ mit $T_0 = (L(T), \emptyset)$, die zu T gehörige logische Theorie. Jede Theorie ist also darstellbar als: logische Theorie + Axiomensystem.

5.4.2 Deduktionstheorem Es sei T eine Theorie und B ein Satz aus $L(T)$. Dann gilt:

$$T + \{B\} \vdash \Gamma : \Delta \iff T \vdash B, \Gamma : \Delta.$$

Beweis von $\Leftarrow$. Gegeben sei eine Herleitung von $B, \Gamma : \Delta$ in T. Diese Herleitung ist dann auch eine Herleitung in $T + \{B\}$. Also haben wir

$$T + \{B\} \vdash B, \Gamma : \Delta.$$

Hieraus schließen wir mit einem $T + \{B\}$-Schluss auf $T + \{B\} \vdash \Gamma : \Delta$.

Beweis von $\Rightarrow$. Wir zeigen durch Induktion nach n:

$$T + \{B\} \stackrel{n}{\vdash} \Gamma : \Delta \Rightarrow T \stackrel{n}{\vdash} B, \Gamma : \Delta.$$

1. $n = 0$. Wenn $\Gamma : \Delta$ ein logisches Axiom ist, ist auch $B, \Gamma : \Delta$ ein logisches Axiom.

2. $T + \{B\} \stackrel{n+1}{\vdash} \Gamma : \Delta$, und der letzte Schluss ist kein $T + \{B\}$-Schluss, der gerade das Axiom B aus dem Antezedens herausschneidet. Hat dieser Schluss die Prämissen $T + \{B\} \stackrel{n}{\vdash} \Gamma_i : \Delta_i$ ($i = 1, 2$ oder $i = 1$), so ist nach Induktionsvoraussetzung

$$T \stackrel{n}{\vdash} B, \Gamma_i : \Delta_i \qquad (i = 1, 2 \text{ oder } i = 1).$$

Mit einem Schluss nach derselben Regel folgt dann

$$T \overset{n}{\vdash} B, \Gamma : \Delta,$$

denn im Fall eines $(\forall S)$-Schlusses verletzt das geschlossene B keine Variablenbedingung, und im Fall eines $T + \{B\}$-Schlusses liegt nach unserer Voraussetzung ein T-Schluss vor.

3. $T + \{B\} \overset{n+1}{\vdash\!\!\!\!-} \Gamma : \Delta$, und der letzte Schluss ist ein $T + \{B\}$-Schluss mit der Prämisse $T + \{B\} \overset{n}{\vdash} B, \Gamma : \Delta$. Nach Induktionsvoraussetzung ist dann

$$T \overset{n}{\vdash} B, B, \Gamma : \Delta \equiv B, \Gamma : \Delta,$$

und das ist bereits die Behauptung.

Mit Induktion nach n folgt die Richtung $\Rightarrow$. Damit ist das Deduktionstheorem vollständig bewiesen.

Herleitungen sind endliche baumartige Figuren. Jede einzelne Herleitung besteht aus endlich vielen Sequenzen, sie ist aus endlich vielen logischen Axiomen mit endlich vielen Grundschlüssen gebildet. Insbesondere wird in einer Herleitung die T-Grundschlussregel nur auf endlich viele Axiome von T angewandt. Diesen einfachen, aber grundlegenden Gedanken kann man wie folgt formulieren:

5.4.3 Endlichkeitslemma Es sei T eine Theorie und Σ eine (eventuell unendliche) Menge von Sätzen aus $L(T)$.

$$T + \Sigma \vdash \Gamma : \Delta$$

ist gleichwertig mit: Es gibt endlich viele Sätze $B_1, \ldots, B_k$ aus Σ, so dass

$$T + \{B_1, \ldots, B_k\} \vdash \Gamma : \Delta.$$

Beweis. Wir betrachten eine Herleitung H von $\Gamma : \Delta$ in $T + \Sigma$. Weil H endlich ist, werden in H auch nur endlich viele Sätze $B_1, \ldots, B_k$ aus Σ durch $T + \Sigma$-Schlüsse abgetrennt. Alle in H auftretenden $T + \Sigma$-Schlüsse sind daher auch $T + \{B_1, \ldots, B_k\}$-Schlüsse. Weil die Sprache von $T + \Sigma$ ohnehin gleich $L(T)$ ist, ist damit H auch eine Herleitung von $\Gamma : \Delta$ in $T + \{B_1, \ldots, B_k\}$.
Umgekehrt ist jede Herleitung in $T + \{B_1, \ldots, B_k\}$ eine Herleitung in $T + \Sigma$,

wenn die Sätze $B_1, \ldots, B_k \in \Sigma$ sind.

Man erhält mit 5.4.3 nicht nur eine Aussage über die *Herleitbarkeit* in $T + \Sigma$, sondern über die *Herleitungen* in $T + \Sigma$ selbst: Jede Herleitung H in $T + \Sigma$ ist zugleich eine Herleitung in einer Theorie $T + \{B_1, \ldots, B_k\}$ (und umgekehrt), wobei die Wahl der $B_i \in \Sigma$ von der gegebenen Herleitung H bestimmt wird: Verschiedene Herleitungen sogar derselben Sequenz $\Gamma : \Delta$ können verschiedene $T + \Sigma$-Schlüsse und dabei verschiedene Sätze aus Σ verwenden.
Damit lässt sich das Deduktionstheorem auf beliebige Mengen Σ von Zusatzaxiomen erweitern.

5.4.4 Verallgemeinertes Deduktionstheorem Es sei T eine Theorie und Σ eine Menge von Sätzen aus $L(T)$.

$$T + \Sigma \vdash \Gamma : \Delta$$

ist gleichwertig mit: Es gibt endlich viele Sätze $B_1, \ldots, B_k$ aus Σ, so dass

$$T \vdash B_1, \ldots, B_k, \Gamma : \Delta.$$

Beweis. $T + \Sigma \vdash \Gamma : \Delta$ ist nach dem Endlichkeitslemma 5.4.3 gleichwertig mit

$$T + \{B_1, \ldots, B_k\} \vdash \Gamma : \Delta$$

für geeignete Sätze $B_1, \ldots, B_k$ aus Σ. Durch k-fache Anwendung des Deduktionstheorems 5.4.2 erweist sich dies als äquivalent zu

$$T \vdash B_1, \ldots, B_k, \Gamma : \Delta.$$

Hiermit wird auch die Herleitbarkeit in beliebigen Theorien auf die Herleitbarkeit in logischen Theorien zurückgeführt, weil sich jede Theorie als logische Theorie + Axiomensystem schreiben lässt.

5.4.5 Korollar Sei T eine Theorie. $T \vdash \Gamma : \Delta$ ist gleichwertig mit: Es gibt endlich viele Axiome $B_1, \ldots, B_k$ von T, so dass

$$(L(T), \emptyset) \vdash B_1, \ldots, B_k, \Gamma : \Delta.$$

Dies ist der Spezialfall von 5.4.4, in dem die dortige Theorie T die logische Theorie $(L(T), \emptyset)$ ist und Σ das gesamte Axiomensystem $Ax(T)$ bezeichnet.

5.5 Aufgaben

5.5.1 Beweisen sie die Konjunktionsregeln aus 5.2.6

5.5.2 Beweisen Sie Korollar 5.2.7 und 5.2.8.

5.5.3 a. Zeigen Sie: $\vdash \forall x(\mathcal{F}(x) \to G(x)) : \forall x\, \mathcal{F}(x) \to \forall x G(x)$

b. Geben Sie ein Beispiel, für das $\forall x\, \mathcal{F}(x) \to \forall x G(x) : \forall x(\mathcal{F}(x) \to G(x))$ logisch nicht herleitbar ist.

5.5.4 Folgern Sie aus $(\forall SInv)$

$$\vdash \Gamma : \forall x\, \mathcal{F}(x), \Delta \Rightarrow \quad \vdash \Gamma : \forall y\, \mathcal{F}(y), \Delta.$$

Warum gibt es keinen „analogen“ Beweis für

$$\vdash \Gamma, \forall x\, \mathcal{F}(x) : \Delta \Rightarrow \quad \vdash \Gamma, \forall y\, \mathcal{F}(y) : \Delta?$$

5.5.5 Zeigen Sie durch Induktion nach n:

a. $\vdash^n \Gamma : \forall x\, \mathcal{F}(x), \Delta \Rightarrow \vdash^n \Gamma : \forall y\, \mathcal{F}(y), \Delta$

b. $\vdash^n \Gamma, \forall x\, \mathcal{F}(x) : \Delta \Rightarrow \vdash^n \Gamma, \forall y\, \mathcal{F}(y) : \Delta$

5.5.6 Seien $\forall B, \forall' B$ zwei Allabschlüsse einer Formel B. Folgern Sie aus 5.5.5b.:

$$\vdash^n \Gamma, \forall B : \Delta \Rightarrow \vdash^n \Gamma, \forall' B : \Delta.$$

5.5.7 Gegeben sei eine Theorie T und eine Menge Σ von Formeln aus $L(T)$. $\Sigma^\forall$ sei die Menge aller Allabschlüsse von Formeln aus Σ. Ferner sei $\forall\Sigma$ eine Menge von Sätzen, die zu jeder Formel $B \in \Sigma$ genau einen Allabschluss $\forall B \in \forall\Sigma$ (und sonst keine Formeln) enthält. Zeigen Sie:

a. Jede Herleitung in $T + \forall\Sigma$ ist eine Herleitung in $T + \Sigma^\forall$.

b. $T + \Sigma^\forall \vdash^n \Gamma : \Delta \Rightarrow T + \forall\Sigma \vdash^n \Gamma : \Delta$.

c. Geben Sie ein Σ an, für das nicht jede Herleitung in $T + \Sigma^\forall$ auch eine Herleitung in $T + \forall\Sigma$ ist.

5.5.8 Beweisen Sie die Zulässigkeit der Schnittregel in der Form

$$T \overset{k}{\vdash} \Gamma, C : \Delta \text{ und } T \overset{l}{\vdash} \Gamma : C, \Delta \Rightarrow T \vdash \Gamma : \Delta$$

durch Induktion nach dem Aufbau der Formel C, also nach der Anzahl der Auftreten logischer Partikel in C.
Hinweis: Für Primformeln C beweisen Sie die Behauptung, also den Induktionsanfang durch (Neben-)Induktion nach l. In dem Fall, daß C eine Allformel ist, beweisen Sie den Induktionsschritt durch (Neben-)Induktion nach k, der Ordnung der Herleitung von $\Gamma, C : \Delta$. (Für den Fall von Implikationen C ist keine Nebeninduktion nötig.) Verwenden Sie reichlich Struktur- und eventuell auch Inversionsschlüsse.

5.5.9 Folgern Sie aus der Zulässigkeit der Schnittregel 5.5.8 die Zulässigkeit des *modus ponens* in den beiden Formen:

a. $T \vdash A : B$ und $T \vdash A \Rightarrow T \vdash B$.

b. $T \vdash A \to B$ und $T \vdash A \Rightarrow T \vdash B$.

§6 Gleichheit und Äquivalenz

6.1 Gleichheit

6.2 Logische Äquivalenz

6.3 Exkurs: Eine HILBERTartige Formalisierung der Prädikatenlogik

6.4 Aufgaben

6.1 Gleichheit

In diesem Abschnitt leiten wir Gesetze der Gleichheit unter wesentlicher Verwendung der Gleichheitsregeln $(= I)$, $(= F)$ und $(= P)$ her. Bisher haben wir diese Grundschlussregeln allenfalls nebenher betrachtet und bei Induktionsbeweisen, besonders in § 5.1, nur pauschal verarbeitet. Als erstes betrachten wir den Spezialfall von $(= P)$, in dem das Prädikatszeichen p das Gleichheitszeichen $=$ ist. Es handelt sich also um die Schlussregel

$(*)$ $\Gamma, t_1 = t_2 : \Delta \vdash \Gamma, s_1 = t_1, s_2 = t_2, s_1 = s_2 : \Delta.$

Zur Anwendung dieser Regel hat man zu einer gegebenen Gleichung $t_1 = t_2$ zwei Terme $s_1,\ s_2$ zu wählen und dann aus dem „Gleichungsquadrat“

$$\begin{array}{ccc} s_1 & = & s_2 \\ \| & & \| \\ t_1 & = & t_2 \end{array}$$

die „untere Kante“ $t_1 = t_2$ im Antezedens einer Sequenz durch die drei übrigen „Kanten“ zu ersetzen.
Die Regel $(*)$ kann dadurch degenerieren, dass einige der auftretenden Terme, z. B. s_1 und s_2 oder s_1 und t_1, als Zeichenreihen identisch sind. Folgende Schlussregeln entstehen auf diese Weise.

6.1.1 Lemma Folgende Regeln sind (in allen Theorien) zulässig:
Komparativitäts-Regel

$$\vdash \Gamma, t_1 = t_2 : \Delta \Rightarrow \vdash \Gamma, s = t_1, s = t_2 : \Delta$$

Transitivitäts-Regel

$$\vdash \Gamma, r = t : \Delta \Rightarrow \vdash \Gamma, r = s, s = t : \Delta$$

Symmetrie-Regel

$$\vdash \Gamma, s = t : \Delta \Rightarrow \vdash \Gamma, t = s : \Delta$$

Beweis. Für die Komparativitäts-Regel setzen wir in $(*)$ $s_1 \equiv s_2 \equiv s$ und erhalten

$$\vdash \Gamma, t_1 = t_2 : \Delta \Rightarrow \vdash \Gamma, s = t_1, s = t_2, s = s : \Delta,$$

und ein $(= I)$-Schluss liefert die Behauptung.
Für die Transitivitäts-Regel setzen wir in $(*)$ $s_1 \equiv t_1 \equiv r$, $s_2 \equiv s$ und $t_2 \equiv t$ und erhalten

$$\vdash \Gamma, r = t : \Delta \Rightarrow \vdash \Gamma, r = r, s = t, r = s : \Delta,$$

und wieder liefert ein $(= I)$-Schluss die Behauptung.
Für die Symmetrie-Regel identifizieren wir in $(*)$ sogar drei Terme, setzen $t_1 \equiv s$ und $s_1 \equiv s_2 \equiv t_2 \equiv t$ und erhalten mit $(*)$

$$\vdash \Gamma, s = t : \Delta \Rightarrow \vdash \Gamma, t = s, t = t, t = t : \Delta,$$

und auch hier liefert ein $(= I)$-Schluss die Behauptung.

Bemerkung. Die Symmetrie-Regel folgt mit $(= I)$ auch allein aus der Komparativitäts-Regel, wenn wir dort $s \equiv t_2$ setzen. Diese Überlegung enthält einen Beweis der Tatsache, dass jede reflexive rechts-komparative Relation auch symmetrisch ist.

6.1.2 Lemma Folgende Gleichheitsgesetze sind herleitbar:

1. Reflexivität		$: t = t$
2. Komparativität	$s = t_1, s = t_2$	$: t_1 = t_2$
3. Transitivität	$r = s, s = t$	$: r = t$
4. Symmetrie	$t = s$	$: s = t$

Beweis. Sequenzen $t_1 = t_2 : t_1 = t_2$ sind logische Axiome. Wendet man darauf (bei passender Wahl von t_1, t_2) die Identitäts-Regel $(= I)$ bzw. die Regeln aus 6.1.1 an, so erhält man die entsprechenden Sequenzen 1 bis 4.

Selbstverständlich kann man die Sequenzen 2, 3 und 4 aus 6.1.2 unmittelbar

herleiten, ohne zulässige Regeln zu verwenden. Dies sei zur Übung empfohlen. Man wird feststellen, dass man nur die entsprechenden Spezialfälle der Regeln aus 6.1.1 nachrechnet, um diese Herleitungen zu gewinnen.

6.1.3 Lemma Folgende Gleichheitsgesetze sind herleitbar:

1. $s_1 = t_1, \ldots, s_n = t_n : fs_1 \ldots s_n = ft_1 \ldots t_n$
2. $s_1 = t_1, \ldots, s_n = t_n, ps_1 \ldots s_n : pt_1 \ldots t_n$

Beweis. Für $n = 0$ ist 1. ein Fall der Reflexivität 6.1.2, 1 und 2. ein logisches Axiom. Sei $n > 0$. Die Sequenzen

1'. $fs_1 \ldots s_n = ft_1 \ldots t_n : fs_1 \ldots s_n = ft_1 \ldots t_n$ und

2'. $pt_1 \ldots t_n : pt_1 \ldots t_n$

sind logische Axiome. $(= F)$, angewandt auf 1', liefert 1. $(= P)$, angewandt auf 2', liefert 2.

6.1.4 Satz (Gleichheitsregel)

$$\vdash \Gamma, t(r) = t(s) : \Delta \Rightarrow \vdash \Gamma, r = s : \Delta.$$

Beweis durch Induktion nach dem Aufbau des Terms $t(a)$.

1. $t(a)$ ist a, t ist $*_1$. Dann stimmen Voraussetzung und Behauptung überein.
2. $t(a) \equiv t$ ist eine andere Variable oder eine Konstante c. Dann ist die Voraussetzung $\vdash \Gamma, c = c : \Delta$. Ein $(= I)$-Schluss ergibt $\vdash \Gamma : \Delta$, woraus mit einem Strukturschluss die Behauptung folgt.
3. $t(a)$ beginnt mit einem n-stelligen Funktionszeichen $(n > 0)$, t ist $ft_1 \ldots t_n$. Dann folgt aus der Voraussetzung mit $(= F)$

   $$\vdash \Gamma, t_1(r) = t_1(s), \ldots, t_n(r) = t_n(s) : \Delta.$$

   Wenn man hierauf n-mal die Induktionsvoraussetzung anwendet, erhält man die Behauptung.

Mit Induktion folgt nun der Satz.

6.1.5 Korollar

$$\vdash r = s : t(r) = t(s).$$

Dies folgt mit der Gleichheitsregel aus dem logischen Axiom
$t(r) = t(s) : t(r) = t(s)$.

Dieses Korollar ist eine Verallgemeinerung von 6.1.3, 1 (für 1-stelliges Funktionszeichen f), nämlich von $\vdash r = s : fr = fs$. Während hier die Terme r, s unmittelbar als Argumente des Funktionszeichens f auftreten, können sie in 6.1.5 verschachtelt innerhalb der Argumente $t_i(r), t_i(s)$ auftreten.
Eine entsprechende Verallgemeinerung von 6.1.3,2 ist der folgende Gleichheitssatz.

6.1.6 Gleichheitssatz Folgendes Ersetzungsschema ist herleitbar:

(1) $\quad F(r), r = s : F(s)$

Beweis durch Induktion nach dem Aufbau der Formel $F(a)$.

1. $F(a)$ ist eine Primformel $\not\equiv \perp$, F ist $pt_1 \dots t_n$. Dann ist

   $$\vdash F(r), t_1(r) = t_1(s), \dots, t_n(r) = t_n(s) : F(s)$$

   nach Lemma 6.1.3,2. Wenn man hierauf n-mal die Gleichheitsregel 6.1.4 anwendet, erhält man die Behauptung (1).

2. $F(a)$ ist $\perp$. Dann ist (1) ein logisches Axiom.

3. $F(a)$ ist eine Implikation, F ist $F_1 \to F_2$.
   Nach Induktionsvoraussetzung ist

   $$\vdash F_1(s), s = r : F_1(r),$$

   woraus mit der Symmetrie-Regel aus 6.1.1 und einem Strukturschluss

   $$\vdash F_1(s), r = s : F_2(s), F_1(r)$$

   folgt. Ein $(\to S)$-Schluss ergibt

(2) $\vdash r = s : F(s), F_1(r)$
Ferner ist nach Induktionsvoraussetzung

$$\vdash F_2(r), r = s : F_2(s),$$

woraus mit einem Strukturschluss

$$\vdash F_2(r), r = s, F_1(s) : F_2(s)$$

und weiter mit einem $(\rightarrow S)$-Schluss folgt

(3) $\vdash F_2(r), r = s : F(s)$
Ein $(\rightarrow A)$-Schluss, angewandt auf (2) und (3), ergibt nun

$$\vdash F_1(r) \rightarrow F_2(r), r = s : F(s),$$

und das ist die Behauptung (1).

4. $F(a)$ ist eine Allformel, F ist $\forall x G(x, *_1)$. Die freie Variable a trete nicht in r, s, G auf. Nach Induktionsvoraussetzung ist dann

$$\vdash G(a, r), r = s : G(a, s),$$

woraus mit der Verteilungsregel 5.3.3 a

$$\vdash \forall x G(x, r), r = s : \forall x G(x, s)$$

folgt, und das ist die Behauptung (1).

Mit Induktion folgt nun der Satz.

6.1.7 Korollar $\vdash r = s : F(r) \leftrightarrow F(s)$.

Beweis. Der Gleichheitssatz ergibt mit der Symmetrie-Regel aus 6.1.1

$$\vdash r = s, F(s) : F(r).$$

$(\rightarrow S)$-Schlüsse, angewandt auf den Gleichheitssatz bzw. auf diese Sequenz, ergeben

$$\vdash r = s : F(r) \rightarrow F(s) \qquad \text{und}$$
$$\vdash r = s : F(s) \rightarrow F(r).$$

Hieraus ergibt sich mit einem $(\wedge S)$-Schluss die Behauptung.

6.2 Logische Äquivalenz

Welche Gesetze der Gleichheit lassen sich auf die logische Äquivalenz übertragen? Es sind überraschend viele, wie sich hier zeigen soll. Ein gutes Beispiel ist der Gleichheitssatz in der Form

$$\vdash r = s : F(r) \leftrightarrow F(s),$$

dessen Analogon

$$\vdash A \leftrightarrow B : F(A) \leftrightarrow F(B)$$

eine Fassung des Äquivalenzsatzes ist, der im Mittelpunkt dieses Abschnitts steht. Offenbar hat der Äquivalenzsatz mit den Grundschlussregeln für die Gleichheit nichts zu tun – im Gegensatz zum Gleichheitssatz, der die Regeln $(= F)$ und $(= P)$ gewissermaßen zusammenfasst. Trotzdem sind die Beweise beider Sätze streckenweise analog.
Die Äquivalenz $A \leftrightarrow B$ ist definiert als

$$(A \to B) \wedge (B \to A),$$

also als $\neg((A \to B) \to \neg(B \to A))$, ein relativ komplizierter Ausdruck. Aussagenlogisch ist die Äquivalenz aber leicht zu behandeln, weil für jede Wertung V gilt:

$$(*) \qquad V(A \leftrightarrow B) = w \iff V(A) = V(B).$$

Damit übertragen sich (semantische) Eigenschaften der Gleichheit auf die logische Äquivalenz. Als erstes erhalten wir ein Analogon von 6.1.2.

6.2.1 Lemma Folgene Sequenzen sind Tautologien und somit herleitbar:

1. Reflexivität	$A \leftrightarrow A$
2. Komparativität	$(A \leftrightarrow B),\ (A \leftrightarrow C) : (B \leftrightarrow C)$
3. Transitivität	$(A \leftrightarrow B),\ (B \leftrightarrow C) : (A \leftrightarrow C)$
4. Symmetrie	$(A \leftrightarrow B) : (B \leftrightarrow A)$

Beweis. Sei V eine Wertung. Wir haben zu zeigen, dass die vier Sequenzen unter V den Wert w erhalten. Dazu verwenden wir die Äquivalenz $(*)$:

1. Es ist $V(A) = V(A)$ wegen der Reflexivität der (inhaltlichen) Gleichheit.

2. Wenn $V(A) = V(B)$ und $V(A) = V(C)$ ist, ist auch $V(B) = V(C)$ wegen der Komparativität der Gleichheit.

3. Wenn $V(A) = V(B)$ und $V(B) = V(C)$ ist, ist auch $V(A) = V(C)$ wegen der Transitivität der Gleichheit.

4. Wenn $V(A) = V(B)$ ist, ist auch $V(B) = V(A)$ wegen der Symmetrie der Gleichheit.

Die Sequenzen 1. bis 4. erhalten also bei jeder Wertung den Wert w, sind somit Tautologien und daher nach 4.2.16 auch herleitbar.

Inhaltlich selbstverständliche Eigenschaften der Gleichheit von Wahrheitswerten führen hier zu entsprechenden Herleitungseigenschaften für logische Äquivalenzen. Der Nachweis stützt sich natürlich nicht auf 6.1.2, was eine Herleitbarkeitsaussage ist, wogegen wir in 6.2.1 die inhaltliche Gleichheit (Übereinstimmung) von Wahrheitswerten ausnutzen.
Der Rückgriff auf Tautologien ist nicht mehr möglich, wenn wir statt der Sequenzen aus 6.2.1 die entsprechenden Schlussregeln betrachten.

6.2.2 Lemma (Symmetrie-Regel) $\vdash A \leftrightarrow B \Rightarrow \vdash B \leftrightarrow A$

Beweis. Aus der Voraussetzung folgt mit $(\wedge SInv)$

$$\vdash A \rightarrow B \text{ und } \vdash B \rightarrow A,$$

woraus sich umgekehrt mit $(\wedge S)$ die Behauptung ergibt.

Bemerkung. Komparativitäts- und Transitivitäts-Regel

$$\vdash A \leftrightarrow B, \vdash A \leftrightarrow C \Rightarrow \vdash B \leftrightarrow C \text{ und}$$
$$\vdash A \leftrightarrow B, \vdash B \leftrightarrow C \Rightarrow \vdash A \leftrightarrow C$$

ergeben sich aus der Schnittregel, die wir erst später behandeln (vgl. auch Aufgabe 5.5.8).
Wir untersuchen das Analogon zum Gleichheitssatz 6.1.7

6.2.3 Äquivalenzsatz Folgende Schlussregel ist zulässig:

$$\vdash \Gamma : G_1(a_1, \ldots, a_n) \leftrightarrow G_2(a_1, \ldots, a_n)$$
$$\Rightarrow \vdash \Gamma : F(G_1(x_1, \ldots, x_n)) \leftrightarrow F(G_2(x_1, \ldots, x_n)),$$

wobei die $x_1, \ldots, x_n$ Variablen sind, die erst in F gebunden werden, und die $a_1, \ldots, a_n$ nicht in Γ, G_1, G_2, F auftreten.

Beweis. Wir schreiben a für $a_1, \ldots, a_n$ und x für $x_1, \ldots, x_n$. Aus der Voraussetzung folgt mit $(\wedge SInv)$ und $(\rightarrow SInv)$

$$\vdash \Gamma, G_1(a) : G_2(a) \text{ und } \vdash \Gamma, G_2(a) : G_1(a).$$

Wir beweisen durch Induktion nach dem Aufbau von $F(\bot)$:

$$\vdash \Gamma, F(G_1(x)) : F(G_2(x)) \text{ und } \vdash \Gamma, F(G_2(x)) : F(G_1(x)).$$

1. F ist $*_1$. Dann stimmen die Voraussetzungen mit den Behauptungen überein.

2. F ist eine Primformel, $*_1$ tritt in F nicht auf. Dann sind die behaupteten Sequenzen $\Gamma, F : F$ logische Axiome.

3. F ist $F_1 \rightarrow F_2$. Wir schreiben

   $$A_{ik} \text{ für } F_i(G_k(x)) \quad (i, k = 1, 2)$$

   Dann lauten unsere Induktionsvoraussetzungen

   (1) $\vdash \Gamma, A_{11} : A_{12}$ und (2) $\vdash \Gamma, A_{12} : A_{11}$
   (3) $\vdash \Gamma, A_{21} : A_{22}$ und (4) $\vdash \Gamma, A_{22} : A_{21}$

   Die Behauptungen folgen jetzt im wesentlichen mit aussagenlogischen Schlüssen. Mit Strukturschlüssen erhalten wir aus (2) und (3)

   (2') $\vdash \Gamma, A_{12} : A_{22}, A_{11}$ und (3') $\vdash \Gamma, A_{21}, A_{12} : A_{22}$.

   Ein $(\rightarrow A)$-Schluss ergibt hieraus

   $$\vdash \Gamma, A_{11} \rightarrow A_{21}, A_{12} : A_{22},$$

   woraus wir mit einem $(\rightarrow S)$-Schluss zu

   $$\vdash \Gamma, A_{11} \rightarrow A_{21} : A_{12} \rightarrow A_{22}$$

   kommen, und das ist $\vdash \Gamma, F(G_1(x)) : F(G_2(x))$.
   Ebenso erhalten wir aus (1) und (4)

   (1') $\vdash \Gamma, A_{11} : A_{21}, A_{12}$ und (4') $\vdash \Gamma, A_{22}, A_{11} : A_{21}$.

Wieder ergibt ein $(\to A)$-Schluss hieraus

$$\vdash \Gamma, A_{12} \to A_{22}, A_{11} : A_{21},$$

woraus mit einem $(\to S)$-Schluss

$$\vdash \Gamma, A_{12} \to A_{22} : A_{11} \to A_{21}$$

folgt, und das ist $\vdash \Gamma, F(G_2(x)) : F(G_1(x))$.

4. F ist $\forall x_n F'(x_n, *_1)$. a_n trete nicht in Γ, G_1, G_2, F' auf. Nach Induktionsvoraussetzung ist dann

   $$\vdash \Gamma, F'(a_n, G_1(x_1, \ldots, x_{n-1}, a_n)) : F'(a_n, G_2(x_1, \ldots, x_{n-1}, a_n)).$$

   Hieraus folgt mit der Verteilungsregel 5.3.3 – die Variablenbedingung für a_n ist erfüllt –

   $$\vdash \Gamma, \forall x_n F'(x_n, G_1(x)) : \forall x_n F'(x_n, G_2(x)).$$

   Das ist die erste Behauptung. Man erhält die zweite Behauptung, wenn man in diesem Argument G_1 und G_2 vertauscht.

Induktion nach dem Aufbau der Formel $F(\bot)$ ergibt nun die Induktionsbehauptungen. Aus ihnen schließt man mit $(\to S)$ und $(\wedge S)$ auf

$$\vdash \Gamma : F(G_1(x)) \leftrightarrow F(G_2(x)).$$

Damit ist der Satz bewiesen.

In der (allgemein üblichen) Formulierung 6.1.6 des Gleichheitssatzes ist nicht vorgesehen, dass freie Variablen, die in den Termen r, s auftreten, in den Formeln $F(r)$, $F(s)$ gebunden werden. Dem entspricht offenbar der Fall $n = 0$ des Äquivalenzsatzes, in dem eine Umwandlung von freien in gebundene Variablen *innerhalb der* G_i nicht vorkommt.

6.2.4 Korollar $\vdash A \leftrightarrow B : F(A) \leftrightarrow F(B)$.

Beweis. Im Fall $n = 0$ sind die Nennformen G_1, G_2 aus dem Äquivalenzsatz Formeln A, B. Dann ist $A \leftrightarrow B : A \leftrightarrow B$ nach dem Satz 4.1.2 von der Identität herleitbar, und der Äquivalenzsatz liefert die Behauptung.
Auch die allgemeine Fassung von 6.2.3 lässt sich durch die Herleitbarkeit eines Sequenzenschemas ausdrücken.

6.2.5 Korollar Mit x als Abkürzung für $x_1, \ldots, x_n$ ist

$$\vdash \forall x_1 \ldots \forall x_n (G_1(x) \leftrightarrow G_2(x)) : F(G_1(x)) \leftrightarrow F(G_2(x)).$$

Beweis. Der Fall $n = 0$ ist das vorige Korollar. Im Fall $n > 0$ sei auch wieder a für $a_1, \ldots, a_n$ geschrieben, wobei die $a_i (i = 1, \ldots, n)$ *neue* freie Variablen seien, d. h. die a_i treten in F, G_1, G_2 nicht auf. Nach 4.1.2 ist $\vdash G_1(a) \leftrightarrow G_2(a) : G_1(a) \leftrightarrow G_2(a)$. Hieraus erhält man mit n $(\forall A)$-Schlüssen

$$\vdash \forall x_1 \ldots \forall x_n (G_1(x) \leftrightarrow G_2(x)) : G_1(a) \leftrightarrow G_2(a),$$

und mit dem Äquivalenzsatz folgt die Behauptung.

Dieses Korollar enthält als engen, trotzdem typischen Spezialfall folgendes Verteilungsgesetz für $\leftrightarrow$:

$$\vdash \forall x(G_1(x) \leftrightarrow G_2(x)) : \forall x G_1(x) \leftrightarrow \forall x G_2(x).$$

Der Äquivalenzsatz 6.2.3 ist nicht so allgemein formuliert, wie es möglich wäre. Allgemeinere Formulierungen werden durch unsere Unterscheidung von freien und gebundenen Variablen allerdings umständlich.

6.2.6 Beispiel Es sei $\vdash \Gamma : G_1(a) \leftrightarrow G_2(a)$ und a nicht in Γ, G_1, G_2. Dann ist nach 6.2.3

$$\vdash \Gamma : (\exists x G_1(x) \to \forall x(px \to G_1(x))) \leftrightarrow (\exists x G_2(x) \to \forall x(px \to G_2(x))).$$

Ebenfalls ist

$$\vdash \Gamma : (\exists x G_1(x) \to B \to G_1(a)) \leftrightarrow (\exists x G_2(x) \to B \to G_2(a)),$$

falls a nicht in B auftritt; aber dies folgt nicht unmittelbar aus 6.2.3. Denn hier haben wir einmal $G_1(x)$ durch $G_2(x)$ und einmal $G_1(a)$ durch $G_2(a)$ ersetzt, und das wird von der Formulierung 6.2.3 nicht gedeckt.

6.3 Exkurs: Eine Hilbert-artige Formalisierung der Prädikatenlogik

Zum Abschluss dieses Kapitels sind wir mit dem Sequenzen-Kalkül aus § 3 soweit vertraut, dass wir einen Blick auf alternative Herleitungsbegriffe werfen

können. Auf diesen Abschnitt greifen wir später nicht zurück; er erleichtert aber den Übergang zu Darstellungen der Prädikatenlogik, die nicht auf einem Sequenzenkalkül aufbauen.
Eine klassische Alternative zum Sequenzenkalkül sind Hilbert-artige Formalisierungen des Herleitungsbegriffs, wie sie von David Hilbert eingeführt und auch im ältesten Lehrbuch der Prädikatenlogik, den „*Grundzügen der theoretischen Logik*“ von D. Hilbert und W. Ackermann 1928 zu Grunde gelegt wurden. In diesen Kalkülen leitet man unmittelbar Formeln her, Sequenzen treten gar nicht auf. Sie sind nicht schnittfrei, sondern verwenden den modus ponens als Grundschlussregel. Ferner bevorzugen sie Axiomschemata gegenüber Grundschlussregeln; in unserer Fassung stehen nur zwei Grundschlussregeln (*modus ponens* und Allregel) sechs Axiomschemata gegenüber.
Wir präsentieren eine möglichst kompakte Hilbert-artige Formalisierung der klassischen Prädikatenlogik, die wir der Kürze wegen als *Hilbert-Kalkül* bezeichnen.

6.3.1 Axiome und Regeln des Hilbert-Kalküls

1. Aussagenlogische Axiome sind alle Tautologien.

2. Quantorenlogische Axiome:

$$(\forall Ax) \qquad \forall x\, \mathcal{F}(x) \to \mathcal{F}(t)$$

3. Gleichheitsaxiome:

$$\begin{array}{ll} (=IAx) & t = t \\ (=FAx) & s_1 = t_1 \to \ldots \to s_n = t_n \to f s_1 \ldots s_n = f t_1 \ldots t_n \\ (=PAx) & s_1 = t_1 \to \ldots \to s_n = t_n \to p s_1 \ldots s_n \to p t_1 \ldots t_n \end{array}$$

4. Theorie-Axiome einer Theorie T sind alle $B \in Ax(T)$.

5. Grundschlussregeln:

$$\begin{array}{ll} (mp) & A \to B, A \vdash B \\ (\forall R) & B \to \mathcal{F}(a) \vdash B \to \forall x\, \mathcal{F}(x), \text{ falls } a \text{ nicht in } B, \mathcal{F} \text{ auftritt.} \end{array}$$

Herleitungen im Hilbert-Kalkül sind analog zu den Herleitungen im Sequenzenkalkül induktiv definiert. Wir können uns deshalb mit einer beschreibenden Definition begnügen.

6.3.2 Definition Eine *Herleitung* einer Formel C in einer Theorie T im Hilbert-Kalkül ist ein endlicher Formelbaum wie folgt:

1. An den Spitzen des Baumes stehen Axiome nach 6.3.1,1 bis 4.
2. Unter den (ein oder zwei) Prämissen eines Grundschlusses (mp) oder $(\forall R)$ steht die Konklusion dieses Grundschlusses.
3. An der Wurzel des Baumes steht die Formel C.

C heisst *herleitbar* in T im Hilbert-Kalkül, $T \vdash_H C$, wenn es eine Herleitung von C in T im Hilbert-Kalkül gibt.

Bemerkung. Eine triviale, häufig benutzte Eigenschaft des Hilbert-Kalküls ist:
Ist $C \to D$ eine Tautologie und $T \vdash_H C$, so ist $T \vdash_H D$.
Denn es ist $T \vdash_H C \to D$, weil Tautologien Axiome sind, und die Grundschlussregel (mp) ergibt die Behauptung.

Um die Herleitbarkeit im Sequenzen- und im Hilbert-Kalkül zu vergleichen, verwenden wir in diesem Abschnitt das Ergebnis aus Aufgabe 5.5.9, das sich auch in § 8 „nebenbei“ ergibt und in Kapitel 5 in schärferer Form bewiesen wird:

> $(*)$ *Der modus ponens (mp) ist zulässig im Sequenzenkalkül.*

Dann ergibt sich leicht aus den Ergebnissen dieses Kapitels, dass im Sequenzenkalkül mindestens soviel wie im Hilbert-Kalkül herleitbar ist.

6.3.3 Satz Unter Verwendung von $(*)$ gilt:

$$T \vdash_H C \Rightarrow T \vdash C.$$

Beweis durch Herleitungsinduktion im Hilbert-Kalkül.

1. Tautologien sind nach dem Korollar 4.2.16 zum Tautologiesatz herleitbar.
2. Formeln $(\forall Ax)$ sind nach 5.3.2 herleitbar.
3. Gleichheitsaxiome $(= IAx)$ sind nach 6.1.2,1 und $(= FAx)$ und $(= PAx)$ nach 6.1.3 herleitbar, ggf. mit einigen $(\to S)$-Schlüssen.

4. Theorie-Axiome sind aus dem Satz von der Identität 4.1.2 mit einem T-Schluss herleitbar.

5. Der *modus ponens* ist gemäß $(*)$ zulässig. Ist der letzte Schluss der gegebenen Herleitung (im Hilbert-Kalkül) schließlich ein $(\forall R)$-Schluss, so ist nach Induktionsvoraussetzung $\vdash B \to \mathcal{F}(a)$. Daraus folgt mit $(\to SInv)$ aus 5.1.8 und weiter mit $(\forall S)$ und $(\to S)$

$$\vdash B : \mathcal{F}(a) \vdash B : \forall x\, \mathcal{F}(x) \vdash B \to \forall x\, \mathcal{F}(x).$$

Mit Induktion nach dem Herleitungsaufbau im Hilbert-Kalkül folgt nun der Satz.

Für die Umkehrung dieses Satzes steht man vor dem Problem, dass der Sequenzenkalkül Sequenzen herleitet, die im Hilbert-Kalkül nicht vorkommen. Man muss also zunächst jeder Sequenz eine oder mehrere Formeln zuordnen, die die Bedeutung der Sequenz genau genug, nämlich „modulo Tautologien“ wiedergibt.

6.3.4 Definition Sei $C_1, \ldots, C_m$ eine Aufzählung der Formelmenge Γ und $D_1, \ldots, D_n$ eine Aufzählung der Formelmenge Δ, beide möglicherweise mit Wiederholungen. Dann nennen wir die Formel

$$C_1 \to \ldots \to C_m \to D_1 \vee \ldots \vee D_n$$

eine *der Sequenz* $\Gamma : \Delta$ *zugeordnete Formel* und bezeichnen sie oft mit $\Gamma \to \Delta$. Für $m = 0$ ist dies die Formel $D_1 \vee \ldots \vee D_n$, für $n = 0$ die Formel $C_1 \to \ldots \to C_m \to \bot$.

In naheliegender Übertragung der Schreibweise $\Gamma \to \Delta$ ist auch $A \to \Gamma \to \Delta$ der Sequenz $A, \Gamma : \Delta$ und $\Gamma \to \Delta \vee B \equiv \Gamma \to \neg\Delta \to B$ der Sequenz $\Gamma : \Delta, B$ (und zugleich der Sequenz $\Gamma, \neg\Delta : B$) zugeordnet. $\neg\{D_1, \ldots, D_n\}$ steht dabei für $\{\neg D_1, \ldots, \neg D_n\}$.

Beispiel. Der Sequenz $\Gamma : \Delta \equiv A, B : D$ sind hiernach nicht nur $A \to B \to D$ und $B \to A \to D$, sondern auch u. a.

$$A \to B \to A \to \neg D \to D \text{ und } B \to A \to B \to B \to D$$

zugeordnet, nicht dagegen $A \to B \to D \to D$ (falls $D \not\equiv A, B$).

6.3.5 Lemma Sind F und G zwei der Sequenz $\Gamma : \Delta$ zugeordnete Formeln, so ist $F \leftrightarrow G$ eine Tautologie.

Beweis. Seien Γ und Δ aufgezählt wie in 6.3.4, und sei V eine aussagenlogische Wertung. Dann ist

$$\Gamma = \{C_i | i = 1, \ldots, m\} \text{ und } \Delta = \{D_i | i = 1, \ldots, n\},$$

also ist

$$\begin{aligned} V(\Gamma : \Delta) = w \iff & \text{wenn } V(C_i) = w \text{ ist für alle } i = 1, \ldots, m, \\ & \text{so ist } V(D_i) = w \text{ für ein } i = 1, \ldots, n \\ \iff & V(C_1 \to \ldots \to C_m \to D_1 \vee \ldots \vee D_n) = w. \end{aligned}$$

Da F und G beides Formeln dieser Gestalt sind, ist

$$V(F) = V(\Gamma : \Delta) = V(G),$$

und $F \leftrightarrow G$ ist eine Tautologie.

Im Sinne dieses Lemmas ist *die* $\Gamma : \Delta$ zugeordnete Formel „modulo Tautologien“ eindeutig bestimmt.

6.3.6 Satz Es sei $\Gamma \to \Delta$ eine der Sequenz $\Gamma : \Delta$ zugeordnete Formel. Dann gilt:

$$T \vdash \Gamma : \Delta \Rightarrow T \vdash_H \Gamma \to \Delta.$$

Beweis durch Herleitungsinduktion im Sequenzenkalkül.

1. Die den logischen Axiomen zugeordneten Formeln

   $$P \to \Gamma \to \Delta \vee P \text{ und } \bot \to \Gamma \to \Delta$$

   sind offenbar Tautologien. Nach Lemma 6.3.5 sind dann alle ihnen zugeordneten Formeln Tautologien, also Axiome im Hilbert-Kalkül.

2. Der letzte Grundschluss der gegebenen Sequenzenherleitung sei

   $$\Gamma_i : \Delta_i (i = 1 \text{ oder } i = 1, 2) \vdash \Gamma : \Delta.$$

   Nach Induktionsvoraussetzung ist dann *jede* $\Gamma_i : \Delta_i$ zugeordnete Formel in T im Hilbert-Kalkül herleitbar. Nach Lemma 6.3.5 und der Bemerkung nach 6.3.2 genügt es zu zeigen, dass *eine* $\Gamma : \Delta$ zugeordnete Formel ebenso herleitbar ist. Wir zeigen dies für die einzelnen Grundschlussregeln.

$(\to S)$ $\Gamma, A : B, \Delta \vdash \Gamma : A \to B, \Delta$
Nach Induktionsvoraussetzung ist $T \mid\!\!\!\frac{}{H} \Gamma \to A \to \Delta \vee B$, und

$$(\Gamma \to A \to \Delta \vee B) \to (\Gamma \to \Delta \vee (A \to B))$$

ist eine Tautologie. Also ist $T \mid\!\!\!\frac{}{H} \Gamma \to \Delta \vee (A \to B)$, und diese Formel ist $\Gamma : A \to B, \Delta$ zugeordnet.

$(\to A)$ $\Gamma : A, \Delta$ und $\Gamma, B : \Delta \vdash \Gamma, A \to B : \Delta$
Nach Induktionsvoraussetzung sind $T \mid\!\!\!\frac{}{H} \Gamma \to \neg A \to \Delta$ und $T \mid\!\!\!\frac{}{H} \Gamma \to B \to \Delta$, und

$$(\Gamma \to \neg A \to \Delta) \to (\Gamma \to B \to \Delta) \to (\Gamma \to (A \to B) \to \Delta)$$

ist eine Tautologie. Also ist $T \mid\!\!\!\frac{}{H} \Gamma \to (A \to B) \to \Delta$, und diese Formel ist $\Gamma, A \to B : \Delta$ zugeordnet.

$(\forall S)$ $\Gamma : \mathcal{F}(a), \Delta \vdash \Gamma : \forall x\, \mathcal{F}(x), \Delta$, falls a nicht in $\Gamma, \mathcal{F}, \Delta$ auftritt.
Nach Induktionsvoraussetzung ist $T \mid\!\!\!\frac{}{H} \Gamma \to \neg\Delta \to \mathcal{F}(a)$, und (im wesentlichen) ein $(\forall R)$-Schluss ergibt $T \mid\!\!\!\frac{}{H} \Gamma \to \neg\Delta \to \forall x\, \mathcal{F}(x)$. Diese Formel ist $\Gamma : \forall x\, \mathcal{F}(x), \Delta$ zugeordnet.

$(\forall A)$ $\Gamma, \mathcal{F}(t) : \Delta \vdash \Gamma, \forall x\, \mathcal{F}(x) : \Delta$.
Nach Induktionsvoraussetzung ist $T \mid\!\!\!\frac{}{H} \mathcal{F}(t) \to \Gamma \to \Delta$, und

$$(\forall x\, \mathcal{F}(x) \to \mathcal{F}(t)) \to (\mathcal{F}(t) \to \Gamma \to \Delta) \to (\forall x\, \mathcal{F}(x) \to \Gamma \to \Delta)$$

ist eine Tautologie. Aus dem $(\forall Ax)$ $\forall x\, \mathcal{F}(x) \to \mathcal{F}(t)$ mit (mp) schließt man auf $T \mid\!\!\!\frac{}{H} \forall x\, \mathcal{F}(x) \to \Gamma \to \Delta$, und diese Formel ist $\Gamma, \forall x\, \mathcal{F}(x) : \Delta$ zugeordnet.

$(= I)$ $\Gamma, t = t : \Delta \vdash \Gamma : \Delta$
Nach Induktionsvoraussetzung ist $T \mid\!\!\!\frac{}{H} t = t \to \Gamma \to \Delta$. Aus dem Identitätsaxiom $t = t$ mit (mp) folgt $T \mid\!\!\!\frac{}{H} \Gamma \to \Delta$, was $\Gamma : \Delta$ zugeordnet ist.

$(= F)$ $\Gamma, fs_1 \dots s_n = ft_1 \dots t_n : \Delta \vdash \Gamma, s_1 = t_1, \dots, s_n = t_n : \Delta$.
Wir schreiben s für $s_1 \dots s_n$, analog t und $s = t$. Nach Induktionsvoraussetzung ist $T \mid\!\!\!\frac{}{H} fs = ft \to \Gamma \to \Delta$, und

$$(s = t \to fs = ft) \to (fs = ft \to \Gamma \to \Delta) \to (s = t \to \Gamma \to \Delta)$$

ist eine Tautologie. Aus dem Axiom $(= FAx)$ $s = t \to fs = ft$ mit (mp) schließt man auf $T \vdash_H s = t \to \Gamma \to \Delta$, was $\Gamma, s = t : \Delta$ zugeordnet ist.

$(= P)$ $\Gamma, pt_1 \dots t_n : \Delta \vdash \Gamma, s_1 = t_1, \dots, s_n = t_n, ps_1 \dots s_n : \Delta$.
Wir schreiben wieder s für $s_1 \dots s_n$, analog t und $s = t$. Nach Induktionsvoraussetzung ist $T \vdash_H pt \to \Gamma \to \Delta$, und

$$(s = t \to ps \to pt) \to (pt \to \Gamma \to \Delta) \to (s = t \to ps \to \Gamma \to \Delta)$$

ist eine Tautologie. Aus dem Axiom $(= PAx)$ $s = t \to ps \to pt$ mit (mp) schließt man auf $T \vdash_H s = t \to ps \to \Gamma \to \Delta$, und diese Formel ist $\Gamma, s = t, ps : \Delta$ zugeordnet.

(T) $B, \Gamma : \Delta \vdash \Gamma : \Delta$ für $B \in Ax(T)$
Nach Induktionsvoraussetzung ist $T \vdash_H B \to \Gamma \to \Delta$, und aus dem Theorie-Axiom B schließt man mit (mp) auf $T \vdash_H \Gamma \to \Delta$, was $\Gamma : \Delta$ zugeordnet ist.

Mit Induktion nach dem Aufbau der Sequenzenherleitungen folgt nun der Satz. Der Spezialfall $\Gamma = \emptyset$, $\Delta = \{C\}$ des Satzes ergibt

6.3.7 Korollar $T \vdash C \Rightarrow T \vdash_H C$

Satz 6.3.3 und dieses Korollar ergeben zusammen

6.3.8 Satz Äquivalenz von Hilbert- und Sequenzenkalkül

$$T \vdash_H C \iff T \vdash C.$$

In beiden Formalismen sind dieselben Formeln herleitbar, obwohl die Herleitungen völlig verschieden aussehen. Welcher Kalkül „handlicher“ ist, ist nicht eindeutig zu beantworten. Für die Herleitung logischer Gesetze schreibt der Sequenzenkalkül das Verfahren fast zwingend vor; bei komplizierten Herleitungen in mathematischen Theorien bietet der *modus ponens* Möglichkeiten, die Länge der Herleitungen im Hilbert-Kalkül beträchtlich zu reduzieren. Aus beweistheoretischer Sicht ist der Sequenzenkalkül eindeutig zu bevorzugen.

6.4 Aufgaben

6.4.1 Zeigen Sie:

a. $\vdash \forall x \forall y\ x = y, \exists x\, \mathcal{F}(x) : \forall x\, \mathcal{F}(x)$

b. $\vdash \forall x\ s(x) = t_1, \exists x s(x) = t_2 : t_1 = t_2$

6.4.2 Folgern Sie aus 6.1.7 und 6.2.3:

$$\vdash \forall x\ r(x) = s(x) : \forall x F(r(x)) \leftrightarrow \forall x F(s(x))$$

6.4.3 Verallgemeinern Sie den Äquivalenzsatz 6.2.3, indem Sie zeigen (a steht für $a_1, \ldots, a_n$ und x steht für $x_1, \ldots, x_n$):
Wenn $\vdash \Gamma : G_i(a) \leftrightarrow H_i(a)$ für $i = 1, \ldots, k$, so ist

$$\vdash \Gamma :\ \mathcal{F}(G_1(x), \ldots, G_k(x)) \leftrightarrow\ \mathcal{F}(H_1(x), \ldots, H_k(x)),$$

falls die a_i nicht $\Gamma, \mathcal{F}, G_j, H_j$ auftreten.

6.4.4 Zeigen Sie, dass sich die zweite Behauptung im Beispiel 6.2.6 aus Aufgabe 6.4.3 ableiten lässt.

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 3:
Vollständigkeit

Autor: Justus Diller

01301-7-03-S 1

Klassische Prädikatenlogik

Kurseinheit 3: Inhalt

Studienhinweise 147

Verzeichnis der definierten Begriffe und der wichtigen Sätze 149

1. Sprache, Semantik und Syntax der Prädikatenlogik

2. Syntaktische Sätze und Regeln der Prädikatenlogik

3. Vollständigkeit

§7 Saturierte Theorien und ihr kanonisches Modell 151

7.1 Konsistenz 152
7.2 Maximale Konsistenz 154
7.3 Saturiertheit 159
7.4 Aufgaben 163

§8 Vollständigkeit für beliebige Theorien 165

8.1 Zorns Lemma 165
8.2 Einfache Erweiterungen 166
8.3 Saturierung 169
8.4 Semantischer Beweis der Schnittregel 175
8.5 Aufgaben 176

§9 Vollständigkeit für abzählbare identitätsfreie Theorien 180

9.1 Königs Lemma 180
9.2 D-Bäume und D-Fäden 184
9.3 Syntaktisches und semantisches Hauptlemma 190
9.4 Modelle über den natürlichen Zahlen 197
9.5 Aufgaben 199

4. Modelltheorie

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 3: Studienhinweise

1. Lehrziele

Diese Kurseinheit behandelt den Vollständigkeitssatz. Sie stellt in dem Sinne das Kernstück des Kurses dar, als die reine Einführung in die Prädikatenlogik mit dem Vollständigkeitssatz zum Abschluss kommt. Gerade durch den Beweis der Äquivalenz von syntaktischer Herleitbarkeit und semantischer Gültigkeit muss Ihnen das Gegensätzliche von syntaktisch-konstruktiver und semantisch-abstrakter Methode deutlich werden. Sie haben ein wichtiges Kursziel erreicht, wenn Sie Korrektheit und Vollständigkeit der Prädikatenlogik mit dem ganzen begrifflichen Unterbau verstehen und einen Beweis des Vollständigkeitssatzes genau kennen.

Interessante, zum Teil überraschende Konsequenzen dieses Satzes behandeln wir in den nächsten Kurseinheiten. Die Tragweite der Vollständigkeit können Sie also angemessen erst in Kapitel 4 erfassen.

Die Aussage des Vollständigkeitssatzes ist einfach:

> *Jede in T gültige Sequenz ist auch in T herleitbar, bzw.*
> *Jede konsistente Theorie hat ein Modell.*

Für ein Verständnis des Beweises muss man sich die Methoden bewusst machen, die in den Beweis eingehen. Diese Methoden sind notwendig nicht alle konstruktiv, im Gegensatz zu den rein finiten Methoden des Kapitels 2. In einem Logik-Kurs tritt die Trennung der nicht-konstruktiven Methoden von den konstruktiven naturgemäß stärker in den Vordergrund als in anderen Mathematik-Kursen.

Wir führen zwei Beweise des Vollständigkeitssatzes vor, die sich methodisch wesentlich unterscheiden und – bezüglich der Gleichheitsgesetze – auch unterschiedliche Konsequenzen haben.

Der Ansatz ist in beiden Fällen der gleiche:
Zu einer konsistenten Theorie T sucht man nach einem Modell von T. In beiden Fällen „konstruiert“ man dieses Modell ausschließlich aus dem syntaktischen Material der Theorie T, also aus ihrer Sprache, ihrem Axiomensystem und dem Herleitungsbegriff. An zentralen Stellen enthält diese „Konstruktion“ aber nicht-konstruktive Überlegungen.

In §8 erweitern wir beliebige konsistente Theorien zu saturierten Theorien. Dabei verwenden wir Zorns Lemma, eine zum Auswahlaxiom in der Mengenlehre äquivalente Aussage.

In §9 betrachten wir abzählbare identitätsfreie Theorien. Der besonders transparente Beweis hängt wesentlich an der Schnittfreiheit unseres Herleitungsbegriffs. Das nicht-konstruktive Mittel ist hier Königs Lemma. Das ist im Gegensatz zu Zorns Lemma eine elementare Aussage, die sich allerdings nur auf abzählbare Bereiche bezieht.

Der Ertrag dieser Kurseinheit liegt nicht im Technischen – wie in der Kurseinheit 2 in der Beherrschung des Herleitungsbegriffs –, sondern im Methodischen. Sie sollen bekannt werden mit nicht-konstruktiven Methoden und deren abgestuftem Einsatz im Rahmen von Konstruktionen, die Brücken schlagen zwischen Syntax und Semantik.

2. Eingangsvoraussetzungen

Im Gegensatz zur fast voraussetzungslosen Kurseinheit 2 müssen wir hier auf alle bisherigen Begriffe und Ergebnisse zurückgreifen. Eine vorherige Kenntnis von Zorns Lemma oder Königs Lemma erleichtert das Studium dieser Kurseinheit sehr, ist aber nicht nötig: Beide Prinzipien werden ausführlich diskutiert. Zorns Lemma wird auch in den Aufgaben eingehend behandelt. Auch wird ausgeführt, was man hier über Äquivalenzrelationen, partielle Ordnungen und Bäume wissen muss. Außer auf die übliche mengentheoretische Schreibweise greifen wir also auf nichts zurück, was nicht in diesem Kurs schon behandelt ist.

Klassische Prädikatenlogik

Kurseinheit 3: Verzeichnis der definierten Begriffe und der wichtigen Sätze

7.1.1 Konsistente Theorien

7.1.3 Zugeordneter Satz

7.1.6 Äquivalente Formulierungen der Vollständigkeit

7.2.1 Maximal konsistente Theorien

7.2.7 Äquivalenz von geschlossenen Termen

7.2.9 Kanonische Struktur einer (maximal konsistenten) Theorie

7.3.1 Saturierte Theorien

7.3.3 **Satz** über das kanonische Modell

7.3.5 **Vollständigkeitssatz für saturierte Theorien**

8.1.1 **Zorns Lemma**

8.2.1 Erweiterung von Sprachen, Beschränkung von Strukturen

8.2.3 Erweiterung von Theorien, einfache, konservative

8.2.5 **Satz von Lindenbaum** Jede konsistente Theorie besitzt eine maximal konsistente einfache Erweiterung

8.3.1 Konstantenerweiterung

8.3.2 Lemma über neue Konstanten

8.3.6 **Satz von Henkin** Jede konsistente Theorie besitzt eine saturierte Erweiterung

8.3.7 **Vollständigkeitssatz von Gödel und Henkin** Jede konsistente Theorie besitzt ein Modell

8.3.8 Äquivalenz von Herleitbarkeit und Gültigkeit

8.4.1 **Satz** Zulässigkeit der Schnittregel

9.1.1 Endliche Folgen, Anfangsstück, Verkettung, Baum, Knoten, Wurzel, Blatt, binärer Baum

9.1.5 **Königs Lemma**

9.2.1 Abzählbare Sprachen und Theorien

9.2.3 Identitätsfreie Formeln, Sequenzen und Theorien

9.2.5 Geordnete Sequenz, ausgezeichnetes Formelauftreten

9.2.6 D-Bäume

9.2.8 D-Fäden

9.3.1 **Syntaktisches Hauptlemma**

9.3.2 **Semantisches Hauptlemma**

9.3.8 Struktur und Belegung zu unendlichen D-Fäden

9.4.1 **Vollständigkeitssatz von Schütte** für abzählbare identitätsfreie Theorien

9.4.2 Konservativität abzählbarer identitätsfreier Theorien über der Logik ohne Gleichheitsregeln

9.4.3 Satz von Löwenheim und Skolem für abzählbare identitätsfreie Theorien

Kapitel 3

Vollständigkeit

Dieses Kapitel behandelt Beweise des Vollständigkeitssatzes, der uns als Umkehrung des Korrektheitssatzes 3.2.1 bekannt ist. Wir beweisen ihn in der Fassung:

Jede konsistente Theorie hat ein Modell.

Man hat also ein Modell von T zu „konstruieren“, wenn man nur weiß, dass T konsistent, d. h. dass die leere Sequenz in T nicht herleitbar ist. Für diese „Konstruktion“ steht uns nur syntaktisches Material zur Verfügung, nämlich die Theorie T, ihre Sprache und ihr Axiomensystem. Sie verwendet notwendig abstrakte Mittel, die wir noch diskutieren werden.

§7 Saturierte Theorien und ihr kanonisches Modell

7.1 Konsistenz

7.2 Maximale Konsistenz

7.3 Saturiertheit

7.4 Aufgaben

Wir beweisen den Vollständigkeitssatz in diesem Paragraphen für den Spezialfall saturierter Theorien und zeigen im nächsten Paragraphen, dass sich jede konsistente Theorie zu einer saturierten Theorie erweitern lässt.

7.1 Konsistenz

7.1.1 Definition Eine Theorie T ist *konsistent*, wenn die leere Sequenz in T nicht herleitbar ist: $T \nvdash \square$.

Jede Sequenz folgt strukturell offenbar aus der leeren Sequenz $\square = \emptyset : \emptyset$ und ebenso aus $\emptyset : \bot$. Nach der Strukturschlussregel 5.1.7 ist also jede Sequenz aus $L(T)$ in T herleitbar, wenn (und offenbar auch nur wenn) T nicht konsistent ist, und ebenso wenn (und auch nur wenn) $T \vdash \bot$. Dies können wir umgekehrt formulieren:

7.1.2 Lemma Für jede Theorie T sind äquivalent:

(1) T ist konsistent

(2) $T \nvdash \bot$

(3) Es gibt eine Sequenz aus $L(T)$, die in T nicht herleitbar ist.

Mit der Konsistenz einer geeignet gewählten Theorie lässt sich auch charakterisieren, dass eine Sequenz $\Gamma : \Delta$ in einer Theorie T nicht herleitbar ist.

7.1.3 Definition Ist $\Gamma \to \Delta$ eine $\Gamma : \Delta$ zugeordnete Formel gemäß 6.3.4, so heißt jeder Allabschluss $\forall(\Gamma \to \Delta)$ davon ein $\Gamma : \Delta$ *zugeordneter Satz*.

Bemerkung. $\Gamma : \Delta$ ist genau dann geschlossen, wenn $\Gamma \to \Delta$ ein Satz ist, und genau dann ist $\forall(\Gamma \to \Delta) \equiv \Gamma \to \Delta$, d. h. genau dann stimmen die $\Gamma : \Delta$ zugeordneten Sätze mit den $\Gamma : \Delta$ zugeordneten Formeln überein. $\bot$ ist der einzige Satz, der der leeren Sequenz $\square$ zugeordnet ist.

7.1.4 Lemma Sei $\forall(\Gamma \to \Delta)$ ein $\Gamma : \Delta$ zugeordneter Satz. Dann sind äquivalent:

(1) $T \vdash \Gamma : \Delta$

(2) $T \vdash \Gamma \to \Delta$

(3) $T \vdash \forall(\Gamma \to \Delta)$

(4) $T + \{\neg\forall(\Gamma \to \Delta)\}$ ist nicht konsistent.

Beweis. (1) $\Leftrightarrow$ (2). Seien $C_1, \ldots, C_m$ und $D_1, \ldots, D_n$ die gegebenen Aufzählungen von Γ und Δ.
Wir induzieren nach $m+n$. Ist $m = 0$ und $n \leq 1$, so folgen (1) und (2) strukturell auseinander. Ist $n = 0$ und $m > 0$, so folgt mit $(\neg S)$ und $(\neg SInv)$ aus 5.2.1, dass (1), also $T \vdash \Gamma : \emptyset$ äquivalent ist zu

$$T \vdash C_1, \ldots, C_{m-1} : C_m \to \bot.$$

Ist $n = 1$ und $m > 0$, so folgt mit $(\to S)$ und $(\to SInv)$, dass (1), also $T \vdash \Gamma : D_1$ äquivalent ist zu

$$T \vdash C_1, \ldots, C_{m-1} : C_m \to D_1.$$

Ist $n > 1$, so folgt mit $(\vee S)$ und $(\vee SInv)$ aus 5.2.4, dass (1) äquivalent ist zu

$$T \vdash \Gamma : D_1, \ldots, D_{n-1} \vee D_n.$$

In allen drei Fällen folgt mit Induktionsvoraussetzung nun die Behauptung.
(2) $\Leftrightarrow$ (3) steht in 5.3.6. (3) $\Leftrightarrow$ $T \vdash \neg\neg\forall(\Gamma \to \Delta)$ ist ein Spezialfall von $(\neg\neg S)$ in 5.2.2, und nach dem Deduktionstheorem 5.4.2 ist

$$T \vdash \neg\neg\forall(\Gamma \to \Delta) \Leftrightarrow T + \{\neg\forall(\Gamma \to \Delta)\} \vdash \bot,$$

und das ist (4). Also ist auch (3) $\Leftrightarrow$ (4) und damit das Lemma insgesamt bewiesen.

Kontraposition ergibt aus der Äquivalenz von (1) und (4) unmittelbar:

7.1.5 Korollar Die Sequenz $\Gamma : \Delta$ ist in T genau dann nicht herleitbar, wenn $T + \{\neg\forall(\Gamma \to \Delta)\}$ konsistent ist.

Damit können wir verschiedene Fassungen des Vollständigkeitssatzes leicht vergleichen.

7.1.6 Lemma Folgende Formulierungen des Vollständigkeitssatzes sind äquivalent:

(1) Jede in einer Theorie T gültige Sequenz ist in T auch herleitbar.

(2) Ist $T \nvdash \Gamma : \Delta$, so hat T ein Modell, in dem $\Gamma : \Delta$ nicht gilt.

(3) Jede konsistente Theorie hat ein Modell.

Beweis. (2) ist die Kontraposition von (1), also äquivalent zu (1), und (3) ist der Spezialfall $\Gamma : \Delta = \Box$ von (2). Also ist nur noch (3) $\Rightarrow$ (2) zu beweisen. Sei $T \nvdash \Gamma : \Delta$. Nach 7.1.5 ist dann $T + \{\neg\forall(\Gamma \to \Delta)\}$ konsistent. Nach (3) hat diese Theorie dann ein Modell $\mathcal{A}$, d. h. $\mathcal{A} \models T$ und $\mathcal{A}(\neg\forall(\Gamma \to \Delta)) = w$. Daraus folgt $\mathcal{A}(\forall(\Gamma \to \Delta)) = f$, weiter $\mathcal{A} \not\models \Gamma \to \Delta$ und damit $\mathcal{A} \not\models \Gamma : \Delta$. Also folgt (2) aus (3).

Die dritte Formulierung der Vollständigkeit ist offenbar die einfachste, und sie steht im Folgenden im Mittelpunkt. Wenn wir sie bewiesen haben, haben wir nach dem Lemma auch die Formulierungen (1) und (2) bewiesen, und zwar auch für Formeln statt für Sequenzen $\Gamma : \Delta$.

7.2 Maximale Konsistenz

Wir fragen nach Theorien, die zu einem ihrer Modelle eine besonders nahe Beziehung haben. Wenn schon eine Struktur $\mathcal{A}$ zu einer Sprache L gegeben ist, so ist die Theorie dieser Struktur $Th(\mathcal{A})$ hierfür ein guter Kandidat. Wodurch zeichnet sich $Th(\mathcal{A})$ syntaktisch aus? $Th(\mathcal{A})$ ist konsistent und hat ein besonders großes Axiomensystem, das sich innerhalb der Sprache L nicht konsistent erweitern lässt. Solche Theorien nennt man maximal konsistent.

7.2.1 Definition Eine Theorie T ist *maximal konsistent*, wenn

a. T konsistent ist und

b. für jeden Satz C aus $L(T)$ gilt: Ist $T + \{C\}$ konsistent, so ist $C \in Ax(T)$.

In Lemma 7.3.4 wird gezeigt: $Th(\mathcal{A})$ ist maximal konsistent für jede Struktur $\mathcal{A}$. Ist $\mathcal{A}$ irgendein Modell einer maximal konsistenten Theorie T, so ist bereits $T = Th(\mathcal{A})$. Unser Problem ist es, ein solches Modell zu finden.

7.2.2 Lemma T sei eine maximal konsistente Theorie und C ein Satz aus $L(T)$. Dann sind äquivalent:

(1) $C \notin Ax(T)$

(2) $T \vdash C : \emptyset$

(3) $T \vdash \Gamma, C : \Delta$ für jede Sequenz $\Gamma : \Delta$ aus $L(T)$.

Beweis: Sei C kein Axiom von T. Weil T maximal konsistent ist, ist dann $T + \{C\}$ nicht konsistent:

(4) $T + \{C\} \vdash \square$.

Hieraus folgt (2) mit dem Deduktionstheorem.
Umgekehrt folgt (4) aus (2) mit einem $T+\{C\}$-Schluss. Dann ist $T+\{C\} \neq T$, weil T konsistent ist, und es folgt (1).
Aus (2) folgt (3) mit einem Strukturschluss, und umgekehrt ist (2) ein Spezialfall von (3).

7.2.3 Lemma Ist T maximal konsistent und t ein geschlossener Term aus $L(T)$, so ist

$$t = t \in Ax(T).$$

Beweis: Angenommen, $t = t$ wäre kein Axiom von T. Dann wäre $T \vdash t = t : \emptyset$ nach 7.2.2, also $T \vdash \square$ mit $(= I)$:
T wäre nicht konsistent im Widerspruch zur Voraussetzung. Also ist $t = t$ ein Axiom von T.

7.2.4 Lemma Ist T maximal konsistent und sind s, t_1, t_2 geschlossene Terme aus $L(T)$, so gilt:

$$\text{Aus } s = t_1,\ s = t_2 \in Ax(T) \text{ folgt } t_1 = t_2 \in Ax(T).$$

Beweis: Angenommen, $t_1 = t_2$ wäre kein Axiom von T. Dann wäre

$T \vdash t_1 = t_2 : \emptyset$ nach 7.2.2, also

$T \vdash s = t_1,\ s = t_2 : \emptyset$ nach 6.1.1, also

$T \vdash \square$

mit zwei T-Schlüssen ($s = t_i$ sind Axiome von T): T wäre nicht konsistent im Widerspruch zur Voraussetzung. Also ist auch $t_1 = t_2$ ein Axiom von T.

7.2.5 Lemma Ist T maximal konsistent und sind $s_i, t_i (i = 1, \ldots, n)$ geschlossene Terme von T, so gilt:

$$\text{Aus } s_i = t_i \in Ax(T) \text{ für } i = 1, \ldots, n \text{ folgt } f s_1 \ldots s_n = f t_1 \ldots t_n \in Ax(T).$$

Beweis: Angenommen, $fs_1 \ldots s_n = ft_1 \ldots t_n$ wäre kein Axiom von T. Dann wäre

$T \vdash fs_1 \ldots s_n = ft_1 \ldots t_n : \emptyset$ nach 7.2.2, es folgte

$T \vdash s_1 = t_1, \ldots, s_n = t_n : \emptyset$ mit $(= F)$, also

$T \vdash \square$

mit n T-Schlüssen ($s_i = t_i$ sind Axiome von T): T wäre nicht konsistent im Widerspruch zur Voraussetzung. Also ist auch $fs_1 \ldots s_n = ft_1 \ldots t_n$ ein Axiom von T.

7.2.6 Lemma Ist T maximal konsistent und sind $s_i,\ t_i (i = 1, \ldots, n)$ geschlossene Terme von T, so gilt:
Aus $s_i = t_i \in Ax(T)$ für $i = 1, \ldots, n$ und $ps_1 \ldots s_n \in Ax(T)$ folgt $pt_1 \ldots t_n \in Ax(T)$.

Beweis: Angenommen, $pt_1 \ldots t_n$ wäre kein Axiom von T. Dann wäre

$T \vdash pt_1 \ldots t_n : \emptyset$ nach 7.2.2, es folgte

$T \vdash s_1 = t_1, \ldots, s_n = t_n, ps_1 \ldots s_n : \emptyset$ mit $(= P)$, also

$T \vdash \square$

mit $n + 1$ T-Schlüssen ($s_i = t_i, ps_1 \ldots s_n$ sind Axiome von T): T wäre nicht konsistent im Widerspruch zur Voraussetzung. Also ist auch $pt_1 \ldots t_n$ ein Axiom von T.

Im Folgenden wird der Begriff der Äquivalenzrelation mit seinen einfachsten Eigenschaften vorausgesetzt. Gebraucht wird im Wesentlichen: Eine *Äquivalenzrelation* $\sim$ auf einer Menge B ist eine zweistellige reflexive und komparative Relation auf B, d. h. für $i, j, k \in B$ ist stets $i \sim i$ und aus $i \sim j$ und $i \sim k$ folgt stets $j \sim k$. Die Äquivalenzrelationen sind genau die reflexiven, symmetrischen und transitiven Relationen.

Eine Äquivalenzrelation $\sim$ auf B definiert eine Äquivalenzklasseneinteilung von B wie folgt: Die Teilmenge $\overline{k} := \{i \in B \mid i \sim k\}$ ist die *Äquivalenzklasse* von k relativ zu $\sim$. Es ist

$$i \sim k \Leftrightarrow i \in \overline{k} \Leftrightarrow \overline{i} = \overline{k}.$$

Jedes $k \in B$ gehört genau einer Äquivalenzklasse an. Verschiedene Äquivalenzklassen sind also disjunkt.

Wir kehren zu den maximal konsistenten Theorien zurück.

7.2.7 Definition Wir nennen geschlossene Terme s, t *äquivalent*, $s \sim t$, wenn der Satz $s = t \in Ax(T)$ ist.

7.2.8 Korollar T sei maximal konsistent, und es gebe einen geschlossenen Term in $L(T)$. Dann ist $\sim$ eine Äquivalenzrelation auf der Menge der geschlossenen Terme von $L(T)$, für die aus $s_i \sim t_i (i = 1, \ldots, n)$ folgt

(5) $fs_1 \ldots s_n \sim ft_1 \ldots t_n$,

(6) $ps_1 \ldots s_n \in Ax(T) \Leftrightarrow pt_1 \ldots t_n \in Ax(T)$.

Beweis: Nach 7.2.3 und 4 ist $\sim$ eine reflexive und komparative Relation, also eine Äquivalenzrelation. 7.2.5 ergibt (5), und 7.2.6 zusammen mit der Symmetrie von $\sim$ ergibt (6).

Hiernach können wir allein aus den geschlossenen Termen von T eine Struktur $\mathcal{A}$ zu $L(T)$ konstruieren. Eine solche Struktur $\mathcal{A}$ nennt man auch eine Term-Struktur zu $L(T)$, weil es zu jedem $c \in |\mathcal{A}|$ einen (geschlossenen) Term t aus $L(T)$ gibt, der c bezeichnet, $\mathcal{A}t = c$. Wir suchen nach Term-Strukturen, die möglichst auch Modelle von T sind. Durch die Forderung, dass in $\mathcal{A}$ alle in T herleitbaren geschlossenen Primformeln und negierten Primformeln wahr sein sollen, ist $\mathcal{A}$ bemerkenswerterweise schon eindeutig festgelegt.

Denn ist $s \sim t$, also $s = t \in Ax(T)$, so muss $\mathcal{A}s = \mathcal{A}t$ sein.
Sonst ist $s = t \notin Ax(T)$, also $T \vdash \neg s = t$ nach 7.2.2, und es muss $\mathcal{A}s \neq \mathcal{A}t$ sein. Dann ist (bis auf eine Bijektion) $\mathcal{A}t = \bar{t}$. Weiter ist dann auch $f_{\mathcal{A}}$ für jedes f aus $L(T)$ festgelegt wegen

$$f_{\mathcal{A}}(\bar{t}) = f_{\mathcal{A}}(\mathcal{A}t) = \mathcal{A}(ft) = \overline{ft}.$$

Schließlich muss, falls $pt \in Ax(T)$ ist, $\bar{t} \in p_{\mathcal{A}}$ sein, und ist $pt \notin Ax(T)$, so folgt $T \vdash \neg pt$ wieder mit 7.2.2, und es muss $\bar{t} \notin p_{\mathcal{A}}$ sein.

Die Wahl von $\mathcal{A}$ ist also durch das angestrebte Ziel kanonisch bestimmt. Deshalb heißt diese Term-Struktur $\mathcal{A}$ die *kanonische Struktur* von T.

7.2.9 Definition der *kanonischen Struktur* $\mathcal{A}$ einer Theorie T.

1. Für jeden geschlossenen Term t aus $L(T)$ sei

$$\bar{t} = \{s \mid s \text{ geschlossener Term aus } L(T),\ s = t \in Ax(T))\}$$

2. $|\mathcal{A}| = \{\bar{t} \mid t \text{ geschlossener Term aus } L(T)\}$

3. Zu n-stelligen nicht-logischen Grundzeichen f, p aus $L(T)$ sei

$$f_{\mathcal{A}}(\bar{t}_1, \ldots, \bar{t}_n) := \overline{ft_1 \ldots t_n} \text{ und } (\bar{t}_1, \ldots, \bar{t}_n) \in p_{\mathcal{A}} :\Leftrightarrow\ pt_1 \ldots t_n \in Ax(T)$$

Wir prüfen jetzt, wie weit $\mathcal{A}$ die gesteckten Ziele erfüllt. Ein erstes Problem ist die Wohldefiniertheit der $f_{\mathcal{A}}, p_{\mathcal{A}}$ aus 7.2.9. Die Funktionen $f_{\mathcal{A}}$ und Relationen $p_{\mathcal{A}}$ sind für Äquivalenzklassen $\bar{t}_i$ definiert, aber unter Rückgriff auf (beliebige) Repräsentanten t_i dieser Äquivalenzklassen. Ist ihre Definition von der Wahl dieser Repräsentanten unabhängig?

Im Folgenden sei T eine *maximal konsistente* Theorie, deren Sprache mindestens *eine Konstante* enthält, und $\mathcal{A}$ sei die kanonische Struktur von T.

7.2.10 Lemma Die $f_{\mathcal{A}}, p_{\mathcal{A}}$ mit $f, p \in L(T)$ sind wohldefiniert, und $\mathcal{A}$ ist eine Struktur zu $L(T)$.

Beweis. Wir schreiben s für $s_1, \ldots, s_n$. Ist $\bar{s} = \bar{t}$, so ist nach 7.2.7 $s = t \in Ax(T)$. Dann ist nach 7.2.5 auch $fs = ft \in Ax(T)$, also

$$f_{\mathcal{A}}(\bar{s}) = \overline{fs} = \overline{ft} = f_{\mathcal{A}}(\bar{t}).$$

Ebenso ist nach 7.2.6

$$\bar{s} \in p_{\mathcal{A}} \Leftrightarrow ps \in Ax(T) \Leftrightarrow pt \in Ax(T) \Leftrightarrow \bar{t} \in p_{\mathcal{A}}.$$

Die Definition von $f_{\mathcal{A}}$ und $p_{\mathcal{A}}$ in 7.2.9 ist daher unabhängig von der Wahl des Repräsentanten t_i der Äquivalenzklasse $\overline{t_i}$. Da $L(T)$ eine Konstante enthält, ist $|\mathcal{A}| \neq \emptyset$, und die Behauptung folgt.

Die $f_{\mathcal{A}}$ operieren also auf den Äquivalenzklassen $\bar{t} \in |\mathcal{A}|$, und zwar in einer mit der Termbildung verträglichen Weise.

7.2.11 Lemma Für jeden geschlossenen Term t ist $\mathcal{A}(t) = \bar{t}$.

Beweis durch Induktion nach dem Aufbau von t:

t ist $ft_1 \dots t_n$. Dann ist

$$\begin{aligned} \mathcal{A}(t) &= f_{\mathcal{A}}(\mathcal{A}(t_1), \dots, \mathcal{A}(t_n)) \\ &= f_{\mathcal{A}}(\overline{t_1}, \dots, \overline{t_n}) \text{ nach Induktionsvoraussetzung} \\ &= \overline{ft_1 \dots t_n} = \bar{t} \text{ nach 7.2.9.} \end{aligned}$$

(Der Induktionsanfang ist hier der Fall $n = 0$.)

7.2.12 Lemma Ist P eine geschlossene Primformel aus $L(T)$, so ist

$$P \in Ax(T) \Leftrightarrow \mathcal{A}(P) = w.$$

Beweis durch Fallunterscheidung.

1. P ist eine Gleichung $s = t$. Dann ist

$$\begin{aligned} s = t \in Ax(T) &\Leftrightarrow \mathcal{A}(s) = \bar{s} = \bar{t} = \mathcal{A}(t) \text{ nach 7.2.9 und 11} \\ &\Leftrightarrow \mathcal{A}(s = t) = w \end{aligned}$$

2. P ist eine andere Primformel $pt_1 \dots t_n$. Dann ist

$$\begin{aligned} P \in Ax(T) &\Leftrightarrow (\bar{t}_1, \dots, \bar{t}_n) \in p_{\mathcal{A}} \Leftrightarrow (\mathcal{A}t_1, \dots, \mathcal{A}t_n) \in p_{\mathcal{A}} \text{ nach 7.2.11} \\ &\Leftrightarrow \mathcal{A}(P) = w \end{aligned}$$

3. P ist $\perp$. Es ist $\perp \notin Ax(T)$ wegen der Konsistenz von T und $\mathcal{A}(\perp) = f$.

7.3 Saturiertheit

Das letzte Lemma lässt sich nur auf längere Formeln übertragen, wenn $L(T)$ genügend viele geschlossene Terme enthält. Ist G etwa irgendeine Gruppe, so ist die kanonische Struktur von $Th(G)$ nur die triviale Gruppe, weil in der Gruppentheorie nur das neutrale Element durch einen geschlossenen Term bezeichnet wird. Formeln wie $\forall x(e = x)$ oder das kommutative Gesetz gelten also in der kanonischen Struktur, obwohl sie i.a. in G falsch, also in $Th(G)$ nicht herleitbar sind.

In der Struktur $\mathcal{N}$ dagegen wird jedes Element – jede natürliche Zahl – durch

einen geschlossenen Term – z. B. durch eine Ziffer – bezeichnet. Deswegen ist $\mathcal{N}$ auch (bis auf Isomorphie) die kanonische Struktur von $Th(\mathcal{N})$, so dass Lemma 7.2.12 für $Th(\mathcal{N})$ für alle geschlossenen Formeln der Zahlentheorie gilt. Die syntaktische Eigenschaft, die $Th(\mathcal{N})$ etwa gegenüber $Th(G)$ auszeichnet, ist ihre *Saturiertheit*.

7.3.1 Definition Eine Theorie T ist *saturiert*, wenn

a. T maximal konsistent ist und

b. für jeden Satz $\forall x F(x)$ aus $L(T)$ gilt:
Ist $T \vdash F(s)$ für jeden geschlossenen Term s, so ist $T \vdash \forall x F(x)$.

7.3.2 Lemma Die Sprache einer saturierten Theorie enthält geschlossene Terme.

Beweis. T sei eine saturierte Theorie. Wir nehmen an, in $L(T)$ gebe es keine geschlossenen Terme. Dann ist

$$T \vdash s = s \to \bot \text{ für alle geschlossenen Terme } s,$$

weil es keine solchen gibt. Daraus folgt

$$T \vdash \forall x(x = x \to \bot),$$

weil T saturiert ist. $(\forall S)$-Inversion und $(\to S)$-Inversion ergeben

$$T \vdash a = a : \bot,$$

woraus mit $(= I)$ die Inkonsistenz von T folgt, im Widerspruch zur Voraussetzung.
Hiernach gelten 7.2.10 bis 7.2.12 für saturierte Theorien T.

7.3.3 Satz Es sei $\mathcal{A}$ die kanonische Struktur einer saturierten Theorie T. Dann gilt für jeden Satz C aus $L(T)$:

(1) Aus $C \in Ax(T)$ folgt $\mathcal{A}(C) = w$;

(2) Aus $T \nvdash C$ folgt $\mathcal{A}(C) = f$.

Beweis durch Induktion nach dem Aufbau von C.

1. Für Primformeln C folgen beide Behauptungen aus 7.2.12, weil Axiome herleitbar sind.

2. C ist $A \to B$.

   2.1 Sei $A \to B \in Ax(T)$. Angenommen, es wäre $T \vdash A$ und $B \notin Ax(T)$, also $T \vdash B : \emptyset$ nach 7.2.2. Dann folgte $T \vdash A \to B : \emptyset$ mit einem $(\to A)$-Schluss und $T \vdash \square$ mit einem T-Schluss.
   Also ist unsere Annahme falsch, es ist

   $$T \nvdash A \text{ oder } B \in Ax(T).$$

   Dann ist nach Induktionsvoraussetzung $\mathcal{A}(A) = f$ oder $\mathcal{A}(B) = w$, also $\mathcal{A}(A \to B) = w$.

   2.2 Es sei $T \nvdash A \to B$. Wäre $A \notin Ax(T)$, so wäre $T \vdash A : B$ nach 7.2.2. Wäre $T \vdash B$, so folgte mit einem Strukturschluss $T \vdash A : B$. In beiden Fällen folgte $T \vdash A \to B$ im Widerspruch zur Annahme. Also ist $A \in Ax(T)$ und $T \nvdash B$, und mit Induktionsvoraussetzung folgt $\mathcal{A}(A) = w$ und $\mathcal{A}(B) = f$, also $\mathcal{A}(A \to B) = f$.

3. C ist $\forall x F(x)$.

   3.1 Sei $\forall x F(x) \in Ax(T)$. Angenommen, es gäbe einen geschlossenen Term s mit $F(s) \notin Ax(T)$, also $T \vdash F(s) : \emptyset$ nach 7.2.2. Dann folgte $T \vdash \forall x F(x) : \emptyset$ mit einem $(\forall A)$-Schluss und $T \vdash \square$ mit einem T-Schluss.
   Also ist unsere Annahme falsch, es ist $F(s) \in Ax(T)$ für jeden geschlossenen Term s. Dann ist nach 7.2.11, dem Homomorphieprinzip und Induktionsvoraussetzung

   $$\mathcal{A}(F(\overline{s})) = \mathcal{A}(F(s)) = w \text{ für alle } \overline{s} \in |\mathcal{A}|, \text{ also } \mathcal{A}(\forall x F(x)) = w.$$

   3.2 Sei $T \nvdash \forall x F(x)$. Weil T saturiert ist, gibt es dann einen geschlossenen Term s, für den $T \nvdash F(s)$ ist. Dann folgt wie eben

   $$\mathcal{A}(F(\overline{s})) = \mathcal{A}(F(s)) = f \quad \text{für dieses } \overline{s} \in |\mathcal{A}|, \text{ also}$$
   $$\mathcal{A}(\forall x F(x)) = f.$$

Mit Induktion nach dem Aufbau von C folgen die Behauptungen (1) und (2).

Der Beweis verwendet fast bei jedem Schritt die maximale Konsistenz von T, aber nur an einer einzigen Stelle, in 3.2, die Saturiertheit. Sie wird dort aber auch wesentlich verwendet. Nach (1) ist $\mathcal{A}$ ein Modell von T, und wir können jetzt von dem *kanonischen Modell* einer saturierten Theorie sprechen. Während die Theorie einer Struktur bisher nur zur Motivation verwendet wurde, ist dieser Begriff jetzt nützlich, um unsere Ergebnisse zusammenzufassen. Dabei hilft folgende allgemeine Beobachtung, die in 7.2 schon erwähnt wurde.

7.3.4 Lemma Ist $\mathcal{A}$ ein Modell einer Theorie T, so gilt:

$$T \text{ ist maximal konsistent } \Leftrightarrow T = Th(\mathcal{A}).$$

Beweis. Aus $\mathcal{A} \models T$ folgt sofort $Ax(T) \subseteq Ax(Th(\mathcal{A}))$; die rechte Seite ist daher äquivalent zu $Ax(Th(\mathcal{A})) \subseteq Ax(T)$.

Sei nun T maximal konsistent und $C \in Ax(Th(\mathcal{A}))$, also $\mathcal{A}(C) = w$. Dann ist $\mathcal{A} \models T + \{C\}$, und wegen des Korrektheitssatzes ist $T + \{C\}$ konsistent. Dann ist $C \in Ax(T)$, weil T maximal konsistent ist. Also ist $T = Th(\mathcal{A})$.

Sei umgekehrt $T = Th(\mathcal{A})$ und $T + \{C\}$ konsistent. Wäre C nicht wahr in $\mathcal{A}$, so wäre $\mathcal{A}(\neg C) = w$, mithin $\neg C \in Ax(T)$, und $T + \{C\}$ wäre nicht konsistent. Also ist $\mathcal{A}(C) = w$ und daher $C \in Ax(T)$, und T ist maximal konsistent.

Nun sind saturierte Theorien einerseits maximal konsistent, andererseits haben sie ihr kanonisches Modell. Hieraus ergibt sich mit dem Lemma sofort:

7.3.5 Vollständigkeitssatz für saturierte Theorien Jede saturierte Theorie T hat ein kanonisches Modell $\mathcal{A}$, dessen Theorie $Th(\mathcal{A})$ die Theorie T ist.

In diesem Fall ist die Beziehung zwischen Theorie T und Struktur $\mathcal{A}$ optimal: $\mathcal{A}$ ist nicht nur Modell von T, sondern T ist auch die Theorie von $\mathcal{A}$, und zusätzlich ist $\mathcal{A}$ allein aus der Theorie T, nämlich ihren geschlossenen Termen und Axiomen konstruiert: T und $\mathcal{A}$ bestimmen sich gegenseitig eindeutig.

Für saturierte Theorien ist hiermit die Vollständigkeit in verschärfter Form bewiesen: Die Axiome von T sind genau die in einem einzigen Modell von T wahren Sätze, und das sind auch genau die in T herleitbaren Sätze. So eindeutig ist die Situation im Allgemeinen nicht. Für beliebige konsistente Theorien T werden wir saturierte Erweiterungen suchen, und die sind durch T keineswegs eindeutig bestimmt.

7.4 Aufgaben

7.4.1 Sei $\sim$ eine zweistellige Relation auf einer Menge A.

a. Zeigen Sie: Ist $\sim$ symmetrisch, so ist $\sim$ komparativ genau dann, wenn $\sim$ transitiv ist.

b. Zeigen Sie: Ist $\sim$ reflexiv, so ist $\sim$ komparativ genau dann, wenn $\sim$ symmetrisch und transitiv ist.

c. Geben Sie eine Menge A und eine reflexive und transitive Relation auf A an, die nicht symmetrisch ist.

d. Geben Sie eine komparative Relation $\sim$ auf der dreielementigen Menge $\{0, 1, 2\}$ an, die nicht transitiv ist.

7.4.2 Zwei Strukturen $\mathcal{A}, \mathfrak{B}$ zu einer Sprache L heißen *elementar äquivalent*, $\mathcal{A} \equiv \mathfrak{B}$, wenn $\mathcal{A}(C) = \mathfrak{B}(C)$ ist für jeden Satz C von L. Zeigen Sie die Äquivalenz von

(1) $\mathcal{A} \equiv \mathfrak{B}$

(2) $\mathfrak{B} \models Th(\mathcal{A})$

(3) $Th(\mathcal{A}) = Th(\mathfrak{B})$.

7.4.3 Sei K eine Klasse von Strukturen zu L, und die Theorie dieser Klasse sei $Th(K) = (L, \{C \text{ Satz aus } L | \mathcal{A}(C) = w \text{ für jedes } \mathcal{A} \in K\})$. Zeigen Sie:

a. $Th(K)$ ist maximal konsistent $\Leftrightarrow$ $K \neq \emptyset$ und $Th(K) = Th(\mathcal{A})$ für jedes $\mathcal{A} \in K$.

b. $Th(K)$ ist deduktiv abgeschlossen, d. h. für Sätze C aus L folgt aus $Th(K) \vdash C$ stets $C \in Ax(Th(K))$.

7.4.4 $\mathcal{A}$ sei eine Struktur zu L, so dass es zu jedem $c \in |\mathcal{A}|$ einen geschlossenen Term t aus L gibt mit $\mathcal{A}t = c$. Zeigen Sie:

a. $Th(\mathcal{A})$ ist saturiert.

b. $\mathcal{A}$ ist (bis auf Isomorphie) das kanonische Modell von $Th(\mathcal{A})$.

7.4.5 L sei eine Sprache mit mindestens einer Konstanten, genau einem einstelligen Prädikatszeichen p und keinem weiteren nicht-logischen Prädikatszeichen. $\Pi(L)$ sei die Menge aller Sätze pt und $\neg pt$. $Ax(T) \subseteq \Pi(L)$ enthalte keine Primformel zugleich mit ihrer Negation. T sei die Theorie $(L, Ax(T))$. Zeigen Sie für jeden Satz P aus $\Pi(L)$:

$$P \text{ gilt in } T \Leftrightarrow P \in Ax(T).$$

Hinweis: Gehen Sie für die Richtung $\Rightarrow$ analog zu 7.2.9 vor, indem Sie (insgesamt zwei) Modelle aus den geschlossenen Termen von L konstruieren.

§8 Vollständigkeit für beliebige Theorien

8.1 Zorns Lemma

8.2 Einfache Erweiterungen

8.3 Saturierung

8.4 Semantischer Beweis der Schnittregel

8.5 Aufgaben

Eine konsistente Theorie besitzt i. A. kein kanonisches Modell. Wenn man aber eine konsistente Theorie zu einer saturierten Theorie erweitern kann, ist deren kanonisches Modell im Wesentlichen schon ein Modell der Ausgangstheorie. Auf diese Weise wollen wir die Ergebnisse des vorigen Paragraphen für einen allgemeinen Vollständigkeitssatz nutzen.

8.1 Zorns Lemma

Einleitend formulieren wir Zorns Lemma, das das entscheidende nichtkonstruktive Mittel in unseren Überlegungen ist.

Gegeben sei eine beliebige Menge X. Dann bezeichnet

$$Pot(X) := \{Y \mid Y \subseteq X\}$$

die *Potenzmenge* von X. Ist $B \subseteq Pot(X)$, so bezeichnet

$$\cup B := \{x \in X \mid \exists Y \in B\ x \in Y\}$$

die *Vereinigung* von B. $K \subseteq Pot(X)$ ist eine *Kette*, wenn $K \neq \emptyset$ und wenn

$$Y \subseteq Z \text{ oder } Z \subseteq Y \text{ für alle } Y, Z \in K$$

ist. Ist $A \subseteq Pot(X)$ und $Y \in A$, so ist Y *maximal* in A, wenn für $Z \in A$ aus $Y \subseteq Z$ stets $Y = Z$ folgt, wenn also Y in keinem anderen Element von A enthalten ist. Y ist *größtes* Element von A, wenn für alle $Z \in A \quad Z \subseteq Y$ ist, wenn also Y alle Elemente von A enthält.

Ist $X \in A \subseteq Pot(X)$, so ist X sogar größtes Element von A. Ist $\emptyset \neq A \subseteq Pot(X)$ endlich, so ist jede Kette $K \subseteq A$ endlich. Ihre Vereinigung $\cup K$ ist dann ihr Maximum in K und damit in A. Die Maxima längster Ketten, die es wegen der Endlichkeit von A geben muss, sind dann maximal in A, auch wenn A kein größtes Element enthält. Die Übertragung dieses Sachverhalts auf den unendlichen Fall ist

8.1.1 Zorns Lemma Es sei A eine nicht-leere Teilmenge von $Pot(X)$. Enthält A mit jeder Kette $K \subseteq A$ auch deren Vereinigung $\cup K$, so gibt es maximale Elemente in A.

Zorns Lemma ist im Rahmen der Mengenlehre äquivalent zum Auswahlaxiom.

Den Beweis dieser Äquivalenz verlagern wir in die Aufgaben 8.5; er verwendet elementare Kenntnisse der Mengenlehre. Zorns Lemma sehen wir im folgenden als gültiges Prinzip an (wie das Auswahlaxiom). Die maximalen Elemente in A sind wie im endlichen Fall Vereinigungen längster Ketten, die im Allgemeinen aber unendlich lang sind. Die Existenz längster Ketten, die für endliches A selbstverständlich ist, ist für beliebiges (unendliches) A ebenso problematisch wie Zorns Lemma.

8.2 Einfache Erweiterungen

Wir wenden uns konsistenten Theorien und ihren Erweiterungen zu.

8.2.1 Definition Eine Sprache L' heißt *Erweiterung einer Sprache* L, wir schreiben

$$L \subseteq L' \text{ oder } L' \supseteq L,$$

wenn jedes Grundzeichen von L auch zu L' gehört. Ist ferner $\mathcal{A}$ eine Struktur zu L', so ist die *Beschränkung von* $\mathcal{A}$ *auf* L die Struktur

$$\mathcal{A}|L := (|\mathcal{A}|; (f_{\mathcal{A}})_{f \in L}; (p_{\mathcal{A}})_{p \in L}).$$

$\mathcal{A}|L$ ist die Struktur $\mathcal{A}$ ohne die Interpretation der Grundzeichen, die nicht in L, sondern nur in L' auftreten. $\mathcal{A}|L$ ist eine Struktur zu L.

Beispiel. Ist $\mathcal{A} = (R; 0, 1, -, +, \cdot)$ ein Ring mit 1, also ein Modell von T_R und ist L die Sprache der Gruppentheorie, wobei wir die Funktionszeichen $e, {}^{-1}, \circ$ mit $0, -, +$ identifizieren, so ist

$$\mathcal{A}|L = (R; 0, -, +)$$

nicht nur eine Struktur zur Sprache der Gruppentheorie, sondern sogar eine kommutative Gruppe. Dies ist ein Spezialfall des folgenden Lemmas.

8.2.2 Lemma Es sei $L' \supseteq L$ und A eine Struktur zu L'. Eine Formel B aus L gilt in $\mathcal{A}$ genau dann, wenn B in $\mathcal{A}|L$ gilt.

Beweis: Sätze B aus $L(\mathcal{A})$ werden in $\mathcal{A}$ und in $\mathcal{A}|L$ identisch interpretiert, weil die Interpretation von Zeichen aus L', die in B nicht auftreten, keinen Einfluss auf die Wahrheit von B hat. Also ist stets

$$\mathcal{A}|L(B) = \mathcal{A}(B).$$

Mit 2.3.7 folgt hieraus die Behauptung.

8.2.3 Definition Eine Theorie T' heißt *Erweiterung einer Theorie* T, wir schreiben

$$T \prec T' \text{ oder } T' \succ T,$$

wenn

a. $L(T) \subseteq L(T')$ und

b. $Ax(T) \subseteq Ax(T')$ ist.

Eine Erweiterung T' von T heißt *einfach*, wenn

$$L(T) = L(T')$$

ist, und sie heißt *konservativ*, wenn für jede Sequenz $\Gamma : \Delta$ aus $L(T)$

$$\text{aus } T' \vdash \Gamma : \Delta \text{ stets } T \vdash \Gamma : \Delta \text{ folgt.}$$

Eine einfache Erweiterung hat also dieselbe Sprache, aber eventuell mehr Axiome. Eine konservative Erweiterung hat eventuell eine reichere Sprache, aber alle Sequenzen der kleineren Sprache sind schon in der kleineren Theorie herleitbar, wenn sie in der größeren Theorie herleitbar sind.

Beispiel. Mit den Identifikationen wie oben ist die Theorie der Ringe mit 1 eine Erweiterung der Gruppentheorie. Diese Erweiterung ist nicht einfach, weil die Ringtheorie mehr Funktionszeichen enthält als die Gruppentheorie. Sie ist auch nicht konservativ. Denn in T_R sind

$$\exists x \neg x = 0 \text{ und } a + b = b + a$$

herleitbar, die nicht in jeder Gruppe gelten und deshalb auch nicht in der Gruppentheorie herleitbar sind.

8.2.4 Lemma Ist T' eine Erweiterung von T und $\mathcal{A}$ ein Modell von T', so ist $\mathcal{A}|L(T)$ ein Modell von T.

Beweis. Alle Axiome von T sind wegen $T \prec T'$ Axiome von T' und gelten deshalb in $\mathcal{A}$. Nach 8.2.2 gelten sie dann auch in $\mathcal{A}|L(T)$, und dies ist eine Struktur zu $L(T)$.

Wir kommen jetzt zu dem entscheidenden nicht-konstruktiven Schritt im Beweis des Vollständigkeitssatzes. Dies ist die Stelle, an der wir Zorns Lemma verwenden, um unter den einfachen Erweiterungen einer Theorie die maximal konstenten auszuwählen.

8.2.5 Satz von Lindenbaum Jede konsistente Theorie T besitzt eine maximal konsistente einfache Erweiterung T^+.

Beweis. Wir definieren eine Menge von Axiomensystemen

$$A := \{Ax(T') \mid Ax(T) \subseteq Ax(T') \subseteq L(T) \text{ und } T' \nvdash \Box\}.$$

A besteht also aus den Axiomensystemen aller einfachen, konsistenten Erweiterungen von T.

Als erstes zeigen wir, dass A die Voraussetzungen von Zorns Lemma erfüllt, wobei X die Menge der Sätze von $L(T)$ ist.

A ist nicht leer, weil $Ax(T) \in A$ ist.

Sei K eine Kette in A. Dann ist

$$Ax(T) \subseteq \cup K \subseteq L(T).$$

Angenommen, $(L(T), \cup K)$ wäre inkonsistent. Dann gibt es nach 5.4.3 endlich viele Axiome $A_i \in Ax(T_i)$ mit $Ax(T_i) \in K (i = 1, \ldots, n)$, so dass

$$(L(T), \{A_1, \ldots, A_n\}) \vdash \square$$

ist. Die Axiomensysteme $Ax(T_i)$ bilden eine endliche Teilkette von K, in der es ein größtes Axiomensystem, etwa $Ax(T_j)$ gibt. Dieses enthält alle Axiome $A_1, \ldots, A_n$. Dann ist

$$T_j \vdash \square$$

im Widerspruch zu $Ax(T_j) \in A$. Also ist die Annahme falsch, und $\cup K$ gehört zu A.

Mit Zorns Lemma folgt: Es gibt maximale Axiomensysteme $Ax(T^+)$ in A.

Die zugehörigen Theorien $T^+ = (L(T),\ Ax(T^+))$ sind einfache maximal konsistente Erweiterungen von T.
Damit ist der Satz von Lindenbaum bewiesen.

Bei der Anwendung von Zorns Lemma in diesem Beweis spielt die in 5.4 diskutierte Endlichkeit der Herleitungen eine entscheidende Rolle: Jede Herleitung in $(L(T), \cup K)$ verwendet nur endlich viele Axiome aus $\cup K$. Weil K eine Kette ist, gehören diese endlich vielen Axiome alle zu ein und demselben Axiomensystem aus K. Endlichkeitsargumente liegen vielen Anwendungen des Zornschen Lemmas zugrunde.

8.3 Saturierung

Eine saturierte Theorie hat ihr kanonisches Modell. Wenn man dieses Ergebnis für beliebige Theorien ausnutzen will, genügt der Satz von Lindenbaum nicht, der nur maximal konsistente Erweiterungen liefert. Saturierte Erweiterungen können nicht einfach sein, wenn die Ausgangstheorie nicht genügend viele geschlossene Terme enthält. Im Allgemeinen müssen wir die Ausgangssprache um Konstanten erweitern.

8.3.1 Definition Es sei L eine Sprache, T eine Theorie und E eine Menge von Konstanten. Dann bezeichnet $L + E$ die Sprache L, erweitert um die Konstanten aus E, und $T + E$ bezeichnet die Theorie $(L(T) + E,\ Ax(T))$.

8.3.2 Lemma über neue Konstanten Es sei T eine Theorie und E eine Menge von neuen Konstanten (die nicht schon zu $L(T)$ gehören). Dann ist

$$T \overset{n}{\vdash} \Gamma(a_1, \ldots, a_k) : \Delta(a_1, \ldots, a_k) \Leftrightarrow T + E \overset{n}{\vdash} \Gamma(e_1, \ldots, e_k) : \Delta(e_1, \ldots, e_k)$$

wenn $a_i,\ e_i$ nicht in Γ, Δ auftreten, paarweise verschieden sind und die Konstanten e_i aus E sind $(i = 1, \ldots, k)$.

Beweis: Die Richtung von links nach rechts folgt durch k Substitutionen (Subst).
Die Richtung von rechts nach links beweisen wir durch Induktion nach n. Wir schreiben a für $a_1, \ldots, a_k$ und e für $e_1, \ldots, e_k$.

1. $\Gamma(e) : \Delta(e)$ ist ein logisches Axiom. Dann ist auch $\Gamma(a) : \Delta(a)$ ein logisches Axiom.

2. $T + E \overset{n+1}{\vdash\!\!\!-} \Gamma(e) : \Delta(e)$, und der letzte Schluss ist der $(\forall S)$-Schluss

   $$\overset{n}{\vdash} \Gamma(e) : G(e, b), \Delta(e) \vdash \Gamma(e) : \forall x G(e, x), \Delta(e).$$

   Es sei c eine freie Variable, die nicht in Γ, Δ, G, a auftritt.
   $T + E \overset{n}{\vdash} \Gamma(e) : G(e, c), \Delta(e)$ folgt mit (Subst), also
   $T \overset{n}{\vdash} \Gamma(a) : G(a, c), \Delta(a)$ nach Induktionsvoraussetzung, also
   $T \overset{n+1}{\vdash\!\!\!-} \Gamma(a) : \forall x G(a, x), \Delta(a)$ mit einem $\forall S$-Schluss.

3. $T + E \overset{n+1}{\vdash\!\!\!-} \Gamma(e) : \Delta(e)$, und das ist hergeleitet aus $\Gamma_i(e) : \Delta_i(e)$ $(i = 1, 2$ oder $i = 1)$ mit einem anderen letzten Schluss. Nach Induktionsvoraussetzung ist $T \overset{n}{\vdash} \Gamma_i(a) : \Delta_i(a)$, woraus nach derselben Grundschlussregel

   $$T \vdash \Gamma(a) : \Delta(a)$$

   folgt, auch im Falle eines T-Schlusses, weil keine Konstante aus e in $Ax(T) = Ax(T + E)$ auftritt.

Mit Induktion nach n folgt die Behauptung.

Hiernach ist $T + E$ eine konservative Erweiterung von T:
Konstanten, deren Bedeutung durch Axiome von T nicht eingeschränkt wird, kann man gleichwertig durch freie Variablen ersetzen.

Im Folgenden assoziieren wir mit Sätzen $\exists x F(x)$ öfters neue Konstanten $\varepsilon x F(x)$. Als Konstanten sind dies unzerlegbare Grundzeichen einer Sprache. Wir lesen $\varepsilon x F(x)$ als „ein x, auf das F zutrifft“.

8.3.3 Satz Es sei T eine Theorie,

$E := \{\varepsilon x F(x) \mid \exists x F(x) \text{ ist Satz aus } L(T), \varepsilon x F(x) \text{ ist nicht in } L(T)\}$, und
$\Sigma := \{F(\varepsilon x F(x)) \mid \varepsilon x F(x) \in E,\ \exists x F(x) \in Ax(T)\}$.

Dann ist $T + E + \Sigma$ eine konservative Erweiterung von T.

Beweis: $\Gamma : \Delta$ sei eine Sequenz aus $L(T)$, und es sei

$$T + E + \Sigma \vdash \Gamma : \Delta.$$

Nach dem Deduktionstheorem gibt es endlich viele Nennformen $F_1, \ldots, F_k$, so dass

$$T + E \vdash F_1(\varepsilon x_1 F_1(x_1)), \ldots, F_k(\varepsilon x_k F_k(x_k)), \Gamma : \Delta$$

mit verschiedenen Konstanten $\varepsilon x_i F_i(x_i) \in E$, die in $\Gamma : \Delta$ nicht auftreten. Für verschiedene freie Variablen $a_1, \ldots, a_k$, die nicht in $F_1, \ldots, F_k, \Gamma : \Delta$ auftreten, ist dann nach 8.3.2

$$T \vdash F_1(a_1), \ldots, F_k(a_k), \Gamma : \Delta$$

woraus man mit k $(\exists A)$-Schlüssen schließt auf

$$T \vdash \exists x_1 F_1(x_1), \ldots, \exists x_k F_k(x_k), \Gamma : \Delta.$$

Da $\exists x_i F_i(x_i)$ für $i = 1, \ldots, k$ Axiome von T sind, folgt mit k T-Schlüssen

$$T \vdash \Gamma : \Delta.$$

Also ist $\Gamma : \Delta$ schon in T herleitbar, und $T + E + \Sigma$ ist eine konservative Erweiterung von T.

Dieses Ergebnis zeigt, wie man einen Schritt hin zu einer saturierten Theorie tun kann durch Hinzunahme von Konstanten. Allerdings haben wir in 7.3.1 die Saturiertheit über Allformeln definiert, während hier durch die ε-Konstanten in natürlicher Weise Existenzformeln ins Spiel kommen. Dieser Unterschied ist nicht wesentlich.

8.3.4 Lemma T sei maximal konsistent, und $\exists x F(x)$ sei ein Satz aus $L(T)$. Dann sind äquivalent:

(1) Ist $T \vdash F(s) : \emptyset$ für jeden geschlossenen Term s, so ist $T \vdash \exists x F(x) : \emptyset$.

(2) Ist $\exists x F(x) \in Ax(T)$, so gibt es einen geschlossenen Term s, für den $F(s) \in Ax(T)$ ist.

Beweis. Nach 7.2.2 sind die Voraussetzungen von (1) bzw. (2) äquivalent zur Negation der Behauptung von (2) bzw. (1). Also ist (1) äquivalent zur Kontraposition von (2) und damit zu (2).

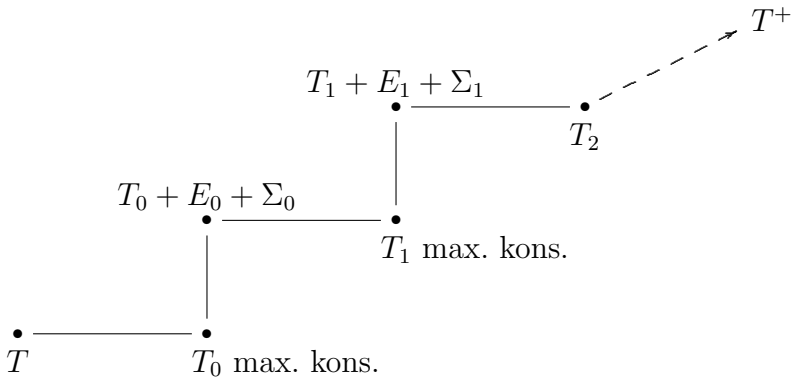
8.3.5 Korollar Eine maximal konsistente Theorie T ist genau dann saturiert, wenn es zu jedem Axiom $\exists x F(x)$ von T einen geschlossenen Term s gibt, für den $F(s)$ Axiom von T ist.

Beweis. Wegen 8.3.4 genügt es zu zeigen: Eine maximal konsistente Theorie T ist genau dann saturiert, wenn 8.3.4 (1) für jede geschlossene Nennform F gilt.
Aus der Saturiertheit folgt 8.3.4 (1) mit $(\neg S)$- und $(\neg A)$-Schlüssen. Umgekehrt folgt aus 8.3.4 (1) die Saturiertheit von T, wenn man zusätzlich beachtet:

$$T \vdash \forall x \neg\neg F(x) \Rightarrow T \vdash \forall x F(x).$$

Wir können nun die Verfahren 8.2.5 von Lindenbaum und aus 8.3.3 so miteinander koppeln, dass wir zu saturierten Erweiterungen kommen.

Wir fangen an mit einer konsistenten Theorie T. 8.2.5 liefert eine maximal konsistente Erweiterung T_0. 8.3.3 liefert die fehlenden Konstanten und führt zu einer konservativen Erweiterung $T_0 + E_0 + \Sigma_0$, die konsistent, aber nicht mehr maximal konsistent ist.



Wir wiederholen diesen Prozess unendlich oft und erhalten als Vereinigung aller dieser Theorien eine saturierte Erweiterung T^+.

8.3.6 Satz von Henkin Jede konsistente Theorie T besitzt eine saturierte Erweiterung T^+.

Beweis. $L(T)$ enthalte o.E. keine Konstante $\varepsilon x F(x)$. Wir definieren rekursiv zu jedem n eine maximal konsistente Erweiterung T_n von T.

1. T_0 sei eine maximal konsistente einfache Erweiterung von T, die es nach 8.2.5 gibt.

2. T_n sei bereits definiert. Dann sei

$$E_n := \{\varepsilon x F(x) \mid \exists x F(x) \text{ ist Satz aus} L(T_n) - L(T_{n-1})\}$$
$$(\text{für } n = 0 \text{ sei } L(T_{n-1}) \text{ leer}),$$

$$\Sigma_n := \{F(\varepsilon x F(x)) \mid \varepsilon x F(x) \in E_n, \exists x F(x) \in Ax(T_n)\}$$

   Nach 8.3.3 ist $T_n + E_n + \Sigma_n$ konsistent, weil dies eine konservative Erweiterung von T_n ist, und T_n ist nach Induktionsvoraussetzung (maximal) konsistent.
   T_{n+1} sei nun eine maximal konsistente einfache Erweiterung von $T_n + E_n + \Sigma_n$ nach 8.2.5.

3. T^+ sei die Vereinigung aller Theorien T_n:

$$T^+ := (\cup\{L(T_n) \mid n \in \mathbb{N}\}, \cup\{Ax(T_n) \mid n \in \mathbb{N}\}).$$

Damit ist die Konstruktion der Erweiterung T^+ abgeschlossen.
Wir wollen zeigen: T^+ ist saturiert.

(a) Wir nehmen an, es gäbe eine Herleitung H von $\square$ in T^+. Diese Herleitung besteht aus endlich vielen Sequenzen und verwendet nur endlich viele Axiome von T^+. Da die Theorien T_n eine aufsteigende Folge bilden,

$$T_n \prec T_{n+1},$$

deren Vereinigung T^+ ist, gibt es eine Theorie T_m mit genügend großem Index m, deren Sprache alle Sequenzen aus H und deren Axiomensystem alle in H verwendeten nicht-logischen Axiomen enthält. Dann ist H eine Herleitung von $\square$ in T_m, im Widerspruch zur Konstruktion von T_m.
Also ist unsere Annahme falsch, T^+ ist konsistent.

(b) Sei C ein Satz aus $L(T^+)$ und sei $T + \{C\}$ konsistent.
Dann gibt es ein n mit C aus $L(T_n)$, und $T_n + \{C\}$ ist konsistent. Da T_n maximal konsistent ist, folgt

$$C \in Ax(T_n) \subseteq Ax(T^+).$$

Also ist auch T^+ maximal konsistent.

(c) Sei $\exists x F(x)$ ein Axiom von T^+. Dann gibt es ein kleinstes n, so dass

$$\exists x F(x) \in Ax(T_n)$$

ist. Auch für $n > 0$ ist dann $\exists x F(x)$ keine Formel aus $L(T_{n-1})$, weil sonst $T_{n-1} + \{\exists x F(x)\}$ konsistent wäre, woraus mit der maximalen Konsistenz von T_{n-1}

$$\exists x F(x) \in Ax(T_{n-1})$$

folgte, im Widerspruch zur Minimalität von n. Also ist $\varepsilon x F(x) \in E_n$ und $F(\varepsilon x F(x)) \in \Sigma_n$, also auch

$$F(\varepsilon x F(x)) \in Ax(T_{n+1}) \subseteq Ax(T^+).$$

Nach 8.3.5 ist dann T^+ saturiert.

8.3.7 Vollständigkeitssatz von Gödel und Henkin
Jede konsistente Theorie besitzt ein Modell.

Beweis. Sei T konsistent. Nach dem Satz von Henkin 8.3.6 besitzt T eine saturierte Erweiterung T^+, und diese hat nach 7.3.5 ein kanonisches Modell $\mathcal{A}^+$.

Sei $\mathcal{A}$ die Struktur $\mathcal{A}^+|L(T)$; $\mathcal{A}$ ist also $\mathcal{A}^+$ ohne die Interpretation der Konstanten $\varepsilon x F(x)$. Da T^+ eine Erweiterung von T ist, ist $\mathcal{A}$ nach 8.2.4 ein Modell von T. Damit ist der Satz bewiesen.

Wegen 7.1.6 sind nun auch die anderen Formulierungen der Vollständigkeit bewiesen, insbesondere auch die Umkehrung der Korrektheit. Wir können also zusammenfassen:

8.3.8 Korollar Zusammenfassung von Korrektheit und Vollständigkeit. Für jede Theorie T und jede Sequenz $\Gamma : \Delta$ aus $L(T)$ gilt:

$$T \vdash \Gamma : \Delta \;\Leftrightarrow\; T \models \Gamma : \Delta.$$

Wir werfen einen Blick zurück auf die Überlegungen, die zu diesem zentralen Ergebnis geführt haben.

Die Voraussetzung, dass T konsistent ist, ist rein syntaktisch. Woher kommt das Modell von T, das nach dem Vollständigkeitssatz existiert? Nach dem Satz von Henkin kann man von T zu einer saturierten Erweiterung T^+ übergehen. Diese ist immer noch ein syntaktisches Objekt, besitzt aber das kanonische Modell, das im Wesentlichen auch ein Modell von T ist. Wie oben diskutiert, wird das kanonische Modell ausschließlich aus dem syntaktischen Material der Theorie T^+ konstruiert.

Eine Grundidee des Vollständigkeitsbeweises liegt also darin, dass wir ein *semantisches* Objekt (nämlich ein Modell von T) aus *syntaktischem* Material (nämlich aus der Theorie T^+) gewinnen.

Der Übergang von T^+ zum kanonischen Modell ist ein eindeutiger, konstruktiver Vorgang. Worin liegt das Abstrakte, Nicht-Konstruktive unserer Überlegung?
Es liegt allein in den Übergängen zu maximal konsistenten Erweiterungen, die nach dem Satz von Lindenbaum möglich sind. Dieser Satz garantiert nur die abstrakte Existenz einer, oft sehr vieler solcher Erweiterungen, gibt aber keine Vorschrift, nach der man auch nur eine einzige konstruieren könnte. Das ist die typische Wirkungsweise von Zorns Lemma. Ein wesentlicher Punkt ist also, dass das zu T existierende Modell i. a. nicht eindeutig bestimmt und *nicht konstruktiv angebbar* ist.

8.4 Semantischer Beweis der Schnittregel

Mit dem Vollständigkeitssatz wird es zwar nicht einfacher, Herleitungen zu konstruieren; wohl aber wird es leichter, die Herleitbarkeit von Formeln festzustellen: Alle semantisch korrekten Schlüsse sind nach 8.3.8 auch syntaktisch zulässig. Als wichtiges Beispiel hierfür zeigen wir die Zulässigkeit der Schnittregel.

8.4.1 Satz Die Schnittregel ist zulässig:

$$\text{Aus } T \vdash \Gamma, B : \Delta \text{ und } T \vdash \Gamma : B, \Delta \text{ folgt } T \vdash \Gamma : \Delta.$$

Beweis. Wegen des Korrektheitssatzes gelten $\Gamma, B : \Delta$ und $\Gamma : B, \Delta$ in jedem Modell $\mathcal{A}$ von T. Weil B unter jeder $\mathcal{A}$-Belegung in $\mathcal{A}$ wahr oder falsch werden

muss, gilt dann auch $\Gamma : \Delta$ in $\mathcal{A}$. Mit dem Vollständigkeitssatz folgt $T \vdash \Gamma : \Delta$.

Wir folgern hieraus rein syntaktisch die Zulässigkeit einiger verwandter Regeln.

8.4.2 Korollar Die *Kettenschlussregel* ist zulässig:

$$\text{Aus } \vdash \Gamma, A : B, \Delta \text{ und } \vdash \Gamma, B : C, \Delta \text{ folgt } \vdash \Gamma, A : C, \Delta.$$

Denn aus den Prämissen folgt mit Strukturschlüssen

$$\vdash \Gamma, A : B, C, \Delta \text{ und } \vdash \Gamma, A, B : C, \Delta.$$

Mit der Schnittregel folgt nun die Behauptung.

8.4.3 Korollar Der *modus ponens* ist zulässig.

$$\text{Aus } \vdash \Gamma, A : \Delta \text{ und } \vdash A \text{ folgt } \vdash \Gamma : \Delta.$$

Denn aus der zweiten Prämisse folgt $\vdash \Gamma : A, \Delta$ mit einem Strukturschluss, so dass mit der Schnittregel die Behauptung folgt.

8.4.4 Korollar Einfache Erweiterungen um herleitbare Sätze sind konservativ:

$$\text{Aus } T + \{C\} \vdash \Gamma : \Delta \text{ und } T \vdash C \text{ folgt } T \vdash \Gamma : \Delta.$$

Denn nach dem Deduktionstheorem 5.4.2 folgt aus der ersten Voraussetzung

$$T \vdash C, \Gamma : \Delta.$$

Mit dem modus ponens folgt nun $T \vdash \Gamma : \Delta$.

8.5 Aufgaben

Zorns Lemma hat zahlreiche Anwendungen in vielen Teilgebieten der Mathematik. Wir führen hier einige Anwendungen auf, die besonders einfache Beweise haben und keine weiterführende Theorie verwenden. In Klammern geben wir das Teilgebiet der Mathematik an, zu dem das Ergebnis gehört. Dessen Grundbegriffe werden als bekannt vorausgesetzt.

8.5.1 (Lineare Algebra)
Beweisen Sie mit Zorns Lemma:

Jeder Vektorraum hat eine Basis.

Hinweis: Zu gegebenem Vektorraum V betrachten Sie die Menge

$$A := \{B \subseteq V \mid B \text{ ist linear unabhängige Menge von Vektoren von } V\}.$$

Die maximalen Elemente von A sind im Wesentlichen die Basen von V.

8.5.2 (Ringtheorie)
Beweisen Sie mit Zorns Lemma:

> Sei I ein Ideal eines kommutativen Ringes R mit 1 und $x \in R - I$. Dann ist I enthalten in einem maximalen Ideal M von R, das x nicht enthält.

Insbesondere ist jedes echte Ideal von R enthalten in einem maximalen Ideal.

Hinweis: Betrachten Sie die Menge

$$A := \{J \subseteq R \mid J \text{ ist Ideal von } R, I \subseteq J, x \notin J\}.$$

Die maximalen Elemente von A sind die gesuchten maximalen Ideale.

8.5.3 (Geordnete Körper)
Sei K ein kommutativer Körper. Ein *Q-Bereich* von K ist eine Teilmenge Q von K mit den Eigenschaften

(1) Für alle Elemente $x \neq 0$ von K ist $x^2 \in Q$.

(2) Sind $x, y \in Q$, so sind auch $x + y \in Q$ und $x \cdot y \in Q$.

(3) $0 \notin Q$.

Beweisen Sie mit Zorns Lemma: Ist Q ein Q-Bereich von K und $z \in K - Q$, so gibt es eine Anordnung $<$ von K, bei der alle $x \in Q$ positiv sind $(0 < x)$, aber $z = 0$ oder $z < 0$ ist.

Folgern Sie daraus: Jeder formal-reelle Körper besitzt eine Anordnung.

Hinweis: Betrachten Sie die Menge

$$A := \{Q' \subseteq K \mid Q' \text{ ist } Q\text{-Bereich von } K, Q \subseteq Q', z \notin Q'\}.$$

Die maximalen Elemente von A sind die Positivbereiche P, die die gesuchten Anordnungen durch $x < y \Leftrightarrow y - x \in P$ definieren.

8.5.4 (Mengenlehre)
Beweisen Sie mit Zorns Lemma die *Vergleichbarkeit von Mächtigkeiten*:

> Zu je zwei Mengen X und Y gibt es eine injektive Abbildung von X in Y oder eine injektive Abbildung von Y in X.

Hinweis: Betrachten Sie die Menge der *partiellen* Injektionen von X in Y, also die Menge

$$A := \{\varphi : X_0 \to Y \mid X_0 \subseteq X, \varphi \text{ injektiv}\},$$

und identifizieren Sie Abbildungen mit ihren Graphen. Die maximalen Elemente von A oder ihre Umkehrabbildungen sind dann die gesuchten Injektionen.

8.5.5 (Mengenlehre)
Das *Auswahlaxiom* ist die Aussage:

(AC) Es sei $R \subseteq X \times Y$, und zu jedem $x \in X$ gebe es ein $y \in Y$, so dass $(x, y) \in R$ ist. Dann gibt es eine Funktion $\varphi : X \to Y$, so dass für alle $x \in X$ $(x, \varphi(x)) \in R$ ist.

Zeigen Sie: Aus Zorns Lemma folgt das Auswahlaxiom AC.

Hinweis: Betrachten Sie die Menge der *partiellen* Auswahlfunktionen in R, also die Menge

$$A := \{\varphi : X_0 \to Y \mid X_0 \subseteq X, \text{ für alle } x \in X_0 \text{ ist } (x, \varphi(x)) \in R\}$$

und identifizieren Sie wieder Abbildungen mit ihren Graphen. Die maximalen Elemente von A sind dann die gesuchten Auswahlfunktionen.

8.5.6 (Mengenlehre)
Sei $\sim$ eine Äquivalenzrelation auf einer Menge X. Ein *Repräsentantensystem* von $\sim$ ist eine Teilmenge Y von X, die von jeder Äquivalenzklasse $\bar{x} = \{y \in X \mid x \sim y\}$ von $\sim$ genau ein Element enthält.
Zeigen Sie: Das Auswahlaxiom AC ist äquivalent dazu, dass jede Äquivalenzrelation ein Repräsentantensystem besitzt.

Um aus dem Auswahlaxiom AC zurück auf Zorns Lemma zu schließen, braucht man den Begriff der Ordinalzahl und das Prinzip der transfiniten Rekursion, die wir hier voraussetzen.

8.5.7 (Mengenlehre)
Eine *Wohlordnung* ist eine lineare Ordnung $(A, \prec)$ mit der Eigenschaft:

> Jede nicht-leere Teilmenge von A besitzt ein bezüglich $\prec$ kleinstes Element.

Man nennt $\prec$ dann auch eine *Wohlordnung von* A. Ein einfaches Ergebnis der Mengenlehre besagt: Eine (nicht-leere) Menge A besitzt eine Wohlordnung genau dann, wenn es eine Ordinalzahl α und eine Bijektion φ von
$(0 \neq)\alpha = \{\beta \mid \beta < \alpha\}$ auf A gibt.
Zeigen Sie: Aus dem Auswahlaxiom AC folgt der Wohlordnungssatz

(WO) Jede nicht-leere Menge besitzt eine Wohlordnung.

8.5.8 (Mengenlehre)
Zeigen Sie: Aus dem Wohlordnungssatz WO folgt das Auswahlaxiom AC.

Mit 8.5.5, 7 und 8 ist die Äquivalenz von Zorns Lemma, Auswahlaxiom AC und Wohlordnungssatz WO im Rahmen der Mengenlehre gezeigt. Der Streit, welches dieser drei Prinzipien die größte anschauliche Überzeugungskraft besitzt, soll hier nicht entschieden werden. Die folgende Aufgabe bietet eine Alternative zu unserem Beweis des Satzes von Lindenbaum.

8.5.9 Beweisen Sie den Satz von Lindenbaum unter Verwendung des Wohlordnungssatzes WO, aber ohne Rückgriff auf Zorns Lemma.
Hinweis: Wegen WO gibt es eine Ordinalzahl α, so dass $\{C_\beta \mid \beta < \alpha\}$ die Menge aller Sätze der Sprache der gegebenen konsistenten Theorie T ist.
Definieren Sie durch transfinite Rekursion eine Folge $(Ax_\beta)_{\beta<\alpha}$ von Axiomensystemen, so dass

$$T^+ := (L(T), \bigcup\{Ax_\beta \mid \beta < \alpha\})$$

eine maximal konsistente einfache Erweiterung von T ist.

Zum Schluss eine leichte Ergänzung von 8.4.4.

8.5.10 Sei C ein Satz aus $L(T)$ und $T \vdash C$.

a. Ist T konsistent, so ist $T + \{C\}$ konsistent.

b. Ist T maximal konsistent, so ist $C \in Ax(T)$.

§9 Vollständigkeit für abzählbare identitätsfreie Theorien

9.1 Königs Lemma

9.2 D-Bäume und D-Fäden

9.3 Syntaktisches und semantisches Hauptlemma

9.4 Modelle über den natürlichen Zahlen

9.5 Aufgaben

Im vorigen Paragraphen haben wir den Vollständigkeitssatz für beliebige mathematische Theorien bewiesen. Dabei haben wir an entscheidender Stelle Zorns Lemma als nicht konstruktives Beweismittel verwendet. Für die Theorien, die in der mathematischen Praxis auftreten (vgl. §1.2), ist dies eine sehr scharfe Waffe.
Wir stellen deshalb einen methodisch anders gearteten Beweis des Vollständigkeitssatzes dar, der sich allerdings nur auf die große Klasse der abzählbaren identitätsfreien Theorien bezieht. In diesem Beweis übernimmt Königs Lemma die Rolle, die Zorns Lemma in §8 hat. Königs Lemma ist zwar auch nicht im strengen Sinne konstruktiv, aber es ist ein elementares Beweismittel, das schon mit klassisch arithmetischen Mitteln leicht zu beweisen ist, wie wir in 9.1 zeigen.
Dieser Beweis des Vollständigkeitssatzes nutzt die genaue Form des Herleitungsbegriffs aus §3.1 optimal aus. Dadurch wird er in seiner Grundstruktur besonders einprägsam.

9.1 Königs Lemma

Wir betrachten endliche Folgen $c = \langle c_0, \ldots, c_{m-1} \rangle$ $(m \geq 0)$ von natürlichen Zahlen c_i $(i < m)$.

9.1.1 Definition Für zwei endliche Folgen $c = \langle c_0, \ldots, c_{m-1} \rangle$ und $d = \langle d_0, \ldots, d_{n-1} \rangle$ ist

$$c \leq d,$$

wenn $m \leq n$ und $c_i = d_i$ ist für $i < m$, wenn also c ein *Anfangsstück* der Folge d ist, und die *Verkettung* von c und d ist die endliche Folge

$$c * d := \langle c_0, \ldots, c_{m-1}, d_0, \ldots, d_{n-1} \rangle,$$

die durch Hintereinanderschreiben der Folgen c und d entsteht. Ein *Baum* ist eine nicht-leere Menge B von endlichen Folgen mit der Eigenschaft:

Aus $d \in B$ und $c \leq d$ folgt $c \in B$.

Die Elemente c eines Baumes B nennt man auch die *Knoten* von B, die leere Folge $\langle\rangle$ seine *Wurzel* und (bzgl. $\leq$) maximale Knoten von B die *Blätter* von B. Ein Baum B ist *endlich verzweigt*, wenn es zu jedem Knoten $c \in B$ nur endlich viele *Nachfolgerknoten* $c * \langle i \rangle \in B$ gibt. B ist ein *binärer Baum*, wenn in allen Knoten $c = \langle c_0, \ldots, c_{m-1} \rangle \in B$ alle $c_i < 2$ sind $(i < m)$.

Bemerkung Die Relation $\leq$ ist offenbar eine partielle Ordnungsrelation. Man kann sie mit Hilfe der Verkettung charakterisieren. Denn man erhält alle endlichen Folgen mit Anfangsstück c, wenn man c mit beliebigen endlichen Folgen verkettet: Es ist

$$c \leq d \iff \text{es gibt eine endliche Folge } e \text{ mit } c * e = d.$$

Ein Blatt eines Baumes B ist dann gerade ein Knoten $c \in B$, für den keine endliche Folge $c * \langle i \rangle$ noch zu B gehört.
Jeder Baum enthält die Wurzel $\langle\rangle$, weil er nicht leer ist und stets $\langle\rangle \leq d$ ist.
Jeder binäre Baum ist endlich verzweigt.

9.1.2 Beispiele Es gibt einen kleinsten Baum $\{\langle\rangle\}$, der nur aus der leeren Folge besteht, und einen größten Baum, der aus allen endlichen Folgen besteht. Der kleinste Baum $\{\langle\rangle\}$ ist in jedem Baum enthalten, er ist endlich, endlich verzweigt und (trivialerweise) auch binär; sein einziger Knoten ist seine Wurzel und zugleich sein einziges Blatt. Der größte Baum enthält jeden Baum, er ist unendlich und unendlich verzweigt.
Dieser größte Baum wächst unendlich in die Breite: Direkt über jedem Knoten $c = \langle c_0, \ldots, c_{m-1} \rangle$ liegen die unendlich vielen Knoten $c * \langle i \rangle = \langle c_0, \ldots, c_{m-1}, i \rangle$ $(i \in \mathbb{N})$. Insbesondere hat er keine Blätter. Er wächst auch unendlich in die Höhe, die meist *Tiefe* genannt wird. Für jede Funktion $\varphi : \mathbb{N} \to \mathbb{N}$ kommen alle Folgen

$$\langle c_0, \ldots, c_{m-1}, \varphi(0), \ldots, \varphi(n-1) \rangle \text{ nach } \langle c_0, \ldots, c_{m-1} \rangle,$$

und die unendliche Menge dieser Folgen ist durch $\leq$ linear geordnet.

9.1.3 Definition Ein *Faden* eines Baumes B ist eine Folge

$$\langle\rangle, \langle c_0\rangle, \langle c_0, c_1\rangle, \langle c_0, c_1, c_2\rangle, \ldots$$

von Elementen von B, die entweder unendlich lang ist oder als letztes Glied ein Blatt von B enthält. Ein Baum ist *fundiert*, wenn alle seine Fäden endlich sind.

9.1.4 Beispiele 1. Wir schreiben 0^m für die endliche Folge $\langle 0, \ldots, 0\rangle$ aus m Nullen. Der *linke Standardfaden*

$$\{0^m \mid m \in \mathbb{N}\}$$

ist ein unverzweigter, also endlich verzweigter, nicht fundierter Baum, der zu $\mathbb{N}$ ordnungsisomorph ist.

2. Der *Kamm*

$$\{0^m \mid m \in \mathbb{N}\} \cup \{0^m * \langle 1\rangle \mid m \in \mathbb{N}\}$$

ist ein binärer, also endlich verzweigter Baum, der unendlich viele endliche Fäden und einen unendlichen Faden enthält und deshalb nicht fundiert ist.

3. Der *universelle binäre Baum* ist

$$F_0 = \{\langle c_0, \ldots, c_{m-1}\rangle \mid m \in \mathbb{N}, c_i < 2\} \text{ für } i < m.$$

Er ist binär verzweigt und enthält kontinuierlich viele, also überabzählbar viele unendlich lange Fäden, ist daher nicht fundiert.

4. Ein Baum der Tiefe 1 ist

$$\{\langle\rangle, \langle i\rangle \mid i \in \mathbb{N}\}.$$

Alle Fäden in diesem Baum haben die Länge 1. Der Baum ist also fundiert, aber er ist im Knoten $\langle\rangle$ unendlich verzweigt.

5. Ein Baum der Tiefe ω ist

$$\{\langle\rangle, \langle i\rangle * 0^m \mid i \in \mathbb{N}, m \leq i\}.$$

Seine sämtlichen Fäden sind endlich, sie wachsen aber unbeschränkt. Der Baum ist fundiert und unendlich verzweigt.

Es liegt auf der Hand, dass jeder nicht fundierte und auch jeder unendlich verzweigte Baum unendlich viele Knoten hat. Die Umkehrung hiervon, dass bei jedem unendlichen Baum einer dieser beiden Fälle vorliegt, ist die Aussage von Königs Lemma.

9.1.5 Königs Lemma Jeder endlich verzweigte, unendliche Baum enthält einen unendlichen Faden.

Beweis: Es sei B ein endlich verzweigter, unendlicher Baum. Wir betrachten die Knoten $c \in B$, die in B unendlich viele Nachfolger haben, für die also gilt:

$$(*) \text{ es gibt unendlich viele } d \in B \text{ mit } c \leq d.$$

a. Da B unendlich ist, gilt $(*)$ für die leere Folge $\langle\rangle$.

b. Wenn ein Knoten c $(*)$ erfüllt, dann gibt es ein i, so dass $c * \langle i\rangle$ $(*)$ erfüllt.
Denn jedes $d \neq c$ mit $c \leq d$ ist ein Nachfolger eines $c * \langle i\rangle \in B$, von denen es nur endlich viele gibt, weil B endlich verzweigt ist. Hätte jedes $c * \langle i\rangle \in B$ nur endlich viele, etwa n_i, Nachfolger, so hätte auch c nur endlich viele Nachfolger, nämlich $1 + \Sigma_i n_i$.

c. Wir definieren eine Funktion $\varphi : \mathbb{N} \to B$ durch

$$\varphi(0) = \langle\rangle,$$
$$\varphi(k+1) = \varphi(k) * \langle i\rangle,$$

wobei i die kleinste Zahl ist, für die $\varphi(k) * \langle i\rangle$ $(*)$ erfüllt, falls es ein solches i gibt.

Wegen a. gilt $(*)$ für $\varphi(0)$, und wegen b. gilt $(*)$ für $\varphi(k+1)$, falls $(*)$ für $\varphi(k)$ gilt. Mit vollständiger Induktion nach k folgt, dass $\varphi(k)$ für jedes k $(*)$ erfüllt. Insbesondere gilt

$$\varphi(k) \in B \text{ für jedes } k \in \mathbb{N};$$
$$\varphi = \varphi(0), \varphi(1), \varphi(2), \ldots$$

ist ein unendlicher Faden in B.

Bemerkung Die Funktion φ im obigen Beweis ist nicht im strengen Sinne konstruktiv definiert, weil die Eigenschaft $(*)$ i. a. nicht entscheidbar ist. Sie ist aber unter Rückgriff auf die Eigenschaft $(*)$ eindeutig definiert, ohne von einer freien, „unbestimmten“ Auswahl Gebrauch zu machen.
Königs Lemma ist eine elementare Aussage, weit schwächer als das Auswahlaxiom oder Zorns Lemma.

9.2 D-Bäume und D-Fäden

Eine Menge M ist *abzählbar*, wenn es eine injektive Abbildung von M in die Menge der natürlichen Zahlen gibt. M ist *abzählbar unendlich*, wenn diese Abbildung bijektiv gewählt werden kann. Eine abzählbare Menge ist also entweder endlich oder abzählbar unendlich. Wenn sich eine Menge genau wie die Menge der natürlichen Zahlen ordnen lässt, ist sie abzählbar unendlich.

9.2.1 Definition Eine Sprache L ist *abzählbar*, wenn sie nur abzählbar viele nicht-logische Grundzeichen enthält. Eine Theorie ist *abzählbar*, wenn ihre Sprache abzählbar ist.

Alle bisher aufgeführten Theorien (vgl. §1.2) sind abzählbar. Die Sprache $L(\mathbb{R})$ (im Anschluss an 2.1.2) ist nicht abzählbar.

9.2.2 Lemma In einer abzählbaren Sprache gibt es abzählbar unendlich viele Terme und abzählbar unendlich viele Formeln.

Beweis: Jede Sprache L enthält abzählbar unendlich viele freie Variablen

$$a_0, a_1, a_2, \ldots$$

und abzählbar unendlich viele gebundene Variablen

$$x_0, x_1, x_2, \ldots.$$

1. Fall. L enthält nur endlich viele nicht-logische Grundzeichen $z_0, \ldots, z_{k-1}$ $(k \geq 0)$. Dann ist

$$\bot, \rightarrow, \forall, =, z_0, \ldots, z_{k-1}, a_0, x_0, a_1, x_1, a_2, x_2, \ldots$$

eine Abzählung aller Grundzeichen.

2. Fall. L enthält abzählbar unendlich viele nicht-logische Grundzeichen $z_0, z_1, z_2, \ldots$. Dann ist

$$\bot, \to, \forall, =, a_0, x_0, z_0, a_1, x_1, z_1, a_2, x_2, z_2, \ldots$$

eine Abzählung aller Grundzeichen.
In jedem Fall gibt es abzählbar unendlich viele Grundzeichen $g_1, g_2, \ldots$.
Jeder Term und jede Formel ist eine endliche Reihe von Grundzeichen g_n. Da es unendlich viele Primzahlen

$$p_0 = 2, p_1 = 3, p_2 = 5, \ldots$$

gibt, können wir jeder endlichen Reihe von Grundzeichen

$$g_{n_0} \ldots g_{n_k} \text{ die Zahl } p_0^{n_0} \cdot \ldots \cdot p_k^{n_k}$$

zuordnen, und diese Zuordnung ist injektiv wegen der Eindeutigkeit der Primfaktorzerlegung und weil $n_i > 0$ ist. Also ist sowohl die Menge der Terme als auch die Menge der Formeln abzählbar, und offenbar sind sie nicht endlich.
Dieser Beweis ist eine Variante des zweiten Cantorschen Diagonalverfahrens, nach dem $\mathbb{N} \times \mathbb{N}$ eine abzählbare Menge ist.

9.2.3 Definition Eine Formel oder eine Sequenz ist *identitätsfrei*, wenn in ihr das Gleichheitszeichen $=$ nicht auftritt. Eine Theorie T ist *identitätsfrei*, wenn alle ihre nicht-logischen Axiome identitätsfrei sind.

Beispiel Lässt man das Axiom $LO3$ aus den Theorien LO und DLO in 1.2.5 fort, so erhält man abzählbare identitätsfreie Theorien. Sonst sind die bisher aufgeführten Theorien nicht identitätsfrei. Man kann allerdings mit einigem Aufwand jede Theorie in eine identitätsfreie Theorie umformen.

9.2.4 Bezeichnungen Im folgenden bezeichne T stets eine abzählbare identitätsfreie Theorie. O. E. sei keine natürliche Zahl ein Term von $L(T)$. Es sei $a_0, a_1, a_2, \ldots$ eine Abzählung (ohne Wiederholungen) der freien Variablen, $t_0, t_1, t_2, \ldots$ eine Abzählung (ohne Wiederholungen) der Terme von $L(T)$ und $A_0, A_1, A_2, \ldots$ eine Abzählung (eventuell mit Wiederholungen) der Axiome von T. Besteht $Ax(T)$ nur aus endlich vielen Sätzen (z. B. wenn $Ax(T)$ leer ist), so braucht es nur endlich viele A_i zu geben.
$\gamma, \delta, \pi, \ldots$ bezeichne endliche *Folgen* von Formeln aus $L(T)$ (im Gegensatz

zu den endlichen Formelmengen $\Gamma, \Delta, \ldots$). Insbesondere verwenden wir π (auch mit Indizes) zur Bezeichnung endlicher Folgen von Primformeln. Ist $\gamma \equiv C_1, \ldots, C_m$ und $\delta \equiv D_1, \ldots, D_n$, so bezeichnet γ, δ die endliche Folge $C_1, \ldots, C_m, D_1, \ldots, D_n$, und $|\gamma|$ steht für die Formelmenge $\{C_1, \ldots, C_m\}$.

Bemerkung In einer endlichen Folge γ von Formeln kann ein und dieselbe Formel offenbar mehrfach auftreten; die verschiedenen Auftreten dieser Formel in γ sind dann voneinander zu unterscheiden. Die Folge C, C ist von der eingliedrigen Folge C verschieden, ebenso ist die Folge γ, δ i. a. von der Folge δ, γ verschieden, obwohl

$$|C, C| = |C| = \{C\} \text{ und } |\gamma, \delta| = |\delta, \gamma| = |\gamma| \cup |\delta|$$

ist.

9.2.5 Definition Eine *geordnete Sequenz* S ist ein Paar $\gamma : \delta$ von endlichen Folgen γ und δ von Formeln aus $L(T)$. $|S|$ bezeichnet dann die (gewöhnliche) Sequenz $|\gamma| : |\delta|$. Ein Auftreten einer Formel C in S heißt *ausgezeichnet*, wenn

1. C keine Primformel ist und

2. rechts von diesem Auftreten von C in S nur noch Primformeln stehen.

Unser Ziel ist es, zu einer gegebenen identitätsfreien geordneten Sequenz S einen Baum aus geordneten Sequenzen zu konstruieren, der

(i) im wesentlichen eine Herleitung von $|S|$ ist, falls $|S|$ überhaupt in T herleitbar ist;

(ii) andernfalls ein Modell von T liefert, in dem $|S|$ nicht gilt.

Dieser Baum wird, ausgehend von S, rekursiv durch Rückschreiten gemäß den Grundschlussregeln $(\rightarrow S), (\rightarrow A), (\forall S), (\forall A)$ konstruiert. Damit der Grundschluss, über den wir zum Aufbau des Baumes zurückschreiten, eindeutig fixiert ist, betrachten wir geordnete Sequenzen, die wir von rechts nach links aufarbeiten. Wegen seiner Ähnlichkeit mit einer Herleitung von $|S|$ in T wird dieser Baum Deduktions-Baum oder D-Baum von S in T genannt.
Formal besteht ein D-Baum aus einem binären Baum B, dessen Knoten geordnete Sequenzen angeheftet oder zugeordnet sind, also aus B und einer Funktion, die jedem Knoten $c \in B$ eine geordnete Sequenz S_c zuordnet.

9.2.6 Rekursive Definition der *D-Bäume*. Zu jeder identitätsfreien geordneten Sequenz S von $L(T)$ definieren wir rekursiv einen binären Baum B, den *Indexbaum* von S, und eine Funktion $c \mapsto S_c$ von B in die Menge der geordneten Sequenzen von $L(T)$, den *D-Baum* von S in T:

0. $\langle\rangle \in B$, und $S_{\langle\rangle}$ ist S.

   Sei nun $c = \langle c_0, \ldots, c_{n-1}\rangle \in B$.

1. Ist $|S_c|$ ein logisches Axiom, so ist c ein Blatt von B, es gibt kein $c * \langle i\rangle \in B$.

2. Ist $|S_c|$ kein logisches Axiom, so ist $c * \langle 0\rangle \in B$ und im Fall 2.3 auch $c * \langle 1\rangle \in B$, und $S_{c*\langle i\rangle}$ hat folgende Gestalt:

   2.1 S_c besteht nur aus Primformeln, $S_c \equiv \pi : \pi'$. Dann ist

   $$S_{c*\langle 0\rangle} \equiv A_n, \pi : \pi'$$

   2.2 S_c ist $\gamma : \delta, C \to D, \pi$ mit ausgezeichnetem $C \to D$. Dann ist

   $$S_{c*\langle 0\rangle} \equiv A_n, \gamma, C : \delta, D, \pi$$

   2.3 S_c ist $\gamma, D \to C, \pi : \pi'$ mit ausgezeichnetem $D \to C$. Dann ist

   $$S_{c*\langle 0\rangle} \equiv A_n, \gamma, \pi : D, \pi' \text{ und } S_{c*\langle 1\rangle} \equiv A_n, \gamma, C, \pi : \pi'.$$

   2.4 S_c ist $\gamma : \delta, \forall x F(x), \pi$ mit ausgezeichnetem $\forall x F(x)$. Dann ist

   $$S_{c*\langle 0\rangle} \equiv A_n, \gamma : \delta, F(a_i), \pi,$$

   wobei a_i die erste freie Variable ist, die in S_c nicht auftritt.

   2.5 S_c ist $\gamma, \forall x F(x), \pi : \pi'$ mit ausgezeichnetem $\forall x F(x)$. Dann ist

   $$S_{c*\langle 0\rangle} \equiv A_n, \forall x F(x), \gamma, F(t_i), \pi : \pi',$$

   wobei t_i der erste Term ist, so dass $F(t_i)$ keine Antezedensformel von

   $$S \equiv S_{\langle\rangle}, S_{\langle c_0\rangle}, S_{\langle c_0, c_1\rangle}, \ldots, S_c$$

   ist (falls $*_1$ in F überhaupt vorkommt).

Falls die Aufzählung der Axiome von T (gemäß 9.2.4) vor dem Index n abbricht, es also kein Axiom A_n gibt, entfällt in 2.1 bis 2.5 jeweils das A_n in der Definition von $S_{c*\langle i\rangle}$.

9.2.7 Beispiel Wir konstruieren den D-Baum der geordneten Sequenz

$$S :\equiv \forall x(px \to qx) : \forall xpx \to \forall xqx$$

in einer logischen Theorie T unter der Voraussetzung, dass a sowohl die erste freie Variable als auch der erste Term nach 9.2.4 ist. Zur Abkürzung setzen wir

$$B :\equiv \forall x(px \to qx) \text{ und } C :\equiv \forall xpx.$$

Dann ist $S_{\langle\rangle} \equiv S$ nach 0. in 9.2.6,

$$\begin{aligned}
&S_{\langle 0\rangle} \equiv B, C : \forall xqx \quad \text{nach 2.2,}\\
&S_{\langle 0,0\rangle} \equiv B, C : qa \quad \text{nach 2.4,}\\
&S_{\langle 0,0,0\rangle} \equiv C, B, pa : qa \quad \text{nach 2.5,}\\
&S_{\langle 0,0,0,0\rangle} \equiv B, C, pa \to qa, pa : qa \quad \text{nach 2.5,}\\
&S_{\langle 0,0,0,0,0\rangle} \equiv B, C, pa : pa, qa \quad \text{und}\\
&S_{\langle 0,0,0,0,1\rangle} \equiv B, C, qa, pa : qa \quad \text{nach 2.3.}
\end{aligned}$$

Nun sind $|S_{\langle 0,0,0,0,i\rangle}|$ für $i = 0, 1$ logische Axiome, und damit sind $\langle 0,0,0,0,i\rangle$ für $i = 0, 1$ Blätter des Indexbaumes von S nach 1. Also ist hiermit der D-Baum von S vollständig angegeben. Offenbar ist

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{|S_{\langle 0,0,0,0,0\rangle}| \qquad |S_{\langle 0,0,0,0,1\rangle}|}{|S_{\langle 0,0,0,0\rangle}|}}{|S_{\langle 0,0,0\rangle}|}}{|S_{\langle 0,0\rangle}|}}{|S_{\langle 0\rangle}|}}{|S_{\langle\rangle}|}$$

eine (logische) Herleitung von S.
Wäre die erste freie Variable a nicht der erste Term t_0 nach 9.2.4, sondern etwa t_1 oder t_3, so würde der D-Baum von S komplizierter: Man hätte mit t_0 zuerst ein „falsches Beispiel“ gewählt.

9.2.8 Definition Ist $\langle\rangle, \langle c_0\rangle, \langle c_0, c_1\rangle, \ldots$ ein Faden im Indexbaum B einer identitätsfreien geordneten Sequenz S, so ist die endliche oder unendliche Folge

$$S_{\langle\rangle}, S_{\langle c_0\rangle}, S_{\langle c_0,c_1\rangle}, \ldots$$

von geordneten Sequenzen ein *D-Faden (Deduktions-Faden)* von S in T.

Der D-Baum aus dem Beispiel 9.2.7 enthält zwei D-Fäden, die beide endlich sind (mit der Länge 5) und beide mit logischen Axiomen enden. Ein D-Faden ist überhaupt nur dann endlich, wenn er ein logisches Axiom enthält, mit dem er nach 9.2.6, 1. dann auch endet. Demnach ist ein D-Baum genau dann fundiert, wenn jeder seiner D-Fäden ein logisches Axiom enthält.
Ein D-Baum kann unendliche D-Fäden enthalten, er kann sogar aus einem einzigen unendlichen D-Faden bestehen.

9.2.9 Beispiel S sei die geordnete Sequenz $pt : \forall xpx$, T sei eine logische Theorie. Dann ist:

$S_{\langle\rangle} \equiv S$ nach 0. in 9.2.6,

$S_{\langle 0\rangle} \equiv pt : pa$ nach 2.4,

wobei a die erste freie Variable ist, die in t nicht auftritt. Nach 2.1 in 9.2.6 ist dann $S_{\langle 0,0\rangle} \equiv S_{\langle 0\rangle}$ und schließlich $S_{0^m} \equiv S_{\langle 0\rangle} \equiv pt : pa$ für alle $m > 0$.
Der Indexbaum B von S besteht also nur aus dem linken Standardfaden $\{0^m \mid m \in \mathbb{N}\}$ (vgl. 9.1.4, 1). Der D-Baum von S ist identisch mit seinem einzigen D-Faden, der unendlich ist und kein logisches Axiom enthält. Die Sequenz $|S|$ ist (logisch) auch nicht herleitbar.

Wir verwenden im nächsten Abschnitt Königs Lemma in folgender Form:

9.2.10 Königs Lemma für D-Bäume Der D-Baum einer identitätsfreien geordneten Sequenz S in T ist endlich, oder S besitzt einen unendlichen D-Faden in T.

Beweis: Der Indexbaum B von S ist als binärer Baum endlich verzweigt. Ist er endlich, so ist der D-Baum von S endlich. Ist er unendlich, so enthält er nach Königs Lemma 9.1.5 einen unendlichen Faden, der nach 9.2.8 einen unendlichen D-Faden von S liefert.

9.3 Syntaktisches und semantisches Hauptlemma

Wir untersuchen die beiden Alternativen, die durch 9.2.10 gegeben sind. Wir wollen zeigen:

Endliche D-Bäume liefern Herleitungen, unendliche D-Fäden liefern Gegenmodelle für die Sequenz $|S|$.

Wir betrachten zuerst endliche D-Bäume.

9.3.1 Syntaktisches Hauptlemma Es sei T eine abzählbare identitätsfreie Theorie und S eine identitätsfreie geordnete Sequenz von $L(T)$. Wenn der D-Baum von S in T endlich ist, so ist $|S|$ in T ohne Gleichheitsschlüsse $(= I)$, $(= F)$, $(= P)$ herleitbar.

Beweis: Nach Voraussetzung ist der Indexbaum B von S in T endlich. Also gibt es zu jedem $c \in B$ nur endlich viele $d \in B$ mit $c \leq d$. Wir beweisen durch Induktion nach der Anzahl der über c liegenden Knoten, dass $T \vdash |S_c|$ ohne Gleichheitsschlüsse ist.

1. Über c liegt kein Knoten von B. Dann ist S_c ein logisches Axiom (vgl. 9.2.6, 1).

2. Über c liegen Knoten $c * \langle i \rangle \in B$ für $i = 0, 1$ bzw. $i = 0$ nach 9.2.6, 2. Da über $c * \langle i \rangle$ mindestens ein Knoten weniger liegt als über c, ist nach Induktionsvoraussetzung

(1) $T \vdash |S_{c*\langle i \rangle}|$ ohne Gleichheitsschlüsse.

Wenn es das Axiom A_n (n sei die Länge von c) in der Aufzählung von $Ax(T)$ gibt, hat $S_{c*\langle i \rangle}$ eine Gestalt $A_n, \gamma : \delta$, und wir bezeichnen die geordnete Sequenz $\gamma : \delta$ mit $S'_{c*\langle i \rangle}$; andernfalls sei $S'_{c*\langle i \rangle} \equiv S_{c*\langle i \rangle}$. Dann folgt aus (1) mit einem T-Schluss oder unmittelbar

(2) $T \vdash |S'_{c*\langle i \rangle}|$ ohne Gleichheitsschlüsse.

Wir betrachten die Fälle 2.1 bis 2.5 aus 9.2.6:

2.1 Es ist $S_c \equiv S'_{c*\langle i \rangle}$;

2.2 $S'_{c*\langle 0\rangle}| \vdash |S_c|$ ist ein $(\to S)$-Schluss;

2.3 $|S'_{c*\langle 0\rangle}|, |S'_{c*\langle 1\rangle}| \vdash |S_c|$ ist ein $(\to A)$-Schluss;

2.4 $|S'_{c*\langle 0\rangle}| \vdash |S_c|$ ist ein $(\forall S)$-Schluss, für den wegen der Wahl von a_i die Variablenbedingung erfüllt ist;

2.5 $S'_{c*\langle 0\rangle}| \vdash |S_c|$ ist ein $(\forall A)$-Schluss.

In allen fünf Fällen folgt aus (2):

$$T \vdash |S_c| \text{ ohne Gleichheitsschlüsse.}$$

Mit Induktion folgt dann

$$T \vdash |S_{\langle\rangle}| \text{ ohne Gleichheitsschlüsse,}$$

und das ist wegen $S \equiv S_{\langle\rangle}$ die Behauptung.

Bemerkung Wenn man den endlichen D-Baum von S in T als Sequenzenbaum betrachtet, also im Knoten c die Sequenz $|S_c|$ liest, so *ist er selbst* eine Herleitung von $|S|$ in T, falls $Ax(T)$ leer ist. Das Beispiel 9.2.7 gibt also die allgemeine Lage wieder. Falls A_n nach 9.2.4 existiert, sind im D-Baum in der Tiefe n ein T-Schluss und ein logischer Grundschluss zu einem Schritt zusammengefasst. In jedem Fall liefert der D-Baum von S, sofern er endlich ist, ganz direkt eine Standard-Herleitung von $|S|$. Dies ist ein besonderer Vorzug schnittfreier Formalismen.

Es ist aber nicht gesagt, dass die Standard-Herleitung auch eine besonders kurze Herleitung wäre. Das hängt insbesondere von der Aufzählung der Terme t_i ab. Ist die Aufzählung „ungeschickt“, so kann es lange dauern, bis man die richtigen Beispiele $F(t_i)$ nach 2.5 in 9.2.6 als Antezedensformeln in den D-Baum von S eingeführt hat.

Wir kommen zur anderen Alternative von 9.2.10.

9.3.2 Semantisches Hauptlemma Es sei T eine abzählbare identitätsfreie Theorie und S eine identitätsfreie geordnete Sequenz von $L(T)$. Wenn es einen unendlichen D-Faden von S in T gibt, so hat T ein Modell mit Individuenbereich $\mathbb{N}$, in dem $|S|$ nicht gilt.

9.3.3 Anfang des Beweises

$$(D) \qquad S \equiv S_{\langle\rangle} \equiv S_0, S_1, \ldots, S_n, \ldots$$

sei ein unendlicher D-Faden von S in T. Ist also $S_n \equiv S_c$ gemäß 9.2.6, so ist

$$S_{n+1} \equiv S_{c*\langle i\rangle} \text{ für } i = 0 \text{ oder } i = 1.$$

Wir beweisen einige Lemmata über den unendlichen D-Faden (D).

9.3.4 Lemma Wenn eine Primformel als Ante- oder als Sukzedensformel, eine Allformel als Antezedensformel in einem S_m in (D) auftritt, so tritt sie ebenso in jedem S_n aus (D) mit $n \geq m$ auf.

Beweis: Es genügt, dies für $n = m + 1$ zu beweisen. (Vollständige Induktion nach $n - m$ erledigt dann den Rest.)
Alle Ante- und Sukzedensformeln von S_m, außer evtl. der ausgezeichneten Formel, sind nach 9.2.6 wieder Ante- bzw. Sukzedensformeln von S_{m+1}. Da Auftreten von Primformeln nie ausgezeichnet sind, gilt die Behauptung daher für Primformeln.
Ist ein Auftreten von $\forall x F(x)$ im Antezedens von S_m ausgezeichnet, so ist $\forall x F(x)$ nach 2.5 in 9.2.6 immer noch Antezedensformel von S_{m+1}. Also gilt die Behauptung auch für Allformeln.

9.3.5 Lemma Ist B keine Primformel und tritt B als Ante- bzw. Sukzedensformel in einem S_m in (D) auf, so gibt es ein S_n in (D) mit $n \geq m$, in dem ein ebensolches Auftreten von B ausgezeichnet ist.

Beweis durch Induktion nach der Summe l_m der Längen der Formelauftreten (mithin nach der Gesamtzahl der Auftreten von $\rightarrow$ und $\forall$), die in S_m rechts vom rechtesten Auftreten von B stehen.

1. Rechts von einem Auftreten von B in S_m stehen nur noch Primformeln. Dann ist B schon in S_m ausgezeichnet. Dies ist insbesondere für $l_m = 0$ der Fall.

2. Rechts von jedem Auftreten von B steht in S_m noch eine zusammengesetzte Formel. Dann steht das ausgezeichnete Formelauftreten in S_m auch rechts von B. Wir unterscheiden nach Gestalt und Position dieses ausgezeichneten Formelauftretens.

2.1 S_m hat eine Gestalt $\gamma : \delta, C \to D, \pi$ mit ausgezeichnetem $C \to D$. Dann ist B in γ oder in δ und

$$S_{m+1} \equiv A_m, \gamma, C : \delta, D, \pi,$$

so dass rechts von B ein Auftreten von $C \to D$ durch D ersetzt ist und, falls B in γ ist, noch ein C hinzukommt, also in jedem Fall $l_{m+1} < l_m$ ist.

2.2 S_m hat eine Gestalt $\gamma, D \to C, \pi : \pi'$ mit ausgezeichnetem $D \to C$. Dann ist B in γ und $S_{m+1} \equiv A_m, \gamma, \pi : D, \pi'$ oder $\equiv A_m, \gamma, C, \pi : \pi'$, so dass rechts von B ein Auftreten von $D \to C$ durch eines von D bzw. von C ersetzt ist, also $l_{m+1} < l_m$ ist.

2.3 S_m hat eine Gestalt $\gamma : \delta, \forall x F(x), \pi$ mit ausgezeichnetem $\forall x F(x)$. Dann ist B in γ oder in δ und

$$S_{m+1} \equiv A_m, \gamma : \delta, F(a_i), \pi,$$

so dass rechts von B ein Auftreten von $\forall x F(x)$ durch $F(a_i)$ ersetzt ist, also $l_{m+1} < l_m$ ist.

2.4 S_m hat eine Gestalt $\gamma, \forall x F(x), \pi : \pi'$ mit ausgezeichnetem $\forall x F(x)$. Dann ist B in γ und

$$S_{m+1} \equiv A_m, \forall x F(x), \gamma, F(t_i), \pi : \pi',$$

so dass rechts von B ein Auftreten von $\forall x F(x)$ durch $F(t_i)$ ersetzt ist, also $l_{m+1} < l_m$ ist.

In jedem Fall gibt es nach Induktionsvoraussetzung ein S_n in (D) mit $n \geq m + 1$, in dem B ausgezeichnet ist.
Mit Induktion folgt aus 1. und 2. die Behauptung.

9.3.6 Lemma Tritt $\forall x F(x)$ im Antezedens eines S_m in (D) auf, so gibt es unendlich viele S_n in (D), in denen $\forall x F(x)$ ausgezeichnet ist.

Beweis: Nach Lemma 9.3.4 tritt $\forall x F(x)$ im Antezedens aller $S_{n'}$ mit $n' \geq m$ auf. Nach Lemma 9.3.5 gibt es dann zu jedem $n' \geq m$ ein $n \geq n'$, so dass $\forall x F(x)$ in S_n ausgezeichnet ist, und das ist die Behauptung.

Bemerkung. Im D-Faden (D) rücken Auftreten von zusammengesetzten Formeln nach 9.3.5 allmählich nach rechts (d. h. die Formeln rechts von ihnen werden nicht weniger, aber kürzer), bis sie ausgezeichnet sind. Dann werden sie gemäß 9.2.6 reduziert und damit i. a. aufgelöst. Am linken Ende vom Ante- oder Sukzedens der S_n treten zusätzliche Formeln auf; insbesondere eine ausgezeichnete Antezedensformel $\forall x F(x)$ taucht links wieder auf, ihr Weg nach rechts beginnt aufs Neue, und sie wird im D-Faden nach 9.3.6 wieder und wieder ausgezeichnet. Rechts im Ante- wie im Sukzedens der S_m sammeln sich Primformeln, die dort in allen späteren S_n wieder auftreten.

9.3.7 Definition Die Ante- und Sukzedensformeln aus den geordneten Sequenzen S_n aus (D) nennen wir *Ante- bzw. Sukzedensformeln von* (D).

9.3.8 Definition einer Struktur $\mathcal{A}$ zu $L(T)$ über $\mathbb{N}$ und einer $\mathcal{A}$-Belegung $'$.

1. Der Individuenbereich $|\mathcal{A}|$ ist $\mathbb{N}$.

   Wir beziehen uns auf die Aufzählung $t_0, t_1, t_2, \ldots$ der Terme von $L(T)$ aus 9.2.4

2. $f_{\mathcal{A}}(i_1, \ldots, i_n) = j \iff f t_{i_1} \ldots t_{i_n}$ ist der Term t_j.

3. $(i_1, \ldots, i_n) \in p_{\mathcal{A}} \iff p t_{i_1} \ldots t_{i_n}$ ist Antezedensformel von (D).

4. Ist die i-te freie Variable a_i der Term t_j, so ist $a_i' := j$.

Die Funktionen $f_{\mathcal{A}}$ und die $\mathcal{A}$-Belegung $'$ sind eindeutig definiert, weil die Terme von $L(T)$ in 9.2.4 ohne Wiederholung aufgezählt sind.

9.3.9 Lemma Für jedes $j \in \mathbb{N}$ ist $\mathcal{A}(t_j') = j$.

Beweis durch Induktion nach dem Aufbau von t_j:

1. t_j ist die freie Variable a_i. Dann ist nach 9.3.8, 4
$$\mathcal{A}(t_j') = \mathcal{A}(a_i') = \mathcal{A}(j) = j.$$

2. t_j ist $f t_{i_1} \ldots t_{i_n}$. Dann ist
$$\begin{aligned} \mathcal{A}(t_j') &= \mathcal{A}(f t_{i_1}' \ldots t_{i_n}') \\ &= f_{\mathcal{A}}(\mathcal{A}(t_{i_1}'), \ldots, \mathcal{A}(t_{i_n}')) \\ &= f_{\mathcal{A}}(i_1, \ldots, i_n) \quad \text{nach Induktionsvoraussetzung} \\ &= j \qquad\qquad \text{nach 9.3.8, 2.} \end{aligned}$$

Es ist also $\mathcal{A}(t_i = t_j)' = w$ genau dann, wenn $i = j$ ist: Alle Gleichungen zwischen Termen verschiedener Gestalt werden in $\mathcal{A}$ unter der Belegung $'$ falsch. Das spielt im folgenden keine Rolle, weil T und S und daher der ganze D-Faden (D) identitätsfrei sind.

Im Beispiel 9.2.9 ist danach $\mathcal{A}(t') \neq \mathcal{A}(a')$, weil a in t nicht auftritt. Ist t der Term t_j, so ist $p_\mathcal{A} = \{j\}$, weil nur pt als Antezedensformel in (D) auftritt. Dann ist $\mathcal{A}(pt') = w$ und $\mathcal{A}(\forall xpx) = f$, und $|S|$ gilt nicht in $\mathcal{A}$. Wir zeigen, dass dem der allgemeine Sachverhalt entspricht.

9.3.10 Lemma Für Primformeln P von $L(T)$, die keine Gleichungen sind, gilt:

$$\mathcal{A}(P') = w \iff P \text{ ist Antezedensformel von } (D).$$

Beweis:

1. P ist $pt_{i_1} \dots t_{i_n}$. Dann ist

$$\begin{aligned} \mathcal{A}(P') = w &\iff (\mathcal{A}(t'_{i_1}), \dots, \mathcal{A}(t'_{i_n})) \in p_\mathcal{A} \\ &\iff (i_1, \dots, i_n) \in p_\mathcal{A} \text{ nach 9.3.9} \\ &\iff pt_{i_1} \dots t_{i_n} \text{ ist Antezedensformel von } (D). \end{aligned}$$

2. $\perp$ ist keine Antezedensformel von (D), weil der unendliche D-Faden (D) kein logisches Axiom enthält.

9.3.11 Lemma

a. Ist C Antezedensformel von (D), so ist $\mathcal{A}(C') = w$;

b. Ist C Sukzedensformel von (D), so ist $\mathcal{A}(C') = f$.

Beweis durch Induktion nach dem Aufbau von C:

1. C ist Primformel. Dann ist C keine Gleichung, weil S und T und deshalb alle S_n aus (D) identitätsfrei sind.

   1a. Ist C Antezedensformel von (D), so ist $\mathcal{A}(C') = w$ nach 9.3.10.

   1b. C sei Sukzedensformel von (D), etwa von S_n. Wäre $\mathcal{A}(C') = w$, so wäre C nach 9.3.10 Antezedensformel von (D), etwa von S_l. Für $m = \max(n, l)$ wäre dann C nach 9.3.4 sowohl Suk- als auch Antezedensformel von S_m, $|S_m|$ wäre ein logisches Axiom, und (D) wäre endlich nach 9.2.6, 1, was nicht der Fall ist. Also ist $\mathcal{A}(C') = f$.

2. C ist $A \to B$. Ist C Ante- oder Sukzedensformel von (D), etwa von S_m, ist C nach 9.3.5 ausgezeichnete Ante- bzw. Sukzedensformel in einem S_n in (D).

   2a. Ist C Antezedensformel von (D), so ist daher A Suk- oder B Antezedensformel in S_{n+1}. Nach Induktionsvoraussetzung ist dann $\mathcal{A}(A') = f$ oder $\mathcal{A}(B') = w$, und in beiden Fällen ist $\mathcal{A}(C') = \mathcal{A}(A' \to B') = w$.

   2b. Ist C Sukzedensformel von (D), so ist ebenso A Ante- und B Sukzedensformel in S_{n+1}. Nach Induktionsvoraussetzung ist dann $\mathcal{A}(A') = w$ und $\mathcal{A}(B') = f$, also $\mathcal{A}(C') = \mathcal{A}(A' \to B') = f$.

3. C ist $\forall x F(x)$.

   3a. Ist C Antezedensformel von (D), so ist C nach 9.3.6 ausgezeichnete Antezedensformel in unendlich vielen S_n in (D). Wegen 2.5 in 9.2.6 ist dann $F(t_i)$ Antezedensformel von (D) für jeden Term t_i. Nach Induktionsvoraussetzung und 9.3.9 ist dann $\mathcal{A}(F'(i)) = \mathcal{A}(F(t_i)') = w$ für alle $i \in \mathbb{N} = |\mathcal{A}|$, also $\mathcal{A}(C') = \mathcal{A}(\forall x F'(x)) = w$.

   3b. Ist C Sukzedensformel von (D), etwa in S_m, so ist C nach 9.3.5 ausgezeichnete Sukzedensformel in einem S_n in (D). Dann ist $F(a_i)$ Sukzedensformel in S_{n+1}. Nach Induktionsvoraussetzung ist für ein $j \in \mathbb{N}$

$$\mathcal{A}(F'(j)) = \mathcal{A}(F(a_i)') = f, \text{ also } \mathcal{A}(C') = \mathcal{A}(\forall x F'(x)) = f.$$

Mit Induktion folgt aus 1., 2. und 3. die Behauptung.

9.3.12 Abschluss des Beweises des semantischen Hauptlemmas.
Jedes nicht-logische Axiom A_n von T ist nach 9.2.6 Antezedensformel von S_n und damit von (D). Nach 9.3.11 ist daher $\mathcal{A}(A_n) = w$, und $\mathcal{A}$ ist ein Modell von T.
Nun ist $S \equiv S_0 \equiv \gamma : \delta$ der Anfang des D-Fadens (D) von S. Daher ist ebenfalls nach 9.3.11 $\mathcal{A}(C') = w$ für alle C aus γ und $\mathcal{A}(D') = f$ für alle D aus δ. Also ist $\mathcal{A}(|S'|) = f$, $|S|$ gilt nicht in $\mathcal{A}$:

$\mathcal{A}$ ist ein Modell von T, in dem die Sequenz $|S|$ nicht gilt. Damit ist 9.3.2 bewiesen.

9.4 Modelle über den natürlichen Zahlen

Wir brauchen nur noch syntaktisches und semantisches Hauptlemma zusammenzusetzen, um den Vollständigkeitssatz zu erhalten.

9.4.1 Vollständigkeitssatz von SCHÜTTE für abzählbare identitätsfreie Theorien.

Es sei T eine abzählbare identitätsfreie Theorie. Jede in T gültige identitätsfreie Sequenz ist in T ohne Gleichheitsschlüsse $(= I)$, $(= F)$, $(= P)$ herleitbar.

Beweis: Die identitätsfreie Sequenz $|S|$ sei gültig in T. Wegen des semantischen Hauptlemmas 9.3.2 hat S dann keinen unendlichen D-Faden. Nach Königs Lemma 9.2.10 ist dann der D-Baum von S endlich. Nach dem syntaktischen Hauptlemma 9.3.1 ist also $|S|$ in T ohne Gleichheitsschlüsse herleitbar.

Aus diesem Satz folgen ebenso wie aus dem allgemeinen Vollständigkeitssatz 8.3.7 rein semantische Aussagen, die wir im nächsten Kapitel behandeln. Aus ihm folgt aber auch ein rein syntaktisches Ergebnis.

9.4.2 Satz Ist in einer abzählbaren identitätsfreien Theorie T eine identitätsfreie Sequenz (mit Gleichheitsschlüssen) herleitbar, so ist sie in T auch ohne Gleichheitsschlüsse herleitbar.

Beweis: Wenn eine Sequenz in T herleitbar ist, gilt sie in T nach dem Korrektheitssatz 3.2.1. Mit 9.4.1 folgt die Behauptung.

Dieser Satz lässt sich offenbar aus dem Gödel-Henkinschen Satz 8.3.7 nicht ohne Weiteres folgern. Wie die Schnittregel 8.4.1 besitzt er auch einen rein syntaktischen, völlig konstruktiven Beweis, der außerdem die Abzählbarkeit der Theorie T nicht verwendet.

Wir haben bis jetzt nicht ausgenutzt, dass das semantische Hauptlemma Modelle mit dem ganz speziellen Individuenbereich $\mathbb{N}$, *Modelle über* $\mathbb{N}$, liefert. Wenn wir dies tun, erhalten wir ein weiteres wichtiges Ergebnis:

9.4.3 Satz von LÖWENHEIM und SKOLEM
Jede konsistente abzählbare identitätsfreie Theorie hat ein Modell über $\mathbb{N}$ (d. h. mit $\mathbb{N}$ als Individuenbereich).

Beweis: Ist T konsistent, so ist $T \nvdash \bot$. Nach dem syntaktischen Hauptlemma 9.3.1 ist dann der D-Baum von $\bot$ in T nicht endlich. Nach Königs Lemma

9.2.10 hat daher $\perp$ einen unendlichen D-Faden in T. Nach dem semantischen Hauptlemma 9.3.2 hat also T ein Modell über $\mathbb{N}$.

Dieser Satz ist in doppelter Hinsicht bemerkenswert.

(i) Bekanntlich gibt es Anzahl-Formeln wie $\forall x \forall y\ x = y$, die nur in einigen endlichen Strukturen wahr sind. Nach 9.4.3 müssen solche Formeln notwendig das Gleichheitszeichen $=$ enthalten: Eine konsistente Theorie, deren sämtliche Modelle endlich sind, kann nicht identitätsfrei sein.

(ii) Die Zermelo-Fraenkelsche Mengenlehre ZF lässt sich leicht identitätsfrei formulieren. Wenn also ZF konsistent ist – was man allgemein annimmt –, dann hat ZF nach 9.4.3 ein Modell über $\mathbb{N}$, in dem es nur abzählbar unendlich viele Mengen gibt. Wir kommen auf dieses Löwenheim-Skolem-Paradoxon im nächsten Kapitel zurück, wenn wir 9.4.3 auf beliebige Theorien verallgemeinern.

9.4.4 Zusammenfassung Ein Prinzip dieses Vollständigkeitsbeweises ist die Aufspaltung in syntaktisches und semantisches Hauptlemma, die durch die Schnittfreiheit der Herleitungen erst ermöglicht wird. Wesentlich für den Erfolg der Konstruktion ist die Beschränkung auf abzählbare Theorien. Gäbe es überabzählbar viele Terme, so könnten in einem (abzählbaren) D-Faden nicht alle Formeln $F(t)$ als Antezedensformeln auftreten, so dass in 9.3.11 der Schluss auf $\mathcal{A}(\forall x F'(x)) = w$ nicht gelingen würde. Überabzählbar lange Fäden wären wiederum mit Königs Lemma nicht zu behandeln.

Die Beschränkung auf identitätsfreie Theorien ist dagegen nicht wesentlich. Sie erleichtert uns jedoch die Arbeit beträchtlich und führt zu dem zusätzlichen Ergebnis 9.4.2 der Konservativität der Prädikatenlogik über der identitätsfreien Logik ebenso wie zu der Teilaussage von 9.4.3, dass identitätsfreie Theorien, wenn überhaupt, dann unendliche Modelle haben.

Die großen methodischen Unterschiede zwischen den Paragraphen 8 und 9 haben wir schon angesprochen: Statt Zorns Lemma wird hier nur Königs Lemma verwendet. Entsprechend werden statt beliebiger nur abzählbare Theorien untersucht. Statt Erweiterungen von Theorien stehen hier D-Bäume und D-Fäden im Vordergrund.

Bei näherem Hinsehen stellt man aber auch Parallelen zwischen beiden Beweisen fest. Insbesondere entsprechen sich

in §§7 und 8	in §9
saturierte Theorien	unendliche D-Fäden (D)
kanonisches Modell	Modell zu (D) nach 9.3.8
Satz 7.3.3	Lemma 9.3.11
Saturierung von Theorien mit Zorns Lemma	Wahl eines unendlichen D-Fadens mit Königs Lemma

9.5 Aufgaben

9.5.1 Konstruieren Sie den D-Baum der geordneten Sequenz

$$S :\equiv \forall xpx : pt$$

in einer logischen Theorie T für den Fall, dass

a. t der erste Term t_0,

b. t der Term t_2 nach 9.2.4 ist.

9.5.2 Beschreiben Sie den D-Baum der geordneten Sequenz

$$S :\equiv \forall x(px \to qx) : \emptyset$$

in einer logischen Theorie T. Wieviele endliche und wieviele unendliche D-Fäden hat S? Benutzen Sie einen der unendlichen D-Fäden von S, um gemäß 9.3.8 eine Struktur zu definieren, in der $|S|$ falsch wird.

9.5.3 Beschreiben Sie den D-Baum der Formel

$$S :\equiv \forall x(px \to qx) : \exists xqx$$

in einer logischen Theorie T. Benutzen Sie den einzigen unendlichen D-Faden von S, um wie in 9.3.8 eine Struktur zu definieren, in der $|S|$ falsch ist.

9.5.4 Eine Struktur $\mathcal{A}$ heißt endlich, wenn ihr Individuenbereich $|\mathcal{A}|$ endlich ist.

a. Folgern Sie aus 9.4.3:
Wenn eine identitätsfreie Formel F in einer endlichen Struktur gilt, dann gibt es auch eine unendliche Struktur, in der F gilt.

b. Zeigen Sie, dass die Umkehrung von a. falsch ist, indem Sie für die Formel

$$F :\equiv \forall x \exists y \forall z (\neg pxx \wedge pxy \wedge (pyz \rightarrow pxz))$$

beweisen: Es gibt unendliche, aber keine endlichen Strukturen, in denen F gilt.

9.5.5 Wir nennen eine geordnete Sequenz *prim*, wenn sie nur aus Primformeln besteht, die keine Gleichungen sind.

a. Zeigen Sie: Ist S prim, so ist $|S|$ ein logisches Axiom, oder es gibt eine endliche Struktur, in der $|S|$ nicht gilt.

b. Beschreiben Sie die D-Bäume von primen geordneten Sequenzen in logischen Theorien.

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 4:
Modelltheorie

Autor: Justus Diller

01301-7-04-S 1

Klassische Prädikatenlogik

Kurseinheit 4: Inhalt

Studienhinweise 203

Verzeichnis der definierten Begriffe und der wichtigen Sätze 205

1. Sprache, Semantik und Syntax der Prädikatenlogik

2. Syntaktische Sätze und Regeln der Prädikatenlogik

3. Vollständigkeit

4. Modelltheorie

§10 Kompaktheit 209

10.1 Der Kompaktheitssatz 210

10.2 Endliche und unendliche Modelle 211

10.3 Zur Charakteristik von Ringen und Körpern 214

10.4 Exkurs: Topologische Deutung des Kompaktheitssatzes 217

10.5 Aufgaben 222

§11 Morphismen 223

11.1 Homomorphismen 223

11.2 Unterstrukturen und Einbettungen 230

11.3 Diagramme 236

11.4 Aufgaben 243

§12 Die Sätze von Löwenheim und Skolem 245

12.1 Kardinalzahlen 245

12.2 Modelle höherer Mächtigkeit 257

12.3 Elementare Untermodelle 259

12.4 Aufgaben 263

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 4: Studienhinweise

1. Lehrziele

Die Kurseinheit gibt eine Einführung in die Modelltheorie. Die Modelltheorie baut die semantischen Aspekte der Prädikatenlogik zu einem eigenen Teilgebiet der mathematischen Logik aus. Während syntaktische Betrachtungen zu Herleitungen, zulässigen Regeln etc. elementar-kombinatorisch verlaufen und ohne alle Vorkenntnisse angegangen werden können, verwendet die Modelltheorie Sprechweisen und Grundkenntnisse der Mengenlehre. Diese werden hier in §11 und in 12.1 in größerer Breite dargestellt, als innerhalb eines solchen Kurses sonst vielleicht üblich ist. Aber auch wer schon Kenntnisse der naiven Mengenlehre mitbringt, sollte diese Abschnitte nicht einfach überspringen, wegen ihrer Ausrichtung auf die Modelltheorie.

Die zentralen Ergebnisse der Kurseinheit sind der Kompaktheitssatz aus 10.1 und die Löwenheim-Skolem-Sätze aus 12.2 und 3. Der Kompaktheitssatz wird hier unmittelbar aus dem Korrektheits- und Vollständigkeitssatz gefolgert. Seine große Bedeutung liegt in seinen Anwendungen. Einige einfache, für die Algebra bedeutsame Anwendungen werden in 10.2 und 3 behandelt. Zwei wichtige Anwendungen in dieser Kurseinheit sind der aufsteigende Satz von Löwenheim und Skolem in 12.2 sowie der Satz von Łoš und Tarski zur Charakterisierung der offen axiomatisierbaren Theorien in 11.3, das dritte relevante modelltheoretische Ergebnis dieser Kurseinheit.

Das Studium der Morphismen in §11 hat allerdings allgemeinere Ziele als den Satz von Łoš und Tarski. Es soll in elementarer, trotzdem effizienter Weise Begriffe wie Unterstruktur, elementare Unterstruktur, Einbettung und Isomorphismus einführen und ihre Bedeutung für den „Wahrheitstransport“ (vgl. Definition 11.3.1) zwischen Strukturen klären. Die elementaren Unterstrukturen sind wesentlich in 12.3. Unterstrukturen, Homomorphismen und Isomorphismen treten überall in der Algebra und ihren Nachbargebieten auf, und die Modelltheorie ist der geeignete Ort, diese Begriffe im allgemeinen Rahmen zu klären.

Die „kleine“ Kardinalzahltheorie in 12.1 behandelt nur das, was für die allgemeine Fassung der Löwenheim-Skolem-Sätze in 12.2 und 3 gebraucht wird. Die Beweise des Abschnitts 12.1 gehören naturgemäß nicht zum Kernbereich

des Kurses, aber die markanten Ergebnisse sind wesentlich für das Folgende; sie sollten wirklich verstanden werden. Es ist leicht zu sehen, dass der klassische Satz 12.3.2 von Löwenheim-Skolem, der sich auf den abzählbaren Fall beschränkt, mit einem kleinen Teil der Kardinalzahltheorie auskommt. Dieser klassische Satz ist auch für das Löwenheim-Skolem-Paradoxon verantwortlich, dessen aufmerksames Studium die überraschende Schlagkraft modelltheoretischer Betrachtungen verdeutlicht.

2. Eingangsvoraussetzungen

Sicherheit im Umgang mit den Grundbegriffen der Semantik (§2) ist durchgängig nötig. Korrektheits- und Vollständigkeitssatz werden nicht nur beim Beweis des Kompaktheitssatzes verwendet, sondern auch später in der Kurseinheit. Wie in der Kurseinheit 3 wird die übliche mengentheoretische Schreibweise verwendet. Vorkenntnisse über (injektive, surjektive) Abbildungen, Homomorphismen etc. erleichtern das Studium von §11, werden aber dort auch erarbeitet. Ähnliches gilt für mengentheoretische Vorkenntnisse in 12.1. Der Exkurs 10.4 setzt Grundkenntnisse der Punktmengen-Topologie voraus. Auf ihn wird später nicht zurückgegriffen.

Klassische Prädikatenlogik

Kurseinheit 4: Verzeichnis der definierten Begriffe und der wichtigen Sätze

10.1.1 Teiltheorie, endlich axiomatisiert

10.1.2, 3 **Kompaktheitssatz**

10.2.1 Expansion einer Struktur

10.2.4 **Satz** Jede Theorie mit beliebig großen endlichen Modellen hat ein unendliches Modell

10.3.1 Charakteristik von Ringen und Körpern

10.3.3 **Satz** Gilt ein Satz in allen Ringen der Charakteristik 0, so gilt er in allen Ringen genügend großer Charakteristik

10.4.1 Hausdorff-Raum, kompakter Raum

10.4.2 Elementare Äquivalenz

10.4.3 Modellklasse $Mod(T)$, elementare Klasse

10.4.7 **Satz** Der Stonesche Raum zu L ist ein kompakter topologischer Raum

11.1.1 Homomorphismus, homomorphes Bild

11.1.4 $\varphi(\mathcal{F})$ für Nennform $\mathcal{F}$

11.1.6 positive und $\exists$-Formeln

11.1.7 Lemma Homomorphismen erhalten die Wahrheit positiver $\exists$-Sätze

11.1.9 **Satz** Die Modellklasse einer positiv axiomatisierten Theorie ist gegen homomorphe Bilder abgeschlossen

11.2.1 Unterstruktur, $\mathcal{A} \subseteq \mathcal{B}$

11.2.3 Einbettung $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$, Isomorphismus $\varphi : \mathcal{A} \cong \mathcal{B}$

11.2.4 Lemma $\mathcal{A} \subseteq \mathcal{B} \iff id : \mathcal{A} \hookrightarrow \mathcal{B}$

11.2.7 Lemma Einbettungen erhalten die Wahrheit von $\exists$-Sätzen

11.2.9 Offene bzw. $\forall$-Theorien

11.2.10 **Satz** Die Modellklasse einer offenen Theorie ist gegen Unterstrukturen abgeschlossen

11.2.12 **Satz** Isomorphe Strukturen sind Modelle derselben Theorie
11.3.1 φ transportiert $\Delta, \varphi : \mathcal{A} \xrightarrow{\Delta} \mathcal{B}$
11.3.4 Elementare Einbettung, elementare Unterstruktur
11.3.7 Expansionen $(\mathcal{B}, im(\varphi)), (\mathcal{A}, A)$
11.3.8 Diagramm $D(\mathcal{A})$, elementares Diagramm $Th((\mathcal{A}, A))$
11.3.9 **Diagramm-Lemma**
11.3.10 Äquivalente Theorien, Theorie $T_\forall$
11.3.11 **Satz von Łoš und Tarski**
12.1.1 gleichmächtig, $\sim$
12.1.2 höchstens so mächtig, $\lesssim$, weniger mächtig, $<$
12.1.3 Potenzmenge, $Pot(A)$
12.1.4 **Satz von Cantor, 1. Cantorsches Diagonalverfahren**
12.1.7 **Satz von Schröder und Bernstein**
12.1.8 Wohlordnung
12.1.10 Abschnitt
12.1.12 Standardeinbettung φ_{st} zu Wohlordnungen $\mathcal{A}, \mathcal{B}$
12.1.13 **Satz** Wohlordnungen sind vergleichbar
12.1.14 **Satz** Mächtigkeiten sind vergleichbar
12.1.16 **Satz** Aus $\mathcal{A} \hookrightarrow \mathcal{B}$ und $\mathcal{B} \hookrightarrow \mathcal{A}$ folgt $\mathcal{A} \cong \mathcal{B}$
12.1.17 Kardinalzahl, Kardinalität $card(A)$
12.1.19 Kardinale Summe, Produkt, Potenz
12.1.20 Hessenberg-Ordnung $<_H, \mathcal{A} \times_H \mathcal{A}$
12.1.22 **Satz von Hessenberg** Für unendliches κ ist $\kappa \times_H \kappa \cong \kappa$
12.1.23 Korollar Kardinale Rechengesetze
12.2.1 Mächtigkeit einer Struktur, $card(\mathcal{A})$
12.2.2 κ-Sprache, κ-Theorie
12.2.4 **Aufsteigender Satz von Löwenheim und Skolem**
12.3.1 **Absteigender Satz von Löwenheim und Skolem**

12.3.2 **Satz von Löwenheim und Skolem**, abzählbarer Fall
12.3.3 Das Löwenheim-Skolem-Paradoxon
12.3.4 **Mächtigkeitssatz von Löwenheim, Skolem und Tarski**

Kapitel 4

Modelltheorie

Modelltheorie ist das Studium beliebiger Modellklassen zu Sprachen der ersten Stufe mit den Methoden der Prädikatenlogik. Da man in Strukturen und Modellen keine Herleitungen, sondern die Wahrheit von Sätzen und die Gültigkeit von Formeln betrachtet, sind es die Methoden der Semantik aus §2, die in der Modelltheorie zum Tragen kommen. Von den Teilgebieten der Logik hat die Modelltheorie ein besonders nahes Verhältnis zur Algebra: Einerseits arbeitet die Modelltheorie intensiv mit algebraischen Begriffen und Methoden (vgl. §11), andererseits hat sie zahlreiche Anwendungen in der Algebra.

Im Rahmen dieses Kurses geben wir eine Einführung in einige klassische Ergebnisse der Modelltheorie. Etwas ausführlicher betrachten wir dabei die Modelle der Zahlentheorie.

§10 Kompaktheit

10.1 Der Kompaktheitssatz

10.2 Endliche und unendliche Modelle

10.3 Zur Charakteristik von Ringen und Körpern

10.4 Exkurs: Topologische Deutung des Kompaktheitssatzes

10.5 Aufgaben

10.1 Der Kompaktheitssatz

Der Kompaktheitssatz ist wohl das wirkungsvollste Werkzeug der Modelltheorie. Er ist eine sehr einfache Folgerung aus Korrektheits- und Vollständigkeitssatz für beliebige Theorien, also aus der Äquivalenz 8.3.8 von Herleitbarkeit und Gültigkeit, die wir hier noch einmal für Formeln B (d. h. für Sequenzen $\emptyset : B$) formulieren:
Für jede Theorie T und jede Formel B aus $L(T)$ ist

$$T \models B \Leftrightarrow T \vdash B.$$

Wir haben diese Äquivalenz für den Beweis eines syntaktischen Ergebnisses ausgenutzt, nämlich für die Zulässigkeit der Schnittregel. Wir wollen sie hier nutzen, um umgekehrt die in 5.4.3 diskutierte Endlichkeit von Herleitungen in eine Gültigkeitsaussage zu übersetzen.

10.1.1 Definition Eine Theorie T' ist *Teiltheorie* einer Theorie T, wenn T eine einfache Erweiterung von T' ist. Eine Theorie T' ist *endlich axiomatisiert*, wenn die Menge $Ax(T')$ endlich ist.

10.1.2 Kompaktheitssatz 1. Fassung
Eine Formel B gilt in einer Theorie T genau dann, wenn B in einer endlich axiomatisierten Teiltheorie T' von T gilt.

Beweis: Wenn B in irgendeiner Teiltheorie T' von T gilt, dann gilt B auch in T. Denn jedes Modell von T ist nach 8.2.4 auch ein Modell von T'.
Umgekehrt gibt es zu jeder Herleitung H von B in T nach 5.4.3 eine endlich axiomatisierte Teiltheorie T' von T, in der H eine Herleitung ist, weil H nur endlich viele Axiome verwendet. Für dieses T' gilt also:

$$T \vdash B \Rightarrow T' \vdash B.$$

Mit der Äquivalenz von Herleitbarkeit und Gültigkeit, genauer mit der Vollständigkeit für T und der Korrektheit für T' folgt hieraus

$$T \models B \Rightarrow T' \models B.$$

Damit ist der Satz bewiesen.

Kurz und einprägsam lässt sich dieses Argument so zusammenfassen: Ist $Ax(T')$ die endliche Menge der Axiome einer Herleitung von B in T, so ist

$$T \models B \Leftrightarrow T \vdash B \Leftrightarrow T' \vdash B \Leftrightarrow T' \models B.$$

10.1.3 Kompaktheitssatz 2. Fassung
Eine Theorie T besitzt genau dann ein Modell, wenn jede endlich axiomatisierte Teiltheorie von T ein Modell besitzt.

Beweis: Die Richtung von links nach rechts ergibt sich aus 8.2.4.

Für die andere Richtung nehmen wir an, dass T kein Modell besitzt. Dann gilt $\perp$ in T, und nach 10.1.2 gibt es eine endlich axiomatisierte Teiltheorie T' von T, in der $\perp$ gilt. Dieses T' hat dann kein Modell, im Widerspruch zur Voraussetzung. Also hat T ein Modell.

Der Kompaktheits*satz* ist eine rein semantische Aussage. Seine beiden Fassungen machen vom Herleitungsbegriff keinen Gebrauch. Allerdings greift unser *Beweis* auf die Äquivalenz von Herleitbarkeit und Gültigkeit zurück. Er ist eines der wichtigsten Hilfsmittel der Modelltheorie. In diesem und den nächsten Paragraphen wollen wir einige überraschende Auswirkungen des Kompaktheitssatzes studieren.

10.2 Endliche und unendliche Modelle

Der Kompaktheitssatz stellt einen Zusammenhang zwischen endlichen und unendlichen Axiomensystemen her. Daraus ergeben sich Zusammenhänge zwischen endlichen und unendlichen Merkmalen von Modellen. Insbesondere können endlich viele Ungleichungen die Existenz von endlich vielen verschiedenen Elementen und unendlich viele Ungleichungen die Existenz von unendlich vielen Elementen ausdrücken.

10.2.1 Definition Eine Struktur $\mathcal{A}$ zur Sprache L' ist *Expansion* einer Struktur $\mathcal{B}$ zur Sprache L, $\mathcal{B}$ wird *expandiert* zu $\mathcal{A}$, wenn $\mathcal{B}$ die Beschränkung von $\mathcal{A}$ auf L ist.

Wie in 8.2.1 ist dann $L \subseteq L'$ und $|\mathcal{A}| = |\mathcal{B}|$. $\mathcal{A}$ ist also keine größere Struktur als $\mathcal{B}$, sondern allenfalls eine „reichere“ Struktur. Während jedoch die Beschränkung von $\mathcal{A}$ auf L eindeutig festgelegt ist, wenn man $\mathcal{A}$ und L kennt,

gibt es i. a. viele Expansionen von $\mathcal{B}$, die Strukturen zur selben Sprache L' sind.

10.2.2 Lemma Sei $L \subseteq L'$. Jede Struktur $\mathcal{B}$ zu L lässt sich zu einer Struktur $\mathcal{A}$ zu L' expandieren.

Beweis: Man setzt $|\mathcal{A}| := |\mathcal{B}|$ und interpretiert die nicht-logischen Grundzeichen aus L in $\mathcal{A}$ genauso wie in $\mathcal{B}$. Dann ist schon $\mathcal{A}|L = \mathcal{B}$.

Da $|\mathcal{A}| = |\mathcal{B}| \neq \emptyset$ ist, gibt es ein Element $0 \in |\mathcal{A}|$. Dann können wir alle Funktionszeichen f aus L', die nicht in L auftreten, in $\mathcal{A}$ durch die konstante Funktion mit dem einzigen Wert 0 interpretieren, $f_{\mathcal{A}}(k_1, \ldots, k_n) = 0$ (übrigens auch für 0-stellige Funktionszeichen: $c_{\mathcal{A}} = 0$). Die nicht-logischen Prädikatszeichen p aus $L' - L$ können wir etwa durch die leere Menge interpretieren, $p_{\mathcal{A}} = \emptyset$. Damit wird $\mathcal{A}$ eine Struktur zu L' und eine Expansion von $\mathcal{B}$.

10.2.3 Korollar Zu jeder nicht-leeren Menge A und jeder Sprache L gibt es eine Struktur $\mathcal{A}$ zu L mit Individuenbereich A.

Denn ist L_0 die Sprache ohne nicht-logische Grundzeichen, die sog. Sprache der Identität, so ist (A) eine Struktur zu L_0, die sich nach dem Lemma zu einer Struktur $\mathcal{A}$ zu L expandieren lässt.

10.2.4 Satz Wenn eine Theorie T beliebig große endliche Modelle hat, so hat T auch unendliche Modelle.

Beweis:

$$E = \{e_0, e_1, e_2, \ldots\}$$

sei eine abzählbar unendliche Menge von neuen Konstanten, d. h. von Konstanten, die in $L(T)$ nicht auftreten. Ferner sei

$$U = \{\neg e_i = e_j \mid i \neq j, i, j \in \mathbb{N}\}.$$

Dann hat $T + E + U$ nur unendliche Modelle, weil wegen der Axiome U alle Konstanten e_i verschieden interpretiert werden müssen.
Jede endlich axiomatisierte Teiltheorie T' von $T + E + U$ enthält nur endlich viele Ungleichungen U' aus U unter ihren Axiomen. Ist m der größte in U' auftretende Index einer Konstanten aus E, so ist

$$T' \prec T + E + \{\neg e_i = e_j \mid i \neq j, i, j \leq m\}.$$

Nach Voraussetzung hat T ein Modell $\mathcal{B}_m$ mit mehr als m Elementen. Dieses Modell wird zu einem Modell $\mathcal{A}_m$ von T' expandiert, wenn man die $\mathcal{A}_m(e_i)$ für $i \leq m$ paarweise verschieden wählt. Dies ist möglich, weil $|\mathcal{A}_m| = |\mathcal{B}_m|$ mindestens $m+1$ verschiedene Elemente enthält. Für $i > m$ kann man e_i beliebig interpretieren, etwa $\mathcal{A}_m(e_i) = \mathcal{A}_m(e_0)$.
Also besitzt jede endlich axiomatisierte Teiltheorie von $T + E + U$ ein Modell. Nach dem Kompaktheitssatz 10.1.3 hat dann auch $T + E + U$ ein Modell $\mathcal{A}$, das notwendig unendlich ist. Die Beschränkung von $\mathcal{A}$ auf $L(T)$ ist dann ein unendliches Modell von T.

10.2.5 Korollar Es sei L eine Sprache. Es gibt keine Theorie, deren Modelle genau die endlichen Strukturen zu L sind.

Beweis: Eine solche Theorie hätte nach 10.2.3 beliebig große endliche Modelle und deshalb nach 10.2.4 auch unendliche Modelle.

10.2.6 Korollar Es gibt keine Theorie, deren Modelle genau die endlichen Gruppen sind.
Denn da es zu jeder natürlichen Zahl $m > 0$ die zyklische Gruppe $\mathbb{Z}/m\mathbb{Z}$ der Ordnung m gibt, die genau m Elemente enthält, hätte eine solche Theorie beliebig große endliche Modelle. Nach 10.2.4 hätte sie dann auch unendliche Modelle.

So einfach diese Ergebnisse sind, sind sie doch erstaunlich: Es gibt mathematisch geläufige Klassen von Strukturen (zur selben Sprache), die sich nicht durch ein Axiomensystem beschreiben lassen.
In unendlichen Strukturen zu Sprachen $L + E$, in denen die unendlich vielen Konstanten $e_i \in E$ verschieden interpretiert werden, gilt die „unendliche Konjunktion“

$$\neg e_0 = e_1 \wedge \neg e_0 = e_2 \wedge \neg e_1 = e_2 \wedge \ldots \wedge \neg e_i = e_j \wedge \ldots (i < j),$$

die zwar keine Formel ist, die sich aber durch die unendliche Ungleichungsmenge U aus 10.2.4 ausdrücken lässt.

Die „unendliche Disjunktion“

$$e_0 = e_1 \vee e_0 = e_2 \vee e_1 = e_2 \vee \ldots \vee e_i = e_j \vee \ldots (i < j),$$

nach der nicht alle e_i voneinander verschieden sind, lässt sich überhaupt nicht durch eine Formelmenge ausdrücken. Dass es höchstens n Elemente gibt, wird ausgedrückt durch

$$(E_{\leq n}) \quad \forall x_0 \ldots \forall x_n (x_0 = x_1 \vee x_0 = x_2 \vee \ldots \vee x_{n-1} = x_n).$$

Die „unendliche Disjunktion“ dieser Formeln

$$(E_{\leq 2}) \vee (E_{\leq 3}) \vee \ldots \vee (E_{\leq n}) \vee \ldots,$$

die besagt, dass es nur endlich viele Elemente gibt, lässt sich wiederum nach 10.2.5 nicht in einem Axiomensystem ausdrücken:
Der Begriff der Endlichkeit ist nicht axiomatisierbar.

10.3 Zur Charakteristik von Ringen und Körpern

In 10.2 haben wir den Kompaktheitssatz nur auf Mengen von Ungleichungen zwischen *neuen* Konstanten angewandt. Die neuen Konstanten werden interpretiert als irgendwelche Elemente, die verschieden sein müssen, damit die Ungleichungen wahr werden. Das in 10.2.4 gesuchte Modell wird deswegen unendlich groß. Schon wenn man den Kompaktheitssatz auf Ungleichungen zwischen Termen der Ausgangssprache anwendet, ergeben sich neue Aspekte. Wir betrachten hier die Sprache $L(T_R)$ der Theorie der Ringe mit 1 (vgl. 1.2.3).

10.3.1 Definition Für jede natürliche Zahl $n > 1$ sei $(n = 0)$ die Formel

$$\underbrace{1 + 1 + \ldots + 1}_{n\text{-mal}} = 0$$

(wobei wir wie stets Rechtsklammerung fortgelassen haben), und $(\mathrm{Char}\, n)$ sei die Formel

$$\neg(2 = 0) \wedge \neg(3 = 0) \wedge \ldots \wedge \neg(n - 1 = 0) \wedge (n = 0).$$

$T_R + \{(\mathrm{Char}\, n)\}$ ist die *Theorie der Ringe mit* 1 *der Charakteristik* n (für $n > 1$), und

$$T_{R,0} = T_R + \{\neg(n = 0) \mid n > 1\}$$

ist die *Theorie der Ringe mit* 1 *der Charakteristik* 0.

Die Körper der Charakteristik n (für $n > 0$) sind offenbar die Modelle von $T_K + \{(\mathrm{Char}\, n)\}$, und die Körper der Charakteristik 0 sind die Modelle von $T_K + \{\neg(n = 0) \mid n > 1\}$.
Während die Theorie der Körper der Charakteristik n für $n > 1$ und alle ihre Teiltheorien endlich axiomatisiert sind, ist die Theorie $T_{R,0}$ nicht endlich axiomatisiert. Wir wollen zeigen, dass dies nicht an unserer speziellen Wahl der Axiome liegt.

10.3.2 Bemerkung Wie in 1.2.3 sei

R10. $a \neq 0 \to \exists y\ a \cdot y = 1$, ferner

R11. $a \cdot b = 0 \to a = 0 \vee b = 0$.

R11 drückt aus, dass ein Ring nullteilerfrei ist. Es gilt

$$T_R + \{R10\} \vdash R11 \quad \text{und}$$
$$T_R + \{R11\} \vdash \neg(\mathrm{Char}\, n), \text{ falls } n \text{ keine Primzahl ist.}$$

Also ist die Theorie der (Schief-) Körper der Charakteristik n nur dann konsistent, wenn $n = 0$ oder n eine Primzahl ist.

Es gibt beliebig große Primzahlen, und zu jeder Primzahl p gibt es Körper der Charakteristik p.

10.3.3 Satz Es sei B eine Formel aus $L(T_R)$. Gilt B in jedem Ring mit 1 der Charakteristik 0, so gibt es eine natürliche Zahl m, so dass B in allen Ringen mit 1 einer Charakteristik $p > m$ gilt.

Beweis: Wenn B in $T_{R,0}$ gilt, gilt B nach dem Kompaktheitssatz 10.1.2 in einer endlich axiomatisierten Teiltheorie T' von $T_{R,0}$. Dann gibt es nur endlich viele Formeln $\neg(n = 0)$ in $Ax(T')$.
Diese endlich vielen Zahlen n besitzen ein Maximum m. Also gilt B in

$$T_R + \{\neg(n = 0) \mid 1 < n \leq m\}$$

und damit in $T_R + \{(\mathrm{Char}\, p)\}$ für $p > m$, weil dann die Formeln $\neg(n = 0)$ Konjunktionsglieder der Formel $(\mathrm{Char}\, p)$ sind.

10.3.4 Korollar Es gibt keine endlich axiomatisierte Theorie, deren Modelle genau die Ringe mit 1 der Charakteristik 0 sind.

Beweis: Angenommen, T wäre eine solche Theorie mit den Axiomen $A_1, \ldots, A_k$. Dann wäre

$$B :\equiv A_1 \wedge \ldots \wedge A_k$$

eine Formel, die in einem Ring mit 1 genau dann gilt, wenn er die Charakteristik 0 hat, im Widerspruch zu 10.3.3.
Das Axiomensystem von $T_{R,0}$ lässt sich also nicht äquivalent durch ein endliches ersetzen, wie nach 10.3.1 bemerkt. Für die Ringe positiver Charakteristik gilt noch mehr:

10.3.5 Satz Es gibt keine mathematische Theorie, deren Modelle genau die Ringe mit 1 der Charakteristik $\neq 0$ sind.

Beweis: Wäre T eine solche Theorie, so hätte $T + \{\neg(n = 0) \mid n > 1\}$ kein Modell. Nach dem Kompaktheitssatz 10.1.3 hätte dann auch eine endlich axiomatisierte Teiltheorie T' dieser Theorie kein Modell. Dann gäbe es nur endlich viele Formeln $\neg(n = 0)$ in $Ax(T')$. Wäre m das Maximum dieser endlich vielen Zahlen n, so hätte auch

$$T + \{\neg(n = 0) \mid 1 < n \leq m\}$$

und damit $T + \{(\mathrm{Char}\, p)\}$ für $p > m$ kein Modell. Dies widerspricht der Annahme, da es zu jeder natürlichen Zahl $p > 1$ Ringe mit 1 der Charakteristik p gibt.
Das Muster ist wieder dasselbe wie in 10.2.4 und 5: In Ringen mit 1 der Charakteristik 0 gilt die „unendliche Konjunktion“

$$\neg(2 = 0) \wedge \neg(3 = 0) \wedge \ldots \wedge \neg(n = 0) \wedge \ldots,$$

die keine Formel ist, sich aber durch die unendliche Formelmenge

$$\{\neg(n = 0) \mid n > 1\}$$

ausdrücken lässt. In Ringen mit 1 der Charakteristik $\neq 0$ gilt die „unendliche Disjunktion“

$$(2 = 0) \vee (3 = 0) \vee \ldots \vee (n = 0) \vee \ldots,$$

die sich nun nicht einmal durch eine unendliche Formelmenge in $L(T_R)$ ausdrücken lässt.
Der Beweis von 10.3.5 ist kürzer, direkter als der von 10.2.4, weil wir hier direkt in der Sprache von T_R argumentieren können und keine Spracherweiterungen, Expansionen und Beschränkungen betrachten müssen.

10.4 Exkurs: Topologische Deutung des Kompaktheitssatzes

In diesem Exkurs erläutern wir, woher der Kompaktheitssatz seinen Namen hat. Dazu konstruieren wir aus den Strukturen zu einer Sprache L einen topologischen Raum, den *Stoneschen Raum* von L. Der Kompaktheitssatz drückt gerade die Kompaktheit dieses Stoneschen Raumes aus.

Wir setzen in diesem Exkurs Grundkenntnisse aus der Punktmengen-Topologie voraus. Auf Ergebnisse dieses Abschnitts wird später nicht zurückgegriffen. Er kann übersprungen werden.

Da hier die abgeschlossenen Mengen einer Topologie im Vordergrund stehen, verwenden wir folgende Charakterisierung der Kompaktheit eines topologischen Raumes $\mathcal{T} = (X, \mathcal{O})$, wobei X die Punktmenge von $\mathcal{T}$ und $\mathcal{O}$ die Familie der *offenen* Mengen von $\mathcal{T}$ bezeichnet.

10.4.1 Definition Ein topologischer Raum $\mathcal{T}$ ist ein *Hausdorff-Raum*, wenn je zwei verschiedene Punkte von $\mathcal{T}$ durch zwei offene Mengen von $\mathcal{T}$ getrennt werden, wenn es also zu Punkten $x \neq y$ aus $\mathcal{T}$ offene Mengen U, V aus $\mathcal{T}$ gibt, so dass

$$x \in U \text{ und } y \in V \text{ und } U \cap V = \emptyset$$

ist. Der Raum $\mathcal{T}$ ist *kompakt*, wenn $\mathcal{T}$

1. ein Hausdorff-Raum ist und
2. für jede Familie $\mathcal{J}$ von abgeschlossenen Mengen von $\mathcal{T}$ gilt:
   Wenn jede endliche Teilfamilie $\mathcal{J}' \subseteq \mathcal{J}$ einen nicht-leeren Durchschnitt hat, dann hat auch $\mathcal{J}$ einen nicht-leeren Durchschnitt.

Schon diese Formulierung stellt eine Beziehung zum Kompaktheitssatz 10.1.3 her. Die Beziehung wird deutlich, wenn wir die Klasse der Modelle einer Theorie als abgeschlossene Menge auffassen können. Allerdings bilden die sämtlichen Modelle einer konsistenten Theorie keine Menge, sondern eine echte Klasse. Es bietet sich an, Strukturen miteinander zu identifizieren, wenn in ihnen dieselben Sätze gelten.

10.4.2 Definition Zwei Strukturen $\mathcal{A}, \mathcal{B}$ zu einer Sprache L heißen *elementar äquivalent*, $\mathcal{A} \equiv \mathcal{B}$, wenn

$$\mathcal{A}(C) = \mathcal{B}(C)$$

ist für jeden Satz C von L.

Die elementare Äquivalenz beschreibt im Grunde dasselbe Konzept wie die Theorie einer Struktur in 2.2.9; vgl. Aufgabe 7.4.2.

10.4.3 Definition Die Klasse aller Modelle einer Theorie T bezeichnet man mit $\mathrm{Mod}(T)$. Ist T endlich axiomatisiert, so heißt $\mathrm{Mod}(T)$ *elementare Klasse*.

Beispiel Die Klasse aller Gruppen ist eine elementare Klasse. Die Klasse aller unendlichen Gruppen ist $\mathrm{Mod}(T)$ für eine geeignete Erweiterung T der Gruppentheorie, ist aber keine elementare Klasse. Die Klasse aller endlichen Gruppen ist keine Klasse $\mathrm{Mod}(T)$ (vgl. 10.2.6).

10.4.4 Lemma Zu Theorien T_1, T_2 mit derselben Sprache L gibt es eine Theorie T mit Sprache L, so dass

$$\mathrm{Mod}(T) = \mathrm{Mod}(T_1) \cup \mathrm{Mod}(T_2).$$

Beweis: Wir setzen

$$Ax(T) := \{B_1 \vee B_2 \mid B_1 \in Ax(T_1), B_2 \in Ax(T_2)\}.$$

a. Ist $\mathcal{A}$ ein Modell von T_1, so ist $\mathcal{A}(B_1) = w$, also erst recht

$$\mathcal{A}(B_1 \vee B_2) = w$$

für alle $B_i \in Ax(T_i)$ $(i = 1, 2)$. Also ist $\mathcal{A}$ auch ein Modell von T. Entsprechend schließt man für die Modelle von T_2.

b. Ist $\mathcal{A}$ weder Modell von T_1 noch Modell von T_2, so gibt es Axiome B_i von T_i, die in $\mathcal{A}$ nicht gelten, und es ist $\mathcal{A}(B_i) = f$ $(i = 1, 2)$. Dann ist auch

$$\mathcal{A}(B_1 \vee B_2) = f,$$

so dass $\mathcal{A}$ auch kein Modell von T ist.

Mit a. und b. ist das Lemma bewiesen.

10.4.5 Lemma Zu jeder Menge $\mathcal{I}$ von Theorien mit derselben Sprache L gibt es eine Theorie T^* mit Sprache L, so dass

$$\mathrm{Mod}(T^*) = \bigcap\{\mathrm{Mod}(T) \mid T \in \mathcal{I}\}.$$

Beweis: Wir setzen

$$Ax(T^*) := \bigcup\{Ax(T) \mid T \in \mathcal{I}\}.$$

Dann gilt: $\mathcal{A}$ ist Modell von T^*

$\Leftrightarrow$ Für jedes $T \in \mathcal{I}$ ist jedes Axiom von T $\mathcal{A}$-gültig

$\Leftrightarrow$ Für jedes $T \in \mathcal{I}$ ist $\mathcal{A}$ Modell von T

$\Leftrightarrow$ $\mathcal{A} \in \bigcap\{\mathrm{Mod}(T) \mid T \in \mathcal{I}\}$.

10.4.6 Definition Es sei L eine feste Sprache der ersten Stufe. Für jede Struktur $\mathcal{A}$ zu L sei

$$\mathcal{A}/\!\equiv \text{ ein Repräsentant der Klasse } \{\mathcal{B} \mid \mathcal{A} \equiv \mathcal{B}\}.$$

Entsprechend sei für jede Theorie T mit Sprache L

$$\mathrm{Mod}(T)/\!\equiv := \{\mathcal{A}/\!\equiv \mid \mathcal{A} \text{ ist Modell von } T\}.$$

(Offenbar ist mit $\mathcal{A}$ auch $\mathcal{A}/\!\equiv$ ein Modell von T.) Dann setzen wir

$X(L) := \{\mathcal{A}/\!\equiv \mid \mathcal{A}$ ist Struktur zu $L\}$,
$\mathcal{A}(L) := \{\mathrm{Mod}(T)/\!\equiv \mid T$ ist Theorie zu $L\}$,
$\mathcal{O}(L) := \{X(L) - \mathrm{Mod}(T)/\!\equiv \mid T$ ist Theorie zu $L\}$.
$(X(L), \mathcal{O}(L))$ ist der *Stonesche Raum* zur Sprache L.

Die Punkte des Stoneschen Raumes zu L sind also die (Repräsentanten der) *Klassen elementar äquivalenter Strukturen* zu L. Die *Modellklassen von Theorien* mit Sprache L liefern die abgeschlossenen Mengen des Raumes, ihre *Komplemente* die offenen Mengen.

Da elementar äquivalente Strukturen zu L denselben Punkt von $X(L)$ definieren, kann $X(L)$ nur so viele Punkte enthalten, wie es Theorien $Th(\mathcal{A})$ zur Sprache L gibt. Man kann $X(L)$, wenn man will, mit der Menge aller maximal konsistenten Theorien zu L identifizieren. Daher ist $X(L)$ eine Menge, im Gegensatz zur Klasse aller Strukturen zu L, die keine Menge ist.

10.4.7 Satz Der Stonesche Raum $(X(L), \mathcal{O}(L))$ zu einer Sprache L ist ein kompakter topologischer Raum.

Beweis. Wir weisen zunächst die drei Eigenschaften abgeschlossener Mengen nach, die topologische Räume definieren.

Zu A1: Die Vereinigung von zwei abgeschlossenen Mengen ist abgeschlossen.
Ist $A_i = \mathrm{Mod}(T_i)/\equiv$ für $i = 1, 2$, so ist

$$A_1 \cup A_2 = (\mathrm{Mod}(T_1) \cup \mathrm{Mod}(T_2))/\equiv,$$

so dass nach 10.4.4 $A_1 \cup A_2 \in \mathcal{A}(L)$ ist.

Zu A2: Der Durchschnitt jeder Menge von abgeschlossenen Mengen ist abgeschlossen.
Ist $A_T = \mathrm{Mod}(T)/\equiv$ für $T \in \mathcal{I}$, so ist

$$\bigcap\{A_T \mid T \in \mathcal{I}\} = \bigcap\{\mathrm{Mod}(T) \mid T \in \mathcal{I}\}/\equiv$$

so dass nach 10.4.5 $\bigcap\{A_T \mid T \in \mathcal{I}\} \in \mathcal{A}(L)$ ist.

Zu A3: $\emptyset$ und $X(L)$ sind abgeschlossen.

$$\begin{aligned} \emptyset &= \mathrm{Mod}(L, \{\bot\})/\equiv \in \mathcal{A}(L), \quad \text{und} \\ X(L) &= \mathrm{Mod}(L, \emptyset)/\equiv \in \mathcal{A}(L). \end{aligned}$$

Also ist $(X(L), \mathcal{O}(L))$ ein topologischer Raum.

Zur Hausdorff-Eigenschaft.
Seien $\mathcal{A}/\equiv$ und $\mathcal{B}/\equiv$ verschiedene Punkte aus $X(L)$. Dann ist nicht $\mathcal{A} \equiv \mathcal{B}$. Es gibt also einen Satz C aus L, der in $\mathcal{A}$ gilt, aber nicht in $\mathcal{B}$. Dann ist

$$\mathcal{A}/\equiv \in \mathrm{Mod}(L, \{C\})/\equiv \text{ und } \mathcal{B}/\equiv \in \mathrm{Mod}(L, \{\neg C\})/\equiv,$$

und diese beiden abgeschlossenen Mengen sind disjunkt.

Zur Kompaktheit.
Es sei $\mathcal{I} \subseteq \mathcal{A}(L)$ und $\bigcap \mathcal{I}_{\mathrm{fin}} \neq \emptyset$ für jede endliche Teilmenge $\mathcal{I}_{\mathrm{fin}} \subseteq \mathcal{I}$. Wir setzen

$$\mathcal{T} := \{T \mid T \text{ ist Theorie mit Sprache } L, \mathrm{Mod}(T)/\equiv \in \mathcal{I}\}.$$

Dann ist

$$\mathcal{I} = \{\mathrm{Mod}(T)/\equiv \mid T \in \mathcal{T}\}.$$

Wegen 10.4.1 ist zu zeigen, dass $\bigcap \mathcal{I}$ nicht leer ist. Zu jeder endlich axiomatisierten Teiltheorie T' von

$$T^* = (L, \bigcup\{Ax(T) \mid T \in \mathcal{T}\})$$

gibt es eine endliche Teilmenge $\mathcal{T}_{\text{fin}}$ von $\mathcal{T}$ mit

$$T' \prec (L, \bigcup\{Ax(T) \mid T \in \mathcal{T}_{\text{fin}}\}).$$

Für

$$\mathcal{I}_{\text{fin}} = \{\text{Mod}(T)/\equiv \mid T \in \mathcal{T}_{\text{fin}}\}$$

ist (vgl. 10.4.5)

$$\emptyset \neq \bigcap \mathcal{I}_{\text{fin}} = \text{Mod}(L, \bigcup\{Ax(T) \mid T \in \mathcal{T}_{\text{fin}}\})/\equiv \subseteq \text{Mod}(T')/\equiv .$$

Also hat jede endlich axiomatisierte Teiltheorie von T^* ein Modell. Nach dem Kompaktheitssatz hat dann auch T^* ein Modell. Mit 10.4.5 folgt

$$\emptyset \neq \text{Mod}(T^*)/\equiv = \bigcap \mathcal{I}.$$

Also ist $(X(L), \mathcal{O}(L))$ kompakt.

Damit ist der Satz bewiesen.

Es ist also gerade der Kompaktheitssatz, der die Kompaktheit des Stoneschen Raumes zu L ausdrückt. Das ist sicherlich ein hinreichendes Motiv für die Namensgebung dieses Satzes. Wir fügen noch eine weitere Eigenschaft des Stoneschen Raumes an.

10.4.8 Definition Ein topologischer Raum $(X, \mathcal{O})$ heißt *total unzusammenhängend*, wenn es zu verschiedenen Punkten $x, y \in X$ stets eine zugleich offene und abgeschlossene Menge U gibt, so dass $x \in U$ und $y \in X - U$ ist.

Beispiel. Der Raum der reellen Zahlen ist zusammenhängend: die einzigen zugleich offenen und abgeschlossenen Mengen sind $\emptyset$ und $\mathbb{R}$. Dagegen ist der Raum der rationalen Zahlen (mit der von den offenen Intervallen erzeugten Topologie) total unzusammenhängend. Z. B. ist $]-\infty, \sqrt{2}[$ eine zugleich offene und abgeschlossene Menge, die 0, aber nicht 2 enthält. Sie ist abgeschlossen, weil $\sqrt{2} \notin \mathbb{Q}$ ist und deshalb ihr Komplement $]\sqrt{2}, +\infty[$ offen ist.

10.4.9 Korollar Der Stonesche Raum $(X(L), \mathcal{O}(L))$ ist total unzusammenhängend.

Beweis. Die im Nachweis der Hausdorff-Eigenschaft angegebenen Theorien $(L, \{C\})$ und $(L, \{\neg C\})$ definieren zugleich abgeschlossene und offene Mengen in $X(L)$, weil sie zueinander komplementär sind.

10.5 Aufgaben

10.5.1 Zeigen Sie: Es gibt keine endlich axiomatisierte Theorie, deren Modelle genau die Körper der Charakteristik 0 sind: Die Theorie $T_{K,O}$ lässt sich nicht endlich axiomatisieren.

10.5.2 Der Begriff der Wohlordnung wurde in Aufgabe 8.5.7 eingeführt. Zeigen Sie:
Es gibt keine Theorie mit einem unendlichen Modell, deren sämtliche Modelle Wohlordnungen sind.

10.5.3 Die Theorie T_{OK} der geordneten Körper hat die Sprache $L(T_K)+\{<\}$, wobei $<$ ein 2-stelliges Relationszeichen ist, und die Axiome von T_K, LO und schließlich

$$\forall x \forall y \forall z (x < y \to x + z < y + z)$$
$$\forall x \forall y \forall z (x < y \to 0 < z \to x \cdot z < y \cdot z).$$

Die *geordneten Körper* sind die Modelle von T_{OK}. Ein geordneter Körper K ist *archimedisch geordnet*, wenn es zu jedem $k \in K$ ein $n := 1+\ldots+1$ (n-mal) gibt, so dass $k < n$ ist. (Alle Unterkörper von $\mathbb{R}$ sind archimedisch geordnet.) Zeigen Sie: Es gibt keine Theorie, deren Modelle genau die archimedisch geordneten Körper sind.

§11 Morphismen

11.1 Homomorphismen

11.2 Unterstrukturen und Einbettungen

11.3 Diagramme

11.4 Aufgaben

Die Modelltheorie ist das Teilgebiet der Logik, das die stärksten Beziehungen zur Algebra hat. Zu ihrer Entwicklung verwendet sie grundlegende algebraische Konstruktionen, wie sie in der sogenannten Universellen Algebra zusammengefasst werden. Man hat gelegentlich die Modelltheorie als

$$\textit{Universelle Algebra + Sprachbewusstsein}$$

charakterisiert. Wie dem auch sei, wir stellen hier, ausgehend vom Homomorphismusbegriff, einige Grundgedanken der Universellen Algebra mit ihrem Bezug zu Sprachen der ersten Stufe zusammen.

11.1 Homomorphismen

In diesem Paragraphen bezeichnen wir häufig Funktionen

$$\varphi : X \to Y \text{ und } \psi : X \to Z$$

mit demselben *Definitionsbereich* $dom(\varphi) = X$, bei denen $\varphi(x) = \psi(x)$ für alle $x \in X$ ist, die also denselben *Graphen*

$$graph(\varphi) = \{(x, \varphi(x)) \mid x \in X\}$$

haben, mit demselben Buchstaben. Wir „identifizieren“ also häufig Funktionen φ mit ihren Graphen $graph(\varphi)$. Dann ist jede Funktion φ surjektiv auf ihr *Bild*

$$im(\varphi) = \{\varphi(x) \mid x \in X\}.$$

Die Surjektivität von $\varphi : X \to Y$ ist eben eine Eigenschaft des *Ziels* Y und jedenfalls nicht des Graphen von φ, im Gegensatz zur Injektivität von φ.

Als *Identität id auf* X bezeichnen wir dann jede Funktion $\varphi : X \to Y$, für die $\varphi(x) = x$ für alle $x \in X$ und Y eine Obermenge von X ist.

Unter Verwendung der *Komposition* $\psi \circ \varphi$ (lies: ψ *nach* φ) von Funktionen $\varphi : X \to Y$ und $\psi : Y \to Z$, gegeben durch

$$\psi \circ \varphi(x) = \psi(\varphi(x)) \text{ für alle } x \in X,$$

können wir dann jede Funktion $\varphi : X \to Y$ als Komposition einer identischen nach einer surjektiven Funktion schreiben:

$$\varphi = id \circ \varphi : X \to Y,$$

wobei φ auf der rechten Seite als surjektive Funktion $\varphi : X \to im(\varphi)$ und $id : im(\varphi) \to Y$ gelesen werden kann.

Sei nun $\varphi : X^n \to X$ und Y eine Teilmenge von X. Die Funktion $\psi : Y^n \to X$, für die

$$\psi(x_1, \ldots, x_n) = \varphi(x_1, \ldots, x_n) \text{ für alle } x_1, \ldots, x_n \in Y$$

ist, ist bekanntlich die *Beschränkung* von φ auf Y^n, bezeichnet mit $\varphi \restriction Y^n$. Ist zusätzlich Y *abgeschlossen unter* φ, d. h. ist

$$\varphi(x_1, \ldots, x_n) \in Y \text{ für alle } x_1, \ldots, x_n \in Y,$$

so ist nach unserer Konvention $\varphi \restriction Y^n$ auch eine Funktion von Y^n in Y.
Sei wieder $\varphi : X \to Y$ eine Abbildung und $X_0 \subseteq X$. Das *Bild von* X_0 *unter* φ ist die Menge

$$im(\varphi \restriction X_0) = \{\varphi(x) \in Y \mid x \in X_0\},$$

die man auch mit $\varphi[X_0]$ bezeichnet. Speziell ist dann $im(\varphi) = \varphi[X]$.

Nach diesem Vorspann kommen wir zum zentralen Begriff dieses Paragraphen.

11.1.1 Definition $\mathcal{A}$ und $\mathcal{B}$ seien Strukturen zu einer Sprache L. Eine Abbildung

$$\varphi : |\mathcal{A}| \to |\mathcal{B}|$$

ist ein *Homomorphismus* von $\mathcal{A}$ in $\mathcal{B}$, wir schreiben

$$\varphi : \mathcal{A} \to \mathcal{B},$$

wenn für jedes n-stellige Funktionszeichen f aus L

(1) $\varphi(f_{\mathcal{A}}(k_1, \ldots, k_n)) = f_{\mathcal{B}}(\varphi(k_1), \ldots, \varphi(k_n))$

und für jedes n-stellige Prädikatszeichen p aus L

(2) $(k_1, \ldots, k_n) \in p_{\mathcal{A}} \Rightarrow (\varphi(k_1), \ldots, \varphi(k_n)) \in p_{\mathcal{B}}$

für alle $k_1, \ldots, k_n \in |\mathcal{A}|$ gilt.
Ist φ surjektiv, so heißt $\mathcal{B}$ auch das *homomorphe Bild von* $\mathcal{A}$ *unter* φ, bezeichnet mit $\varphi(\mathcal{A})$.

Bemerkung. Für Konstanten c aus L bedeutet (1) insbesondere

(1^0) $\varphi(c_{\mathcal{A}}) = c_{\mathcal{B}}$.

11.1.2 Beispiel (Gruppen)
Sind $G_i = (|G_i|; e_i, {}_i^{-1}, \circ_i)$ für $i = 1, 2$ zwei Gruppen, so ist φ ein Homomorphismus von G_1 in G_2, wenn für alle $k, l \in |G_1|$ gilt:

(1.1) $\varphi(k \circ_1 l) = \varphi(k) \circ_2 \varphi(l)$

(1.2) $\varphi(k_1^{-1}) = \varphi(k)_2^{-1}$

(1.3) $\varphi(e_1) = e_2$.

11.1.3 Lemma Für Gruppen G_1, G_2 ist eine Abbildung

$$\varphi : |G_1| \to |G_2|$$

bereits ein Homomorphismus, wenn φ die Gleichung (1.1) erfüllt.

Beweis. Zu (1.3): Für $k \in G_1$ ist $k \circ_1 e_1 = k$, also wegen (1.1)

$$\varphi(k) \circ_2 \varphi(e_1) = \varphi(k \circ_1 e_1) = \varphi(k).$$

Bekanntlich gilt in jeder Gruppe und damit in G_2

$$a \circ b = a \to b = e.$$

Belegt man a mit $\varphi(k)$ und b mit $\varphi(e_1)$, so folgt

$$\varphi(e_1) = e_2.$$

Zu (1.2): Für $k \in G_1$ ist $k \circ_1 k_1^{-1} = e_1$, also wegen (1.1) und (1.3)

$$\varphi(k) \circ_2 \varphi(k_1^{-1}) = \varphi(k \circ_1 k_1^{-1}) = \varphi(e_1) = e_2.$$

Bekanntlich gilt in jeder Gruppe, also auch in G_2

$$a \circ b = e \to b = a^{-1}.$$

Belegt man a mit $\varphi(k)$ und b mit $\varphi(k_1^{-1})$, so folgt

$$\varphi(k_1^{-1}) = \varphi(k)_2^{-1}.$$

In der Algebra werden Gruppenhomomorphismen allein durch (1.1) definiert, und das sind hiernach genau die Homomorphismen zwischen Gruppen nach unserer „universellen“ Definition 11.1.1. Für den Begriff des Gruppenhomomorphismus ist es daher gleichgültig, ob wir die Sprache der Gruppentheorie mit den drei Funktionszeichen e, $^{-1}$ und $\circ$ formulieren wie in 1.2.2, oder ob wir nur das Zeichen $\circ$ für die Gruppenoperation verwenden.

Wir kehren zur allgemeinen Situation zurück und fragen, wie ein Homomorphismus $\varphi : \mathcal{A} \to \mathcal{B}$ auf die Werte geschlossener Terme und auf die Wahrheit von Sätzen der Sprache $L(\mathcal{A})$ wirkt.

11.1.4 Definition Es sei φ ein Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$, die beide Strukturen zu L seien. Ist $\mathcal{F}$ eine Nennform aus $L(\mathcal{A})$, so bezeichne $\varphi(\mathcal{F})$ die Zeichenreihe, die aus $\mathcal{F}$ hervorgeht, wenn man jede Konstante $c \in |\mathcal{A}|$ durch die Konstante $\varphi(c) \in |\mathcal{B}|$ ersetzt.

$\varphi(\mathcal{F})$ ist dann eine Nennform aus $L(\mathcal{B})$, von derselben Länge und mit denselben Variablen und Nennzeichen wie $\mathcal{F}$. Ist $\mathcal{F}$ ein Term oder eine Formel aus $L(\mathcal{A})$, so ist $\varphi(\mathcal{F})$ ein Term bzw. eine Formel aus $L(\mathcal{B})$. Ist $\mathcal{F}$ aus L, so sind $\mathcal{F}$ und $\varphi(\mathcal{F})$ identisch.

11.1.5 Lemma Es sei φ ein Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$. Für geschlossene Terme t von $L(\mathcal{A})$ ist

$$\varphi(\mathcal{A}(t)) = \mathcal{B}(\varphi(t)).$$

Beweis durch Induktion nach dem Aufbau von t:

1. t ist eine Konstante $c \in |\mathcal{A}|$. Dann ist

$$\varphi(\mathcal{A}(c)) = \varphi(c) = \mathcal{B}(\varphi(c)).$$

2. t ist $ft_1 \dots t_n$. Dann ist

$$\begin{aligned}\varphi(\mathcal{A}(t)) &= \varphi(f_{\mathcal{A}}(\mathcal{A}(t_1), \dots, \mathcal{A}(t_n))) \\ &= f_{\mathcal{B}}(\varphi(\mathcal{A}(t_1)), \dots, \varphi(\mathcal{A}(t_n))) \\ &= f_{\mathcal{B}}(\mathcal{B}(\varphi(t_1)), \dots, \mathcal{B}(\varphi(t_n))) \quad \text{nach IV} \\ &= \mathcal{B}(f\varphi(t_1) \dots \varphi(t_n)) = \mathcal{B}(\varphi(t)).\end{aligned}$$

Mit Induktion folgt aus 1. und 2. die Behauptung.

Damit ist der Wertetransport für geschlossene Terme befriedigend geklärt. Der Wahrheitstransport für Sätze ist komplizierter. Ist $c = d$ in $\mathcal{A}$, so ist selbstverständlich $\varphi(c) = \varphi(d)$ in $\mathcal{B}$, wenn φ überhaupt eine Abbildung von $|\mathcal{A}|$ in $|\mathcal{B}|$ ist. Aus $c \neq d$ in $\mathcal{A}$ folgt aber $\varphi(c) \neq \varphi(d)$ für beliebige $c, d \in |\mathcal{A}|$ nur genau dann, wenn φ injektiv ist. Eigenschaften von φ beeinflussen also die Klasse der Sätze, die unter φ wahr bleiben.

11.1.6 Definition Eine Formel heißt *positiv*, wenn sie aus Primformeln allein mit $\wedge, \vee, \forall, \exists$ aufgebaut ist. Sie heißt $\exists$-*Formel*, wenn sie eine Gestalt

$$\exists x_1 \dots \exists x_n G(x_1, \dots, x_n)$$

mit quantorenfreiem Kern $G(a_1, \dots, a_n)(n \geq 0)$ hat.

Diese Formulierung soll nicht andeuten, dass jetzt $\wedge, \vee, \exists$ etwa als Grundzeichen zu lesen wären. Diese Partikel sind wie bisher als definierte Partikel zu lesen (vgl. 1.1.12). Ein positiver $\exists$-Satz ist dann ein $\exists$-Satz, dessen Kern positiv und quantorenfrei ist.

11.1.7 Lemma Ist φ ein Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$ und C ein positiver $\exists$-Satz von $L(\mathcal{A})$, so gilt:

$$\text{Aus } \mathcal{A}(C) = w \text{ folgt } \mathcal{B}(\varphi(C)) = w.$$

Beweis durch Induktion nach dem Aufbau von C:

1. C ist eine Gleichung $s = t$. Dann gilt

$$\begin{aligned}\mathcal{A}(C) = w &\Rightarrow \mathcal{A}(s) = \mathcal{A}(t) \\ &\Rightarrow \mathcal{B}(\varphi(s)) = \varphi(\mathcal{A}(s)) = \varphi(\mathcal{A}(t)) = \mathcal{B}(\varphi(t)) \text{ nach 11.1.5} \\ &\Rightarrow \mathcal{B}(\varphi(s) = \varphi(t)) = w \Rightarrow \mathcal{B}(\varphi(C)) = w.\end{aligned}$$

2. C ist eine Primformel $pt_1 \dots t_n$. Dann gilt

$$\begin{aligned}\mathcal{A}(C) = w \quad &\Rightarrow (\mathcal{A}(t_1), \dots, \mathcal{A}(t_n)) \in p_{\mathcal{A}} \\ &\Rightarrow (\varphi(\mathcal{A}(t_1)), \dots, \varphi(\mathcal{A}(t_n))) \in p_{\mathcal{B}} \\ &\Rightarrow (\mathcal{B}(\varphi(t_1)), \dots, \mathcal{B}(\varphi(t_n))) \in p_{\mathcal{B}} \text{ nach } 11.1.5 \\ &\Rightarrow \mathcal{B}(p\varphi(t_1) \dots \varphi(t_n)) = w \Rightarrow \mathcal{B}(\varphi(C)) = w.\end{aligned}$$

3. $\mathcal{A}(\bot) = \mathcal{B}(\bot) = f$.

4. C ist $C_1 \wedge C_2$ bzw. $C_1 \vee C_2$. Dann gilt

$$\begin{aligned}\mathcal{A}(C) = w \quad &\Rightarrow \mathcal{A}(C_1) = w \text{ und (oder) } \mathcal{A}(C_2) = w \\ &\Rightarrow \mathcal{B}(\varphi(C_1)) = w \text{ und (oder) } \mathcal{B}(\varphi(C_2)) = w \text{ nach IV} \\ &\Rightarrow \mathcal{B}(\varphi(C)) = w.\end{aligned}$$

5. C ist $\exists x\, \mathcal{F}(x)$. Dann gilt

$$\begin{aligned}\mathcal{A}(C) = w \quad &\Rightarrow \mathcal{A}(\mathcal{F}(c)) = w \text{ für ein } c \in |\mathcal{A}| \\ &\Rightarrow \mathcal{B}(\varphi(\mathcal{F})(\varphi(c))) = w \text{ für ein } c \in |\mathcal{A}| \text{ nach IV} \\ &\Rightarrow \mathcal{B}(\varphi(\mathcal{F})(d)) = w \text{ für ein } d \in |\mathcal{B}|, \text{ nämlich für } d = \varphi(c) \\ &\Rightarrow \mathcal{B}(\varphi(C)) = w.\end{aligned}$$

Mit Induktion folgt aus 1. bis 5. die Behauptung.

Wir haben hier etwas mehr bewiesen als behauptet: In dem Satz C brauchen die Existenzquantoren nicht alle am Anfang von C zu stehen, sondern C kann aus Primformeln beliebig mit $\wedge, \vee, \exists$ aufgebaut sein. Warum nicht auch mit $\forall$?

Für einen Existenzsatz, der in $\mathcal{A}$ gilt, muss es ein Beispiel $c \in |\mathcal{A}|$ geben, und dessen Bild $\varphi(c) \in |\mathcal{B}|$ ist ein Beispiel dafür, dass das φ-Bild des Existenzsatzes in $\mathcal{B}$ gilt.

Wenn ein Allsatz $\forall x\, \mathcal{F}(x)$ in $\mathcal{A}$ gilt, ist $\mathcal{A}(\mathcal{F}(c)) = w$ für alle $c \in |\mathcal{A}|$. Auch wenn φ die Wahrheit aller dieser Sätze $\mathcal{F}(c)$ erhält, folgt $\mathcal{B}(\varphi(\mathcal{F})(d)) = w$ nur für die d aus dem Bild von φ, $d \in im(\varphi)$, und das sind erst dann alle $d \in |\mathcal{B}|$, wenn $im(\varphi) = |\mathcal{B}|$, wenn also φ surjektiv ist.

11.1.8 Lemma Ist φ ein surjektiver Homomorphismus von $\mathcal{A}$ auf $\mathcal{B}$ und C ein positiver Satz von $L(\mathcal{A})$, so gilt

$$\text{Aus } \mathcal{A}(C) = w \text{ folgt } \mathcal{B}(\varphi(C)) = w.$$

Der Beweis erfolgt durch Induktion nach dem Aufbau der positiven Sätze C von $L(\mathcal{A})$ und übernimmt alle Induktionsschritte des Beweises von Lemma 11.1.7. Zu ergänzen bleibt nur der Schritt

6. C ist $\forall x\, \mathcal{F}(x)$. Dann gilt

$$\begin{aligned} \mathcal{A}(C) = w \quad &\Rightarrow \mathcal{A}(\mathcal{F}(c)) = w \text{ für alle } c \in |\mathcal{A}| \\ &\Rightarrow \mathcal{B}(\varphi(\mathcal{F})(\varphi(c))) = w \text{ für alle } c \in |\mathcal{A}| \text{ nach IV} \\ &\Rightarrow \mathcal{B}(\varphi(\mathcal{F})(d)) = w \text{ für alle } d \in im(\varphi), \\ &\qquad \text{also für alle } d \in |\mathcal{B}|, \text{ weil } \varphi \text{ surjektiv ist} \\ &\Rightarrow \mathcal{B}(\varphi(C)) = w. \end{aligned}$$

Mit Induktion nach C folgt wieder die Behauptung.

Mit diesem Lemma können wir nun eine Aussage über Modellklassen geeigneter Theorien gewinnen.

11.1.9 Satz Es sei T eine Theorie, deren nicht-logische Axiome positiv sind. Dann sind die homomorphen Bilder von Modellen von T wieder Modelle von T:

Die Klasse der Modelle von T ist abgeschlossen gegen Homomorphismen.

Beweis. Es sei $\mathcal{A}$ ein Modell von T und $\mathcal{B}$ homomorphes Bild von $\mathcal{A}$, also $\mathcal{B} = \varphi(\mathcal{A})$ für einen Homomorphismus φ. Dann ist $\varphi : \mathcal{A} \to \mathcal{B}$ surjektiv nach 11.1.1. Ist $C \in Ax(T)$, so ist C ein positiver Satz aus $L(T)$, der in $\mathcal{A}$ gilt. Nach Lemma 11.1.8 gilt C dann auch in $\mathcal{B}$. Also ist $\mathcal{B}$ ein Modell von T.

11.1.10 Beispiele

1. Selbstverständlich gibt es Homomorphismen φ von einer kommutativen Gruppe G in eine nicht-kommutative Gruppe G'. Für surjektives φ ist aber $G' = \varphi(G)$ immer eine kommutative Gruppe, weil die Axiome der Theorie der kommutativen Gruppen nur Allabschlüsse von Gleichungen, also positive Sätze sind.

2. Sei $\varphi : R_1 \to R_2$ ein surjektiver Homomorphismus von R_1, einem Ring mit 1, auf R_2, eine Struktur zur Sprache L_R der Ringe mit 1. Dann gibt es zwei Fälle:

   (1) In R_2 gilt $0 \neq 1$. Dann ist auch R_2 ein Ring mit 1, weil alle *übrigen* Axiome von T_R positiv sind und sich nach Satz 11.1.9 von R_1 auf R_2 übertragen und $0 \neq 1$ in R_2 vorausgesetzt ist.

(2) In R_2 gilt $0 = 1$. Dann ist R_2 der Nullring mit dem einzigen Element 0, ein Ring ohne 1. Denn in R_1 gilt $r \cdot 0 = 0$ und $r \cdot 1 = r$ für alle $r \in R_1$; mit Lemma 11.1.7 folgt dann wegen $\varphi(0) = 0 = 1 = \varphi(1)$ in R_2 auch

$$\varphi(r) = \varphi(r \cdot 1) = \varphi(r) \cdot 1 = \varphi(r) \cdot 0 = \varphi(r \cdot 0) = 0$$

für alle $\varphi(r) \in R_2$, was wegen der Surjektivität von φ alle Elemente von R_2 sind.

3. Körper besitzen keinen „echten“ Homomorphismus, d. h. ist $\varphi : K_1 \to K_2$ ein Homomorphismus zwischen Körpern, so ist φ bereits injektiv. Denn sind $c, d \in |K_1|$ und $c \neq d$, so ist $c - d \neq 0$, und das ist wegen des Körperaxioms äquivalent zu $\exists y\ (c - d) \cdot y = 1$ in K_1, was ein positiver $\exists$-Satz ist. Nach Lemma 11.1.7 gilt dann $\exists y\ (\varphi(c) = \varphi(d)) \cdot y = 1$ in K_2, woraus $\varphi(c) - \varphi(d) \neq 0$ und $\varphi(c) \neq \varphi(d)$ in K_2 folgt.

11.2 Unterstrukturen und Einbettungen

Die surjektiven Homomorphismen stellen kaum eine Einschränkung des Homomorphie-Begriffs dar. Denn jeder Homomorphismus $\varphi : \mathcal{A} \to \mathcal{B}$ ist auch ein surjektiver Homomorphismus von $\mathcal{A}$ auf eine *Unterstruktur* von $\mathcal{B}$.

11.2.1 Definition $\mathcal{A}$ und $\mathcal{B}$ seien Strukturen zu L. $\mathcal{A}$ ist *Unterstruktur* von $\mathcal{B}$, wir schreiben $\mathcal{A} \subseteq \mathcal{B}$, wenn für $A := |\mathcal{A}|$

(3) $A \subseteq |\mathcal{B}|$ und

(4) $\mathcal{A} = (A, (f_{\mathcal{B}} \restriction A^n)_{f \in L}, (p_{\mathcal{B}} \cap A^n)_{p \in L})$

ist, wobei n die Stellenzahl von f bzw. p angibt.

Beispiele.

1. Für Gruppen G_1, G_2 ist $G_1 \subseteq G_2$ genau dann der Fall, wenn G_1 Untergruppe von G_2 ist. Entsprechendes gilt für Ringe mit 1 und Körper.

2. Ist $(X, <)$ eine lineare Ordnung und Y eine Teilmenge von X, so ist $(Y, <') \subseteq (X, <)$, wenn $<'$ die Beschränkung von $<$ auf Y ist, d. h. wenn $<' = < \cap (Y \times Y)$ ist. Jede Unterstruktur von $(X, <)$ erhält man so.

3. Es gibt Einbettungen vom Ring $\mathbb{Z}$ der ganzen Zahlen in den Körper $\mathbb{Q}$ der rationalen Zahlen und von $\mathbb{Q}$ in den Körper $\mathbb{R}$ der reellen Zahlen. Fasst man ganze Zahlen als spezielle rationale Zahlen und rationale Zahlen als spezielle reelle Zahlen auf, so ist $\mathbb{Z}$ Unterstruktur von $\mathbb{Q}$ und $\mathbb{Q}$ Unterstruktur von $\mathbb{R}$.

Bemerkung. Ist $\mathcal{A} \subseteq \mathcal{B}$, so ist $\mathcal{A}$ durch $\mathcal{B}$ und seinen Individuenbereich $|\mathcal{A}|$ eindeutig festgelegt. Die Gleichung (4) kann man als Definition von $\mathcal{A}$ lesen, wenn die Struktur $\mathcal{B}$ und die Menge A gemäß (3) gegeben sind. Wir sagen dann auch, A *definiert* oder *trägt* die Unterstruktur $\mathcal{A}$ von $\mathcal{B}$. $\mathcal{A}$ und $\mathcal{B}$ sind dann Strukturen zur selben Sprache: $\mathcal{A}$ ist vielleicht „kleiner“ als $\mathcal{B}$, aber sprachlich ebenso „reich“ wie $\mathcal{B}$. Das ist in gewisser Weise dual zum Begriff der Beschränkung: Ist $\mathcal{A}$ Beschränkung von $\mathcal{B}$ (auf L), so ist $\mathcal{A}$ vielleicht sprachlich „ärmer“ als $\mathcal{B}$, aber ebenso „groß“ wie $\mathcal{B}$.

11.2.2 Lemma Ist φ ein Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$, so ist das Bild

$$B_0 := im(\varphi)$$

von φ eine Teilmenge von $|\mathcal{B}|$, die eine Unterstruktur $\mathcal{B}_0$ von $\mathcal{B}$ definiert. Dann ist φ ein surjektiver Homomorphismus von $\mathcal{A}$ auf $\mathcal{B}_0$, und es ist

$$\mathcal{B}_0 = \varphi(\mathcal{A}).$$

Beweis. Zu zeigen ist, dass für n-stelliges f aus L und $l_1, \ldots, l_n \in B_0$ auch $f_{\mathcal{B}}(l_1, \ldots, l_n) \in B_0$ ist. Zu $l_i \in B_0 = im(\varphi)$ gibt es $k_i \in |\mathcal{A}|$, so dass $l_i = \varphi(k_i)$ ist $(i = 1, \ldots, n)$. Also ist

$$f_{\mathcal{B}}(l_1, \ldots, l_n) = f_{\mathcal{B}}(\varphi(k_1), \ldots, \varphi(k_n)) = \varphi(f_{\mathcal{A}}(k_1, \ldots, k_n)) \in B_0,$$

weil φ (1) erfüllt. Also ist die Beschränkung $f_{\mathcal{B}} \restriction B_0^n$ eine Funktion von B_0^n in B_0, und B_0 trägt die Unterstruktur

$$\mathcal{B}_0 = (B_0, (f_{\mathcal{B}} \restriction B_0^n)_{f \in L}, (p_{\mathcal{B}} \cap B_0^n)_{p \in L})$$

von $\mathcal{B}$. Der Rest ist nach Definition 11.1.1 klar.

Ist $\mathcal{A} \subseteq \mathcal{B}$, so ist die identische Abbildung id auf $|\mathcal{A}|$, wie man leicht sieht, ein Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$, der stets injektiv, i.a. aber nicht surjektiv ist. id ist das einfachste, aber auch besonders typische Beispiel einer *Einbettung*.

11.2.3 Definition $\mathcal{A}$ und $\mathcal{B}$ seien Strukturen zu L. Ein Homomorphismus $\varphi : \mathcal{A} \to \mathcal{B}$ ist eine *Einbettung* von $\mathcal{A}$ in $\mathcal{B}$, wir schreiben $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$, wenn

(5) φ injektiv ist und

(6) für jedes n-stellige Prädikatszeichen p aus L gilt:

$$(\varphi(k_1), \ldots, \varphi(k_n)) \in p_{\mathcal{B}} \Rightarrow (k_1, \ldots, k_n) \in p_{\mathcal{A}} \text{ für alle } k_1, \ldots, k_n \in |\mathcal{A}|.$$

Ist eine Einbettung $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ surjektiv, also wegen (5) sogar bijektiv, so heißt φ *Isomorphismus von* $\mathcal{A}$ *auf* $\mathcal{B}$, $\varphi : \mathcal{A} \cong \mathcal{B}$.

$\mathcal{A}$ ist *einbettbar in* $\mathcal{B}$, $\mathcal{A} \hookrightarrow \mathcal{B}$, wenn es eine Einbettung von $\mathcal{A}$ in $\mathcal{B}$ gibt. $\mathcal{A}$ ist *isomorph zu* $\mathcal{B}$, wenn es einen Isomorphismus von $\mathcal{A}$ auf $\mathcal{B}$ gibt.

Jede Einbettung $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ ist nach Lemma 11.2.2 also zugleich ein Isomorphismus von $\mathcal{A}$ auf $\varphi(\mathcal{A})$. Und Unterstrukturen lassen sich trivial, nämlich mit der identischen Abbildung einbetten:

11.2.4 Lemma $\mathcal{A} \subseteq \mathcal{B}$ ist äquivalent zu $id : \mathcal{A} \hookrightarrow \mathcal{B}$.

Beweis.

1. Sei $\mathcal{A} \subseteq \mathcal{B}$. Dann ist $A := |\mathcal{A}| \subseteq |\mathcal{B}|$, und $id : A \to |\mathcal{B}|$ ist durch $id(k) = k$ für alle $k \in A$ definiert. Wegen (3) und (4) in 11.2.1 ist

   $$f_{\mathcal{A}}(k_1, \ldots, k_n) = f_{\mathcal{B}}(k_1, \ldots, k_n) \text{ und}$$
   $$(k_1, \ldots, k_n) \in p_{\mathcal{A}} \iff (k_1, \ldots, k_n) \in p_{\mathcal{B}} \text{ für } k_1, \ldots, k_n \in A,$$

   so dass die injektive Abbildung $\varphi = id$ die Eigenschaften (1), (2) und (6) hat. Also ist id Einbettung von $\mathcal{A}$ in $\mathcal{B}$.

2. Sei $id : \mathcal{A} \hookrightarrow \mathcal{B}$. Dann ist $|\mathcal{A}| \subseteq |\mathcal{B}|$, und die Eigenschaften (1), (2) und (6) sagen für $\varphi = id$ gerade dass $f_{\mathcal{A}}, p_{\mathcal{A}}$ die Beschränkungen von $f_{\mathcal{B}}, p_{\mathcal{B}}$ auf $|\mathcal{A}|^n$ sind. Also ist $\mathcal{A} \subseteq \mathcal{B}$.

Hiernach und nach Lemma 11.2.2 lässt sich jeder Homomorphismus $\varphi : \mathcal{A} \to \mathcal{B}$ als Komposition aus einem surjektiven Homomorphismus $\varphi : \mathcal{A} \to \varphi(\mathcal{A})$ und (danach) der identischen Einbettung $id : \varphi(\mathcal{A}) \hookrightarrow \mathcal{B}$ schreiben, ebenso jede Einbettung $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ als Komposition aus einem Isomorphismus $\varphi : \mathcal{A} \cong \varphi(\mathcal{A})$ und (danach) wieder $id : \varphi(\mathcal{A}) \hookrightarrow \mathcal{B}$.

Die Forderung (6) in 11.2.3 ist offenbar die Umkehrung der Eigenschaft (2) aus der Definition 11.1.1. Sie kommt nur für nicht-algebraische Sprachen und Strukturen zum Tragen. Denn für das Gleichheitszeichen $=$ anstelle von p wird sie zu

$$\varphi(k_1) = \varphi(k_2) \Rightarrow k_1 = k_2 \text{ für alle } k_1, k_2 \in |\mathcal{A}|,$$

und das ist die unter (5) geforderte Injektivität von φ.
Damit ist gezeigt:

11.2.5 Lemma Sind $\mathcal{A}$, $\mathcal{B}$ algebraische Strukturen (zu einer algebraischen Sprache L), so ist jeder injektive Homomorphismus $\varphi : \mathcal{A} \to \mathcal{B}$ schon eine Einbettung.

Für nicht-algebraische Sprachen und Strukturen stellt (6) aber eine echte Forderung dar, wie folgendes Beispiel zeigt:

Beispiel. L sei gegeben durch das 2-stellige Prädikatszeichen $<$. Für natürliche Zahlen $k, l \neq 0$ bedeute $k <_1 l$, dass k echter Teiler von l ist, und $k <_2 l$, dass k kleiner als l ist. Da positive echte Teiler einer positiven Zahl l kleiner sind als l, folgt aus $k <_1 l$ stets $k <_2 l$, aber nicht umgekehrt, weil etwa 2 kleiner als 3, aber kein Teiler von 3 ist. Also ist

$$id : (\mathbb{N} - \{0\}, <_1) \to (\mathbb{N} - \{0\}, <_2)$$

ein injektiver (sogar bijektiver) Homomorphismus, aber keine Einbettung.

Wir kehren zur allgemeinen Situation zurück und fragen wieder, von welchen Sätzen die Wahrheit unter Einbettungen erhalten bleibt.

11.2.6 Lemma Ist $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ eine Einbettung, so gilt für jeden quantorenfreien Satz C aus $L(\mathcal{A})$:

$$\mathcal{A}(C) = w \;\Leftrightarrow\; \mathcal{B}(\varphi(C)) = w.$$

Beweis durch Induktion nach dem Aufbau von C.

1. C sei $s = t$. Dann ist

$$\begin{aligned} \mathcal{A}(C) = w \quad &\Leftrightarrow\; \mathcal{A}(s) = \mathcal{A}(t) \\ &\Leftrightarrow\; \mathcal{B}(\varphi(s)) = \varphi(\mathcal{A}(s)) = \varphi(\mathcal{A}(t)) = \mathcal{B}(\varphi(t)) \\ &\qquad \text{nach Lemma 11.1.5 und weil } \varphi \text{ injektiv ist} \\ &\Leftrightarrow\; \mathcal{B}(\varphi(C)) = w. \end{aligned}$$

2. C sei $pt_1 \dots t_n$. Wir schreiben t für $t_1 \dots t_n$. Dann ist

$$\begin{aligned} \mathcal{A}(C) = w \quad &\Leftrightarrow \mathcal{A}(t) \in p_{\mathcal{A}} \\ &\Leftrightarrow \mathcal{B}(\varphi(t)) = \varphi(\mathcal{A}(t)) \in p_{\mathcal{B}} \\ &\quad \text{nach Lemma 11.1.5 und weil } \varphi \text{ Einbettung ist} \\ &\Leftrightarrow \mathcal{B}(\varphi(C)) = w. \end{aligned}$$

3. Für $C \equiv \bot$ ist nichts zu beweisen.

4. C sei $C_1 \to C_2$. Dann ist

$$\begin{aligned} \mathcal{A}(C) = w \quad &\Leftrightarrow \text{ aus } \mathcal{A}(C_1) = w \text{ folgt } \mathcal{A}(C_2) = w \\ &\Leftrightarrow \text{ aus } \mathcal{B}(\varphi(C_1)) = w \text{ folgt } \mathcal{B}(\varphi(C_2)) = w \text{ nach IV} \\ &\Leftrightarrow \mathcal{B}(\varphi(C)) = \mathcal{B}(\varphi(C_1) \to \varphi(C_2)) = w. \end{aligned}$$

Dieses Lemma lässt sich in einer Richtung verschärfen:

11.2.7 Lemma Ist $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ eine Einbettung, so gilt für jeden $\exists$-Satz C von $L(\mathcal{A})$:

$$\text{Aus } \mathcal{A}(C) = w \text{ folgt } \mathcal{B}(\varphi(C)) = w.$$

Beweis. Sei C der $\exists$-Satz $\exists x_1 \dots \exists x_n \mathcal{F}(x_1, \dots, x_n)$ mit quantorenfreiem $\mathcal{F}$. Wir induzieren nach n.

1. Für $n = 0$ ist C quantorenfrei, und 11.2.6 liefert die Behauptung.

2. Der Schritt von n nach $n+1$ ist der Induktionsschritt 5. aus dem Beweis von Lemma 11.1.7

Mit Kontraposition folgt hieraus:

11.2.8 Korollar Ist $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ eine Einbettung, so gilt für jeden $\forall$-Satz C von $L(\mathcal{A})$:

$$\text{Aus } \mathcal{B}(\varphi(C)) = w \text{ folgt } \mathcal{A}(C) = w.$$

Beweis. Angenommen, der $\forall$-Satz $C \equiv \forall x_1 \dots \forall x_n \mathcal{F}(x_1, \dots, x_n)$ sei falsch in $\mathcal{A}$. Dann ist $\mathcal{A}(\exists x_1 \dots \exists x_n \neg \mathcal{F}(x_1, \dots, x_n)) = w$. Daraus folgt mit dem vorigen Lemma $\mathcal{B}(\varphi(\exists x_1 \dots \exists x_n \neg \mathcal{F}(x_1, \dots, x_n))) = w$. Dann ist auch $\mathcal{B}(\varphi(C)) = f$, und Kontraposition ergibt die Behauptung.

Aus dieser Formulierung erhalten wir ein Ergebnis für offene Theorien.

11.2.9 Definition Eine Theorie T nennen wir *offen* oder eine $\forall$-*Theorie*, wenn alle ihre nicht-logischen Axiome $\forall$-Sätze, also Allabschlüsse von quantorenfreien Formeln sind.

11.2.10 Satz Die Unterstrukturen von Modellen einer offenen Theorie T sind wieder Modelle von T.

Beweis. Sei T eine offene Theorie, $C \in Ax(T)$, $\mathcal{B}$ Modell von T und $\mathcal{A} \subseteq \mathcal{B}$. Dann ist C ein $\forall$-Satz von L (also auch von $L(\mathcal{A})$), $\mathcal{B}(C) = w$, und nach Lemma 11.2.4 ist id eine Einbettung von $\mathcal{A}$ in $\mathcal{B}$. Nach dem Korollar 11.2.8 ist dann auch $\mathcal{A}(C) = w$. Das gilt für jedes $C \in Ax(T)$. Also ist $\mathcal{A} \models T$.

Beispiele.

1. Die Unterstrukturen von Gruppen sind wieder Gruppen, von Ringen mit 1 sind wieder Ringe mit 1.

2. Dagegen sind die Unterstrukturen von Körpern i.a. keine Körper, sondern nur Ringe mit 1, weil das Körper-Axiom kein $\forall$-Satz ist.

3. Jede Unterstruktur einer linearen Ordnung ist wieder eine lineare Ordnung; dagegen sind die Unterstrukturen von dichten linearen Ordnungen i.a. nicht dicht: das Dichte-Axiom $LO4$ lässt sich nicht durch einen $\forall$-Satz ausdrücken.

Die Aussage von Lemma 11.2.7 können wir für Isomorphismen (also im surjektiven Fall) wesentlich verschärfen.

11.2.11 Lemma Ist $\varphi : \mathcal{A} \cong \mathcal{B}$ ein Isomorphismus, so gilt für jeden Satz C aus $L(\mathcal{A})$:

$$\mathcal{A}(C) = w \;\Leftrightarrow\; \mathcal{B}(\varphi(C)) = w.$$

Insbesondere sind isomorphe Strukturen elementar äquivalent.

Beweis durch Induktion nach dem Aufbau von C. Die Induktionsschritte 1. bis 4. übernehmen wir aus dem Beweis von Lemma 11.2.6. Zu ergänzen bleibt nur:

5. C sei $\forall x\, \mathcal{F}(x)$. Die Richtung $\Rightarrow$ der Behauptung ist in Lemma 11.1.8 unter 6. bewiesen. Die Richtung $\Leftarrow$ folgt wie im Beweis von Lemma 11.1.7 unter 5. (mit Kontraposition, vgl. 11.2.8).

Mit Induktion nach C folgt die behauptete Äquivalenz. Beschränkt man sich auf Sätze C aus L, so ist $\varphi(C) \equiv C$, und aus dem Bewiesenen folgt $\mathcal{A} \equiv \mathcal{B}$.

Dieses Lemma erscheint eher selbstverständlich als die Lemmata 11.2.6, 11.1.7, 11.1.8. Es muss aber notwendig induktiv nach dem Formelaufbau bewiesen werden, von dem geeignete Teile in den genannten Lemmata abgearbeitet werden.

Aus diesem Lemma folgt unmittelbar:

11.2.12 Satz Isomorphe Strukturen sind Modelle derselben Theorien.

Beweis. Ist $\mathcal{A} \cong \mathcal{B}$, so ist nach dem Lemma $\mathcal{A} \equiv \mathcal{B}$. Ist $\mathcal{A} \models T$, so ist $\mathcal{A}(C) = w$ für jedes $C \in Ax(T)$. Wegen $\mathcal{A} \equiv \mathcal{B}$ ist dann auch $\mathcal{B}(C) = w$ für jedes $C \in Ax(T)$, und es ist $\mathcal{B} \models T$.

11.3 Diagramme

In den letzten Abschnitten handelten einige Ergebnisse davon, dass bestimmte Morphismen $\varphi : \mathcal{A} \to \mathcal{B}$ die Wahrheit geeigneter $L(\mathcal{A})$-Sätze von $\mathcal{A}$ nach $\mathcal{B}$ transportieren. Das ist der Anlass für folgende Begriffsbildung.

11.3.1 Definition Seien $\mathcal{A}$, $\mathcal{B}$ Strukturen zu L, Δ eine Menge von Formeln von L und $\varphi : |\mathcal{A}| \to |\mathcal{B}|$ eine Abbildung. φ *transportiert* Δ *von* $\mathcal{A}$ *nach* $\mathcal{B}$, $\varphi : \mathcal{A} \xrightarrow{\Delta} \mathcal{B}$, wenn für jede Formel $\mathcal{F}(a_1, \ldots, a_n) \in \Delta$ ($\mathcal{F}$ geschlossen) gilt:

$$\text{Aus } \mathcal{A}(\mathcal{F}(k_1, \ldots, k_n)) = w \quad \text{folgt } \mathcal{B}(\mathcal{F}(\varphi(k_1), \ldots, \varphi(k_n))) = w \text{ für alle } k_1, \ldots, k_n \text{ aus } |\mathcal{A}|.$$

Bemerkung. Hier ist $\mathcal{F}(a_1, \ldots, a_n)$ stets eine Formel aus L, deren freie Variablen unter $a_1, \ldots, a_n$ auftreten. Dann ist $\mathcal{F}(k_1, \ldots, k_n)$ ein Satz C aus $L(\mathcal{A})$, und $\mathcal{F}(\varphi(k_1), \ldots, \varphi(k_n))$ ist der Satz $\varphi(C)$ aus $L(\mathcal{B})$. Mit dem Begriff der $\mathcal{A}$-Belegung kann man das Definiens auch so formulieren: Für jede Formel $D \in \Delta$ gilt:

$$\text{Aus } \mathcal{A}(D') = w \text{ folgt } \mathcal{B}(\varphi(D')) = w \text{ für jede } \mathcal{A}\text{-Belegung.}'$$

11.3.2 Beispiele Sei $\varphi : |\mathcal{A}| \to |\mathcal{B}|$ eine Abbildung.

1. In jedem Fall transportiert φ $\{a = b\}$, weil aus $k = l$ (in $\mathcal{A}$) immer $\varphi(k) = \varphi(l)$ (in $\mathcal{B}$) folgt.

2. φ transportiert $\{a \neq b\}$ (mit verschiedenen Variablen a, b) genau dann, wenn φ injektiv ist. Denn $k \neq l \Rightarrow \varphi(k) \neq \varphi(l)$ (für alle $k, l \in |\mathcal{A}|$) charakterisiert die Injektivität von φ:

$$\varphi : \mathcal{A} \xrightarrow{a \neq b} \mathcal{B} \Leftrightarrow \varphi \text{ ist injektiv.}$$

3. Sei *prim* die Menge der Primformeln von L. Wenn φ *prim* transportiert, transportiert φ insbesondere Formeln

$$fa_1 \dots a_n = b \text{ und } pa_1 \dots a_n$$

mit paarweise verschiedenen freien Variablen $a_1, \dots, a_n, b$. Dann erfüllt φ aber (1) und (2) aus Definition 11.1.1 und ist ein Homomorphismus.
Wenn umgekehrt φ ein Homomorphismus ist, transportiert φ nach Lemma 11.1.7 alle positiven $\exists$-Formeln von L, insbesondere alle Primformeln:

$$\varphi : \mathcal{A} \xrightarrow{prim} \mathcal{B} \Leftrightarrow \varphi \text{ ist Homomorphismus.}$$

4. Sei *qf* die Menge der quantorenfreien Formeln von L. Wenn φ *qf* transportiert, transportiert φ insbesondere alle Primformeln und alle negierten Primformeln, auch $a \neq b$. Nach Beispiel 2 und 3 ist φ dann ein injektiver Homomorphismus, und φ transportiert Formeln $\neg pa_1 \dots, a_n$ (mit paarweise verschiedenen $a_1, \dots, a_n$):

$$\text{Aus } \mathcal{A}(\neg pk_1 \dots k_n) = w \text{ folgt} \quad \mathcal{B}(\neg p\varphi(k_1) \dots \varphi(k_n)) = w \text{ für alle } k_1, \dots, k_n \in |\mathcal{A}|.$$

Das ergibt (nach Kontraposition) (6) aus Definition 11.2.3: φ ist eine Einbettung.
Wenn φ umgekehrt eine Einbettung ist, transportiert φ nach Lemma 11.2.6 *qf*. Damit ist gezeigt:

11.3.3 Lemma $\varphi : \mathcal{A} \xrightarrow{qf} \mathcal{B} \Leftrightarrow \varphi$ ist Einbettung.

5. Bezeichne L die Menge aller Formeln aus L. Dann folgt aus Lemma 11.2.11:

$$\varphi : \mathcal{A} \cong \mathcal{B} \Rightarrow \varphi : \mathcal{A} \xrightarrow{L} \mathcal{B}.$$

Die Umkehrung gilt nicht, wie sich in §12 zeigen wird.

Allgemein kann eine Voraussetzung, dass φ ein Δ transportiert, nicht erzwingen, dass φ surjektiv ist. Ein Argument im Beweis von Lemma 11.2.11 war (vgl. Beweis von 11.1.8): Weil φ surjektiv ist, transportiert φ auch Allsätze. Umgekehrt könnte φ „zufällig“ Allsätze transportieren, auch wenn φ nicht surjektiv ist. Zwischen den Begriffen *Einbettung* und *Isomorphismus* tut sich eine Lücke auf, wenn man Homomorphismen nach ihrem Wahrheitstransport klassifizieren will. Und zwar fehlt gerade ein Begriff für die Abbildungen, die die gesamte „elementare“ Sprache L transportieren.

11.3.4 Definition Eine Einbettung $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$ ist eine *elementare Einbettung* von $\mathcal{A}$ in $\mathcal{B}$, wenn

$$\mathcal{A}(C) = w \;\Leftrightarrow\; \mathcal{B}(\varphi(C)) = w$$

ist für alle Sätze C aus $L(\mathcal{A})$. Ist $id : \mathcal{A} \subseteq \mathcal{B}$ eine elementare Einbettung, so heißt $\mathcal{A}$ *elementare Unterstruktur* von $\mathcal{B}$, man schreibt $\mathcal{A} \prec \mathcal{B}$.

11.3.5 Lemma Jede elementare Einbettung φ von $\mathcal{A}$ in $\mathcal{B}$ ist ein Isomorphismus von $\mathcal{A}$ auf die elementare Unterstruktur $\varphi(\mathcal{A})$ von $\mathcal{B}$.

Beweis. Nach Lemma 11.2.2 ist $\varphi : \mathcal{A} \cong \varphi(\mathcal{A})$ und $\varphi(\mathcal{A}) \subseteq \mathcal{B}$. Für alle Sätze C aus $L(\mathcal{A})$ ist nach Lemma 11.2.11 $\mathcal{A}(C) = w$ äquivalent zu $\varphi(\mathcal{A})(\varphi(C)) = w$ und auch äquivalent zu $\mathcal{B}(\varphi(C)) = w$, weil φ elementare Einbettung ist. Weil jeder Satz aus $L(\varphi(\mathcal{A}))$ eine Gestalt $\varphi(C)$ (C aus $L(\mathcal{A})$) hat, ist dann $\varphi(\mathcal{A}) \prec \mathcal{B}$.

Beispiel 5, also Lemma 11.2.11, wird ergänzt zu:

11.3.6 Lemma $\varphi : \mathcal{A} \xrightarrow{L} \mathcal{B} \;\Leftrightarrow\; \varphi$ ist elementare Einbettung.

Beweis. Die Richtung $\Leftarrow$ ist trivial. Transportiere nun φ ganz L. Nach Lemma 11.3.3 ist φ eine Einbettung. Ferner gilt für alle Sätze C aus $L(\mathcal{A})$:

$$\mathcal{A}(C) = w \Rightarrow \mathcal{B}(\varphi(C)) = w, \text{ also auch}$$

$$\mathcal{A}(C) = f \Rightarrow \mathcal{A}(\neg C) = w \Rightarrow \mathcal{B}(\varphi(\neg C)) = w \Rightarrow \mathcal{B}(\varphi(C)) = f.$$

Das ergibt mit Kontraposition:

$$\mathcal{B}(\varphi(C)) = w \Rightarrow \mathcal{A}(C) = w.$$

Damit ist auch die Richtung $\Rightarrow$ des Lemmas bewiesen.

Wahrheitstransporte von $\mathcal{A}$ nach $\mathcal{B}$ verwenden wesentlich die Sprache $L(\mathcal{A})$. Wenn $\mathcal{A}$ und $\mathcal{B}$ Strukturen zu L sind, lassen sie sich immerhin expandieren zu Strukturen zu $L(\mathcal{A})$.

11.3.7 Definition Seien $\mathcal{A}$, $\mathcal{B}$ Strukturen zu L und φ eine Abbildung von $A := |\mathcal{A}|$ in $|\mathcal{B}|$. Die *Expansion von $\mathcal{B}$ um die Konstanten aus $im(\varphi)$* bezeichnen wir mit $(\mathcal{B}, im(\varphi))$. Das ist eine Struktur zu $L(\mathcal{A})$, wobei jedes $k \in A$ durch $\varphi(k) \in |\mathcal{B}|$ interpretiert wird, $k_{(\mathcal{B},im(\varphi))} = \varphi(k)$.

Der einfachste Spezialfall hiervon ist der Fall $\varphi = id$ und $\mathcal{B} = \mathcal{A}$. Dann ist $im(\varphi) = A$, und man erhält $(\mathcal{A}, A)$ als Struktur zu $L(\mathcal{A})$.

11.3.8 Definition Sei $\mathcal{A}$ Struktur zu L. Das *Diagramm von $\mathcal{A}$* ist die Theorie

$$D(\mathcal{A}) := (L(\mathcal{A}), \{C \text{ quantorenfreier Satz aus } L(\mathcal{A}) | \mathcal{A}(C) = w\}).$$

Das *elementare Diagramm von $\mathcal{A}$* ist die Theorie

$$Th((\mathcal{A}, A)) = (L(\mathcal{A}), \{C \text{ Satz aus } L(\mathcal{A}) | \mathcal{A}(C) = w\}).$$

Beide Diagramme sind also Theorien mit Sprache $L(\mathcal{A})$. Beide haben offenbar das Modell $(\mathcal{A}, A)$. Der Name *Diagramm* rührt daher, dass die Axiomensysteme eine genaue Beschreibung von $\mathcal{A}$ in der Sprache $L(\mathcal{A})$ enthalten in folgendem Sinne:

$$\begin{aligned}
&(k_1, \ldots, k_n) \in p_{\mathcal{A}} \Leftrightarrow pk_1 \ldots k_n \in Ax(D(\mathcal{A})) \\
&(k_1, \ldots, k_n) \notin p_{\mathcal{A}} \Leftrightarrow \neg pk_1 \ldots k_n \in Ax(D(\mathcal{A})) \\
&f_{\mathcal{A}}(k_1, \ldots, k_n) = l \Leftrightarrow fk_1 \ldots k_n = l \in Ax(D(\mathcal{A})).
\end{aligned}$$

Wie man sich leicht überlegt, folgen alle Axiome von $D(\mathcal{A})$ bereits aus diesen Primformeln und negierten Primformeln. Das elementare Diagramm enthält als Axiome alle in der elementaren Sprache $L(\mathcal{A})$ formulierbaren Sätze, die in $\mathcal{A}$ wahr sind.

Die sämtlichen Modelle dieser Theorien sind uns im Grunde schon bekannt:

11.3.9 Diagramm-Lemma

1. $\mathcal{A}$ ist einbettbar in $\mathcal{B}$ genau dann, wenn eine $L(\mathcal{A})$-Expansion von $\mathcal{B}$ Modell von $D(\mathcal{A})$ ist.

2. $\mathcal{A}$ ist elementar einbettbar in $\mathcal{B}$ genau dann, wenn eine $L(\mathcal{A})$-Expansion von $\mathcal{B}$ Modell von $Th((\mathcal{A}, A))$ ist.

Beweis von 1.
Sei φ eine Abbildung von $|\mathcal{A}|$ in $|\mathcal{B}|$. Dann folgt: $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$

$$\begin{aligned}
&\Leftrightarrow \quad \varphi : \mathcal{A} \xrightarrow{qf} \mathcal{B} \text{ nach Lemma 11.3.3} \\
&\Leftrightarrow \quad \mathcal{B}(C) = w \text{ für jedes Axiom } C \text{ von } D(\mathcal{A}), \text{ wenn man jede} \\
&\qquad\quad \text{Konstante } k \in |\mathcal{A}| \text{ durch } \varphi(k) \text{ interpretiert} \\
&\Leftrightarrow \quad (\mathcal{B}, im(\varphi)) \models D(\mathcal{A})
\end{aligned}$$

und $(\mathcal{B}, im(\varphi))$ ist die gesuchte Expansion von $\mathcal{B}$. Ist umgekehrt eine Expansion $\mathcal{B}^+$ von $\mathcal{B}$ ein Modell von $D(\mathcal{A})$, so ist, weil $\mathcal{B}$ und $\mathcal{A}$ Strukturen zu L sind, $\mathcal{B}^+$ eine Struktur $(\mathcal{B}, B_0)$ zu $L(\mathcal{A})$, wobei B_0 gerade aus den Interpretationen der Konstanten $k \in |\mathcal{A}|$ besteht:

$$B_0 = \{k_{(\mathcal{B},B_0)} \mid k \in |\mathcal{A}|\}.$$

Dann ist $\varphi : k \mapsto k_{(\mathcal{B},B_0)}$ eine Abbildung von $|\mathcal{A}|$ in $|\mathcal{B}|$, und aufgrund der oben aufgeführten Äquivalenzenkette ist φ die gesuchte Einbettung von $\mathcal{A}$ in $\mathcal{B}$.

Der Beweis von 2. verläuft ebenso. Dabei ist nur *Einbettung* durch *elementare Einbettung*, *qf* durch L, Lemma 11.3.3 durch Lemma 11.3.5 und $D(\mathcal{A})$ durch $Th((\mathcal{A}, |\mathcal{A}|))$ zu ersetzen.

Sicherlich ist in Umkehrung von Satz 11.2.10 eine Theorie, von der jede Unterstruktur eines Modells wieder ein Modell ist, nicht automatisch offen. Denn man könnte dann ein triviales Axiom wie $\exists x\ x = x$ zu der Theorie hinzufügen, ohne ihre Modellklasse zu ändern, und sie wäre dann nicht mehr offen. Man wird sich damit begnügen, die Theorie T durch eine offene Theorie mit denselben Modellen ersetzen zu können. Welche offene Theorie bietet sich an? Ihre Axiome sollen einerseits $\forall$-Sätze und andererseits in T gültig sein. Wir definieren also:

11.3.10 Definition Zwei Theorien (zu derselben Sprache) sind *äquivalent*, wenn sie dieselben Modelle haben. Zu einer Theorie T bezeichne $T_\forall$ die Theorie zur Sprache $L(T)$, deren Axiome die sämtlichen in T gültigen $\forall$-Sätze (Allabschlüsse von quantorenfreien Formeln) sind:

$$T_\forall = (L(T), \{C \in L(T) \mid C \text{ ist } \forall\text{-Satz}, T \models C\}).$$

Nun zeigt sich, dass $T_\forall$ die für die Umkehrung von Satz 11.2.10 zu T gesuchte Theorie ist.

11.3.11 Satz von Łoš und Tarski Wenn jede Unterstruktur eines Modells von T wieder ein Modell von T ist, so ist T äquivalent zu einer offenen Theorie.

Beweis. Unter der Voraussetzung des Satzes zeigen wir, dass T und $T_\forall$ dieselben Modelle haben. Weil jedes Axiom von $T_\forall$ in T gilt, ist auch jedes Modell von T ein Modell von $T_\forall$.

Umgekehrt nehmen wir an, es gebe ein Modell $\mathcal{A}$ von $T_\forall$, das kein Modell von T ist. Was für Modelle $\mathcal{B}^+$ hat dann die Theorie $D(\mathcal{A}) + Ax(T)$? Die Beschränkung $\mathcal{B}$ von $\mathcal{B}^+$ auf $L(T)$ ist dann offenbar ein Modell von T, in das sich $\mathcal{A}$ nach dem Diagramm-Lemma 11.3.9 einbetten lässt mit einer Einbettung φ. Dann ist nach Lemma 11.2.2

$$\mathcal{A} \cong \varphi(\mathcal{A}) \subseteq \mathcal{B}.$$

Nun sind Unterstrukturen von Modellen von T wieder Modelle von T. Da $\mathcal{B}$ Modell von T ist, ist also auch $\varphi(\mathcal{A})$ Modell von T und wegen Satz 11.2.12 auch $\mathcal{A}$ selbst Modell von T, im Widerspruch zu unserer Annahme.

Also hat $D(\mathcal{A}) + Ax(T)$ überhaupt kein Modell. Nach dem Kompaktheitssatz gibt es dann eine endliche Menge $C_1, \ldots, C_m$ von Axiomen von $D(\mathcal{A})$ und damit den quantorenfreien Satz $C \equiv C_1 \wedge \ldots \wedge C_m$, der auch ein Axiom von $D(\mathcal{A})$ ist, so dass die Theorie

$$(L(\mathcal{A}), Ax(T) \cup \{C\})$$

kein Modell hat, also inkonsistent ist. Mit dem Deduktionstheorem folgt

$$(L(\mathcal{A}), Ax(T)) \vdash \neg C.$$

Die Konstanten aus $\mathcal{A}$ treten jetzt nur noch in C, nicht in $Ax(T)$ auf. Ist etwa $C \equiv \mathcal{F}(k_1, \ldots, k_n)$ ($\mathcal{F}$ Nennform aus $L, k_1, \ldots, k_n \in |\mathcal{A}|$), so folgt mit dem Lemma über neue Konstanten

$$T \vdash \neg\, \mathcal{F}(a_1, \ldots, a_n)$$

für paarweise verschiedene freie Variablen $a_1, \ldots, a_n$, also auch

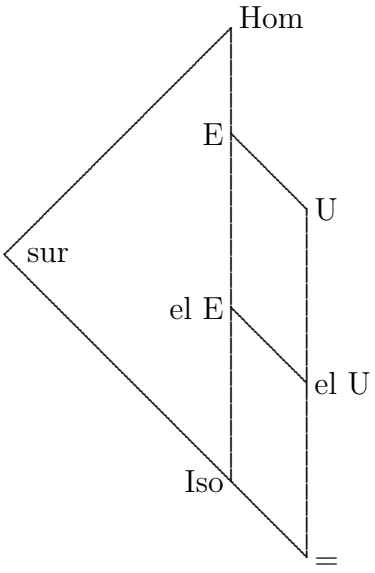
$$T \vdash \forall x_1 \ldots \forall x_n \neg\, \mathcal{F}(a_1, \ldots, a_n).$$

Das ist aber ein $\forall$-Satz, weil C, somit auch $\mathcal{F}$ quantorenfrei ist. Also ist dieser Satz ein Axiom von $T_\forall$ und gilt daher in dem Modell $\mathcal{A}$ von $T_\forall$, im Widerspruch zu $\mathcal{A}(\mathcal{F}(k_1, \ldots, k_n)) = \mathcal{A}(C) = w$, weil C zum Diagramm von $\mathcal{A}$ gehört.

Also war unsere Annahme falsch, und jedes Modell $\mathcal{A}$ von $T_\forall$ ist ein Modell von T. Damit ist der Satz bewiesen.

Der Satz von Łoš und Tarski ist eine Anwendung des Diagramm-Lemmas, die deutlich macht, wie sich universell-algebraische Eigenschaften von Modellklassen in der Gestalt der Axiome einer Theorie widerspiegeln können. Er ist damit ein typischer Satz der Modelltheorie, soweit sie vom Zusammenwirken von universeller Algebra und dem Studium elementarer Sprachen handelt.

Wir fassen die verschiedenen Begriffe von Abbildungen, die wir in diesem Paragraphen eingeführt haben, in einer Figur zusammen:



Legende

Hom	Homomorphismus
sur	surjektiver Homomorphismus
E	Einbettung
U	Unterstruktur
el E	elementare Einbettung
el U	elementare Unterstruktur
Iso	Isomorphismus
=	Identität (von Strukturen)

Oben steht der allgemeine Homomorphie-Begriff, längs der Linien nach unten werden die Begriffe spezieller. Die senkrechte Linie von Hom bis Iso ist die Hauptentwicklungslinie, die dazu parallele Linie enthält die Spezialfälle $\varphi = id$, also die Fälle der Unterstrukturen. Die Schräge von links nach unten enthält die surjektiven Homomorphismen. Von E abwärts wird qf transportiert, von el E abwärts ganz L.

11.4 Aufgaben

11.4.1 (Lineare Algebra) Sei K ein kommutativer Körper. Als Sprache der K-Vektorräume bezeichnen wir die Sprache, die neben den Grundzeichen $0, -, +$ der Sprache der additiven Vektorgruppe zu jedem $\alpha \in K$ ein 1-stelliges Funktionszeichen $\alpha\cdot$ enthält. Für Vektoren v bezeichnet $\alpha\cdot v$ das skalare α-fache von v, αv.

a. Wie schreibt man in dieser Sprache das Assoziativgesetz

$$\alpha(\beta v) = (\alpha \cdot \beta)v?$$

b. Zeigen Sie: Sind V, W K-Vektorräume, so sind die Homomorphismen $\varphi : V \to W$ genau die linearen Abbildungen von V in W.

11.4.2 (Lineare Algebra) Als Sprache der Vektorräume (allgemein) bezeichnen wir die Sprache, die neben den Grundzeichen $0_V, -_V, +_V$ der Sprache der additiven Vektorgruppe und den Grundzeichen aus $L(T_K)$ (für den Koordinatenkörper) noch ein Zeichen $\cdot$ für die skalare Multiplikation enthält.
Zeigen Sie: Sind V, W Vektorräume über demselben Körper, so sind die Homomorphismen $\varphi : V \to W$ genau die semilinearen Abbildungen von V in W. (Hinweis: Verwenden Sie Beispiel 3 aus 11.1.10.)

11.4.3 Wir nennen eine Theorie *positiv axiomatisiert*, wenn alle ihre Axiome positive Sätze sind. Zeigen Sie:

a. Jede positiv axiomatisierte Theorie hat ein 1-elementiges Modell.

b. Es gibt keine zu T_R äquivalente, positiv axiomatisierte Theorie.

11.4.4 Seien $\mathcal{A}, \mathcal{B}$ lineare Ordnungen, also Modelle von LO. Zeigen Sie:

Jeder Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$ ist eine Einbettung.

11.4.5 $<$ sei die natürliche Anordnung auf der Menge $\mathbb{N}$ der natürlichen Zahlen.

a. Wieviele Einbettungen von $(\mathbb{N}, <)$ in sich selbst gibt es?

b. Zeigen Sie: Die einzige elementare Einbettung von $(\mathbb{N}, <)$ in sich selbst ist die Identität.

11.4.6 Zeigen Sie: Die Isomorphie ist eine Äquivalenzrelation (auf der Klasse der Strukturen zu L), also

a. $\mathcal{A} \cong \mathcal{A}$

b. Aus $\mathcal{A} \cong \mathcal{B}$ folgt $\mathcal{B} \cong \mathcal{A}$

c. Aus $\mathcal{A} \cong \mathcal{B}$ und $\mathcal{B} \cong \mathcal{C}$ folgt $\mathcal{A} \cong \mathcal{C}$.

11.4.7 Sei $\mathcal{A}$ eine Struktur zu L. D_0 sei die Menge der Primformeln

$$\begin{array}{l} \neg\ k = l, \text{ falls in } |\mathcal{A}| \quad k \neq l \text{ ist} \\ pk_1 \ldots k_n, \text{ falls } (k_1, \ldots, k_n) \in p_{\mathcal{A}} \\ \neg pk_1 \ldots k_n, \text{ falls } (k_1, \ldots, k_n) \notin p_{\mathcal{A}} \\ fk_1 \ldots k_n = l, \text{ falls } f_{\mathcal{A}}(k_1, \ldots, k_n) = l \end{array}$$

für alle $k_1, \ldots, k_n, k, l \in |\mathcal{A}|$ und alle nicht-logischen Prädikatszeichen p und Funktionszeichen f aus L. Ferner sei D_1 die Menge aller in $\mathcal{A}$ wahren $\exists$-Sätze von $L(\mathcal{A})$. Zeigen Sie:

Die Theorien $(L(\mathcal{A}), D_0)$ und $(L(\mathcal{A}), D_1)$ sind äquivalent zum Diagramm $D(\mathcal{A})$.

11.4.8 Sei $\mathcal{A}$ eine Struktur zu L und φ ein Isomorphismus von $\mathcal{A}$ auf sich selbst. Zeigen Sie:

> φ lässt L-definierbare Teilmengen von $|\mathcal{A}|$ fest, d. h. für jede Formel $\mathcal{F}(a)$ von L ($\mathcal{F}$ geschlossen) ist
>
> $$\{\varphi(k) \mid \mathcal{A}(\mathcal{F}(k)) = w\} = \{k \mid \mathcal{A}(\mathcal{F}(k)) = w\}.$$

§12 Die Sätze von Löwenheim und Skolem

12.1 Kardinalzahlen

12.2 Modelle höherer Mächtigkeit

12.3 Elementare Untermodelle

12.4 Aufgaben

In 10.2 haben wir Beziehungen zwischen endlichen und unendlichen Modellen von Theorien untersucht. Nun gibt es unendliche Mengen und Modelle von wesentlich verschiedener Größe, von, wie man sagt, unterschiedlicher Mächtigkeit. Die Sätze von Löwenheim und Skolem setzen solche Modelle verschiedener unendlicher Mächtigkeiten miteinander in Beziehung in Analogie, aber auch in Erweiterung der Ergebnisse von 10.2. Sie geben einen tiefen Einblick in die Grenzen dessen, was durch mathematische Theorien ausdrückbar ist. Ihre Anwendung auf die Mengenlehre liefert das verblüffende Löwenheim-Skolem-Paradoxon.

12.1 Kardinalzahlen

Was ist die Mächtigkeit einer Menge, und wie rechnet man mit Kardinalzahlen? Wir erläutern diese Begriffe auf der Basis einer Mengenlehre mit Wohlordnungssatz. Für die Ergebnisse, die wir über Kardinalzahlen beweisen, müssen wir ein Stück weit in die Theorie wohlgeordneter Mengen einsteigen. Die Theorie der Mächtigkeiten und Wohlordnungen wird hier nur soweit entwickelt, wie sie für die allgemeinen Sätze von Löwenheim und Skolem unbedingt nötig ist. Das ist allerdings zu den Themen Mächtigkeiten und Wohlordnungen ziemlich genau das, was man beim weiterführenden Studium der Reinen Mathematik ohnehin irgendwann benötigt. Ohne den Formalismus der axiomatischen Mengenlehre einzusetzen, stellen wir dieses Material mathematisch vollständig dar.

12.1.1 Definition Zwei Mengen A, B heißen *gleichmächtig*, $A \sim B$, wenn es eine Bijektion von A auf B gibt.

Beispiele

a) Zwei endliche Mengen sind genau dann gleichmächtig, wenn sie dieselbe endliche Anzahl von Elementen haben.

b) Jede abzählbar unendliche Menge ist zur Menge $\mathbb{N}$ der natürlichen Zahlen gleichmächtig.

12.1.2 Definition Eine Menge A ist *höchstens so mächtig* wie eine Menge $B, A \lesssim B$, wenn es eine injektive Abbildung $\varphi : A \hookrightarrow B$ gibt. A ist *weniger mächtig* als $B, A < B$, wenn $A \lesssim B$, aber nicht $A \sim B$ ist.

Beispiele

a) Jede Teilmenge einer Menge A ist höchstens so mächtig wie A.

b) Ist $m < n \in \mathbb{N}$, so ist jede Menge mit m Elementen weniger mächtig als jede Menge mit n Elementen.

c) Jede endliche Menge ist weniger mächtig als jede unendliche Menge.

Dass auch unendliche Mengen verschieden mächtig sein können, ist eines der frühesten Ergebnisse der Mengenlehre.

12.1.3 Definition Die *Potenzmenge* $Pot(A)$ einer Menge A ist die Menge aller Teilmengen von A:

$$Pot(A) = \{X \mid X \subseteq A\}.$$

12.1.4 Satz von Cantor 1. Cantorsches Diagonalverfahren
Jede Menge ist weniger mächtig als ihre Potenzmenge:

$$A < Pot(A).$$

Beweis. Offenbar ist $\psi : a \mapsto \{a\}$ eine Injektion von A in $Pot(A)$. Also ist $A \lesssim Pot(A)$. Sei nun φ irgendeine Abbildung von A in $Pot(A)$ und sei dazu

$$D = \{a \in A \mid a \notin \varphi(a)\}.$$

(Das ist die Cantorsche Diagonale von φ.) Es ist

$$a \in D \Leftrightarrow a \notin \varphi(a) \text{ (für alle } a \in A).$$

Wäre nun D im Bild von φ, so wäre $D = \varphi(b)$ für ein $b \in A$. Dann wäre

$$a \in D \;\Leftrightarrow\; a \in \varphi(b) \text{ (für alle } a \in A).$$

Für $a = b$ folgt

$$b \in \varphi(b) \;\Leftrightarrow\; b \in D \;\Leftrightarrow\; b \notin \varphi(b),$$

und das ist ein Widerspruch. Also ist die Diagonale $D \subseteq A$ nicht im Bild von φ; φ kann nicht surjektiv und erst recht nicht bijektiv sein.

Also gibt es keine Bijektion von A auf $Pot(A)$, und es ist $A < Pot(A)$.

Beispiele

a) Hat eine endliche Menge n Elemente, so hat ihre Potenzmenge 2^n Elemente, wie man mit Induktion nach n nachrechnet.

b) $Pot(\mathbb{N})$, die Menge aller Mengen von natürlichen Zahlen hat die Mächtigkeit des Kontinuums $\mathbb{R}$:

$$Pot(\mathbb{N}) \sim \mathbb{R}\,.$$

c) $Pot(\mathbb{R})$ hat die Mächtigkeit der Menge ${}^{\mathbb{R}}\mathbb{R}$ aller reellen Funktionen, es folgt

$$Pot^2(\mathbb{N}) := Pot(Pot(\mathbb{N})) \sim Pot(\mathbb{R}) \sim {}^{\mathbb{R}}\mathbb{R}\,.$$

Es gibt also auch zu jeder unendlichen Menge eine mächtigere Menge. Wenn man die Potenzmengenbildung iteriert, etwa in Verallgemeinerung des letzten Beispiels

$$Pot^n(A) := Pot(\ldots Pot(A)\ldots)$$

für jedes $n \in \mathbb{N}$ bildet, so ist die Vereinigung aller dieser Mengen $\bigcup\{Pot^n(A) \mid n \in \mathbb{N}\}$ offenbar mächtiger als jedes einzelne $Pot^n(A)$, und die Potenzmengenbildung kann weitere, noch mächtigere Mengen schaffen: Man beginnt etwas von der gewaltigen Vielfalt der unendlichen Mengen zu ahnen.

Nicht-leere Mengen A, genauer Strukturen $(A, \emptyset, \emptyset) = (A)$ sind Strukturen zur reinen Sprache der Identität, zu der Sprache ohne nicht-logische Grundzeichen. Dann sind Bijektionen $\varphi : A \to B$ Isomorphismen von (A) auf (B), und Injektionen werden Einbettungen. So werden einige Aussagen über Mächtigkeiten zu (trivialen) Spezialfällen von Ergebnissen des vorigen Paragraphen:

12.1.5 Lemma Die Gleichmächtigkeit $\sim$ ist eine Äquivalenzrelation auf der Klasse aller Mengen. Sie ist verträglich mit der Relation $\lesssim$:

$$A \sim A' \lesssim B' \sim B \Rightarrow A \lesssim B.$$

Die Relation $\lesssim$ ist reflexiv und transitiv

$$A \lesssim A$$
$$A \lesssim B \text{ und } B \lesssim C \Rightarrow A \lesssim C.$$

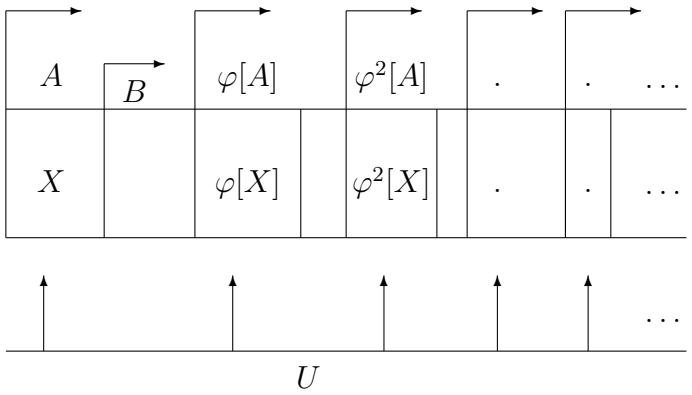
Die Beweise hierzu überlegt man sich leicht selbst. Es drängt sich sofort die Frage auf, ob $\lesssim$ auch antisymmetrisch ist. Diese Frage wird vom Satz von Schröder-Bernstein positiv beantwortet, den wir zunächst in einer scheinbar speziellen Form beweisen.

12.1.6 Lemma Sei $\varphi : A \hookrightarrow A$ injektiv und B eine Teilmenge von A, die $im(\varphi)$ enthält. Dann gibt es eine Bijektion von A auf B.

Beweis. Wir setzen $X = A - B$. Für jedes $n \in \mathbb{N}$ ist $\varphi^n := \varphi \circ \ldots \circ \varphi$ (n-mal) eine Injektion von A in A ($\varphi^0 = id$, $\varphi^1 = \varphi$). Da $X \subseteq A - \varphi[A]$ ist, folgt mit Induktion nach n wegen der Injektivität von φ:

$$\varphi^n[X] \subseteq \varphi^n[A] - \varphi^{n+1}[A].$$

Daher sind alle $\varphi^n[X]$ ($n \in \mathbb{N}$) paarweise disjunkt.



φ bildet jedes $\varphi^n[X]$ bijektiv auf $\varphi[\varphi^n[X]] = \varphi^{n+1}[X]$ ab.
Sei $U := \bigcup\{\varphi^n[X] \mid n \in \mathbb{N}\}$ die Vereinigung aller dieser $\varphi^n[X]$. Dann ist

$$\varphi \restriction U : U \to \varphi[U] = \bigcup\{\varphi^{n+1}[X] \mid n \in \mathbb{N}\} = U - X$$

eine Bijektion. Definieren wir ψ auf A durch

$$\begin{array}{ll} \psi(a) = \varphi(a) & \text{für } a \in U \\ \psi(a) = a & \text{für } a \in A - U, \end{array}$$

so ist ψ hiernach eine Bijektion von A auf

$$\varphi[U] \cup (A - U) = (U - X) \cup (A - U) = A - X = B.$$

Also sind A und B gleichmächtig.

Hieraus folgt der allgemeine Fall sofort:

12.1.7 Satz von Schröder und Bernstein
Lässt sich A injektiv in B und B injektiv in A abbilden, so sind A und B gleichmächtig.

Beweis. Seien $\varphi_1 : A \to B$ und $\varphi_2 : B \to A$ injektiv, ferner $B' := \mathit{im}(\varphi_2)$. Dann gilt für $\varphi = \varphi_2 \circ \varphi_1 : A \to A$:

$$\mathit{im}(\varphi) = \varphi_2[\varphi_1[A]] \subseteq \mathit{im}(\varphi_2) = B' \subseteq A.$$

Also gibt es nach dem Lemma eine Bijektion $\psi : A \to B'$. Dann ist $\varphi_2^{-1} \circ \psi$ eine Bijektion von A auf B, es ist $A \sim B$.

Beispiel. Wenn man jeder Teilmenge $X \subseteq \mathbb{N}$ ihre *charakteristische Funktion* χ_X zuordnet, gegeben durch

$$\begin{array}{ll} \chi_X(n) = 0 & \text{für } n \in X \\ \chi_X(n) = 1 & \text{für } n \notin X, \end{array}$$

so erhält man eine Injektion von $\mathit{Pot}(\mathbb{N})$ in ${}^{\mathbb{N}}\mathbb{N}$, die Menge aller Funktionen von $\mathbb{N}$ in $\mathbb{N}$ (und eine Bijektion von $\mathit{Pot}(\mathbb{N})$ auf ${}^{\mathbb{N}}2$).
Wenn man umgekehrt jedem $\varphi : \mathbb{N} \to \mathbb{N}$ einen Code seines Graphen, z. B. die Menge $\{2^x \cdot 3^{\varphi(x)} \mid x \in \mathbb{N}\}$ zuordnet, so erhält man eine Injektion von ${}^{\mathbb{N}}\mathbb{N}$ in

$Pot(\mathbb{N})$. Nach dem Satz von Schröder-Bernstein sind dann $Pot(\mathbb{N})$ und ${}^{\mathbb{N}}\mathbb{N}$ gleich mächtig. Eine Bijektion von $Pot(\mathbb{N})$ auf ${}^{\mathbb{N}}\mathbb{N}$ ist direkt nicht ganz so einfach anzugeben.

Sind Mengen hinsichtlich ihrer Mächtigkeit stets vergleichbar? Gibt es zu Mengen A, B stets eine Injektion von A in B oder eine Injektion von B in A? An dieser Stelle greifen wir auf den Wohlordnungssatz zurück:

Jede Menge lässt sich wohlordnen.

12.1.8 Definition Eine *Wohlordnung* ist eine lineare Ordnung $\mathcal{A} = (A, <_{\mathcal{A}})$, in der jede nicht-leere Teilmenge $X \subseteq A$ ein bezüglich $<_{\mathcal{A}}$ kleinstes Element besitzt.

Zur Veranschaulichung.
Jede (nicht-leere) Wohlordnung $\mathcal{A} = (A, <)$ enthält ein kleinstes Element, und zu jedem $\alpha \in A$, das nicht größtes Element von $\mathcal{A}$ ist, gibt es einen Nachfolger $\alpha + 1 \in A$, nämlich das kleinste Element der Menge $\{\beta \in A \mid \alpha < \beta\}$. Also ist jede endliche Wohlordnung isomorph zu einem Anfangsabschnitt $\{0, \ldots, n\}$ von $\mathbb{N}$, versehen mit der natürlichen Anordnung, und jede unendliche Wohlordnung fängt an mit einer Folge

$$\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots (n \in \mathbb{N}),$$

wobei α_{n+1} jeweils der Nachfolger von α_n ist, die mithin isomorph zu $(\mathbb{N}, <)$ ist. Schöpfen diese α_n die Menge A noch nicht aus, so gibt es in $\mathcal{A}$ ein kleinstes *Limes*-Element, etwa α_ω. Nachfolgerbildung und Limesbildung wechseln sich nun in einem transfiniten Zählprozess ab, bis die Menge A erschöpft ist.

Wir wenden uns wieder der systematischen Behandlung zu.

12.1.9 Lemma Jede Unterstruktur einer Wohlordnung ist wieder eine Wohlordnung.

Auf der Basis von Satz 11.2.10 ist das leicht einzusehen.

12.1.10 Definition Sei $\mathcal{A} = (A, <)$ eine Wohlordnung. Eine Teilmenge X von A ist ein *Abschnitt* von $\mathcal{A}$, wenn aus $\beta \in X$ und $\gamma < \beta$ stets $\gamma \in X$ folgt. Zu $\alpha \in A$ ist $\mathcal{A}_\alpha$ der Abschnitt $\{\beta \in A \mid \beta < \alpha\}$.

12.1.11 Lemma Die sämtlichen Abschnitte einer Wohlordnung $\mathcal{A} = (A, <)$ sind einerseits A selbst und andererseits die echten Abschnitte $\mathcal{A}_\alpha$ mit $\alpha \in A$.

Beweis. Ist X ein Abschnitt $\neq A$ von $\mathcal{A}$, so gibt es in $\mathcal{A}$ ein kleinstes $\alpha \in A - X$. Dann ist $\mathcal{A}_\alpha \subseteq X$, und wäre ein $\beta > \alpha$ in X, so wäre auch $\alpha \in X$, weil X ein Abschnitt ist. Also ist $\mathcal{A}_\alpha = X$.

Für Wohlordnungen $\mathcal{A}, \mathcal{B}$ ist nach 11.4.4 jeder Homomorphismus von $\mathcal{A}$ in $\mathcal{B}$ schon eine Einbettung. Wir suchen nach einem kanonischen Verfahren, $\mathcal{A}$ in $\mathcal{B}$ oder $\mathcal{B}$ in $\mathcal{A}$ einzubetten.

12.1.12 Definition Seien $\mathcal{A}, \mathcal{B}$ Wohlordnungen. Als *Standardeinbettung zu* $\mathcal{A}, \mathcal{B}$ bezeichnen wir die Abbildung φ_{st} eines Abschnitts von $\mathcal{A}$ in $\mathcal{B}$ mit

(1) $\varphi_{st}(\alpha) =$ das kleinste $\beta \in |\mathcal{B}| - \varphi_{st}[\mathcal{A}_\alpha]$, falls es ein solches β gibt, und $\varphi_{st}(\alpha)$ ist nicht definiert sonst.

Diese Definition durch „transfinite Rekursion“ legt φ_{st} eindeutig fest. Denn sonst müsste es ein kleinstes α in $\mathcal{A}$ geben, für das (1) $\varphi_{st}(\alpha)$ nicht eindeutig festlegt. Dann wäre $\varphi_{st}(\alpha')$ für $\alpha' \in \mathcal{A}_\alpha$, also auch $\varphi_{st}[\mathcal{A}_\alpha]$ und damit $\varphi_{st}(\alpha)$ festgelegt, im Widerspruch zur Annahme.

Ferner ist der Definitionsbereich von φ_{st} ein Abschnitt von $\mathcal{A}$. Denn ist φ_{st} nicht auf ganz $|\mathcal{A}|$ definiert, so gibt es ein kleinstes α, zu dem es kein β mit (1) gibt. Dann ist φ_{st} auf ganz $\mathcal{A}_\alpha$ definiert und es folgt $|\mathcal{B}| = \varphi_{st}[\mathcal{A}_\alpha]$, so dass $\mathcal{A}_\alpha$ der Definitionsbereich von φ_{st} ist.

12.1.13 Satz Wohlordnungen sind vergleichbar: Ist φ_{st} die Standardeinbettung zu $\mathcal{A}, \mathcal{B}$, so ist φ_{st} ein Isomorphismus von $\mathcal{A}$ auf einen Abschnitt von $\mathcal{B}$, oder φ_{st}^{-1} ist ein Isomorphismus von $\mathcal{B}$ auf einen Abschnitt von $\mathcal{A}$.

Beweis. Wenn $\varphi_{st}(\alpha)$ durch (1) definiert ist, gilt

(2) $\varphi_{st}[\mathcal{A}_\alpha]$ ist Abschnitt von $\mathcal{B} \Leftrightarrow \forall \alpha' < \alpha \quad \varphi_{st}(\alpha') < \varphi_{st}(\alpha)$,

weil $\varphi_{st}(\alpha)$ sich gemäß (1) eine Lücke in $\varphi_{st}[\mathcal{A}_\alpha]$ suchen würde und damit $< \varphi(\alpha')$ wäre für ein $\alpha' < \alpha$, wenn es eine Lücke gäbe. Angenommen, φ_{st} wäre kein Homomorphismus, also nicht streng monoton. Dann gäbe es ein kleinstes α, für das $\varphi_{st}(\alpha) < \varphi_{st}(\alpha')$ wäre für ein $\alpha' < \alpha$. ($\varphi_{st}(\alpha) = \varphi_{st}(\alpha')$ widerspräche (1) wegen $\alpha' \in \mathcal{A}_\alpha$.) Wegen der Minimalität von α wäre nach (2) $\varphi_{st}[\mathcal{A}_{\alpha'}]$ ein Abschnitt, also

$$\varphi_{st}(\alpha) \in \varphi_{st}[\mathcal{A}_{\alpha'}] \subseteq \varphi_{st}[\mathcal{A}_\alpha],$$

wieder im Widerspruch zu (1).

Also ist φ_{st} ein Homomorphismus, mithin eine Einbettung. Nun gilt folgende Alternative:

a. Entweder φ_{st} ist auf ganz $\mathcal{A}$ definiert. Dann ist wegen (2) jedes $\varphi_{st}[\mathcal{A}_\alpha]$ ein Abschnitt von $\mathcal{B}$, und $im(\varphi_{st})$ ist wegen (1) ebenfalls ein Abschnitt von $\mathcal{B}$. In diesem Fall ist φ_{st} ein Isomorphismus von $\mathcal{A}$ auf diesen Abschnitt von $\mathcal{B}$.

b. Oder es gibt ein kleinstes α in $\mathcal{A}$, für das $\varphi_{st}(\alpha)$ nicht definiert ist. Mit dem oben Gezeigten folgt dann

$$\varphi_{st} : \mathcal{A}_\alpha \cong \mathcal{B},$$

und φ_{st}^{-1} ist ein Isomorphismus von $\mathcal{B}$ auf einen (echten) Abschnitt von $\mathcal{A}$.

Nach dem Wohlordnungssatz kann man jede Menge A zu einer Wohlordnung $\mathcal{A} = (A, <_{\mathcal{A}})$ expandieren. Deswegen überträgt sich die Vergleichbarkeit von Wohlordnungen unmittelbar auf die Vergleichbarkeit von Mächtigkeiten.

12.1.14 Satz Mächtigkeiten sind vergleichbar: Zu beliebigen Mengen A, B gibt es eine Injektion von A in B oder von B in A.

Beweis. Ist A oder B leer, so ist die Behauptung trivial. Sonst gibt es Wohlordnungen $\mathcal{A} = (A, <_{\mathcal{A}})$ und $\mathcal{B} = (B, <_{\mathcal{B}})$. Dann ist die Standardeinbettung φ_{st} zu $\mathcal{A}, \mathcal{B}$ nach 12.1.13 eine Injektion von A in B oder φ_{st}^{-1} ist eine Injektion von B in A.

Man hat mit 12.1.13 viel mehr bewiesen als mit 12.1.14, aber die entscheidende Definition von φ_{st} durch (1) ist nur möglich, wenn Wohlordnungen vorliegen. Wir ziehen weitere Folgerungen aus 12.1.13 mit dem Ziel, mit Mächtigkeiten rechnen zu können.

12.1.15 Lemma Seien $\mathcal{A}, \mathcal{B}$ Wohlordnungen. Ist $\psi : \mathcal{A} \hookrightarrow \mathcal{B}$ eine Einbettung, so ist

(3) $\varphi_{st}(\alpha) \leq \psi(\alpha)$ für alle $\alpha \in |\mathcal{A}|$:
φ_{st} ist die „kleinste“ Einbettung von $\mathcal{A}$ in $\mathcal{B}$.

Beweis. Wir setzen (3) für alle $\alpha < \alpha'$ voraus. Weil ψ streng monoton ist, ist $\varphi_{st}(\alpha) < \psi(\alpha')$ für alle $\alpha < \alpha'$. Wegen (1) ist dann auch $\varphi_{st}(\alpha') \leq \psi(\alpha')$. Es kann also kein kleinstes α' geben, das (3) widerspricht.

12.1.16 Satz Für Wohlordnungen $\mathcal{A}, \mathcal{B}$ gilt: Aus $\mathcal{A} \hookrightarrow \mathcal{B}$ und $\mathcal{B} \hookrightarrow \mathcal{A}$ folgt $\mathcal{A} \cong \mathcal{B}$.

Das folgt unmittelbar aus 12.1.15 und 12.1.13, weil φ_{st}^{-1} die Standardeinbettung zu $\mathcal{B}, \mathcal{A}$ ist.

Die Klasse aller Wohlordnungen ist also durch die Einbettbarkeit (reflexiv) linear geordnet, sogar wohlgeordnet, wie man sich leicht überlegt.

12.1.17 Definition Eine Wohlordnung $\mathcal{A} = (A, <)$ ist eine *Kardinalzahl*, wenn A zu keinem echten Abschnitt $\mathcal{A}_\alpha (\alpha \in |A|)$ gleichmächtig ist. $\mathcal{A}$ heißt dann auch die *Mächtigkeit* oder *Kardinalität von* A, bezeichnet mit $card(A)$. Ferner ist $\emptyset$ die Kardinalzahl 0. Wir identifizieren isomorphe Kardinalzahlen und bezeichnen (geeignete Repräsentanten) ihrer Isomorphieklassen mit κ, λ. Für $\kappa \hookrightarrow \lambda$ schreibt man oft $\kappa \leq \lambda$.

Beispiel. $\aleph_0 := (\mathbb{N}, <)$ ist die kleinste unendliche Kardinalzahl und gleich der Kardinalität jeder abzählbar unendlichen Menge.

12.1.18 Lemma Jede Menge hat eine Kardinalität, und gleichmächtige Mengen haben gleiche Kardinalitäten.

Beweis. Sei $\mathcal{A} = (A, <)$ eine Wohlordnung der gegebenen Menge A. Ist $\mathcal{A}$ nicht selbst eine Kardinalzahl, so ist $A \sim \mathcal{A}_\alpha$ für ein $\alpha \in A$. Unter allen solchen α gibt es dann ein kleinstes α_0. Dann ist $(\mathcal{A}_{\alpha_0}, <)$ eine Kardinalzahl und gleich der Kardinalität von A.

Seien nun $\kappa = (A, <_\kappa)$ und $\lambda = (B, <_\lambda)$ Kardinalzahlen und $A \sim B$. Dann gibt es überhaupt keine Injektion von A in einen echten Abschnitt λ_β von λ, geschweige denn eine Einbettung von κ in $(\lambda_\beta, <_\lambda)$, und umgekehrt. Nach 12.1.13 ist dann $\kappa \cong \lambda$.

12.1.19 Definition Seien $\kappa = (A, <_\kappa)$ und $\lambda = (B, <_\lambda)$ Kardinalzahlen. Die *kardinale Summe* von κ, λ ist

$$\kappa + \lambda := card(\{(a, 0) \mid a \in A\} \cup \{(b, 1) \mid b \in B\}),$$

die Kardinalität der *disjunkten Vereinigung* von A und B.
Das *kardinale Produkt* von κ, λ ist

$$\kappa \cdot \lambda := card(\{(a,b) \mid a \in A, b \in B\}),$$

die Kardinalität des *cartesischen Produkts* von A und B.
Die *kardinale κ-te Potenz* ist

$$2^\kappa := card(Pot(A)),$$

die Kardinalität der Potenzmenge von A.

Bemerkung. Nach dem Satz von Cantor ist $\kappa < 2^\kappa$. Es gibt also einen kürzesten Abschnitt von 2^κ, der mächtiger ist als κ. Die Kardinalität dieses Abschnitts bezeichnet man mit κ^+. Per Definition ist $\kappa < \kappa^+ \leq 2^\kappa$. Die Aussage, dass

(GCH) $\kappa^+ = 2^\kappa$ für alle Kardinalzahlen κ

ist, ist die allgemeine Kontinuumshypothese. Sie ist im Rahmen der Mengenlehre auch mit Wohlordnungssatz nicht zu entscheiden. Auch der Spezialfall $\kappa = \aleph_0$ von (GCH), die spezielle Kontinuumshypothese

(CH) $\aleph_1 = 2^{\aleph_0}$

(man schreibt $\aleph_1$ für $\aleph_0^+$, die kleinste überabzählbare Mächtigkeit) ist unabhängig von der Mengenlehre mit Wohlordnungssatz.

In jeder linearen Ordnung $\mathcal{A} = (A, <_\mathcal{A})$ haben je zwei Elemente $\alpha, \beta \in A$ ein *Maximum* $\max(\alpha, \beta)$, nämlich

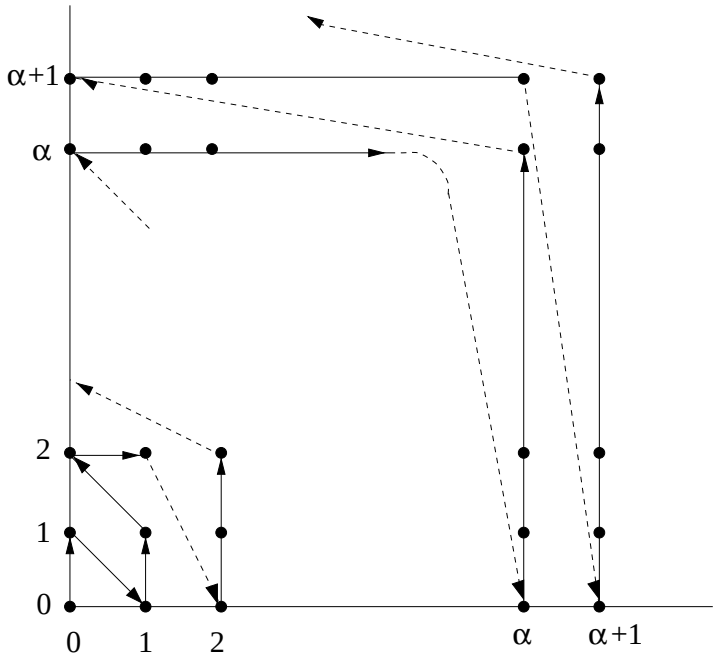
$$\begin{aligned} &\max(\alpha,\beta) = \beta, \quad \text{falls } \alpha <_\mathcal{A} \beta, \\ &\max(\alpha,\beta) = \alpha \quad \text{sonst.} \end{aligned}$$

Es gibt mehrere Möglichkeiten, das cartesische Produkt $A \times A$ wieder zu ordnen. Wir untersuchen eine von Hessenberg angegebene Möglichkeit, die das Maximum bez. $<_\mathcal{A}$ verwendet.

12.1.20 Definition Sei $\mathcal{A} = (A, <_\mathcal{A})$ eine lineare Ordnung. Die *Hessenberg-Ordnung* $<_H$ auf $A \times A$ ist wie folgt definiert:

$$\begin{aligned} (\alpha_0,\alpha_1) <_H (\beta_0,\beta_1) :\Leftrightarrow \quad & \max(\alpha_0,\alpha_1) <_\mathcal{A} \max(\beta_0,\beta_1), \text{ oder} \\ & \max(\alpha_0,\alpha_1) = \max(\beta_0,\beta_1) \text{ und } \alpha_0 <_\mathcal{A} \beta_0, \text{ oder} \\ & \max(\alpha_0,\alpha_1) = \max(\beta_0,\beta_1) \text{ und } \alpha_0 = \beta_0 \text{ und } \alpha_1 <_\mathcal{A} \beta_1. \end{aligned}$$

Die Struktur $(A \times A, <_H)$ bezeichnen wir auch mit $\mathcal{A} \times_H \mathcal{A}$.



Durch $<_H$ sind also die Paare $\in A \times A$ zuerst nach ihrem Maximum, danach lexikographisch geordnet.

12.1.21 Lemma Ist $\mathcal{A}$ eine Wohlordnung, so ist auch $\mathcal{A} \times_H \mathcal{A}$ eine Wohlordnung.

Den anschaulich einfachen Beweis verlegen wir in die Aufgaben.

Bemerkung. $\mathcal{A}$ lässt sich immer in $\mathcal{A} \times_H \mathcal{A}$ mittels $\alpha \mapsto (\alpha, \alpha)$ einbetten. Ist $\mathcal{A}$ endlich, so ist auch $\mathcal{A} \times_H \mathcal{A}$ endlich, und zwar mit $(card(\mathcal{A}))^2$ Elementen. $(\mathbb{N}, <) \times_H (\mathbb{N}, <)$ ist dann isomorph zu $(\mathbb{N}, <)$. Diese „Absorption“ setzt sich auf alle unendlichen Kardinalzahlen fort:

12.1.22 Satz von Hessenberg Für jede unendliche Kardinalzahl κ ist

(4) $\quad \kappa \times_H \kappa \cong \kappa.$

Beweis. Wie bemerkt, ist $\alpha \mapsto (\alpha, \alpha)$ eine Einbettung von κ in $\mathcal{B} := \kappa \times_H \kappa$. Wegen Satz 12.1.16 brauchen wir also nur noch $\mathcal{B}$ in κ einzubetten. Weil κ eine Kardinalzahl ist, genügt es dafür, zu zeigen, dass jeder eigentliche Abschnitt

$\mathcal{B}_\gamma$ von $\mathcal{B}$ weniger mächtig ist als κ. Denn dann ist jedes $(\mathcal{B}_\gamma, <_H)$ isomorph zu einem eigentlichen Abschnitt $(\kappa_{\varphi(\gamma)}, <_\kappa)$ von κ, und die so entstehende Zuordnung φ bettet $\mathcal{B}$ in κ ein.

Sei nun κ die kleinste unendliche Kardinalzahl, für die (4) bezweifelt wird, und sei $\gamma = (\gamma_0, \gamma_1)$ aus $\mathcal{B}$ und $\alpha = \max(\gamma_0, \gamma_1)$. Dann ist α in κ, also $\kappa_\alpha \sim \lambda$ für eine Kardinalzahl $\lambda < \kappa$. Es folgt

$$(\mathcal{B}_\gamma, <_H) \subseteq (\mathcal{B}_{(\alpha,\alpha)}, <_H) = \kappa_\alpha \times_H \kappa_\alpha \sim \lambda \times_H \lambda.$$

Ist λ endlich, so ist auch $\mathcal{B}_\gamma$ endlich und deshalb weniger mächtig als κ. Ist λ unendlich, so ist nach unserer (Induktions-) Voraussetzung über κ

$$\lambda \times_H \lambda \sim \lambda < \kappa,$$

und wieder ist $\mathcal{B}_\gamma$ weniger mächtig als κ. Damit ist der Satz von Hessenberg bewiesen.

Dieser Satz ist eine Verallgemeinerung des 2. Cantorschen Diagonalverfahrens, mit dem man die Abzählbarkeit von $\mathbb{N} \times \mathbb{N}$ und ebenso von $\mathbb{Q}$ beweist, auf beliebige unendliche Kardinalzahlen.

12.1.23 Korollar Für jede unendliche Kardinalzahl κ und jede Kardinalzahl $\lambda \neq \emptyset$ gelten die Rechengesetze

$$\begin{aligned} \kappa \cdot \kappa &= \kappa \\ \kappa \cdot \lambda &= \max(\kappa, \lambda) \\ \kappa + \lambda &= \max(\kappa, \lambda) \end{aligned}$$

Beweis. Sei o. E. $\lambda \subseteq \kappa$, also $\max(\kappa, \lambda) = \kappa$. Dann ist

$$\kappa \hookrightarrow \kappa \cdot \lambda \hookrightarrow \kappa \times_H \lambda \hookrightarrow \kappa \times_H \kappa$$

und ähnlich

$$\kappa \hookrightarrow \kappa + \lambda \hookrightarrow \kappa \times_H 2 \hookrightarrow \kappa \times_H \kappa,$$

und mit (4) folgen daraus alle drei Rechengesetze.

12.2 Modelle höherer Mächtigkeit

Welche Rolle spielen Mächtigkeitsfragen für die Modelle einer Theorie?

12.2.1 Definition Die *Mächtigkeit* einer Struktur $\mathcal{A}$, bezeichnet mit $card(\mathcal{A})$, ist die Mächtigkeit ihres Individuenbereichs $|\mathcal{A}|$. Strukturen der Mächtigkeit $\aleph_0$ heißen auch *abzählbar unendliche Strukturen*.

Beispiele. Die Struktur $\mathcal{N}$ der natürlichen Zahlen ist abzählbar unendlich, ebenso wie der Körper $\mathbb{Q}$ der rationalen Zahlen. Dagegen ist der Körper $\mathbb{R}$ der reellen Zahlen nicht abzählbar, er hat die Mächtigkeit $2^{\aleph_0}$. Die Modelle über den natürlichen Zahlen aus 9.4 sind abzählbar unendlich. Also hat nach 9.4.3 jede konsistente abzählbare identitätsfreie Theorie ein abzählbar unendliches Modell.

12.2.2 Definition Es sei κ eine unendliche Kardinalzahl. Eine Sprache L heißt κ-*Sprache*, wenn L höchstens κ nicht-logische Grundzeichen enthält. Eine Theorie T heißt κ-*Theorie* wenn die Sprache $L(T)$ eine κ-Sprache ist.

Dies erweitert die Definition 9.2.1 der abzählbaren Sprachen und Theorien auf beliebige unendliche Kardinalzahlen. Die abzählbaren Sprachen und Theorien (insbesondere alle Theorien aus §1) sind gerade die $\aleph_0$-Sprachen und $\aleph_0$-Theorien. Die Sprache $L(\mathbb{R})$ ist eine $2^{\aleph_0}$-Sprache. Eine Sprache oder Theorie legt ihre Kardinalität nicht eindeutig fest. Ist $\kappa \leq \lambda$, so ist jede κ-Sprache zugleich eine λ-Sprache. Jede $\aleph_0$-Sprache ist also zugleich λ-Sprache für jede unendliche Kardinalzahl λ. Entsprechendes gilt für Theorien.

Wir verallgemeinern 9.2.2 durch Anwendung von 12.1.23.

12.2.3 Lemma Jede κ-Sprache hat höchstens κ Terme und höchstens κ Formeln.

Beweis. Es sei L eine κ-Sprache, also $\kappa \geq \aleph_0$. L enhält höchstens κ nicht-logische Grundzeichen. Da L nur $\aleph_0$ logische Grundzeichen enthält, enthält L insgesamt höchstens $\kappa + \aleph_0 = \kappa$ Grundzeichen. Dann enthält L höchstens

$$\kappa^n = \underbrace{\kappa \cdot \ldots \cdot \kappa}_{n\text{-fach}} = \kappa$$

Zeichenreihen aus n Grundzeichen. Also enthält L höchstens $\kappa \cdot \aleph_0 = \kappa$ Zeichenreihen (endlicher Länge).

12.2.4 Aufsteigender Satz von Löwenheim und Skolem Jede κ-Theorie, die ein unendliches Modell besitzt, besitzt ein Modell einer Mächtigkeit $\geq \kappa$.

Beweis. Es sei T eine κ-Theorie und E eine Menge von κ neuen Konstanten. Wie in 10.2.4 definieren wir

$$T_\kappa := T + E + \{\neg e = e' \mid e, e' \text{ verschiedene Konstanten aus } E\}.$$

T_κ ist dann wegen $\kappa + \kappa = \kappa$ eine κ-Theorie. Jedes Modell $\mathcal{A}_\kappa$ von T_κ enthält mindestens κ verschiedene Elemente, weil in $\mathcal{A}_\kappa$ alle Konstanten aus E verschieden interpretiert werden müssen. Die Beschränkung $\mathcal{A}$ von $\mathcal{A}_\kappa$ auf $L(T)$ ist dann ein Modell von T einer Mächtigkeit $\geq \kappa$.

Also bleibt zu zeigen, dass T_κ überhaupt ein Modell besitzt. Es sei T' eine endlich axiomatisierte Teiltheorie von T_κ, und $e_1, \ldots, e_m$ seien die endlich vielen neuen Konstanten, die in $Ax(T')$ auftreten. Dann lässt sich das unendliche Modell $\mathcal{B}$ von T, das es nach Voraussetzung gibt, wie folgt zu einem Modell $\mathcal{B}'$ von T' expandieren:

$$\mathcal{B}'(e_i) \neq \mathcal{B}'(e_j) \in |\mathcal{B}| \text{ für } 1 \leq i < j \leq m,$$
$$\mathcal{B}'(e) = \mathcal{B}'(e_1) \text{ für alle anderen } e \in E.$$

Diese Festsetzung ist möglich, weil $|\mathcal{B}|$ unendlich ist.

Dann ist $\mathcal{B}'$ eine Struktur zur Sprache $L(T) + E = L(T')$. $\mathcal{B}'|L(T) = \mathcal{B}$ ist Modell von T, weil $\mathcal{B}'$ eine Expansion von $\mathcal{B}$ ist. Ferner ist

$$\mathcal{B}'(\neg e = e') = w$$

für alle endlich vielen Ungleichungen $\neg e = e'$ zwischen Konstanten e, e' aus E, die Axiome von T' sind. Also ist $\mathcal{B}'$ ein Modell von T'. Hiernach besitzt jede endlich axiomatisierte Teiltheorie T' von T_κ ein Modell. Der Kompaktheitssatz 10.1.3 ergibt dann, dass T_κ ein Modell $\mathcal{A}_\kappa$ besitzt. Damit ist der Satz bewiesen.

Alle in diesem Beweis mit $\mathcal{B}'$ bezeichneten Modelle sind Expansionen ein und desselben gegebenen unendlichen Modells $\mathcal{B}$ von T. Ist $\mathcal{B}$ abzählbar, so sind auch alle $\mathcal{B}'$ abzählbar. Das mit dem Kompaktheitssatz gewonnene Modell $\mathcal{A}_\kappa$ hat aber mindestens die Mächtigkeit κ, kann also sehr viel größer sein. $\mathcal{A}_\kappa$ wird also keineswegs in einer direkten oder anschaulichen Weise aus den Modellen $\mathcal{B}'$ gewonnen.

Beispiel. Die gewöhnliche Zahlentheorie Z besitzt außer dem Standardmodell $\mathcal{N}$ noch Modelle beliebig großer Mächtigkeit. Dasselbe gilt für die (saturierte) Theorie $Th(\mathcal{N})$. In der Sprache von Z ist es also nicht möglich, die natürlichen Zahlen durch ein Axiomensystem eindeutig bis auf Isomorphie zu beschreiben. Ähnliches gilt für alle unendlichen Strukturen.

12.3 Elementare Untermodelle

Im aufsteigenden Satz von Löwenheim-Skolem 12.2.4 konnten wir die Mächtigkeit λ des dort konstruierten Modells der κ-Theorie T nicht vorschreiben. Zu jeder (noch so großen) Kardinalzahl κ ließ sich dort nur ein $\lambda \geq \kappa$ finden, so dass T ein Modell der Mächtigkeit λ besitzt. Wir wollen dieses Ergebnis dahin verschärfen, dass $\lambda = \kappa$ gewählt werden kann, sofern nur T eine κ-Theorie mit unendlichem Modell ist. Dazu sondern wir aus einem eventuell größeren Modell eine elementare Unterstruktur der Mächtigkeit κ aus (vgl. 11.3.4), die dann wieder ein Modell von T ist.

12.3.1 Absteigender Satz von Löwenheim und Skolem Es sei κ eine unendliche Kardinalzahl und L eine κ-Sprache. Jede Struktur $\mathcal{B}$ zu L einer Mächtigkeit $\geq \kappa$ besitzt eine elementare Unterstruktur der Mächtigkeit κ.

Beweis. Wir wählen eine Teilmenge A_0 von $B := |\mathcal{B}|$ von der Mächtigkeit κ. Das ist wegen $card(B) \geq \kappa$ stets möglich. Sei nun A_i schon definiert mit $card(A_i) = \kappa$. Dann wählen wir eine Teilmenge A_{i+1} von B wie folgt: A_{i+1} enthalte A_i, und zu jedem Existenzsatz $\exists x F(x)$ aus $L + A_i$, für den

$$\mathcal{B}(\exists x F(x)) = w$$

ist, enthalte A_{i+1} ein (weiteres) $c \in B$ mit

$$\mathcal{B}(F(c)) = w.$$

Wegen 12.2.3 gibt es höchstens κ viele solche Sätze in $L + A_i$, weil dies wegen $\kappa + \kappa = \kappa$ eine κ-Sprache ist. Also kommen zu A_i höchstens κ viele solche $c \in B$ hinzu, so dass

$$\kappa = card(A_i) \leq card(A_{i+1}) \leq \kappa + \kappa = \kappa,$$

mithin $card(A_{i+1}) = \kappa$ ist. Wir setzen

$$A = \bigcup \{A_i \mid i \in \mathbb{N}\}.$$

Dann ist nach 12.1.23

$$\kappa \leq card(A_0) \leq card(A) \leq \kappa \cdot \aleph_0 = \kappa :$$

(1) $card(A) = \kappa$.

Wir behaupten:

(2) A ist abgeschlossen unter allen Funktionen $f_{\mathcal{B}}$ mit $f \in L$. Denn ist $k_1, \ldots, k_n \in A$, so gilt für hinreichend große i:

$$k_1, \ldots, k_n \in A_i.$$

Dann ist $\exists y\, f k_1 \ldots k_n = y$ ein Existenzsatz aus $L + A_i$, der in $\mathcal{B}$ wahr ist. Nach Wahl von A_{i+1} gibt es also ein $c \in A_{i+1}$, so dass

$$\mathcal{B}(f k_1 \ldots k_n = c) = w$$

ist. Nach Definition von A ist $c \in A$, also

$$f_{\mathcal{B}}(k_1, \ldots, k_n) \in A.$$

Nach (1) und (2) ist

$$\mathcal{A} := (A, (f_{\mathcal{B}} \restriction A^n)_{f \in L}, (p_{\mathcal{B}} \cap A^n)_{p \in L})$$

eine Unterstruktur von $\mathcal{B}$ mit der Mächtigkeit κ.

Zu zeigen bleibt:

(3) $\mathcal{A}(C) = w \Leftrightarrow \mathcal{B}(C) = w$ für alle Sätze $C \in L(\mathcal{A}) = L + A$.

Das folgt durch Induktion nach dem Aufbau von C.

1. Für quantorenfreies C gilt (3) nach 11.2.6.
2. Auch für Implikationen C können wir den Induktionsschritt aus dem Beweis von 11.2.6 übernehmen.

3. C sei $\forall x F(x)$.

3.1 $\mathcal{B}(C) = w \quad \Rightarrow \mathcal{B}(\mathcal{F}(c)) = w$ für alle $c \in |\mathcal{B}|$
$\Rightarrow \mathcal{B}(\mathcal{F}(c)) = w$ für alle $c \in A \subseteq |\mathcal{B}|$
$\Leftrightarrow \mathcal{A}(\mathcal{F}(c)) = w$ für alle $c \in A$ nach IV
$\Leftrightarrow \mathcal{A}(C) = w$.

3.2 $\mathcal{B}(C) = f \Rightarrow \mathcal{B}(\exists x \neg \mathcal{F}(x)) = w$.
Da nur endlich viele Konstanten aus A in $\mathcal{F}$ auftreten, ist $\exists x \neg \mathcal{F}(x)$ ein Existenzsatz aus $L + A_i$, wenn man i groß genug wählt. Nach Wahl von A_{i+1} ist dann

$$\mathcal{B}(\neg \mathcal{F}(c)) = w \text{ für ein } c \in A_{i+1} \subseteq A$$

woraus mit den Äquivalenzen aus 3.1 folgt:

$$\mathcal{A}(C) = f.$$

3.1 und 3.2 ergeben (3) für $C \equiv \forall x \mathcal{F}(x)$. Mit Induktion folgt (3) allgemein. Damit ist der Satz vollständig bewiesen.

Ist in diesem Beweis $card(\mathcal{B}) = \kappa$, so kann man $A_0 = B$ wählen, und man erhält $A = A_0 = B$ und $\mathcal{A} = \mathcal{B}$. Ein neues, kleineres Modell erhält man also nur im Fall $card(\mathcal{B}) > \kappa$. Der Fall $\kappa = \aleph_0$ spielt eine besondere Rolle:

12.3.2 Satz von Löwenheim und Skolem Jede abzählbare Theorie, die ein unendliches Modell besitzt, hat auch ein abzählbares Modell.

Beweis. Es sei T eine abzählbare Theorie, also $L(T)$ eine abzählbare Sprache. $\mathcal{B}$ sei ein unendliches Modell von T. Dann ist $\mathcal{B}$ eine Struktur zu $L(T)$ von einer Mächtigkeit $\geq \aleph_0$. Nach 12.3.1 besitzt $\mathcal{B}$ eine abzählbare elementare Unterstruktur $\mathcal{A}$. Dann ist auch $\mathcal{A}$ ein Modell von T, und zwar ein abzählbares.

12.3.3 Das Löwenheim-Skolem-Paradoxon
Die Zermelo-Fraenkelsche Mengenlehre ZF ist eine abzählbare Theorie (vgl. 1.2.6). Wenn ZF konsistent ist – was wir wie üblich voraussetzen wollen –, hat ZF nach dem Vollständigkeitssatz ein Modell $(V, \in)$.

Dies allein ist schon überraschend, weil V eine Menge ist, die das Universum aller Mengen darstellt. Man könnte meinen, dass die Menge V selbst im Modell

$(V, \in)$ auftreten müsste, so dass $V \in V$ wäre. Das ist keineswegs der Fall. Das Modell $(V, \in)$ simuliert nur die Verhältnisse des naiven Mengenuniversums, soweit sie in der Sprache von ZF formulierbar sind.

Die Situation wird vollends paradox angesichts des Satzes von Löwenheim und Skolem. ZF besitzt kein endliches Modell, was auch ohne Kenntnis der Axiome von ZF naheliegt. Also ist das Modell, das wegen der Konsistenz von ZF existiert, unendlich, und nach 12.3.2 besitzt ZF ein abzählbar unendliches Modell $\mathcal{A} = (V_0, \in)$. In ZF ist die Menge $\mathbb{N}$ der natürlichen Zahlen definierbar. Weil $\mathcal{A}$ abzählbar ist, treten in $\mathcal{A}$ nur abzählbar viele Teilmengen von $\mathbb{N}$ auf, so dass die Potenzmenge von $\mathbb{N}$ im Modell $\mathcal{A}$, $\mathcal{A}(Pot(\mathbb{N}))$, abzählbar unendlich ist:

$$\textit{es gibt eine Bijektion } \varphi \textit{ von } \mathcal{A}(\mathbb{N}) \textit{ auf } \mathcal{A}(Pot(\mathbb{N})).$$

Andererseits ist in ZF herleitbar (vgl. den Satz von Cantor 12.1.4), dass $Pot(\mathbb{N})$ nicht abzählbar ist:

$$ZF \vdash \forall\varphi(\varphi : \mathbb{N} \to Pot(\mathbb{N}) \to \neg\varphi \text{ ist bijektiv}).$$

Diese Formel gilt nach dem Korrektheitssatz in $\mathcal{A}$:

$$\textit{Keine Abbildung } \varphi \in V_0, \varphi : \mathbb{N} \to \mathcal{A}(Pot(\mathbb{N})), \textit{ ist bijektiv.}$$

Einerseits ist also $\mathcal{A}(Pot(\mathbb{N}))$ abzählbar; andererseits gilt in $\mathcal{A}$, das $Pot(\mathbb{N})$ nicht abzählbar ist. Das ist zwar paradox, aber nur scheinbar ein Widerspruch:

Zwar gibt es eine Bijektion $\varphi : \mathbb{N} \to \mathcal{A}(Pot(\mathbb{N}))$, aber diese Bijektion (als Menge aufgefasst) gehört nicht zu $V_0 = |\mathcal{A}|$. Das Universum V_0 und die Menge $\mathcal{A}(Pot(\mathbb{N}))$ sind zwar *von außen* abzählbar, aber nicht *von innen*.

Ob eine Menge abzählbar ist oder nicht, ist hiernach nicht absolut zu beantworten, sondern hängt entscheidend daran, welche Klasse von Bijektionen man für eine Abzählung in Betracht zieht. Das Problem, ob es überabzählbare Mengen „objektiv gibt“, ist im Rahmen der Theorien der ersten Stufe nicht zu lösen.

12.3.4 Mächtigkeitssatz von Tarski, Löwenheim und Skolem Jede κ-Theorie mit einem unendlichen Modell besitzt ein Modell der Mächtigkeit κ.

Beweis. Es sei T eine κ-Theorie. Nach 12.2.4 besitzt T ein Modell einer Mächtigkeit $\geq \kappa$. Dieses Modell besitzt nach 12.3.1 eine elementare Unterstruktur der Mächtigkeit κ, die offenbar auch ein Modell von T ist.

Beispiele. Die abzählbaren Theorien aus 1.2 besitzen unendliche Modelle. Also besitzen sie Modelle jeder unendlichen Mächtigkeit: Es gibt Gruppen, Ringe, Körper, (dichte) lineare Ordnungen, Modelle der Zahlentheorie Z und der Mengenlehre ZF von jeder unendlichen Mächtigkeit.

Der Mächtigkeitssatz 12.3.4 fasst den aufsteigenden und den absteigenden Satz von Löwenheim und Skolem zusammen. Als Spezialfall $\kappa = \aleph_0$ enthält er den klassischen Satz 12.3.2 von Löwenheim-Skolem, der allerdings allein aus dem absteigenden Satz 12.3.1 folgt und vom Kompaktheitssatz unabhängig ist.

12.4 Aufgaben

12.4.1 Zeigen Sie für injektives $\varphi : A \to A$ und $X \subseteq A - \varphi[A]$:

a. Für alle $n \in \mathbb{N}$ ist $\varphi^n[X] \subseteq \varphi^n[A] - \varphi^{n+1}[A]$.

b. Für $m < n \in \mathbb{N}$ ist $\varphi^m[X] \cap \varphi^n[X] = \emptyset$.

12.4.2 Zeigen Sie: $\mathbb{R} \sim Pot(\mathbb{N})$.

12.4.3 Beweisen Sie Lemma 12.1.21.

12.4.4 $L(T)$ enthalte als einziges nicht-logisches Grundzeichen das einstellige Funktionszeichen f.

$$Ax(T) = \{\forall x \neg fx = x,\ \forall x ffx = x\}.$$

a. Folgern Sie aus 12.1.23 und ohne Rückgriff auf die Löwenheim-Skolem-Sätze, dass T ein Modell jeder unendlichen Mächtigkeit hat.

b. Für welche $n \in \mathbb{N}$ hat T ein n-elementiges Modell?

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 5:
Modelltheorie (Fortsetzung)

Autor: Justus Diller

01301-7-05-S 1

Klassische Prädikatenlogik

Kurseinheit 5: Inhalt

Studienhinweise 267

Verzeichnis der definierten Begriffe und der wichtigen Sätze 269

4. Modelltheorie

§10 Kompaktheit 209

§11 Morphismen 223

§12 Die Sätze von Löwenheim und Skolem 245

§13 Kategorizität und die Theorie der dichten linearen Ordnung 271

13.1 Kategorizität und Vollständigkeit 271
13.2 $\aleph_0$-Kategorizität der Theorie DLO 276
13.3 Nicht-Kategorizität von DLO in der Mächtigkeit $2^{\aleph_0}$ 278
13.4 Zusammenfassung und Aufgaben 280

§14 Nicht-Standard-Modelle der Zahlentheorie 283

14.1 Enderweiterungen 283
14.2 Existenz und Anzahl von abzählbaren Nicht-Standard-Modellen 292
14.3 Overspill 297
14.4 Anordnung in Nicht-Standard-Modellen 303
14.5 Aufgaben 309

§15 Zur Zahlentheorie der zweiten Stufe: Übersetzungen 311

15.1 Die Zahlentheorie der zweiten Stufe als mathematische Theorie Z^2 311
15.2 Relativierung und Übersetzung 318
15.3 Modelle von Z^2 322
15.4 Aufgaben 327

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 5: Studienhinweise

1. Lehrziele

Diese Kurseinheit setzt das Studium der Modelltheorie fort. Anders als in der vorigen Kurseinheit, in der eine allgemeine Einführung in fundamentale Ergebnisse und Techniken der Modelltheorie im Mittelpunkt stand, sollen hier wichtige modelltheoretische Konzepte und Verfahren an einzelnen mathematischen Theorien vorgeführt werden. Diese Theorien sind

- in §13 die Theorie DLO, an der der Begriff der Kategorizität untersucht wird.
- in §14 die Zahlentheorien Z und $Th(\mathcal{N})$. Das Studium ihrer Nicht-Standard-Modelle liefert überraschende Ergebnisse und entfaltet großen mathematischen Reichtum.
- in §15 die Zahlentheorie Z^2 der zweiten Stufe. An ihr soll die elementare Technik des Relativierens und Übersetzens, bezogen auf die Theorie Z, beispielhaft vorgeführt werden.

Nach den Löwenheim-Skolem-Sätzen leuchtet die Frage nach der κ-Kategorizität 13.1.1 unmittelbar ein. Neben trivialen Beispielen kategorischer Theorien wird in 13.2 die $\aleph_0$-Kategorizität von DLO gezeigt. Der Beweis verwendet ein externes Element, nämlich Abzählungen der beteiligten Strukturen. Er ist der einfachste Fall der back-and-forth-Methode der Modelltheorie. Die Wichtigkeit dieses externen Elements wird in 13.3 gerade dadurch verdeutlicht, dass eine anschauliche Überlegung DLO als nicht $2^{\aleph_0}$-kategorisch nachweist.

Schwerpunkt dieser Kurseinheit ist das Studium der abzählbaren Modelle der Zahlentheorie. Einerseits hat sogar $Th(\mathcal{N})$ $2^{\aleph_0}$ paarweise nicht-isomorphe abzählbare Modelle (14.2.9); andererseits sind sogar alle Modelle von Z^- Enderweiterungen von $\mathcal{N}$ (14.1.10), und alle abzählbaren Nicht-Standard-Modelle von Z^- haben denselben Ordnungstyp (14.4.10). Man beachte insgesamt die Uniformität der Ergebnisse über beliebige Modelle der schwachen Theorie Z^- (14.1.4) und gleichzeitig die Vielfalt unter den Modellen der starken, saturierten Theorie $Th(\mathcal{N})$. Dazwischen rückt die gewöhnliche Zahlentheorie Z, für die das Induktionsschema charakteristisch ist, bei den Overspill-Eigenschaften (14.3) in den Mittelpunkt.

Die Zahlentheorie Z^2 der zweiten Stufe wird in 15.1 ausführlich motiviert und als Theorie der ersten Stufe formuliert. Sie ist wesentlich ausdrucksstärker als Z, und an ihrem Verhältnis zur Theorie Z wird in 15.2 die allgemeine, elementare Technik des Relativierens und Übersetzens illustriert (15.2.10 und 12). Die kurze Einführung in die Modelle von Z^2 soll klären, wie man das Standardmodell approximieren kann: im erststufigen Bereich durch die ω-Modelle, im zweitstufigen Bereich durch die Vollständigkeit (15.3.3). Der Satz von Peano 15.3.8 bietet hier einen gewissen Abschluss.

2. Eingangsvoraussetzungen

Die wesentlichen Voraussetzungen sind in den Paragraphen 10 bis 12 enthalten. Ständig verwendet werden der Kompaktheitssatz aus 10.1 und die Löwenheim-Skolem-Sätze aus §12, aber auch die Morphismen-Terminologie und einfache Ergebnisse über Einbettungen (11.2) und Wahrheitstransport (11.3). Vor dem Hintergrund des Vollständigkeitssatzes werden Herleitbarkeit und Gültigkeit in Theorien äquivalent verwendet. Sicherheit im Umgang mit den Grundbegriffen und -eigenschaften der Semantik aus §2 wird vorausgesetzt. Die geringen Kenntnisse über Primzahlen, die in 14.2 verwendet werden, sind in 14.2.4 bis 6 zusammengetragen.

Klassische Prädikatenlogik

Kurseinheit 5: Verzeichnis der definierten Begriffe und der wichtigen Sätze

13.1.1 κ-kategorische Theorien

13.1.3 Reine Theorie $T_=$ der Identität

13.1.4 Lemma. $T_=$ ist κ-kategorisch für jedes $\kappa > 0$.

13.1.5 Lemma. Es gibt λ-Theorien, die nicht λ-kategorisch, aber κ-kategorisch sind für alle $\kappa > \lambda$.

13.1.7 Vollständige Theorien

13.1.8 **Satz** κ-kategorische Theorien, die nur unendliche Modelle besitzen, sind vollständig.

13.2.1 **Satz** DLO ist $\aleph_0$-kategorisch

13.2.2 Korollar. DLO ist vollständig

13.3.1 $\leq_{\mathcal{A}}$, obere Schranke, Supremum

13.3.3 **Satz** DLO ist nicht $2^{\aleph_0}$-kategorisch.

13.4.1 **Satz von Morley** (ohne Beweis)

14.1.2 $\leq$ und $<$ in $L(Z)$

14.1.4 Teiltheorie Z^- von Z

14.1.8 Nicht-Standard-Modelle, Standardzahlen, Nicht-Standard-Zahlen; Enderweiterung, Anfangsabschnitt, $\mathcal{A} \subseteq_e \mathcal{B}$

14.1.10 **Satz** Jedes Modell von Z^- ist Enderweiterung von $\mathcal{N}$.

14.1.12 beschränkte Quantoren, Δ_0- und Σ_1-Formeln

14.1.14 **Satz** Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so ist $id : \mathcal{A} \xrightarrow{\Sigma_1} \mathcal{B}$.

14.1.15 **Satz** Für $\mathcal{A} \models Z^-$ ist $\mathcal{A}{\restriction}\mathbb{N} : \mathcal{N} \xrightarrow{\Sigma_1} \mathcal{A}$.

14.2.2 **Satz** $Th(\mathcal{N})$ ist nicht $\aleph_0$-kategorisch

14.2.3 Lemma. Abzählbare Theorien haben $\leq 2^{\aleph_0}$ abzählbare Modelle

14.2.4 a teilt b, $a \mid b$; Primzahl, Prim(a)

14.2.6 i-te Primzahl p_i; e kodiert X in $\mathcal{A}$

14.2.8 **Satz** Zu $X \subseteq \mathbb{N}$ gibt es abzählbares $\mathcal{A} \equiv \mathcal{N}$, das X kodiert.

14.2.9 **Satz** $Th(\mathcal{N})$ hat $2^{\aleph_0}$ nicht-isomorphe Modelle.

14.3.1 Schnitt I in $\mathcal{A}$

14.3.2 in $\mathcal{A}$ definierbare Teilmenge

14.3.3 **Satz** Kein echter Schnitt in $\mathcal{A} \models Z$ ist definierbar in $\mathcal{A}$.

14.3.5 **Overspill-Lemma**

14.3.10 **Satz** Ist $\mathcal{A} \models Th(\mathcal{N})$ Nicht-Standard, so gilt:
$\mathcal{A} \models \mathcal{F}(e)$ für ein Nicht-Standard $e \Leftrightarrow \mathcal{N} \models \forall x \exists y(x < y \wedge \mathcal{F}(y))$

14.4.3 Größenordnung $[e]$ in $\mathcal{A}$.

14.4.6 Struktur der unendlichen Größenordnungen

14.4.8 $\mathcal{A} + \mathcal{B}$, $\mathcal{A} \times \mathcal{B}$ für $\mathcal{A}, \mathcal{B} \models LO$

14.4.9 **Satz** Zu Nicht-Standard $\mathcal{A} \models Z^-$ gibt es $\mathcal{B} \models DLO$ mit
$(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + \mathcal{B} \times (\mathbb{Z}, <)$

14.4.10 Korollar. Ist $\mathcal{A} \models Z^-$ abzählbar, so ist $\mathcal{A} \cong \mathcal{N}$ oder
$(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + (\mathbb{Q}, <) \times (\mathbb{Z}, <)$

15.1.7 Zahlentheorie Z^2 der zweiten Stufe

15.1.8 Komprehensions-Gleichung

15.2.1 Erststufiger Anteil

15.2.2 Relativierung ^{U}

15.2.3 Übersetzung von L in T'

15.2.7 Übersetzung von T in T'. Dafür gilt:

15.2.9 **Satz** Ist $\mathcal{A} \models T'$, so ist $\mathcal{B} \subseteq \mathcal{A}|L$ mit $|\mathcal{B}| = U_{\mathcal{A}}$ Modell von T.

15.2.11 **Satz** Aus $T \models C$ folgt $T' \models C^{(U)}$.

15.3.1 Standardmodell $\mathcal{N}^2$ von Z^2

15.3.3 Reguläre, ω- und vollständige Modelle von Z^2

15.3.7 **Satz** Jedes Modell von Z^2 ist isomorph zu einem regulären Modell.

15.3.8 Satz von Peano

§13 Kategorizität und die Theorie der dichten linearen Ordnung

13.1 Kategorizität und Vollständigkeit

13.2 $\aleph_0$-Kategorizität der Theorie *DLO*

13.3 Nicht-Kategorizität von *DLO* in der Mächtigkeit $2^{\aleph_0}$

13.4 Zusammenfassung und Aufgaben

13.1 Kategorizität und Vollständigkeit

Wenn man sich für die Anzahl der Modelle einer Theorie interessiert, wird man isomorphe Modelle nicht unterscheiden. Man interessiert sich dann für die Anzahl der Modelle der Theorie *modulo* Isomorphie. Der Nutzen mancher Theorien liegt gerade darin, dass sie viele nicht-isomorphe Modelle haben und ihre Ergebnisse in vielen mathematisch relevanten Situationen angewandt werden können. Gruppentheorie, Ringtheorie, Körpertheorie, Theorie der linearen Ordnung sind Theorien dieser Art. Es gibt aber auch Theorien, die entworfen wurden, um eine einzige Struktur möglichst eindeutig zu beschreiben. Sie sollten *kategorisch* sein in dem Sinne, dass sie bis auf Isomorphie nur ein einziges Modell besitzen. Die Zahlentheorie und die Mengenlehre waren solche Theorien, denen man gern eine kategorische Formulierung gegeben hätte. Das ist nun nach den Sätzen von Löwenheim und Skolem nicht möglich:

Theorien, die ein unendliches Modell haben, können nicht (absolut) kategorisch sein, weil sie sogar Modelle jeder (genügend großen) unendlichen Mächtigkeit besitzen.

Die Frage nach der Kategorizität muss man also auf die Kardinalzahlen einzeln beziehen: Hat eine Theorie bis auf Isomorphie u. U. genau ein Modell der Mächtigkeit κ?

13.1.1 Definition Sei κ eine Kardinalzahl. Eine Theorie T heißt *κ-kategorisch* oder *kategorisch in der Mächtigkeit* κ, wenn

(1) T ein Modell der Mächtigkeit κ besitzt und

(2) alle Modelle von T, die die Mächtigkeit κ haben, zueinander isomorph sind.

Beispiele.

1. Die Theorie LO der linearen Ordnung ist n-kategorisch für jede natürliche Zahl $n > 0$. Denn jede lineare Ordnung von n Elementen ist isomorph zum Abschnitt $\{0, \ldots, n-1\}$ von $\mathbb{N}$, versehen mit der natürlichen Anordnung.

   LO ist aber nicht $\aleph_0$-kategorisch. Denn sowohl $(\mathbb{N}, <)$ als auch $(\mathbb{Q}, <)$ sind abzählbar unendliche lineare Ordnungen, und sie sind nicht isomorph, weil sie nicht einmal elementar äquivalent sind.

2. Die Gruppentheorie T_G ist p-kategorisch für jede Primzahl p. Denn die zyklische Gruppe der Ordnung p, $\mathbb{Z}/p\mathbb{Z}$, ist bis auf Isomorphie die einzige Gruppe mit p Elementen, wenn p eine Primzahl ist.

   Dagegen ist die Gruppentheorie nicht 4-kategorisch. Denn es gibt zwei nicht-isomorphe Gruppen mit 4 Elementen, nämlich außer der zyklischen Gruppe der Ordnung 4 noch die Kleinsche Vierergruppe

   $$(\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z}).$$

   Die Gruppentheorie ist auch nicht κ-kategorisch für irgendein unendliches κ. Denn es gibt bekanntlich für jedes $\kappa \geq \aleph_0$ sowohl kommutative als auch nicht-kommutative Gruppen der Mächtigkeit κ.

3. Die endlichen Körper sind die Galois-Felder. Sie haben eine endliche Charakteristik p und p^m Elemente, wobei p eine Primzahl und m eine natürliche Zahl > 0 ist. Bis auf Isomorphie gibt es auch nur ein Galois-Feld mit p^m Elementen. Also ist die Körpertheorie T_K p^m-kategorisch, wenn p eine Primzahl und $m > 0$ ist.

4. Wir betrachten die abzählbaren Körper $\mathbb{Q}$ der rationalen Zahlen und $\mathbb{Q}(\sqrt{-1})$ der komplex-rationalen Zahlen, ferner die Körper $\mathbb{R}$ der reellen und $\mathbb{C}$ der komplexen Zahlen, die beide die Mächtigkeit $2^{\aleph_0}$ haben. Da der Satz

   $$\exists x \quad x \cdot x = -1$$

in $\mathbb{Q}(\sqrt{-1})$ und in $\mathbb{C}$ wahr, in $\mathbb{Q}$ und $\mathbb{R}$ aber falsch ist, ist weder $\mathbb{Q}$ zu $\mathbb{Q}(\sqrt{-1})$ noch $\mathbb{R}$ zu $\mathbb{C}$ elementar äquivalent. Diese Körper sind daher auch nicht isomorph.

Also ist die Theorie $T_{K,0}$ der Körper der Charakteristik 0 weder $\aleph_0$- noch $2^{\aleph_0}$-kategorisch. Dies gilt dann umso mehr für alle Teiltheorien von $T_{K,0}$ wie T_K, $T_{R,0}$, T_R.

Aus endlich-kategorischen Theorien, für die es hiernach viele Beispiele gibt, gewinnt man in trivialer Weise absolut-kategorische Theorien, indem man zu ihnen einen Satz $(E_{=n})$ als Axiom hinzufügt, der besagt, dass es genau n Elemente gibt. Wir setzen

$$(E_{=n}) \equiv (E_{\leq n}) \wedge (E_{\geq n}).$$

$(E_{\leq n})$ wurde am Ende von 10.2 eingeführt, und analog wird „es gibt mindestens n Elemente“ ausgedrückt durch

$$(E_{\geq n}) \qquad \exists x_0 \ldots \exists x_{n-1}(x_0 \neq x_1 \wedge x_0 \neq x_2 \wedge \ldots \wedge x_{n-2} \neq x_{n-1}).$$

13.1.2 Lemma Sei n eine natürliche Zahl > 0. Ist T n-kategorisch, so ist $T+(E_{=n})$ absolut-kategorisch: $T+(E_{=n})$ hat bis auf Isomorphie nur ein einziges Modell, und dieses Modell hat n Elemente.

Beweis. Wegen des Axioms $(E_{=n})$ hat jedes Modell von $T + (E_{=n})$ genau n Elemente, und weil T n-kategorisch ist, gibt es von diesen bis auf Isomorphie nur genau eines.

Es gibt eine abzählbare Theorie, die in jeder Kardinalzahl $\kappa > 0$ kategorisch ist.

13.1.3 Definition Die *reine Theorie der Identität* $T_=$ ist die logische Theorie, deren Sprache keine nicht-logischen Grundzeichen enthält.

Bemerkung. Die einzigen Primformeln von $T_=$ sind Gleichungen $a = b$ zwischen Variablen.

13.1.4 Lemma $T_=$ ist κ-kategorisch für jede endliche oder unendliche Kardinalzahl $\kappa > 0$.

Beweis. Die Modelle von $T_=$ sind die Strukturen $\mathcal{A} = (A)$ „ohne Struktur“, nämlich ohne Funktionen und Relationen. Die Isomorphismen sind dann einfach die Bijektionen zwischen den Individuenbereichen. $\mathcal{A} \cong \mathcal{B}$ bedeutet dann nur, dass $|\mathcal{A}|$ und $|\mathcal{B}|$ gleichmächtig sind.

Aus $T_=$ gewinnen wir leicht λ-Theorien, die nicht λ-kategorisch sind, wohl aber κ-kategorisch für größere κ.

13.1.5 Lemma Sei λ eine unendliche Kardinalzahl und E eine Menge von λ vielen Konstanten. Dann ist

$$T \;:=\; T_= + E + \{\neg e = e' \mid e, e' \in E,\; e \not\equiv e'\}$$

eine λ-Theorie, die nicht λ-kategorisch, wohl aber κ-kategorisch ist für jedes $\kappa > \lambda$.

Beweis. Es sei M irgendeine zu E disjunkte Menge. Dann ist

$$\mathcal{A}_M := (E \cup M, (e)_{e \in E})$$

ein Modell dieser Theorie T, und jedes Modell $\mathcal{B} = (B, (e_\mathcal{B})_{e \in E})$ von T ist zu einem solchen $\mathcal{A}_M$ isomorph, weil die durch

$$\begin{aligned} \varphi(e) &= e_\mathcal{B} \quad \text{für } e \in E \\ \varphi(k) &= k \quad \text{für } k \in B - \mathcal{B}[E] \end{aligned}$$

definierte Abbildung φ ein Isomorphismus von $A_{B-\mathcal{B}[E]}$ auf $\mathcal{B}$ ist.

Zwei Modelle $\mathcal{A}_M$ und $\mathcal{A}_{M'}$ von T sind hiernach genau dann isomorph, wenn M und M' gleichmächtig sind. Zu jeder Kardinalzahl κ gibt es also eine Isomorphieklasse von Modellen $\mathcal{A}_M$ mit $card(M) = \kappa$. Solange $\kappa \leq \lambda$ ist – und dafür gibt es mindestens die unendlich vielen Möglichkeiten $\kappa \in \mathbb{N}$, $\kappa = \aleph_0$ –, hat $\mathcal{A}_M$ die Mächtigkeit $\lambda + \kappa = \lambda$, und T ist nicht λ-kategorisch. Für $\kappa > \lambda$ hat $\mathcal{A}_M$ aber die Mächtigkeit $\lambda + \kappa = \kappa$, so dass in diesem Fall T κ-kategorisch ist.

Für $\lambda = \aleph_0$ ergibt dieses Lemma das

13.1.6 Korollar Es gibt abzählbare Theorien, die nicht $\aleph_0$-kategorisch, wohl aber κ-kategorisch sind für jede Kardinalzahl $\kappa > \aleph_0$.

Eine Theorie, die n-kategorisch ist für ein natürliches n, ist entweder absolut kategorisch wie in 13.1.2, oder ihre Modelle sind nicht einmal alle elementar äquivalent. Denn wenn sie außer dem n-elementigen Modell $\mathcal{A}$ noch ein Modell $\mathcal{B}$ hat, gilt die Anzahlaussage $(E_{=n})$ in $\mathcal{A}$, aber nicht in $\mathcal{B}$: Die Theorie ist nicht *vollständig*.

13.1.7 Definition Eine Theorie T ist *vollständig*, wenn für jeden Satz C von $L(T)$ entweder C oder $\neg C$ in T gilt.

Diese Vollständigkeit von Theorien hat offenbar mit der Vollständigkeit des Herleitbarkeitsbegriffs, wie sie sich im Vollständigkeitssatz ausdrückt, wenig zu tun.

Beispiele. Jede Theorie $Th(\mathcal{A})$ ist vollständig. Denn für Sätze C aus $L(Th(\mathcal{A}))$ ist

$$\mathcal{A} \models C \;\Leftrightarrow\; C \in Ax(Th(\mathcal{A})) \;\Leftrightarrow\; Th(\mathcal{A}) \models C,$$

und in $\mathcal{A}$ gilt entweder C oder $\neg C$.

Also ist nach 7.3.4 jede maximal konsistente Theorie vollständig. Man überlegt sich leicht, dass jede vollständige Theorie im Sinne von Definition 11.3.10 äquivalent ist zu einer maximal konsistenten Theorie.

13.1.8 Satz Sei κ eine unendliche Kardinalzahl und T eine κ-Theorie, die nur unendliche Modelle besitzt. Ist T κ-kategorisch, so ist T vollständig.

Beweis. Sei T κ-kategorisch und $\mathcal{A}$ ein Modell von T der Mächtigkeit κ. Wir behaupten für Sätze C aus $L(T)$:

$$T \models C \;\Leftrightarrow\; \mathcal{A} \models C.$$

Die Richtung $\Rightarrow$ ist klar. Wenn $T \not\models C$, hat T ein Modell $\mathcal{B}_0$, in dem C nicht gilt. Dieses Modell $\mathcal{B}_0$ ist – wie alle Modelle von T – unendlich. Nach dem Satz von Löwenheim-Skolem-Tarski hat also T ein Modell $\mathcal{B}$ der Mächtigkeit κ, in dem C nicht gilt. Da T κ-kategorisch ist, ist $\mathcal{B}$ isomorph zu dem gegebenen Modell $\mathcal{A}$. Nach 11.2.11 gilt C dann auch in $\mathcal{A}$ nicht. Damit ist auch die Richtung $\Leftarrow$ bewiesen.

Da in jeder Struktur $\mathcal{A}$ zu $L(T)$ entweder C oder $\neg C$ gilt, ist damit T vollständig.

13.2 $\aleph_0$-Kategorizität von *DLO*

Wir wenden uns der Theorie *DLO* der dichten linearen Ordnung ohne erstes und letztes Element zu (vgl. 1.2.5). Sie ist ein besonders einfaches Beispiel einer Theorie, die aus nicht-trivialen Gründen $\aleph_0$-kategorisch ist. Im Gegensatz zur Theorie *LO* hat *DLO* nur unendliche Modelle. Die Struktur $(\mathbb{Q}, <)$, die Menge der rationalen Zahlen mit ihrer natürlichen Anordnung, ist ein abzählbar unendliches Modell von *DLO* (vgl. 2.2.7). Wir wollen zeigen, dass $(\mathbb{Q}, <)$ bis auf Isomorphie das einzige abzählbare Modell von *DLO* ist.

13.2.1 Satz Jedes abzählbare Modell von *DLO* ist isomorph zu $(\mathbb{Q}, <)$:

$$DLO \text{ ist } \aleph_0\text{-kategorisch.}$$

Beweis. Es sei $\mathcal{A} = (A, <)$ ein abzählbares Modell von *DLO*. Gegeben seien eine Abzählung von $\mathbb{Q}$ und eine Abzählung von A. Wir definieren eine Abbildung

$$\varphi : \mathbb{Q} \to A$$

rekursiv wie folgt:

φ sei für n rationale Zahlen

$$q_0 < q_1 < \ldots < q_{n-1}$$

bereits definiert, und es sei

$$\varphi(q_0) < \varphi(q_1) < \ldots < \varphi(q_{n-1}) \text{ in } \mathcal{A}.$$

Nun sei q das erste Element von $\mathbb{Q} - \{q_0, \ldots q_{n-1}\}$ in der gegebenen Aufzählung von $\mathbb{Q}$. Da $\mathbb{Q}$ durch $<$ linear geordnet ist, ist

$$q < q_0 \text{ oder } q_i < q < q_{i+1} \text{ für ein } i < n-1 \text{ oder } q_{n-1} < q.$$

In jedem Fall sei $\varphi(q)$ dasjenige Element aus dem entsprechenden offenen Intervall in $\mathcal{A}$, das in der gegebenen Aufzählung von A das erste ist. Dann ist

$$\begin{array}{ll} \varphi(q) < \varphi(q_0), & \text{falls } q < q_0 \text{ ist,} \\ \varphi(q_i) < \varphi(q) < \varphi(q_{i+1}), & \text{falls } q_i < q < q_{i+1} \text{ ist, und} \\ \varphi(q_{n-1}) < \varphi(q), & \text{falls } q_{n-1} < q \text{ ist.} \end{array}$$

Eine solche Wahl von $\varphi(q)$ ist möglich, weil LO 4 bis 6 in $\mathcal{A}$ gelten und deshalb alle diese Intervalle nicht leer sind.

Sei nun $r_0 < \ldots < r_n$ eine Anordnung von $q_0, \ldots, q_{n-1}, q$ nach der Größe. Dann ist nach Konstruktion auch

$$\varphi(r_0) < \ldots < \varphi(r_n),$$

und die Eingangsvoraussetzungen sind (für $n+1$ Zahlen aus $\mathbb{Q}$) wieder hergestellt.

1. φ ist auf ganz $\mathbb{Q}$ definiert. Denn wenn $q_0, \ldots, q_{n-1}$ die n ersten Elemente in der Aufzählung von $\mathbb{Q}$ sind, ist das erste Element q von $\mathbb{Q} - \{q_0, \ldots, q_{n-1}\}$ offenbar das $n+1$-te Element von $\mathbb{Q}$. $\varphi(q)$ wird also per Rekursion für alle $q \in \mathbb{Q}$ nacheinander in ihrer gegebenen Aufzählung definiert.

2. φ ist ein Homomorphismus von $(\mathbb{Q}, <)$ in $(A, <)$ nach der rekursiven Konstruktion von φ. Deshalb ist φ auch injektiv und nach 11.4.4 eine Einbettung.

3. φ ist surjektiv. Denn angenommen, dies wäre nicht so. Dann gäbe es ein erstes Element a in der gegebenen Aufzählung von A, das nicht im Bild von φ läge.

   $$b_0 < b_1 < \ldots < b_{n-1}$$

   sei eine Anordnung der n in der Aufzählung vorhergehenden Elemente von A. Zu diesen gäbe es dann $q_i = \varphi^{-1}(b_i)$, und nach 2. wäre

   $$q_0 < q_1 < \ldots < q_{n-1}.$$

   Weil $\mathcal{A}$ linear geordnet ist, wäre

   $$a < b_0 \text{ oder } b_i < a < b_{i+1} \text{ für ein } i \text{ oder } b_{n-1} < a.$$

   Das entsprechende Intervall in $\mathbb{Q}$ wäre nicht leer, weil $(\mathbb{Q}, <)$ dicht geordnet ohne erstes und letztes Element ist. Wäre q die erste rationale Zahl in der gegebenen Aufzählung von $\mathbb{Q}$, die in dem entsprechenden Intervall liegt, etwa

   $$q_i < q < q_{i+1}, \text{ falls } b_i < a < b_{i+1},$$

   so wäre $\varphi(q) = a$ nach Konstruktion von φ, im Widerspruch dazu, dass a nicht im Bild von φ liegt. Also ist φ surjektiv.

Nach 1., 2. und 3. ist φ ein Isomorphismus von $(\mathbb{Q}, <)$ auf $(A, <)$ (vgl. 11.2.3).

Bemerkung. Dieser Beweis ist die einfachste Anwendung der back-and-forth-Methode aus der Modelltheorie. In unserer Darstellung liegt der Vorwärts-Teil in der Konstruktion der Abbildung φ. Aus ihr folgt unmittelbar, dass φ auf ganz $\mathbb{Q}$ definiert ist. Der Rückwärts-Teil liegt im Beweis der Surjektivität von φ. Dort wird de facto gezeigt, dass φ^{-1} auf ganz A definiert ist.

Zur Konstruktion von φ haben wir die Gültigkeit von LO 1 bis 3 in $(\mathbb{Q}, <)$ und von LO 4 bis 6 in $\mathcal{A}$ ausgenutzt. Zum Beweis der Surjektivität von φ brauchten wir umgekehrt, dass LO 1 bis 3 in $\mathcal{A}$ und LO 4 bis 6 in $(\mathbb{Q}, <)$ gelten.

13.2.2 Korollar Die Theorie DLO ist vollständig.

Denn DLO ist eine $\aleph_0$-kategorische $\aleph_0$-Theorie, die nur unendliche Modelle hat. Also ist DLO nach 13.1.8 vollständig.

13.3 Nicht-Kategorizität von *DLO* in der Mächtigkeit $2^{\aleph_0}$

Hängt der Beweis von Satz 13.2.1 wirklich an der Abzählbarkeit des Modells $\mathcal{A}$, oder lässt sich mit einem geschickteren Argument auch die Kategorizität von DLO in größeren Mächtigkeiten nachweisen? Wir zeigen, dass das nicht der Fall ist.

$(\mathbb{R}, <)$ und $(\mathbb{R} - \{0\}, <)$ sind Modelle von DLO derselben überabzählbaren Mächtigkeit $2^{\aleph_0}$. Wir zeigen, dass es zwischen diesen beiden Strukturen keinen Isomorphismus geben kann, weil unter einem solchen Isomorphismus das Urbild der negativen Zahlen in $(\mathbb{R}, <)$ ein Supremum hätte, dessen Bild die Null sein müsste, die aber nicht zu $(\mathbb{R} - \{0\}, <)$ gehört.

13.3.1 Definition Sei $\mathcal{A} = (A, <_{\mathcal{A}})$ eine lineare Ordnung. Man schreibt $a \leq_{\mathcal{A}} b$ für $a <_{\mathcal{A}} b \vee a = b$. $s \in A$ ist *obere Schranke* von $X \subseteq A$, wenn $k \leq_{\mathcal{A}} s$ für alle $k \in X$ gilt. s ist *Supremum* von X (in $\mathcal{A}$), wenn s obere Schranke von X und $s \leq_{\mathcal{A}} s'$ ist für jede obere Schranke s' von X.

Bemerkung. Jede Teilmenge X von $\mathcal{A}$ hat höchstens ein Supremum in $\mathcal{A}$. Denn sind s, s' Suprema von X, so ist $s \leq_{\mathcal{A}} s'$ und $s' \leq_{\mathcal{A}} s$, also $s = s'$. Man

schreibt daher auch $s = \sup X$, wenn s Supremum von X ist.

Beispiele.

1. Jede nicht-leere nach oben beschränkte Teilmenge von $\mathbb{R}$ hat ein Supremum in $(\mathbb{R}, <)$. Das ist die (Ordnungs-) Vollständigkeit der reellen Zahlen. Dagegen hat die Menge der negativen reellen Zahlen kein Supremum in $(\mathbb{R} - \{0\}, <)$. Denn da es keine größte negative reelle Zahl gibt, sind die oberen Schranken dieser Menge (in $(\mathbb{R} - \{0\}, <)$) genau die positiven reellen Zahlen, und unter denen gibt es keine kleinste.

2. Ist $\alpha \in \mathbb{R}$ irrational, etwa $\alpha = \sqrt{2}$, so ist die Menge $\{q \in \mathbb{Q} \mid q < \alpha\} \subseteq \mathbb{Q}$ zwar nicht-leer und beschränkt, besitzt aber kein Supremum in $(\mathbb{Q}, <)$, aus einem analogen Grund wie eben.

13.3.2 Lemma Sei $\mathcal{A}, \mathcal{B} \models LO, s \in |\mathcal{A}|$ und $X \subseteq |\mathcal{A}|$.

(1) Ist $\varphi : \mathcal{A} \hookrightarrow \mathcal{B}$, so ist s obere Schranke von X genau dann, wenn $\varphi(s)$ obere Schranke von $\varphi[X]$ ist.

(2) Ist $\varphi : \mathcal{A} \cong \mathcal{B}$ und $s = \sup X$, so ist $\varphi(s) = \sup \varphi[X]$.

Beweis. Für Einbettungen φ von $\mathcal{A}$ in $\mathcal{B}$ folgt aus $k \leq_{\mathcal{A}} s$ stets $\varphi(k) \leq_{\mathcal{B}} \varphi(s)$, und umgekehrt. Damit folgt (1).
Ist s' obere Schranke von $\varphi[X]$ (in $\mathcal{B}$), so ist bei surjektivem φ $s' = \varphi(s'')$, wobei wegen (1) s'' obere Schranke von X ist. Dann ist $s = \sup X \leq_{\mathcal{A}} s''$, also $\varphi(s) \leq_{\mathcal{B}} s'$, und es gilt (2).

Damit ist die Behauptung klar:

13.3.3 Satz Es gibt keinen Isomorphismus von $(\mathbb{R}, <)$ auf $(\mathbb{R} - \{0\}, <)$:

$$DLO \text{ ist nicht } 2^{\aleph_0}\text{-kategorisch.}$$

Beweis. Angenommen, φ wäre ein solcher Isomorphismus. Wir setzen $X = \{\alpha \in \mathbb{R} \mid \varphi(\alpha) < 0\}$. Nach 13.3.2 (1) ist $\varphi^{-1}(1)$ obere Schranke von X. Also hat X ein Supremum s in $(\mathbb{R}, <)$. Nach 13.3.2 (2) wäre

$$\varphi(s) = \sup \varphi[X] = \sup\{\alpha \in \mathbb{R} \mid \alpha < 0\}.$$

Aber, wie in Beispiel 1 gezeigt, hat $\{\alpha \in \mathbb{R} \mid \alpha < 0\}$ kein Supremum in $(\mathbb{R} - \{0\}, <)$. Also kann φ kein Isomorphismus sein: $(\mathbb{R}, <)$ und $(\mathbb{R} - \{0\}, <)$

sind nicht isomorph.

Da, wie eingangs erwähnt, $(\mathbb{R}, <)$ und $(\mathbb{R} - \{0\}, <)$ beides Modelle von DLO der Mächtigkeit $2^{\aleph_0}$ sind, ist DLO nicht $2^{\aleph_0}$-kategorisch.

Trotz dieses Ergebnisses sind $(\mathbb{R}, <)$ und $(\mathbb{R} - \{0\}, <)$ elementar äquivalent, weil DLO nach 13.2.2 vollständig ist. Beide Modelle unterscheiden sich also durch eine nicht-elementare Eigenschaft, die sich nicht durch eine Formel aus $L(LO)$ ausdrücken lässt: Die Eigenschaft, dass jeder nicht-leere echte Abschnitt ein Supremum besitzt, lässt sich nach 13.3.3 und 13.2.2 nicht in der (elementaren) Sprache der linearen Ordnung formulieren.

13.4 Zusammenfassung und Aufgaben

Wir betrachten zum Abschluss dieses Paragraphen nur noch abzählbare Theorien und stellen zusammen, was wir über ihre Kategorizität in unendlichen Mächtigkeiten wissen.

(1) Die Theorie $T_=$ ist nach 13.1.4 aus trivialen Gründen κ-kategorisch für jedes (unendliche) κ.

(2) Gruppen-, Ring- und Körpertheorie T_G, T_R und T_K sind ebenso wie die Theorie LO der linearen Ordnung für kein unendliches κ κ-kategorisch, wie in den Beispielen in 13.1 skizziert. Das gleiche gilt für die Zahlentheorie Z, worauf wir noch im nächsten Paragraphen eingehen werden.

(3) Die Theorie DLO ist zwar $\aleph_0$-kategorisch nach 13.2.1, aber nicht $2^{\aleph_0}$-kategorisch nach 13.3.3.

(4) Umgekehrt ist die Theorie $T_= + E + \{\neg e = e' \mid e, e' \in E, e \not\equiv e'\}$ mit abzählbar unendlichem E nach 13.1.5 zwar nicht $\aleph_0$-kategorisch, wohl aber κ-kategorisch für alle $\kappa > \aleph_0$.

Zu jeder Möglichkeit, $\aleph_0$-Kategorizität und κ-Kategorizität für geeignete $\kappa > \aleph_0$ positiv oder negativ zu kombinieren, haben wir also abzählbare Theorien gefunden, die diese Möglichkeiten realisieren. Nicht gefunden haben wir eine abzählbare Theorie, die zwar κ-, aber nicht λ-kategorisch ist für geeignete $\kappa, \lambda > \aleph_0$. Das kann auch nicht gelingen wegen:

13.4.1 Satz von Morley
Ist eine abzählbare Theorie κ-kategorisch für *ein* $\kappa > \aleph_0$, so ist sie schon κ-kategorisch für *jedes* $\kappa > \aleph_0$.

Dieser Satz ist ein tiefliegendes Ergebnis der Modelltheorie. Sein Beweis würde den Rahmen dieser Einführung bei weitem sprengen. Für uns liefert er zu 13.3.3 das

13.4.2 Korollar *DLO* ist für kein $\kappa > \aleph_0$ κ-kategorisch.

Der Satz von Morley vereinfacht auch die Unterscheidungen in folgendem Diagramm, in das wir die in (1) bis (4) genannten Theorien einordnen.

abzählbare Theorien	κ-kategorisch für $\kappa > \aleph_0$	nicht κ-kategorisch für $\kappa > \aleph_0$
$\aleph_0$-kategorisch	$T_=$	DLO
nicht $\aleph_0$-kategorisch	$T_= + E + \{\neg e = e' \mid e \not\equiv e'\}$	T_G, T_R, T_K, LO, Z

Durch ihre Kategorizität kann also eine abzählbare Theorie Kardinalzahlen oberhalb von $\aleph_0$ nicht voneinander trennen. In unserem Diagramm ist es deshalb gleichgültig, ob wir in den oberen Eingängen nach der κ-Kategorizität für *ein* $\kappa > \aleph_0$ oder für *alle* $\kappa > \aleph_0$ unterscheiden.

Aufgaben

13.4.3 Zeigen Sie für konsistente Theorien T die Äquivalenz von

a) T ist vollständig.

b) Alle Modelle von T sind elementar äquivalent.

c) T ist äquivalent zu einer maximal konsistenten Theorie.

d) T hat eine maximal konsistente konservative Erweiterung.

13.4.4 Zeigen Sie: Eine vollständige Theorie, die ein endliches Modell besitzt, hat bis auf Isomorphie nur dieses eine endliche Modell.

13.4.5 Für welche κ ist die Theorie aus 12.4.4 κ-kategorisch?

13.4.6 a. Sei $\mathcal{A} = (A, <_{\mathcal{A}})$ ein Modell von $LO + \{LO4, \exists x \forall y \neg y < x\}$ und sei $\min \in A$ das kleinste Element von $\mathcal{A}$. Zeigen Sie: Die Unterstruktur $\mathcal{B}$ von $\mathcal{A}$ mit Individuenbereich $A - \{\min\}$ ist ein Modell von $LO + \{LO4, LO5\}$.

b. Zeigen Sie: $LO + \{LO4\}$ hat bis auf Isomorphie genau 4 abzählbare Modelle. (Verwenden Sie a.)

13.4.7 Wir betrachten das Hessenberg-Produkt $\times_H$ aus 12.1.20.

a. Zeigen Sie: Ist $\mathcal{A}$ ein Modell von DLO, so auch $\mathcal{A} \times_H \mathcal{A}$.

b. Ist $(\mathbb{Q}, <) \times_H (\mathbb{Q}, <) \cong (\mathbb{Q}, <)$?

c. Ist $(\mathbb{R}, <) \times_H (\mathbb{R}, <) \cong (\mathbb{R}, <)$?

§14 Nicht-Standard-Modelle der Zahlentheorie

14.1 Enderweiterungen

14.2 Existenz und Anzahl von abzählbaren Nicht-Standard-Modellen

14.3 Overspill

14.4 Anordnung in Nicht-Standard-Modellen

14.5 Aufgaben

Die Struktur $\mathcal{N} = (\mathbb{N}; 0, S, +, \cdot)$ lässt sich nach dem Satz von Löwenheim-Skolem-Tarski nicht bis auf Isomorphie eindeutig durch eine mathematische Theorie beschreiben: Jede Theorie T, die $\mathcal{N}$ als Modell besitzt, ist abzählbar und hat deshalb Modelle jeder unendlichen Mächtigkeit. Was kann man über Struktur und Anzahl dieser Modelle Genaueres sagen?

Die Anzahl der Modelle einer Zahlentheorie T wird eher kleiner sein, wenn wir für T eine möglichst starke Theorie wählen, etwa $Th(\mathcal{N})$, die Theorie des Standardmodells $\mathcal{N}$. Die innere Struktur wird eher besonders uneinheitlich sein, wenn wir für T eine möglichst schwache Theorie wählen, etwa eine Teiltheorie von Z mit möglichst wenig Induktionsaxiomen. Es wird sich zeigen, dass wir in beiden Extremfällen überraschend einheitliche und starke Ergebnisse gewinnen können.

14.1 Enderweiterungen

Bisher haben wir die gewöhnliche Zahlentheorie Z in 1.2.4 formuliert und in 2.2 festgehalten, dass die Struktur $\mathcal{N}$ der natürlichen Zahlen ein Modell von Z ist, das sog. Standardmodell. In der Theorie Z haben wir aber noch nichts als gültig oder herleitbar nachgewiesen. Das holen wir jetzt – in minimalem Umfang – nach.

Dabei können wir uns wegen des Vollständigkeitssatzes auf Gültigkeitsnachweise beschränken. Das erlaubt ein Argumentieren im üblichen axiomatischen Rahmen und verdeutlicht die benötigten mathematischen Ideen.

Wir wollen eine Ordnungsrelation in der Sprache von Z definieren und für sie einige Eigenschaften in Z nachweisen. Dazu müssen wir auch eine Reihe von

Additionsgesetzen in Z beweisen. Da die Multiplikation in diesen Überlegungen nicht auftritt, gelten die hier bewiesenen Formeln schon in der *additiven Zahlentheorie*, dem (viel schwächeren) Fragment von Z, das nur Null, Nachfolger und Addition, nicht aber die Multiplikation behandelt.

14.1.1 Lemma In Z gelten folgende Formeln:

1. $a = 0 \vee \exists x\ a = Sx$
2. $(a + b) + c = a + (b + c)$
3. $0 + b = b$
4. $Sa + b = S(a + b)$
5. $a + b = b + a$
6. $b + a = c + a \to b = c$
7. $a + b = a + c \to b = c$
8. $\exists x(x + x = a \vee S(x + x) = a)$

Beweis. Alle Beweise verwenden Induktionsaxiome (*Ind*).

1. folgt mit Abschwächungen und (*Ind*) aus $0 = 0$ und $\exists x\ Sa = Sx$.
2. In Z gilt $(a + b) + 0 = a + b = a + (b + 0)$ und

$$\begin{aligned}(a + b) + c \ &= a + (b + c) \ \to \ (a + b) + Sc = S((a + b) + c)\\ &= S(a + (b + c)) = a + S(b + c) = a + (b + Sc).\end{aligned}$$

   Mit (*Ind*) folgt 2.

3. folgt mit (*Ind*) aus $0 + 0 = 0$ und

$$0 + b = b \to 0 + Sb = S(0 + b) = Sb.$$

4. folgt mit (*Ind*) aus $Sa + 0 = Sa = S(a + 0)$ und

$$Sa + b = S(a + b) \to Sa + Sb = S(Sa + b) = SS(a + b) = S(a + Sb).$$

5. Aus 3. folgt $0 + b = b = b + 0$; aus 4. folgt

$$a + b = b + a \to Sa + b = S(a + b) = S(b + a) = b + Sa.$$

Mit (*Ind*) folgt hieraus 5.

6. folgt mit (*Ind*) aus $b + 0 = c + 0 \to b = b + 0 = c + 0 = c$ und

$(b+a = c+a \to b = c) \to S(b+a) = b+Sa = c+Sa = S(c+a) \to b = c,$

was wegen $S(b + a) = S(c + a) \to b + a = c + a$ in Z gilt.

7. folgt aus 5. und 6.

8. folgt mit (*Ind*) aus $0 + 0 = 0$,

$$b + b = a \to S(b + b) = Sa \text{ und}$$
$$S(b + b) = a \to Sb + Sb = S(Sb + b) = SS(b + b) = Sa$$

was nach 4. in Z gilt.

Bemerkung. Nach 1. ist jede Zahl Null oder Nachfolger: Anders als bei den Ordinalzahlen gibt es in den Modellen von Z keine Limeszahlen. 2. und 5. sind das *assoziative* und das *kommutative Gesetz* für die Addition. 6. und 7. sind Kürzungsregeln für die Addition. Die Zahl, deren Existenz in 8. behauptet wird, ist der (eindeutig bestimmte) *ganzzahlige Anteil von* $\frac{a}{2}$.

14.1.2 Definition

$$a \leq b :\Leftrightarrow \exists x(a + x = b)$$
$$a < b :\Leftrightarrow \exists x(a + Sx = b).$$

Damit sind Zeichen für *kleiner-gleich* und für *kleiner* als Abkürzungen in Z eingeführt.

14.1.3 Lemma In Z gelten folgende Formeln:

1. $\neg a < a$
2. $a < b \to b < c \to a < c$
3. $a < b \vee a = b \vee b < a$

4. $0 \leq b$
5. $a \leq b \leftrightarrow a = b \vee a < b$
6. $a < b \leftrightarrow Sa \leq b$
7. $b < c \rightarrow a + b < a + c$.

Beweis.

1. Aus $a + Sb = a = a + 0$ folgt $Sb = 0$ nach 14.1.1, 7, woraus $\perp$ folgt. Also gilt 1.
2. Aus $a + Sa' = b$ und $b + Sb' = c$ folgt mit 14.1.1, 2
$$a + S(Sa' + b') = a + (Sa' + Sb') = (a + Sa') + Sb' = c,$$
also $a < c$, und es gilt 2.
3. beweisen wir zum Schluss.
4. folgt aus 14.1.1, 3.
5. folgt aus 14.1.1, 1.
6. folgt aus 14.1.1, 4.
7. folgt aus 14.1.1, 2 wegen
$$b + Sd = c \rightarrow (a + b) + Sd = a + (b + Sd) = a + c.$$

Zu 3. Es ist $0 = a \vee 0 < a$ nach 4. und 5. Ferner gilt

$$a \leq b \vee b < a \rightarrow a < Sb \vee Sb \leq a$$

nach 6. Mit 5. und (*Ind*) folgt 3.

14.1.4 Definition Mit Z^- bezeichnen wir die Theorie Z ohne das Induktionsschema (*Ind*), dafür ergänzt um Allabschlüsse der Formeln aus 14.1.1 und 14.1.3, wobei $\leq$ und $<$ Abkürzungen gemäß 14.1.2 sind.

Diese schwache Zahlentheorie Z^- ist endlich axiomatisiert, im Gegensatz zu Z. Die neuen Axiome sind nicht alle unabhängig voneinander; aus 14.1.1, 1, 2, 5, 6 und 8, sowie aus 14.1.3, 3 folgen bereits die übrigen Gesetze. Mit der Einführung von Z^- soll verdeutlicht werden, dass sich das Folgende beweisen lässt unter Verwendung ganz weniger Induktionsaxiome aus Z, die auch nur für die schon genannten Ergebnisse benutzt werden.

Wir hatten die Ziffern $0, S0, SS0, \ldots$ als spezielle geschlossene Terme von $L(Z)$ in 1.1 eingeführt. In 2.2 hatten wir die Ziffern mit den natürlichen Zahlen, den Elementen des Standardmodells $\mathcal{N}$, identifiziert. Dann ist

$$0_{\mathcal{N}} = 0 \text{ und } S_{\mathcal{N}} : k \mapsto Sk \text{ (für } k \in \mathbb{N}).$$

14.1.5 Lemma Für beliebige Ziffern $k, l, m \in \mathbb{N}$ gilt:

1. Aus $\mathcal{N} \models k = l$ folgt $Z^- \vdash k = l$
2. Aus $\mathcal{N} \models k + l = m$ folgt $Z^- \vdash k + l = m$
3. Aus $\mathcal{N} \models k \cdot l = m$ folgt $Z^- \vdash k \cdot l = m$
4. Aus $\mathcal{N} \models k < l$ folgt $Z^- \vdash k < l$
5. Aus $\mathcal{N} \models \neg k = l$ folgt $Z^- \vdash \neg k = l$.

Beweis.

1. Sind k, l dieselbe Ziffer, so gilt $k = l$ logisch, also auch in Z^-.
2. Wir induzieren nach der Länge der Ziffer l. Ist $l = 0$ und gilt $k+0 = m \in \mathcal{N}$, so ist $k \equiv m$, und $k+0 = m$ folgt in Z^- aus dem Axiom $\forall x\, x+0 = x$. Ist $l = Sl_0$ und gilt $k + Sl_0 = m$ in $\mathcal{N}$, so ist m der Nachfolger Sm_0 von $m_0 = k + l_0$. Nach IV gilt dann $k + l_0 = m_0$ in Z^-, und aus dem Axiom $\forall x \forall y\, x + Sy = S(x + y)$ folgt $k + Sl_0 = S(k + l_0) = Sm_0$ in Z^-.
3. wird entsprechend unter Rückgriff auf die Rekursionsgleichungen für die Multiplikation und 2. bewiesen.
4. Ist in $\mathcal{N}$ $k < l$, so gibt es ein $n \in \mathbb{N}$, so dass $k + Sn = l$ ist. Wegen 2. gilt dann $k + Sn = l$ in Z^-, also auch $\exists x\, k + Sx = l$, und das ist $k < l$.

5. Ist $k \neq l$, so ist $k < l$ oder $l < k$. Wegen 4. gilt dann $k < l$ oder $l < k$ in Z^-, also wegen des Axioms $\forall x \neg x < x$ auch $\neg k = l$.

14.1.6 Lemma Für jede Ziffer k gilt:

$$Z^- \vdash a < k \to a = 0 \vee a = S0 \vee \ldots \vee a = k - 1.$$

(Dabei ist $\bot$ für die leere Disjunktion (im Fall $k = 0$) zu lesen.)

Beweis durch Induktion nach der Länge der Ziffer k.

1. $Z^- \vdash a < 0 \to \bot$, weil

$$Z^- \vdash S(a + c) = a + Sc = 0 \to \bot.$$

2. $Z^- \vdash a < Sk \to a \leq k$, weil

$$Z^- \vdash S(a + c) = a + Sc = Sk \to a + c = k.$$

$Z^- \vdash a \leq k \to a < k \vee a = k$ folgt aus 14.1.3, 5. Die IV ergibt

$$Z^- \vdash a < k \vee a = k \to a = 0 \vee \ldots \vee a = k - 1 \vee a = k,$$

und mit zwei Kettenschlüssen folgt die Behauptung für Sk.

Aus 1. und 2. folgt mit Induktion nach k das Lemma.

14.1.7 Satz Sei $\mathcal{A} \models Z^-$. Dann ist die Interpretation $\mathcal{A}$ eine Einbettung von $\mathcal{N}$ in $\mathcal{A}$, und $\mathcal{A}[\mathbb{N}]$ ist ein Abschnitt der linearen Ordnung $(|\mathcal{A}|, <_{\mathcal{A}})$.

(Die Definition des Abschnitts 12.1.10 wird hier allgemein auf lineare Ordnungen bezogen.)

Beweis. Sei $\mathcal{A} \models Z^-$. Die Interpretation $\mathcal{A}$ ordnet jedem geschlossenen Term von $L(Z^-)$, insbesondere also jeder Ziffer $k \equiv S^k 0$ ein Element $\mathcal{A}(k) \in |\mathcal{A}|$ zu. Da wir die Ziffern k mit den natürlichen Zahlen $k \in \mathbb{N} = |\mathcal{N}|$ identifizieren (vgl. 14.1.5, 1), ist die Interpretation $\mathcal{A}$ (beschränkt auf $\mathbb{N}$) eine Abbildung von $\mathbb{N}$ in $|\mathcal{A}|$.

Weil $\mathcal{A}$ ein Modell von Z^- ist, gilt in $\mathcal{A}$ alles, was in Z^- gilt. Also ist diese Abbildung $\mathcal{A} \restriction \mathbb{N}$ nach 14.1.5, 5 injektiv, und nach 14.1.5, 2 bis 4 ist sie ein Homomorphismus von $\mathcal{N}$ in $\mathcal{A}$, sogar von $(\mathcal{N}, <) = (\mathbb{N}; 0, S, +, \cdot, <)$ in die

Expansion $(\mathcal{A}, <_{\mathcal{A}})$. Nach 14.1.3, 1 bis 3 ist ebenso wie $(\mathbb{N}, <)$ auch $(|\mathcal{A}|, <_{\mathcal{A}})$ eine lineare Ordnung. Dann ist der injektive Homomorphismus $\mathcal{A} \restriction \mathbb{N}$ nach 11.2.5 und nach 11.4.4 eine Einbettung von $(\mathcal{N}, <)$ in $(\mathcal{A}, <_{\mathcal{A}})$.

Nun gilt auch die Aussage von 14.1.6 in $\mathcal{A}$. Ist k eine Ziffer, so gilt deshalb für jedes $e \in |\mathcal{A}|$: Ist $e <_{\mathcal{A}} \mathcal{A}(k)$, so ist $e = \mathcal{A}(i)$ für eine Ziffer $i < k$. Alle $e \notin \mathcal{A}[\mathbb{N}]$ sind also in $\mathcal{A}$ größer als alle $\mathcal{A}(k)$, so dass $\mathcal{A}[\mathbb{N}]$ ein Abschnitt von $(|\mathcal{A}|, <_{\mathcal{A}})$ ist. Damit ist der Satz bewiesen.

Dieses allgemeine, instruktive Ergebnis über Modelle von Z^- legt folgende Terminologie nahe.

14.1.8 Definition Die Struktur $\mathcal{N} = (\mathbb{N}; 0, S, +, \cdot)$ ist das *Standardmodell* der Zahlentheorie. Jedes Modell von Z^-, das nicht isomorph zu $\mathcal{N}$ ist, nennt man ein *Nicht-Standard-Modell* (von Z^-). Ist $\mathcal{A} \models Z^-$, so heißen die Elemente von $\mathcal{A}[\mathbb{N}]$ die *Standardzahlen* von $\mathcal{A}$, und die Elemente von $|\mathcal{A}| - \mathcal{A}[\mathbb{N}]$ heißen die *Nicht-Standard-Zahlen* von $\mathcal{A}$.

Sind $\mathcal{A}, \mathcal{B} \models Z^-$ und ist $\mathcal{A} \subseteq \mathcal{B}$, so heißt $\mathcal{B}$ *Enderweiterung* von $\mathcal{A}$, $\mathcal{A}$ heißt *Anfangsabschnitt* von $\mathcal{B}$, man schreibt $\mathcal{A} \subseteq_e \mathcal{B}$, wenn $|\mathcal{A}|$ ein Abschnitt von $(|\mathcal{B}|, <_{\mathcal{B}})$ ist, wenn also aus $c <_{\mathcal{B}} d$ und $d \in |\mathcal{A}|$ stets $c \in |\mathcal{A}|$ folgt.

14.1.9 Lemma Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so ist auch

$$(|\mathcal{A}|, <_{\mathcal{A}}) \subseteq (|\mathcal{B}|, <_{\mathcal{B}}),$$

d. h. für $c, d \in |\mathcal{A}|$ ist $c <_{\mathcal{A}} d \Leftrightarrow c <_{\mathcal{B}} d$.

Beweis.

1. Ist $c <_{\mathcal{A}} d$, so ist $c +_{\mathcal{A}} S_{\mathcal{A}} c' = d$ für ein $c' \in |\mathcal{A}| \subseteq |\mathcal{B}|$. Wegen $\mathcal{A} \subseteq \mathcal{B}$ folgt dann $c <_{\mathcal{B}} d$.

2. Sei $c <_{\mathcal{B}} d$, also $c +_{\mathcal{B}} S_{\mathcal{B}} c' = d$ für ein $c' \in |\mathcal{B}|$. Weil $\mathcal{B}$ Modell von Z^- ist, ist
$$c' +_{\mathcal{B}} S_{\mathcal{B}} c = c +_{\mathcal{B}} S_{\mathcal{B}} c' = d, \text{ also } c' <_{\mathcal{B}} d.$$
Wegen $\mathcal{A} \subseteq_e \mathcal{B}$ ist dann auch $c' \in |\mathcal{A}|$, und es folgt $c <_{\mathcal{A}} d$.

Bemerkung. Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so ist hiernach

$$<_{\mathcal{A}} = <_{\mathcal{B}} \cap |\mathcal{A}|^2 = <_{\mathcal{B}} \cap (|\mathcal{B}| \times |\mathcal{A}|).$$

Nach Satz 14.1.7 ist in jedem Modell von Z^- jede Nicht-Standard-Zahl größer als jede Standardzahl. Insgesamt erhält der Satz mit diesen Begriffen folgende Fassung:

14.1.10 Satz Jedes Modell von Z^- ist (bis auf Isomorphie) eine Enderweiterung des Standardmodells $\mathcal{N}$.

Im Spezialfall, dass die Interpretation $\mathcal{A}\restriction\mathbb{N}$ surjektiv ist, ergibt sich:

14.1.11 Korollar Ein Modell $\mathcal{A}$ von Z^- ist genau dann isomorph zum Standardmodell $\mathcal{N}$, wenn $|\mathcal{A}| = \mathcal{A}[\mathbb{N}]$ ist.

Denn in genau dem Fall ist die Einbettung $\mathcal{A}\restriction\mathbb{N}$ surjektiv, also ein Isomorphismus.

Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so ist nach 11.3.3 jedenfalls $id : \mathcal{A} \xrightarrow{qf} \mathcal{B}$. Setzt man $\mathcal{A} = \mathcal{N}$, so folgt wegen 14.1.10: Alle quantorenfreien, sogar alle $\exists$-Sätze, die in $\mathcal{N}$ gelten, gelten in jedem Modell von Z^-. Weil $\mathcal{B}$ sogar *End*erweiterung von $\mathcal{N}$ ist, lässt sich dieses Ergebnis hinsichtlich der $<$-Relation verschärfen.

14.1.12 Definition Wir schreiben abkürzend

$$\begin{array}{lll} \forall x < t\ F(x) & \text{für} & \forall x(x < t \to F(x)) \\ \exists x < t\ F(x) & \text{für} & \exists x(x < t \wedge F(x)) \end{array}$$

und nennen $\forall x < t$ und $\exists x < t$ *beschränkte Quantoren*. Formeln aus $L(Z)$, in denen nur beschränkte Quantoren auftreten, heißen *beschränkte Formeln* oder Δ_0-*Formeln*. Σ_1-*Formeln* sind die Formeln $\exists x_1 \ldots \exists x_n F(x_1, \ldots, x_n)$, deren Kern $F(a_1, \ldots, a_n)$ Δ_0 ist.

Bemerkungen.

1. Im Rahmen von $L(Z)$ sind die Δ_0-Formeln eine naheliegende Verallgemeinerung der quantorenfreien Formeln, ebenso die Σ_1-Formeln eine Verallgemeinerung der $\exists$-Formeln.

2. $\exists x < t\ F(x)$ ist die Formel $\neg\forall x\neg\neg(x < t \to \neg F(x))$, also bis auf eine doppelte Negation die Formel $\neg\forall x < t\neg F(x)$.

3. Wir interessieren uns für diese Formeln und Formelklassen nur im Rahmen der Theorie Z^-. Die Formel $a < b$, die an sich die Σ_1-Formel $\exists x\ a + Sx = b$ ist, ist in Z^- äquivalent zu $\exists x < b\ a + Sx = b$ und damit eine Δ_0-Formel. Ebenso kann man in Z^- $\forall x \leq t$ und $\exists x \leq t$ als $\forall x < St$ bzw. $\exists x < St$ lesen.

14.1.13 Lemma Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so gilt

$$\mathcal{A}(C) = w \Leftrightarrow \mathcal{B}(C) = w$$

für alle $\Delta_0(\mathcal{A})$-Sätze C, also auch $id : \mathcal{A} \xrightarrow{\Delta_0} \mathcal{B}$.

Beweis durch Induktion nach C. Für Primformeln und Implikationen übernehmen wir die Induktionsschritte 1. bis 4. aus 11.2.6. Zu ergänzen bleibt nur:

5. C sei $\forall x < t\ F(x)$. Es ist $\mathcal{A}(t) = \mathcal{B}(t)$ nach 11.1.5 und $\mathcal{A}(F(d)) = \mathcal{B}(F(d))$ für $d \in |\mathcal{A}|$ nach IV. Es folgt:

   $$\mathcal{A}(C) = w \Leftrightarrow \text{ für alle } d \in |\mathcal{A}| \text{ mit } d <_{\mathcal{A}} \mathcal{A}(t) \text{ ist } \mathcal{A}(F(d)) = w.$$

   Für diese d ist $d <_{\mathcal{A}} \mathcal{A}(t)$ nach 14.1.9 äquivalent zu $d <_{\mathcal{B}} \mathcal{A}(t)$, und das sind wegen $\mathcal{A} \subseteq_e \mathcal{B}$ auch die sämtlichen $d \in |\mathcal{B}|$ mit $d <_{\mathcal{B}} \mathcal{A}(t)$. Also folgt:

   $$\begin{aligned} \mathcal{A}(C) = w \Leftrightarrow\ & \text{für alle } d \in |\mathcal{B}| \text{ mit } d <_{\mathcal{B}} \mathcal{A}(t) = \mathcal{B}(t) \text{ ist} \\ & \mathcal{A}(F(d)) = \mathcal{B}(F(d)) = w \\ \Leftrightarrow\ & \mathcal{B}(C) = w. \end{aligned}$$

Mit Induktion nach C folgt das Lemma. Zum Wahrheitstransport $id : \mathcal{A} \xrightarrow{\Delta_0} \mathcal{B}$ vgl. 11.3.1.

Mit demselben Argument, mit dem 11.2.7 aus 11.2.6 folgt, ergibt sich hieraus:

14.1.14 Satz Ist $\mathcal{A} \subseteq_e \mathcal{B}$, so ist:

$$id : \mathcal{A} \xrightarrow{\Sigma_1} \mathcal{B}.$$

Die beiden Sätze 14.1.7 und 14.1.14 liefern nun unmittelbar:

14.1.15 Satz Für $\mathcal{A} \models Z^-$ ist

$$\mathcal{A}\restriction\mathbb{N} : \mathcal{N} \xrightarrow{\Sigma_1} \mathcal{A} :$$

In jedem Modell von Z^- gelten alle in $\mathcal{N}$ wahren Σ_1-Sätze.

Beweis. Für jedes $\mathcal{A} \models Z^-$ ist $\mathcal{A}\restriction\mathbb{N}$ ein Isomorphismus von $\mathcal{N}$ auf einen Anfangsabschnitt $\mathcal{N}'$ von $\mathcal{A}$. Dann ist

$$\mathcal{A}\restriction\mathbb{N} : \mathcal{N} \xrightarrow{L(Z)} \mathcal{N}' \text{ und } id : \mathcal{N}' \xrightarrow{\Sigma_1} \mathcal{A}$$

nach 11.2.11 bzw. 14.1.14. Daraus folgt die Behauptung.

Aus dem 1. Gödelschen Unvollständigkeitssatz ergibt sich: In geeigneten Modellen von Z^- und ebenso von Z können *mehr* Σ_1-Sätze gelten als in $\mathcal{N}$. Diese Tatsache spielt im Folgenden aber keine Rolle.

14.2 Existenz und Anzahl von abzählbaren Nicht-Standard-Modellen

Alle überabzählbaren Modelle von Z^- sind notwendig Nicht-Standard-Modelle. Nach dem Satz von Löwenheim-Skolem-Tarski besitzt daher Z^-, sogar $Th(\mathcal{N})$ Nicht-Standard-Modelle jeder überabzählbaren Mächtigkeit. Gibt es auch abzählbare Nicht-Standard-Modelle?

14.2.1 Lemma Sei e eine von 0 verschiedene Konstante. Dann besitzt $Th(\mathcal{N}) + \{e\}$ abzählbare Modelle, in denen $\neg e = k$ für jede Ziffer k gilt.

Beweis. Sei T' eine endlich axiomatisierte Teiltheorie von

$$T^e := Th(\mathcal{N}) + \{e\} + \{\neg e = k | k \text{ Ziffer}\}.$$

Da T' nur endlich viele Axiome $\neg e = k$ enthält, gibt es eine Ziffer m, die größer ist als alle diese Ziffern k. Nun sei $\mathcal{N}_m$ die Expansion von $\mathcal{N}$ zu der Struktur zu $L(T^e) = L(Z) + \{e\}$, in der $\mathcal{N}_m(e) = m$ ist. Wegen $m > k$ ist

$$\mathcal{N}_m(\neg e = k) = w$$

für die Axiome $\neg e = k$ von T'. Also ist $\mathcal{N}_m$ ein Modell von T'.

Also hat jede endlich axiomatisierte Teiltheorie von T^e ein Modell. Nach dem Kompaktheitssatz hat dann auch T^e ein Modell, und das ist unendlich. Nun ist T^e eine abzählbare Theorie. Nach dem Satz von Löwenheim und Skolem hat daher T^e ein abzählbares Modell, und das ist die Behauptung.

14.2.2 Satz Die Theorie von $\mathcal{N}$ ist nicht $\aleph_0$-kategorisch: $Th(\mathcal{N})$ besitzt abzählbare Nicht-Standard-Modelle.

Beweis. Nach 14.2.1 besitzt $Th(\mathcal{N}) + \{e\}$ ein abzählbares Modell $\mathcal{B}$, in dem $e_\mathcal{B} \neq \mathcal{B}(k)$ ist für jede Ziffer k. Nach 8.2.4 ist dann

$$\mathcal{A} := \mathcal{B}|L(Z)$$

ein abzählbares Modell von $Th(\mathcal{N})$, und es ist

$$e_\mathcal{B} \in |\mathcal{A}| \text{ und } e_\mathcal{B} \neq \mathcal{A}(k) \text{ für alle Ziffern } k,$$

auch wenn $e_\mathcal{B}$ nicht durch einen Term von $L(Z)$ bezeichnet wird. Nach 14.1.11 ist dann $\mathcal{A}$ nicht isomorph zu $\mathcal{N}$. Also ist $\mathcal{A}$ ein abzählbares Nicht-Standard-Modell von $Th(\mathcal{N})$.

Dieses Ergebnis gilt offenbar erst recht für die Teiltheorien Z und Z^- von $Th(\mathcal{N})$. Es wirft die Frage auf, wieviele abzählbare Nicht-Standard-Modelle diese Zahlentheorien haben, wobei man isomorphe Modelle nur einmal zählen wird. Eine allgemeine Abschätzung kann man leicht finden.

14.2.3 Lemma Jede abzählbare Theorie hat höchstens $2^{\aleph_0}$ paarweise nicht-isomorphe abzählbare Modelle.

Beweis. Jede (endliche oder unendliche) abzählbare Struktur ist offenbar isomorph zu einer Struktur $\mathcal{A}$, deren Individuenbereich $|\mathcal{A}|$ gleich $\mathbb{N}$ oder eine endliche Teilmenge von $\mathbb{N}$ ist.

Sei nun T eine abzählbare Theorie mit Sprache L und $\mathcal{A}$ ein solches Modell von T. Dann ist auch

$$L(\mathcal{A}) \subseteq L + \mathbb{N}$$

eine abzählbare Sprache, und das Diagramm $D(\mathcal{A})$ (vgl. 11.3.8) ist eine abzählbare Theorie. Das abzählbare Axiomensystem von $\mathcal{D}(\mathcal{A})$ legt nun $\mathcal{A}$ eindeutig fest, wie im Anschluss an 11.3.8 diskutiert. Also gibt es höchstens so viele paarweise nicht-isomorphe Modelle $\mathcal{A}$ von T, wie es Theorien $\mathcal{D}(\mathcal{A})$ mit $|\mathcal{A}| \subseteq \mathbb{N}$

gibt. Davon gibt es aber höchstens $2^{\aleph_0}$ viele, weil $Ax(\mathcal{D}(\mathcal{A}))$ stets eine Teilmenge der abzählbaren Menge der Sätze von $L + \mathbb{N}$ ist. Damit ist das Lemma bewiesen.

Das Lemma gibt eine Abschätzung im ungünstigsten Fall, die im Einzelfall sehr schlecht sein kann. Z.B. trifft es auch auf die Theorie DLO zu, die $\aleph_0$-kategorisch ist. Wie sieht die Abschätzung nun für $Th(\mathcal{N})$ aus? Wir greifen auf elementare Eigenschaften der natürlichen Zahlen zurück.

14.2.4 Definition $a|b$, lies: *a teilt b*, *a ist Teiler von b*, steht abkürzend für die Formel $\exists y\, a \cdot y = b$. $\mathrm{Prim}(a)$, lies: *a ist Primzahl*, steht abkürzend für

$$1 < a \wedge \forall x(x|a \to x = 1 \vee x = a).$$

14.2.5 Bemerkungen Wir betrachten diese Formeln in beliebigen Modellen $\mathcal{A}$ von $Th(\mathcal{N})$, die wir als Enderweiterungen von $\mathcal{N}$ auffassen. In diesen Modellen $\mathcal{A}$ gilt:

1. $1|a$: 1 teilt jede Zahl.
2. $a|b \wedge b|c \to a|c$: Die Teiler-Relation ist transitiv.
3. $\mathrm{Prim}(a) \wedge a < 10 \leftrightarrow a = 2 \vee a = 3 \vee a = 5 \vee a = 7$: Die kleinsten Primzahlen sind 2, 3, 5, 7,. . ..
4. $\mathrm{Prim}(a) \wedge a|b \cdot c \to a|b \vee a|c$: Teilt eine Primzahl ein Produkt von zwei Zahlen, so teilt sie einen der beiden Faktoren des Produkts. Hieraus ergibt sich auch die Eindeutigkeit der Primfaktorzerlegung.
5. $\forall x \exists y(x < y \wedge \mathrm{Prim}(y))$: Es gibt unendlich viele Primzahlen.
   Das gilt in $\mathcal{N}$. Denn für jede natürliche Zahl $n > 0$ lässt $n! + 1$ ($n!$, lies: *n Fakultät*, bezeichnet das Produkt $1 \cdot 2 \cdot \ldots \cdot n$ der n natürlichen Zahlen von 1 bis n) bei Teilung durch $2, 3, \ldots, n$ den Rest 1, ist also nicht durch natürliche Zahlen zwischen 2 und n teilbar. Der kleinste Teiler > 1 von $n! + 1$ ist also $> n$, und er ist wegen 2. eine Primzahl. Also gibt es zu jedem $n \in \mathbb{N}$ eine Primzahl $p > n$ in $\mathcal{N}$. Damit gilt die Behauptung in $\mathcal{N}$, ist also ein Axiom von $Th(\mathcal{N})$ und gilt daher auch in jedem Modell $\mathcal{A}$ von $Th(\mathcal{N})$.

Im Folgenden interessieren wir uns für Primzahlen nur noch im Standardmodell $\mathcal{N}$, für Teilbarkeit aber noch in beliebigen Modellen von $Th(\mathcal{N})$.

14.2.6 Definition Zu $i \in \mathbb{N}$ bezeichne p_i die i-te Primzahl $\in \mathbb{N}$. Es ist also $p_0 = 2$, und p_{i+1} ist die kleinste Primzahl $> p_i$. Sei $\mathcal{A}$ Modell von $Th(\mathcal{N})$ und $e \in |\mathcal{A}|$. Wir sagen, e *kodiert in* $\mathcal{A}$ die Teilmenge $X \subseteq \mathbb{N}$, wenn

$$i \in X \Leftrightarrow \mathcal{A} \models p_i | e.$$

Dabei setzen wir $\mathcal{A}$ nach 14.1.10 als Enderweiterung von $\mathcal{N}$ voraus.

Jede Standardzahl $e \in \mathcal{A}[\mathbb{N}](e > 0)$ kodiert offenbar eine endliche Teilmenge von $\mathbb{N}$. So kodiert 1 die leere Menge, die Zahlen 2, 4, 8, 16,... kodieren alle die Menge $\{0\}$, die Zahl 70 kodiert $\{0, 2, 3\}$. Nicht-Standard-Zahlen e können aber auch unendliche Teilmengen von $\mathbb{N}$ kodieren. Das ergibt sich wieder aus dem Kompaktheitssatz. Wir argumentieren ähnlich wie in 14.2.1.

14.2.7 Lemma Sei e eine von 0 verschiedene Konstante und X eine Teilmenge von $\mathbb{N}$. Dann besitzt $Th(\mathcal{N})+\{e\}$ abzählbare Modelle $\mathcal{B}$, in denen $e_{\mathcal{B}}$ die Menge X kodiert.

Beweis. Wir betrachten die abzählbare Theorie

$$T^e := Th(\mathcal{N}) + \{e\} + (\{p_i | e \mid i \in X\} \cup \{\neg p_j | e \mid j \in \mathbb{N} - X\}).$$

Sei T' eine endlich axiomatisierte Teiltheorie von T^e. Dann gibt es eine endliche Teilmenge $X_0 \subseteq X$, so dass

$$T' \prec T_0 := Th(\mathcal{N}) + \{e\} + (\{p_i | e \mid i \in X_0\} \cup \{\neg p_j | e \mid j \in \mathbb{N} - X\})$$

ist. Weil X_0 endlich und jedes $p_i \in \mathbb{N}$ ist, ist auch das Produkt

$$m := \prod_{i \in X_0} p_i \in \mathbb{N}.$$

Dieses $m \in \mathbb{N}$ kodiert wegen der Eindeutigkeit der Primfaktorzerlegung gerade die Menge X_0.

Nun sei $\mathcal{N}_m$ die Expansion von $\mathcal{N}$ zu der Struktur zu $L(T^e) = L(Z) + \{e\}$, in der $\mathcal{N}_m(e) = m$ ist. Weil m X_0 kodiert, ist $\mathcal{N}_m$ ein Modell von T_0 und daher erst recht von T'.

Also hat jede endlich axiomatisierte Teiltheorie von T^e ein Modell. Nach dem Kompaktheitssatz hat dann auch T^e ein Modell, und das ist unendlich. Nun ist T^e eine abzählbare Theorie. Nach dem Satz von Löwenheim und Skolem hat daher T^e abzählbare Modelle $\mathcal{B}$, und in allen Modellen $\mathcal{B}$ von T^e kodiert $e_{\mathcal{B}}$ gerade die Menge X.

14.2.8 Satz Zu jeder Teilmenge $X \subseteq \mathbb{N}$ gibt es ein abzählbares Modell $\mathcal{A}$ von $Th(\mathcal{N})$, in dem ein Element von $|\mathcal{A}|$ die Menge X kodiert.

Beweis. Nach 14.2.7 besitzt $Th(\mathcal{N}) + \{e\}$ ein abzählbares Modell $\mathcal{B}$, in dem $e_{\mathcal{B}}$ die Menge X kodiert. Nach 8.2.4 ist dann

$$\mathcal{A} := \mathcal{B}|L(Z)$$

ein abzählbares Modell von $Th(\mathcal{N})$, dessen Element $e_{\mathcal{B}} \in |\mathcal{A}|$ die Menge X kodiert.

Jedes Element eines Modells von $Th(\mathcal{N})$ kodiert genau eine Teilmenge von $\mathbb{N}$. Daher können die abzählbar unendlich vielen Elemente eines abzählbaren Modells von $Th(\mathcal{N})$ auch nur abzählbar viele Teilmengen von $\mathbb{N}$ kodieren. Das ergibt im Umkehrschluss:

14.2.9 Satz Die Theorie $Th(\mathcal{N})$ hat genau $2^{\aleph_0}$ paarweise nicht-isomorphe abzählbare Modelle.

Beweis. Da $Th(\mathcal{N})$ abzählbar ist, kann $Th(\mathcal{N})$ nach 14.2.3 höchstens $2^{\aleph_0}$ nicht-isomorphe abzählbare Modelle haben.

Angenommen, $Th(\mathcal{N})$ hätte nur $\kappa < 2^{\aleph_0}$ nicht-isomorphe abzählbare Modelle. Da in isomorphen Modellen offenbar dieselben Teilmengen von $\mathbb{N}$ kodiert werden und von den Elementen eines einzelnen abzählbaren Modells höchstens $\aleph_0$ Teilmengen von $\mathbb{N}$ kodiert werden, würden dann nach den Sätzen von Hessenberg und Cantor insgesamt höchstens

$$\kappa \cdot \aleph_0 = \max(\kappa, \aleph_0) < 2^{\aleph_0}$$

Teilmengen von $\mathbb{N}$ von Elementen beliebiger abzählbarer Modelle von $Th(\mathcal{N})$ kodiert. Das widerspricht 14.2.8, weil es $2^{\aleph_0}$ Teilmengen von $\mathbb{N}$ gibt. Also ist die Annahme falsch und der Satz bewiesen.

Im schroffen Gegensatz zur Theorie DLO nutzt $Th(\mathcal{N})$ den Rahmen des Lemmas 14.2.3 voll aus und hat die für abzählbare Theorien maximale Anzahl von abzählbaren Modellen (modulo Isomorphie). Der Satz 14.2.9 gilt offenbar auch für alle Teiltheorien von $Th(\mathcal{N})$, speziell für Z und Z^-.

Für die Ergebnisse 14.2.7 und 14.2.8 spielt die Abzählbarkeit der Modelle $\mathcal{B}$ bzw. $\mathcal{A}$ keine Rolle. Im Beweis von 14.2.7 braucht man statt des Satzes von

Löwenheim-Skolem nur den Satz von Löwenheim-Skolem-Tarski auf die Theorie T^e anzuwenden, und man erhält Modelle $\mathcal{B}$ von T^e und $\mathcal{A} = \mathcal{B}|L(Z)$ jeder unendlichen Mächtigkeit λ, in denen ein Element die gegebene Teilmenge X von $\mathbb{N}$ kodiert.

Wesentlich ist die Abzählbarkeit der Modelle für das Endergebnis 14.2.9. Jedenfalls muss ihre Mächtigkeit $\lambda < 2^{\aleph_0}$ sein, um die Abschätzung $\kappa \cdot \lambda < 2^{\aleph_0}$ im Beweis von 14.2.9 zu ermöglichen. Tatsächlich hat $Th(\mathcal{N})$ ein Modell der Mächtigkeit $2^{\aleph_0}$, dessen Elemente alle Teilmengen von $\mathbb{N}$ kodieren, wie man sich analog zum Beweis von 14.2.7 überlegt. In der Situation kann man aus der Kodierbarkeit der Teilmengen von $\mathbb{N}$ nicht mehr auf die Existenz mehrerer nicht-isomorpher Modelle von $Th(\mathcal{N})$ der Mächtigkeit $2^{\aleph_0}$ schließen.

14.3 Overspill

Jedes Modell $\mathcal{A}$ von Z^- ist durch $<_{\mathcal{A}}$ linear geordnet. Zu jedem Element $e \in |\mathcal{A}|$ bilden die Zahlen

$$(1)\quad e, e +_{\mathcal{A}} 1, e +_{\mathcal{A}} 2, \ldots, e +_{\mathcal{A}} n, \ldots \qquad (n \in \mathbb{N})$$

eine zu $(\mathbb{N}, <)$ isomorphe Folge von aufeinanderfolgenden Elementen von $|\mathcal{A}|$. Das ergibt sich aus 14.1.3, wenn man $\mathcal{A}$ gemäß 14.1.10 als Enderweiterung von $\mathcal{N}$ betrachtet. Ist e eine Nicht-Standard-Zahl, so gilt in $\mathcal{A}$

$$e + n < e + e \text{ für alle } n \in \mathbb{N},$$

und die Folge (1) definiert einen echten Abschnitt von $(|\mathcal{A}|, <_{\mathcal{A}})$, der gegen die Nachfolgerfunktion $S_{\mathcal{A}}$ abgeschlossen ist. Solche Abschnitte sind spezielle *Schnitte in* $\mathcal{A}$.

14.3.1 Definition Sei $\mathcal{A} \models Z^-$. Ein Abschnitt I von $(|\mathcal{A}|, <_{\mathcal{A}})$ ist ein *Schnitt in* $\mathcal{A}$, wenn I nicht leer ist und mit einer Zahl e stets deren Nachfolger $S_{\mathcal{A}}(e)$ enthält. Ein Schnitt in $\mathcal{A}$ ist *echt*, wenn er $\neq |\mathcal{A}|$ ist, also wenn er ein echter Abschnitt ist.

Die Schnitte in $\mathcal{A}$ sind gerade die Abschnitte von $\mathcal{A}$, die kein größtes Element enthalten. Dieser Begriff von Schnitt entspricht also dem des Dedekindschen Schnitts in $\mathbb{Q}$ oder $\mathbb{R}$. Mit den syntaktischen Schnitten, von denen die Schnittregel handelt, hat er offenbar nichts zu tun.

Wie kann es in Modellen von Z – in denen das Induktionsschema gilt – überhaupt echte Schnitte geben? Jede Eigenschaft, die auf die Null zutrifft und mit jedem $e \in |\mathcal{A}|$, auf das sie zutrifft, auch auf dessen Nachfolger $S_{\mathcal{A}}(e)$ zutrifft, trifft doch auf alle $e \in |\mathcal{A}|$ zu – aber eben nur, wenn diese Eigenschaft durch eine arithmetische Formel in $\mathcal{A}$, also durch eine Formel von $L(Z)(\mathcal{A})$ definiert ist.

14.3.2 Definition Sei $\mathcal{A} \models Z^-$. Eine Teilmenge $X \subseteq |\mathcal{A}|$ heißt *definierbar in* $\mathcal{A}$, wenn es eine Formel $\mathcal{F}(a)$ aus $L(Z)(\mathcal{A})$ gibt ($\mathcal{F}$ geschlossen), so dass

$$X = \{e \in |\mathcal{A}| \mid \mathcal{A} \models \mathcal{F}(e)\}.$$

Beispiele.

1. Jeder Abschnitt von $(|\mathcal{A}|, <_{\mathcal{A}})$ *mit* einem größten Element, allgemeiner jedes abgeschlossene Intervall $[d_0, d_1]$ von $\mathcal{A}$ ist definierbar in $\mathcal{A}$ wegen

$$[d_0, d_1] = \{e \in |\mathcal{A}| \mid \mathcal{A} \models d_0 \leq e \wedge e \leq d_1\}.$$

2. Die Menge der Teiler einer beliebigen Zahl $d \in |\mathcal{A}|$ ist definierbar in $\mathcal{A}$ wegen

$$\{e | e \text{ teilt } d\} = \{e \in |\mathcal{A}| \mid \mathcal{A} \models \exists y\, e \cdot y = d\}.$$

3. Auch die Menge aller Primzahlen von $\mathcal{A}$ ist definierbar in $\mathcal{A}$, weil $\mathrm{Prim}(a)$ (vgl. 14.2.4) eine Formel von $L(Z)$ ist.

Nun folgt sofort:

14.3.3 Satz Kein echter Schnitt in einem Modell $\mathcal{A}$ von Z ist definierbar in $\mathcal{A}$.

Beweis. Sei $\mathcal{A} \models Z$ and I ein Schnitt in $\mathcal{A}$, der durch die Formel $F(a)$ aus $L(Z)(\mathcal{A})$ definiert ist, also

$$\mathcal{A} \models F(e) \Leftrightarrow e \in I.$$

Dann folgt $\mathcal{A} \models F(0)$, weil I ein nicht-leerer Abschnitt ist und 0 das kleinste Element von $\mathcal{A}$ ist, weiter

$$\mathcal{A} \models \forall x(F(x) \to F(Sx)),$$

weil I ein Schnitt in $\mathcal{A}$ ist, schließlich

$$\mathcal{A} \models F(0) \to \forall x(F(x) \to F(Sx)) \to \forall x F(x),$$

weil $\mathcal{A} \models Z$ ist. Daraus folgt

$$\mathcal{A} \models \forall x F(x), \text{ also } I = |\mathcal{A}|,$$

und I ist kein echter Schnitt. Damit ist der Satz bewiesen.

14.3.4 Korollar Ist $\mathcal{A}$ ein Nicht-Standard-Modell von Z und Enderweiterung von $\mathcal{N}$, so ist $\mathbb{N}$ in $\mathcal{A}$ nicht definierbar.

Das ist der Spezialfall von 14.3.3, in dem der echte Schnitt $I = \mathbb{N}$ ist.

„Wissen“ die Nicht-Standard-Modelle von Z also überhaupt nicht, dass sie echte Schnitte besitzen und dass sie echte Enderweiterungen von $\mathcal{N}$ sind? Ja, wenn man unter diesem „Wissen“ die Fähigkeit des Modells versteht, den Sachverhalt in seiner Sprache zu formulieren. Können die Nicht-Standard-Modelle von Z überhaupt „wissen“, dass sie sich vom Standardmodell $\mathcal{N}$ unterscheiden? Soweit sie Modelle von $Th(\mathcal{N})$ sind, sicher nicht. Aber Z ist nach dem 1. Gödelschen Unvollständigkeitssatz eine unvollständige Theorie. Es gibt also Modelle von Z, die nicht Modelle von $Th(\mathcal{N})$ sind. Diese Modelle sind notwendig Nicht-Standard-Modelle von Z, sie halten andere Aussagen für wahr als $\mathcal{N}$, möglicherweise mehr Σ_1-Sätze. Aber dass sie deswegen Nicht-Standard-Modelle sind, kann man nur von außen, im Rahmen einer stärkeren Sprache feststellen.

Den Satz 14.3.3 kann man positiv wenden:

14.3.5 Overspill-Lemma Sei I ein echter Schnitt in einem Modell $\mathcal{A}$ von Z. Sei $\mathcal{F}(a)$ eine Formel aus $L(Z)(\mathcal{A})$ ($\mathcal{F}$ geschlossen), so dass

$$\mathcal{A} \models \mathcal{F}(e) \quad \text{für alle } e \in I.$$

Dann gibt es ein $e^+ > I$ (d. h. $i <_{\mathcal{A}} e^+$ für alle $i \in I$), so dass

$$\mathcal{A} \models \forall x \leq e^+ \mathcal{F}(x). \tag{2}$$

Beweis. Angenommen, es gäbe kein $e^+ > I$ mit (2), so dass aus (2) also stets $e^+ \in I$ folgt. Dann gilt

$$e \in I \;\Leftrightarrow\; \mathcal{A} \models \forall x \leq e\, \mathcal{F}(x),$$

weil I ein Abschnitt ist. Dann wäre der echte Schnitt I in $\mathcal{A}$ durch $\forall x \leq a\, \mathcal{F}(x)$ definiert, im Widerspruch zu 14.3.3. Also ist die Annahme falsch, und das war zu zeigen.

Die Eigenschaft $\mathcal{F}$, genauer die Menge

$$\mathcal{A}(\mathcal{F}) := \{e \in |\mathcal{A}| \mid \mathcal{A} \models \mathcal{F}(e)\},$$

die den ganzen Schnitt I ausfüllt, „läuft über“; sie beschränkt sich nicht auf I, sondern sie füllt noch (mindestens) einen größeren Abschnitt $\mathcal{A}_{e^+}$ von $\mathcal{A}$ aus. Dieses „Überlaufen“ heißt auf Englisch *overspill*.

14.3.6 Korollar Ist $\mathcal{A}$ ein Nicht-Standard-Modell von Z und $\mathcal{F}(a)$ eine Formel aus $L(Z)(\mathcal{A})$ ($\mathcal{F}$ geschlossen), so dass

$$\mathcal{A} \models \mathcal{F}(k) \text{ für alle Standardzahlen } k,$$

so gibt es eine Nicht-Standard-Zahl e^+, für die (2) gilt.

Das ist der Spezialfall $I = \mathbb{N}$ des Overspill-Lemmas.

Beispiel. Oberhalb von jeder natürlichen Zahl $k \in \mathbb{N}$ gibt es Primzahlen, wie wir in 14.2.5 gesehen haben:

$$\mathcal{N} \models \exists y(y > k \wedge \mathrm{Prim}(y)) \quad \text{für jedes } k \in \mathbb{N}.$$

Weil $\mathrm{Prim}(a)$ sich als Δ_0-Formel schreiben lässt, gilt dies nach 14.1.15 auch für echte Enderweiterungen $\mathcal{A}$ von $\mathcal{N}$:

$$\mathcal{A} \models \exists y(y > k \wedge \mathrm{Prim}(y)).$$

Ist $\mathcal{A}$ ein Nicht-Standard-Modell von Z, so gibt es nach 14.3.6 eine Nicht-Standard-Zahl $e^+ \in |\mathcal{A}|$ mit

$$\mathcal{A} \models \forall x \leq e^+ \exists y(y > x \wedge \mathrm{Prim}(y)) :$$

Es gibt Nicht-Standard-Primzahlen. Es gilt allerdings weit mehr, weil man $\forall x \exists y(y > x \wedge \mathrm{Prim}(y))$ in Z beweisen kann. Dieser Satz gilt also – ohne Beschränkung auf e^+ – auch in dem Modell $\mathcal{A}$ von Z: Es gibt beliebig große Nicht-Standard-Primzahlen.

Das Beispiel lässt sich auch in der Richtung zu kleineren Nicht-Standard-Zahlen verschärfen. Dazu gibt man dem Overspill-Lemma eine andere Wendung.

14.3.7 Lemma Sei I ein echter Schnitt in einem Modell $\mathcal{A}$ von Z. Sei $\mathcal{F}(a)$ eine Formel aus $L(Z)(\mathcal{A})$ ($\mathcal{F}$ geschlossen), so dass es zu jedem $e \in I$ ein $d \in I$ gibt mit

$$\mathcal{A} \models e < d \wedge \mathcal{F}(d).$$

Dann gibt zu jedem $d^+ > I$ ein $e^+ > I$ mit

$$\mathcal{A} \models e^+ < d^+ \wedge \mathcal{F}(e^+) : \tag{3}$$

Wenn es in I beliebig große Zahlen gibt, auf die $\mathcal{F}$ zutrifft, dann gibt es oberhalb von I beliebig kleine Zahlen, auf die $\mathcal{F}$ zutrifft.

Beweis. Seien $e \in I$, $d^+ > I$ beliebig in $|\mathcal{A}|$. Die Voraussetzung über d ergibt dann

$$\mathcal{A} \models \exists y(e < y < d^+ \wedge \mathcal{F}(y)).$$

Nach dem Overspill-Lemma gibt es dann ein $e^+ > I$ mit

$$\mathcal{A} \models \exists y(e^+ < y < d^+ \wedge \mathcal{F}(y)),$$

und daraus folgt die Behauptung, wenn man das hier als existent erkannte y wieder mit e^+ bezeichnet.

Wir notieren hiervon wieder den Spezialfall $I = \mathbb{N}$.

14.3.8 Korollar Ist $\mathcal{A}$ ein Nicht-Standard-Modell von Z und $\mathcal{F}(a)$ eine Formel aus $L(Z)(\mathcal{A})$ ($\mathcal{F}$ geschlossen), so dass es zu jedem $k \in \mathbb{N}$ ein $l \in \mathbb{N}$ gibt mit

$$\mathcal{A} \models k < l \wedge \mathcal{F}(l),$$

dann gibt es zu jeder Nicht-Standard-Zahl d^+ eine Nicht-Standard-Zahl e^+, für die (3) gilt:

Wenn es beliebig große Standardzahlen gibt, auf die $\mathcal{F}$ zutrifft, gibt es auch beliebig kleine Nicht-Standard-Zahlen, auf die $\mathcal{F}$ zutrifft.

Beispiel. Es gibt beliebig kleine Nicht-Standard-Primzahlen – und das kann man offenbar nicht direkt in Z beweisen.

Dual zum *overspill* gibt es auch einen sogenannten *underspill*, der nichts anderes als das Überlaufen einer Eigenschaft von oberhalb eines Schnittes in den Schnitt hinein ist.

14.3.9 Lemma Sei I ein echter Schnitt in einem Modell $\mathcal{A}$ von Z und sei $\mathcal{F}(a)$ eine Formel aus $L(Z)(\mathcal{A})$ mit geschlossenem $\mathcal{F}$.

1. Wenn $\mathcal{A} \models \mathcal{F}(e)$ für alle $e > I$, so gibt es ein $e^- \in I$ mit

$$\mathcal{A} \models \forall x \geq e^- \mathcal{F}(x).$$

2. Wenn es zu jedem $e > I$ ein $d > I$ gibt mit

$$\mathcal{A} \models d < e \wedge \mathcal{F}(d),$$

dann gibt es zu jedem $d^- \in I$ ein $e^- \in I$ mit

$$\mathcal{A} \models d^- < e^- \wedge \mathcal{F}(e^-).$$

Diese Behauptung führt man leicht auf 14.3.7 und 14.3.5 zurück.

Die *overspill*- und *underspill*-Ergebnisse dieses Abschnitts gelten für beliebige Nicht-Standard-Modelle von Z, weil in diesen Modellen wegen der Induktionsaxiome von Z die echten Schnitte nicht definierbar sind. Noch einen überraschenden Schritt weiter kommt man, wenn man sich auf den Vergleich des Standardmodells $\mathcal{N}$ mit dazu elementar äquivalenten Nicht-Standard-Modellen, also mit Nicht-Standard-Modellen von $Th(\mathcal{N})$ konzentriert.

14.3.10 Satz Sei $\mathcal{A}$ ein Nicht-Standard-Modell von $Th(\mathcal{N})$ und $\mathcal{F}(a)$ eine Formel von $L(Z)$ ($\mathcal{F}$ geschlossen). Dann sind äquivalent:

(i) $\mathcal{A} \models \mathcal{F}(e)$ für eine Nicht-Standard-Zahl $e \in |\mathcal{A}|$

(ii) $\mathcal{N} \models \forall x \exists y (x < y \wedge \mathcal{F}(y))$.

Beweis von (i) $\Rightarrow$ (ii).
Wir suchen Folgerungen aus (i), die die Nicht-Standard-Zahl e nicht erwähnen, um die elementare Äquivalenz von $\mathcal{A}$ und $\mathcal{N}$ ausnutzen zu können

Sei n eine Ziffer. Dann gilt $n < e$ in $\mathcal{A}$ für alle Nicht-Standard-Zahlen $e \in |\mathcal{A}|$. Deshalb folgt aus (i)

$$\mathcal{A} \models \exists y(n < y \wedge \mathcal{F}(y)).$$

Das ist nun ein Satz aus $L(Z)$, der wegen $\mathcal{A} \models Th(\mathcal{N})$ auch in $\mathcal{N}$ für jede Ziffer $n \in \mathbb{N} = |\mathcal{N}|$ gilt. Also folgt (ii).

(ii) $\Rightarrow$ (i). Aus (ii) folgt wie gehabt

$$\mathcal{A} \models \forall x \exists y(x < y \wedge \mathcal{F}(y)).$$

Weil $\mathcal{A}$ Nicht-Standard-Modell ist, gibt es eine Nicht-Standard-Zahl $d \in |\mathcal{A}|$. Für diese gilt:

$$\mathcal{A} \models \exists y(d < y \wedge \mathcal{F}(y)).$$

Also gibt es ein $e \in |\mathcal{A}|$ mit $\mathcal{A} \models d < e \wedge \mathcal{F}(e)$. Dann ist auch e Nicht-Standard-Zahl, und es folgt (i).

Wir betonen, dass dieser Satz nicht für alle Nicht-Standard-Modelle von Z gilt, sondern nur für die von $Th(\mathcal{N})$. Nach dem 1. Gödelschen Unvollständigkeitssatz gibt es tatsächlich Nicht-Standard-Modelle $\mathcal{A}$ von Z, eine Formel $\mathcal{F}(a)$ mit geschlossenem $\mathcal{F}$ und eine Nicht-Standard-Zahl $e \in |\mathcal{A}|$, so dass $\mathcal{A} \models \mathcal{F}(e)$ und es keine einzige natürliche Zahl k gibt, für die $\mathcal{F}(k)$ in $\mathcal{N}$ oder in $\mathcal{A}$ gilt.

14.4 Anordnung in Nicht-Standard-Modellen

Jedes Modell von Z^- zerfällt nach 14.1.7 in einen zu $\mathcal{N}$ isomorphen Standard-Teil der endlichen Zahlen und einen Nicht-Standard-Teil der unendlich großen Zahlen. Im Standardmodell ist der Nicht-Standard-Teil leer. Wir untersuchen die Anordnung der Nicht-Standard-Zahlen genauer.

14.4.1 Definition Es sei $\mathcal{A}$ ein Modell von Z^-. Elemente $e, d \in |\mathcal{A}|$ heißen *von gleicher Größenordnung*, $e \sim f$, wenn in $\mathcal{A}$ zwischen e und f nur endlich viele Elemente von $|\mathcal{A}|$ liegen:

$$e \sim f :\Leftrightarrow \{z \in |\mathcal{A}| \mid \mathcal{A} \models (e < z \wedge z < f) \vee (f < z \wedge z < e)\} \text{ ist endlich.}$$

14.4.2 Lemma Die Relation $\sim$ ist eine Äquivalenzrelation auf $|\mathcal{A}|$.

Beweis. Zwischen e und e liegen 0 Elemente, also $e \sim e$.
Liegen k Elemente zwischen e und f und l Elemente zwischen e und g, so liegen höchstens $k + l + 1$ Elemente zwischen f und g, also $e \sim f \to e \sim g \to f \sim g$.

14.4.3 Definition Es sei $\mathcal{A}$ ein Modell von Z^- und $\sim$ die in 14.4.1 definierte Äquivalenzrelation auf $|\mathcal{A}|$. Dann bezeichne $[e]$ die Äquivalenzklasse von e bezüglich $\sim$. Wir nennen $[e]$ *die Größenordnung von* e in $\mathcal{A}$.

Nach der Bemerkung vor 7.2.7 ist $[e] = [f]$ genau dann, wenn $e \sim f$ ist, wenn also e und f von gleicher Größenordnung sind.

14.4.4 Lemma Für Elemente $e \neq f$ eines Modells $\mathcal{A}$ von Z gilt:

Zwischen e und f liegen k Zahlen $\Leftrightarrow$ In $\mathcal{A}$ gilt $f = e + Sk$ oder $e = f + Sk$.

Beweis durch Induktion nach k:
Gelte $e < f$ in $\mathcal{A}$, also auch $Se \leq f$.

1. Nichts liegt zwischen e und f $\Leftrightarrow$ $f = Se = e + 1$ nach 14.1.3, 6.
2. Zwischen e und f liegen $k + 1$ Zahlen

   $\Leftrightarrow$ zwischen Se und f liegen k Zahlen, nach 14.1.3, 6

   $\Leftrightarrow$ $f = Se + Sk = e + SSk$ nach IV und 14.1.1, 4.

Mit Induktion folgt aus 1. und 2. für $e < f$:

$$\text{Zwischen } e \text{ und } f \text{ liegen } k \text{ Zahlen} \Leftrightarrow \text{In } \mathcal{A} \text{ gilt } f = e + Sk.$$

Vertauscht man hierin e und f, so folgt insgesamt die Behauptung.

Bemerkung. Setzen wir

$$a -_{\mathcal{A}} k = \begin{cases} f, \text{ falls } f +_{\mathcal{A}} k = e \\ 0_{\mathcal{A}}, \text{ falls es kein solches } f \text{ in } \mathcal{A} \text{ gibt,} \end{cases}$$

so ergibt 14.4.4 gerade

$$[e] = \{e \pm_{\mathcal{A}} k | k \text{ ist Standardzahl in } \mathcal{A}\}.$$

Die Fallunterscheidung in der Definition von $e -_{\mathcal{A}} k$ trennt nun die Standardzahlen von den Nicht-Standard-Zahlen.

14.4.5 Lemma Ist $\mathcal{A}$ ein Modell von Z^-, so ist

$$([e], <_{\mathcal{A}}) \cong (\mathbb{N}, <) \quad \text{für Standardzahlen } e \text{ und}$$
$$([e], <_{\mathcal{A}}) \cong (\mathbb{Z}, <) \quad \text{für Nicht-Standard-Zahlen } e.$$

Beweis. Ist e eine Standardzahl von $\mathcal{A}$, so ist $[e] = \mathcal{A}[\mathbb{N}]$ die Menge der Standardzahlen nach 14.1.7. Nach 14.1.10 ist die Interpretation $\mathcal{A}$ auch ein Isomorphismus von $(\mathbb{N}, <)$ auf $(\mathcal{A}[\mathbb{N}], <_{\mathcal{A}})$.

Ist e eine Nicht-Standard-Zahl, so ist nach 14.1.1, 1 jedes Element von $[e]$ Nachfolger, und zwar ist $e -_{\mathcal{A}} k$ der Nachfolger von $e -_{\mathcal{A}} Sk$ wegen 14.1.1, 4. Zu jeder Standardzahl k gibt es daher genau ein $f \in [e]$ mit $f +_{\mathcal{A}} k = e$. Für $k \in \mathbb{Z}$, $k < 0$ identifizieren wir $-k$ mit der entsprechenden Standardzahl in $\mathcal{A}$ und setzen $e +_{\mathcal{A}} k := e -_{\mathcal{A}} (-k)$. Dann ist

$$\varphi : \mathbb{Z} \to [e], k \mapsto e +_{\mathcal{A}} k$$

eine Abbildung, die wegen 14.4.4 surjektiv, und wegen 14.1.3, 7 (für $0 \leq k$ oder $l < 0$) bzw. wegen 14.1.3, 2 (für $k < 0 \leq l$) ein Homomorphismus von $(\mathbb{Z}, <)$ auf $([e], <_{\mathcal{A}})$ ist. Da $<_{\mathcal{A}}$ auf $[e]$ eine lineare Ordnungsrelation ist, ist dann φ ein Isomorphismus (vgl. 11.4.4).

In der Anordnung $<_{\mathcal{A}}$ eines Modells $\mathcal{A}$ von Z^- kommen zuerst die Standardzahlen. Danach kommen die Nicht-Standard-Zahlen, gemäß 14.4.5 eingeteilt in Exemplare von $\mathbb{Z}$, ihre Größenordnungen. Wenn $\mathcal{A}$ das Standardmodell ist, kommen null Exemplare von $\mathbb{Z}$, sonst mehrere. Wir untersuchen die Anordnung dieser (unendlichen) Größenordnungen.

14.4.6 Definition Es sei $\mathcal{A}$ ein Modell von Z^-. Wir setzen

$$B := \{[e] | e \text{ ist Nicht-Standard-Zahl von } \mathcal{A}\}$$
$$[e] <_{\mathcal{B}} [f] :\Leftrightarrow e <_{\mathcal{A}} f \text{ und nicht } e \sim f.$$

Dann ist $\mathcal{B} := (B, <_{\mathcal{B}})$ die *Struktur der unendlichen Größenordnungen* von $\mathcal{A}$.

Bemerkung. $<_{\mathcal{B}}$ ist auf B wohldefiniert. Denn ist

$$e' \sim e <_{\mathcal{A}} f \sim f' \text{ und nicht } e \sim f,$$

dann gibt es unendlich viele Zahlen z mit

$$e <_{\mathcal{A}} z \text{ und } z <_{\mathcal{A}} f,$$

aber nur endlich viele Zahlen zwischen e' und e und zwischen f und f'. Dann gibt es immer noch unendlich viele z mit

$$e' <_{\mathcal{A}} z \text{ und } z <_{\mathcal{A}} f', \text{ also nicht } e' \sim f',$$

und es ist $e' <_{\mathcal{A}} f'$.

14.4.7 Satz Ist $\mathcal{A}$ ein Nicht-Standard-Modell von Z^-, so ist die Struktur $\mathcal{B} = (B, <_{\mathcal{B}})$ der unendlichen Größenordnungen von $\mathcal{A}$ ein Modell von DLO.

Beweis. B ist nicht leer, weil $\mathcal{A}$ als Nicht-Standard-Modell mindestens eine Nicht-Standard-Zahl erhält. Wir prüfen die Axiome LO 1 bis LO 6 aus 1.2.5 in $\mathcal{B}$ nach.

LO 1. Ist $[e] <_{\mathcal{B}} [f] <_{\mathcal{B}} [g]$, so gilt in $\mathcal{A}$ $e < f < g$, also $e < g$, und alle unendlich vielen Zahlen, die zwischen e und f oder zwischen f und g liegen, liegen auch zwischen e und g wegen 14.1.3,2. Also ist $[e] <_{\mathcal{B}} [g]$.

LO 2 gilt in $\mathcal{B}$ wegen $e \sim e$.

LO 3. Aus $e <_{\mathcal{A}} f$ folgt $[e] <_{\mathcal{B}} [f]$ oder $[e] = [f]$. Also gilt LO 3 in $\mathcal{B}$, weil es nach 14.1.3,3 in $\mathcal{A}$ gilt.
Damit ist $\mathcal{B}$ eine lineare Ordnung.

LO 6. Für jede Nicht-Standard-Zahl e ist $[e] <_{\mathcal{B}} [e +_{\mathcal{A}} e]$. Denn für alle Ziffern $k \neq 0$ gilt in $\mathcal{A}$ $0 < k < e$ nach 14.1.7, also nach 14.1.3,7 auch

$$e < e + k < e + e.$$

Also hat $\mathcal{B}$ kein letztes Element $[e]$.

LO 5. Zur Nicht-Standard-Zahl e sei d eine nach 14.1.1, 8 existierende Zahl, für die $d + d = e$ oder $d + d + 1 = e$ in $\mathcal{A}$ gilt. Dann ist auch d eine Nicht-Standard-Zahl, weil $\mathcal{A}[\mathbb{N}]$ nach 14.1.5,2 gegen Addition abgeschlossen ist. Wie unter LO 6 folgt, dass in $\mathcal{A}$ für alle Ziffern $k \neq 0$ gilt

$$d < d + k < d + d \leq e.$$

Also ist $[d] <_{\mathcal{B}} [e]$, und $\mathcal{B}$ hat kein erstes Element.

LO 4. e, f seien Nicht-Standard-Zahlen, und es sei $[e] <_{\mathcal{B}} [f]$. d sei eine nach 14.1.1,8 existierende Zahl, für die $d+d = e+f$ oder $d+d+1 = e+f$ in $\mathcal{A}$ gilt. Wäre $[d] \leq_{\mathcal{B}} [e]$, so würde $d \leq e+k$ in $\mathcal{A}$ für eine Standardzahl k gelten, also auch

$$e + f \leq d + d + 1 \leq e + e + k + k + 1 < e + f,$$

im Widerspruch zu 14.1.3,1. Also ist

$$[e] <_{\mathcal{B}} [d] \text{ und entsprechend } [d] <_{\mathcal{B}} [f],$$

und $\mathcal{B}$ ist eine dichte Ordnung.

Damit ist der Satz bewiesen.

Jedes Nicht-Standard-Modell $\mathcal{A}$ von Z^- bestimmt hiernach durch seine Kleiner-Relation ein Modell $\mathcal{B}$ von *DLO*, dessen Elemente die unendlichen Größenordnungen von $\mathcal{A}$ sind. Umgekehrt legt $\mathcal{B}$ zwar nicht das ganze Nicht-Standard-Modell $\mathcal{A}$ fest, wohl aber den Individuenbereich $|\mathcal{A}|$ und die Kleiner-Relation $<_{\mathcal{A}}$. Um dies deutlich zu machen, führen wir *Summe* und *Produkt* von linearen Ordnungen ein.

14.4.8 Definition $\mathcal{A} = (A, <_{\mathcal{A}})$ und $\mathcal{B} = (B, <_{\mathcal{B}})$ seien Modelle von *LO*. Dann ist $\mathcal{A} + \mathcal{B} := (A \cup B, <_+)$, falls $A \cap B$ leer ist, mit

$$\begin{aligned} a <_+ b :\Leftrightarrow\ & a \in A \text{ und } b \in B, \text{ oder} \\ & a, b \in A \text{ und } a <_{\mathcal{A}} b, \text{ oder} \\ & a, b \in B \text{ und } a <_{\mathcal{B}} b. \end{aligned}$$

$\mathcal{A} \times \mathcal{B} := (A \times B, <.)$ mit

$$\begin{aligned} (a, b) < .(a', b') :\Leftrightarrow\ & a <_{\mathcal{A}} a', \text{ oder} \\ & a = a' \text{ und } b <_{\mathcal{B}} b'. \end{aligned}$$

$\mathcal{A} + \mathcal{B}$ entsteht aus $\mathcal{A}$ und $\mathcal{B}$ einfach durch Hintereinanderlegen von $\mathcal{A}$ und $\mathcal{B}$: erst $\mathcal{A}$, dann $\mathcal{B}$. $\mathcal{A} \times \mathcal{B}$ entsteht durch lexikographische Anordnung von $A \times B$: die Paare (k, l) werden in erster Linie nach dem ersten Element gemäß $<_{\mathcal{A}}$ geordnet, und nur, wenn bei zwei Paaren die ersten Elemente übereinstimmen, richtet man sich nach dem zweiten Element gemäß $<_{\mathcal{B}}$. Mit $\mathcal{A}$ und $\mathcal{B}$ sind daher auch $\mathcal{A} + \mathcal{B}$ und $\mathcal{A} \times \mathcal{B}$ lineare Ordnungen.

14.4.9 Satz Zu jedem Nicht-Standard-Modell $\mathcal{A}$ von Z^- gibt es ein Modell $\mathcal{B}$ von *DLO* von derselben Mächtigkeit wie $\mathcal{A}$, so dass

$$(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + \mathcal{B} \times (\mathbb{Z}, <).$$

Beweis. Nach 14.1.7 ist $(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + (|\mathcal{A}| - \mathcal{A}[\mathbb{N}], <_{\mathcal{A}})$. Nach 14.4.5 zerfällt $|\mathcal{A}| - \mathcal{A}[\mathbb{N}]$ in Größenordnungen, die alle isomorph zu $(\mathbb{Z}, <)$ sind. Wählen wir aus jeder Größenordnung $[e]$ ein Element e_0 als Repräsentanten aus, so hat jede Nicht-Standard-Zahl e eine eindeutige Darstellung

$$e = e_0 +_{\mathcal{A}} k \text{ mit } e_0 \text{ Repräsentant von } [e], k \in \mathbb{Z}.$$

Dann ist die Abbildung

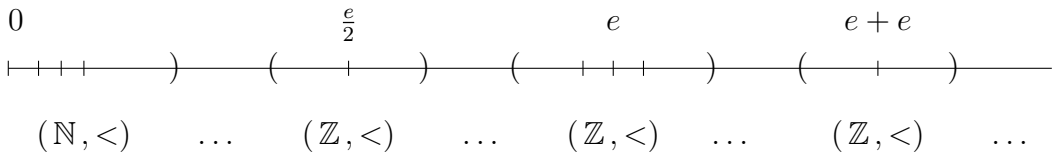
$$\varphi : |\mathcal{A}| - \mathcal{A}[\mathbb{N}] \to B \times \mathbb{Z}, e \mapsto ([e], k)$$

ein Isomorphismus von $(|\mathcal{A}| - \mathcal{A}[\mathbb{N}], <_{\mathcal{A}})$ auf $\mathcal{B} \times (\mathbb{Z}, <)$, wobei $\mathcal{B}$ die Struktur der unendlichen Größenordnungen aus 14.4.6 bezeichnet. Nach 14.4.7 ist $\mathcal{B}$ ein Modell von *DLO*.

Hat $\mathcal{B}$ die Mächtigkeit κ, so hat auch $\mathcal{A}$ die Mächtigkeit

$$\aleph_0 + \kappa \cdot \aleph_0 = \kappa,$$

weil $\kappa \geq \aleph_0$ ist (*DLO* hat nur unendliche Modelle). Damit sind alle Behauptungen bewiesen.



Die $\aleph_0$-Kategorizität von *DLO* 13.2.1 liefert zusammen mit diesem Satz eine überraschende Folgerung für abzählbare Modelle der Zahlentheorie.

14.4.10 Korollar Ist $\mathcal{A}$ ein abzählbares Modell von Z^-, so ist

entweder $\mathcal{A} \cong \mathcal{N}$
oder $(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + (\mathbb{Q}, <) \times (\mathbb{Z}, <)$.

Bezüglich der Kleiner-Relation hat Z^- bis auf Isomorphie genau zwei abzählbare Modelle.

Beweis. Entweder ist $|\mathcal{A}| = \mathcal{A}[\mathbb{N}]$; dann ist $\mathcal{A} \cong \mathcal{N}$ nach 14.1.11. Oder $\mathcal{A}$ ist ein Nicht-Standard-Modell; dann gilt 14.4.9 für ein abzählbares Modell $\mathcal{B}$ von DLO. Nach 13.2.1 ist $\mathcal{B} \cong (\mathbb{Q}, <)$, so dass 14.4.9 übergeht in

$$(|\mathcal{A}|, <_{\mathcal{A}}) \cong (\mathbb{N}, <) + (\mathbb{Q}, <) \times (\mathbb{Z}, <).$$

Dieser zweite Fall tritt wegen 14.2.2 auch tatsächlich ein.

Obwohl schon die vollständige (sogar saturierte) Theorie $Th(\mathcal{N})$ $2^{\aleph_0}$ paarweise nicht-isomorphe abzählbare Modelle hat, tragen alle diese Modelle, sogar alle abzählbaren Modelle von Z^- – außer dem Standardmodell – dieselbe leicht zu beschreibende, leicht zu veranschaulichende Ordnung

$$(\mathbb{N}, <) + (\mathbb{Q}, <) \times (\mathbb{Z}, <).$$

Die Unterschiede der verschiedenen Modelle liegen offenbar in tieferen Eigenschaften der Interpretationen von Addition und Multiplikation.

14.5 Aufgaben

14.5.1 Zeigen Sie 14.1.5, 3: $\mathcal{N} \models k \cdot l = m \Rightarrow Z^- \vdash k \cdot l = m$.

14.5.2 Zeigen Sie: $Z^- \vdash a + Sc = b \to c < b$ und folgern Sie:

$$Z^- \vdash a < b \leftrightarrow \exists x < b \;\; a + Sx = b.$$

14.5.3 Zeigen Sie: $Z^- \vdash b \cdot Sc = a \to b \leq a$ und folgern Sie:

$$Z^- \vdash b|a \wedge 0 < a \to b \leq a \qquad \text{und weiter:}$$
$$Z^- \vdash \mathrm{Prim}(a) \leftrightarrow 1 < a \wedge \forall x < a(x|a \to x = 1).$$

14.5.4 Zeigen Sie für Δ_0-Formeln C (u. U. mit freien Variablen) und alle Modelle $\mathcal{A}$ von Z^-:

$$\mathcal{A} \models C \Rightarrow \mathcal{N} \models C.$$

14.5.5 Aus welchem Gesetz für die Multiplikation (das nicht zu den Axiomen von Z^- gehört) folgt unmittelbar $a|b \wedge b|c \to a|c$?

14.5.6 Zeigen Sie: $Z \vdash \forall x(\forall y < x\, \mathcal{F}(y) \to \mathcal{F}(x)) \to \forall x\, \mathcal{F}(x)$.
Hinweis: Wenden Sie (Ind) auf die Formel $\forall y < a\, \mathcal{F}(y)$ an.

14.5.7 Zeigen Sie: $Th(\mathcal{N})$ hat ein Modell $\mathcal{A}$ der Mächtigkeit $2^{\aleph_0}$, in dem es zu jeder Teilmenge $X \subseteq \mathbb{N}$ eine Zahl e_X gibt, die X kodiert.

14.5.8 Beweisen Sie die Aussagen 1. und 2. des Underspill-Lemmas 14.3.9.

14.5.9 Zeigen Sie:

a. Mit $\mathcal{A}$ und $\mathcal{B}$ sind auch $\mathcal{A} + \mathcal{B}$ und $\mathcal{A} \times \mathcal{B}$ Modelle von LO.

b. Mit $\mathcal{A}$ und $\mathcal{B}$ sind auch $\mathcal{A} + \mathcal{B}$ und $\mathcal{A} \times \mathcal{B}$ Modelle von DLO.

Entscheiden Sie:

c. Ist $(\mathbb{N}, <) \times (\mathbb{N}, <) \cong (\mathbb{N}, <)$?

d. Ist $(\mathbb{Q}, <) + (\mathbb{Q}, <) \cong (\mathbb{Q}, <) \times (\mathbb{Q}, <) \cong (\mathbb{Q}, <)$?

e. Ist $(\mathbb{R}, <) + (\mathbb{R}, <) \cong (\mathbb{R}, <)$?

14.5.10 Zeigen Sie: In jedem abzählbaren Nicht-Standard-Modell von Z^- gibt es genau $2^{\aleph_0}$ Schnitte.
Hinweis: Benutzen Sie, dass es in $(\mathbb{Q}, <)$ zu jeder reellen Zahl genau einen Dedekindschen Schnitt gibt.

§15 Zur Zahlentheorie der zweiten Stufe: Übersetzungen

15.1 Die Zahlentheorie der zweiten Stufe als mathematische Theorie Z^2

15.2 Relativierung und Übersetzung

15.3 Modelle von Z^2

15.4 Aufgaben

In der Mathematik begegnet man häufig der Situation, dass eine Theorie T' mindestens so viel leistet, mindestens so ausdrucksstark ist wie eine Theorie T, ohne dass T' eine Erweiterung von T wäre. Es bedarf eventuell einiger Definitionen in T', um die Formeln von T in der Sprache von T' in *relativierter* Form wiederzufinden und die entsprechend *relativierten* Axiome von T in T' nachzuweisen. Dieses Verfahren, T in T' zu *übersetzen* oder T in T' *syntaktisch zu interpretieren*, ist besonders geläufig, wenn T' ein System der Mengenlehre ist. Die gewaltige Ausdruckskraft der Mengenlehre äußert sich gerade darin, dass sich jede mathematische Theorie in sie übersetzen lässt. Wir haben das in informeller Weise durch die Einführung der Semantik und den Beweis des Korrektheitssatzes nachgewiesen.

Hier wollen wir die Technik des Relativierens und Übersetzens allgemein einführen und an der besonders naheliegenden Übersetzung der gewöhnlichen Zahlentheorie Z in die Zahlentheorie der zweiten Stufe Z^2 illustrieren.

15.1 Die Zahlentheorie der zweiten Stufe als mathematische Theorie Z^2

Die Zahlentheorie der zweiten Stufe basiert auf einer möglichst weitgehenden Erfassung des Prinzips der vollständigen Induktion. Wir betrachten deshalb noch einmal dieses Prinzip und verschiedene Grade seiner Formalisierung.

In 1.1 haben wir die Ziffern $0, S0, SS0, \ldots$ induktiv definiert. Diese induktive Definition begründet – wenn man die Ziffern mit den natürlichen Zahlen identifiziert – das Prinzip der vollständigen Induktion (vgl. 1.1.6):

15.1.1 Induktionsprinzip inhaltliche Fassung Jede Eigenschaft E, die

1. auf die 0 zutrifft und

2. immer, wenn sie auf eine natürliche Zahl n zutrifft, auch auf deren Nachfolger Sn zutrifft,

trifft auf jede natürliche Zahl zu.

Das Induktionsschema (*Ind*) der gewöhnlichen Zahlentheorie Z erfasst hiervon nur die arithmetisch definierbaren Eigenschaften E (vgl. 14.3.2). Auch wenn (*Ind*) ein unendliches Axiomschema ist, ist es doch eine starke Einschränkung der inhaltlichen Fassung. Es gibt eben viele Eigenschaften natürlicher Zahlen, die nicht in $L(Z)$ ausdrückbar sind. Solange kein fester sprachlicher Rahmen vorliegt, wird man 15.1.1 eher so formulieren:

15.1.2 Induktionsprinzip teil-formalisiert

$$\begin{aligned}\forall E \quad & (\text{Eigenschaft } (E) \wedge E(0) \wedge \forall x(\text{Zahl } (x) \rightarrow E(x) \rightarrow E(Sx)) \\ & \rightarrow \forall x(\text{Zahl } (x) \rightarrow E(x)))\end{aligned}$$

Hierin treten die Prädikate „(natürliche) Zahl“, „Eigenschaft (von natürlichen Zahlen)“ und als zweistelliges Prädikat „trifft zu auf“ auf. Eine Eigenschaft von natürlichen Zahlen können wir mit der *Menge der Zahlen* identifizieren, auf die diese Eigenschaft zutrifft:

$$E \sim \{x \in N \mid E(x)\}.$$

Das Zutreffen-auf wird dabei zur *Element-Beziehung*:

$$E(k) \Leftrightarrow k \in \{x \in N \mid E(x)\}.$$

Zur Formalisierung von 15.1.1 braucht man also neben dem Null- und dem Nachfolgerzeichen

> ein *Zahlen*-Prädikatszeichen N für die Eigenschaft, eine natürliche Zahl zu sein
>
> ein *Mengen*-Prädikatszeichen M für die Eigenschaft, eine Menge (Eigenschaft) von natürlichen Zahlen zu sein,

und ein zweistelliges *Element*-Prädikatszeichen $\in$ für die Relation, Element (einer Menge) zu sein.

Dann erhalten wir:

15.1.3 Induktionsprinzip formalisiert

$$\forall y(My \wedge 0 \in y \wedge \forall x(Nx \to x \in y \to Sx \in y) \to \forall x(Nx \to x \in y)).$$

Dies ist inhaltlich ein Prinzip zweiter Stufe, weil in ihm über Mengen (Eigenschaften) natürlicher Zahlen quantifiziert wird. Formal ist es aber als eine Formel aus einer Sprache der ersten Stufe hingeschrieben. Für die glatte Formulierung 15.1.3 – statt des unendlichen Axiomschemas aus Z hat man nur eine einzige Formel – handelt man sich das Problem ein, was man unter einer Menge von natürlichen Zahlen verstehen soll. Intendiert sind alle Teilklassen von $\mathbb{N}$, aber diese lassen sich nicht erzeugen oder konstruieren wie die natürlichen Zahlen. Zunächst betrachten wir Mengen als durch ihre Elemente bestimmt:

15.1.4 Extensionalitäts-Prinzip Mengen, die dieselben Elemente enthalten, sind gleich:

$$Ma \wedge Mb \to \forall x(x \in a \leftrightarrow x \in b) \to a = b.$$

Nach diesem Prinzip soll es bei Mengen nicht auf ihr Zustandekommen, ihre Definitionsgeschichte o. ä. ankommen, sondern nur auf ihre Elemente oder auf ihren *Umfang*. Mit diesem Prinzip kann man aber noch keine Mengen konstruieren.

Wir kehren notgedrungen zu dem Ansatz zurück, der schon in 15.1.1 steckt: Mengen (von natürlichen Zahlen) sind *Umfänge von Eigenschaften* (natürlicher Zahlen). Eigenschaften werden in vielen Fällen durch sprachliche Ausdrücke bestimmt.

15.1.5 Komprehensions-Prinzip Jede Formel $F(a)$ bestimmt eine Menge, deren Elemente genau die Zahlen sind, auf die die Formel zutrifft:

$$\exists y(My \wedge \forall x(x \in y \leftrightarrow Nx \wedge F(x))).$$

Die Extensionalität 15.1.4 bewirkt, dass es nicht zu viele Mengen gibt: Auch verschieden entstandene Teilklassen von $\mathbb{N}$ sind gleich, wenn sie dieselben Zahlen enthalten. Dagegen garantiert die Komprehension 15.1.5 die Existenz vieler Mengen:

> Zu jeder Formel $F(a)$ gibt es die Menge $\{x \in \mathbb{N} \mid F(x)\}$.

Die Komprehension ordnet jeder Formel wieder ein Objekt (nämlich eine Menge) zu, dessen Name seinerseits wieder in Formeln auftreten kann. Die Komprehension ist damit charakteristisch für Systeme höherer Stufe. Durch sie wird der Mengenbegriff abhängig von der zu Grunde gelegten Sprache:

Je ausdrucksstärker eine Sprache, desto mehr Eigenschaften sind durch ihre Formeln ausdrückbar, und desto mehr Mengen entstehen durch Komprehension.

Wir haben schließlich noch die Verteilung der Objekte auf das Zahlen- und das Mengen-Prädikat zu regeln.

15.1.6 Sortierungs-Prinzipien Jedes Objekt ist entweder eine Zahl oder eine Menge:

$$Na \leftrightarrow \neg Ma.$$

Elemente sind Zahlen, Elementbesitzer sind Mengen:

$$a \in b \rightarrow Na \wedge Mb.$$

Damit haben wir alle Prinzipien erläutert, in denen sich die Zahlentheorie der zweiten Stufe von der gewöhnlichen Zahlentheorie unterscheidet. Wir können nun die Theorie Z^2 zusammenfassend formulieren.

15.1.7 Definition der Zahlentheorie Z^2 der zweiten Stufe
Die *Sprache* $L(Z^2)$ ist gegeben durch die folgenden nicht-logischen Zeichen:

> die zahlentheoretischen Funktionszeichen $0, S, +, \cdot$;
>
> die einstelligen Prädikatszeichen N, M;
>
> das zweistellige Prädikatszeichen $\in$.

Das *Axiomensystem* $Ax(Z^2)$ besteht einerseits aus Allabschlüssen der oben besprochenen Prinzipien:

$(Sort)$ $Na \leftrightarrow \neg Ma$ und $a \in b \rightarrow Na \wedge Mb$

(Ext) $Ma \wedge Mb \rightarrow \forall x(x \in a \leftrightarrow x \in b) \rightarrow a = b$

(CA) $\exists y(My \wedge \forall x(x \in y \leftrightarrow Nx \wedge F(x)))$
für jede Formel $F(a)$, in der x, y nicht auftreten;

andererseits aus den zahlentheoretischen Axiomen, allerdings relativiert auf das Prädikat N, also aus Allabschlüssen von

$1. N0$	$Na \rightarrow NSa$
$2. Na \rightarrow Sa = 0 \rightarrow \bot$	$Na \rightarrow Nb \rightarrow Sa = Sb \rightarrow a = b$
$3. Na \rightarrow a + 0 = a$	$Na \rightarrow Nb \rightarrow a + Sb = S(a + b)$
$4. Na \rightarrow a \cdot 0 = 0$	$Na \rightarrow Nb \rightarrow a \cdot Sb = a \cdot b + a$
$5. Ma \rightarrow Sa = 0$	$Ma \vee Mb \rightarrow a + b = 0 \wedge a \cdot b = 0$

(IND) $0 \in b \wedge \forall x(x \in b \rightarrow Sx \in b) \rightarrow \forall x(Nx \rightarrow x \in b)$.

Bemerkung. Die zahlentheoretischen Axiome erscheinen gegenüber den Axiomen von Z leicht verändert. Weil $0, S, +, \cdot$ Funktionszeichen sind, enthält jede Struktur zu $L(Z^2)$ – ebenso wie zu $L(Z)$ – eine Null und ist gegen Nachfolger, Addition und Multiplikation abgeschlossen. Das genügt für Z, aber nicht für Z^2, weil der Individuenbereich in Zahlen *und Mengen* zerfällt. Es muss also sichergestellt werden, das in jedem Modell $\mathcal{A}$ von Z^2 die Menge $N_{\mathcal{A}}$ gegen diese Funktionen abgeschlossen ist. Für 0 und S wird dies von 1. direkt geleistet, für $+$ und $\cdot$ folgt es aus den Axiomen, wie noch gezeigt wird.

In jeder Struktur zu $L(Z^2)$ sind Nachfolger, Addition und Multiplikation notwendig auch für Mengen definiert, obwohl uns z.B. der Wert von $e + f$ nicht interessiert, wenn e oder f eine Menge ist. Durch 5. werden diese „nicht interessierenden“ Werte einheitlich auf Null festgelegt. Dadurch wird verhindert, dass Modelle von Z^2 nur deshalb verschieden sind, weil sie sich in diesen „nicht interessierenden“ Funktionswerten unterscheiden. 2., 3. und 4. legen dagegen die „interessierenden“ Werte dieser Funktionen ebenso wie in Z fest.

(IND) unterscheidet sich von 15.1.3 nur unwesentlich. Die in 15.1.3 aufgeführten Prämissen My und Nx (im Induktionsschritt) folgen mit $(Sort)$ aus den daneben stehenden Prämissen $0 \in y$ bzw. $x \in y$ und können deshalb fortgelassen werden. Die Gültigkeit von (IND) ist deshalb äquivalent zur Wahrheit von 15.1.3, wenn $(Sort)$ gilt.

Damit ist die Formulierung und Erläuterung der Zahlentheorie der zweiten Stufe als eine *Theorie Z^2 der ersten Stufe* abgeschlossen.

15.1.8 Definition Die *Komprehensions-Gleichung*

$$b = \{x \in N \mid F(x)\}$$

(lies: b ist die Menge der Zahlen, auf die F zutrifft) steht abkürzend für die Formel

$$Mb \wedge \forall x(x \in b \leftrightarrow Nx \wedge F(x)).$$

Ferner steht $b = \{a\}$ für $b = \{x \in N \mid x = a\}$.

Die Schreibweise $b = \{x \in N \mid F(x)\}$ suggeriert, dass b durch F eindeutig bestimmt ist. Das ist in Z^2 wegen der Extensionalität auch der Fall.

15.1.9 Lemma In Z^2 gilt

$$b = \{x \in N \mid F(x)\} \rightarrow c = \{x \in N \mid F(x)\} \rightarrow b = c.$$

Beweis.

$$\forall x(x \in b \leftrightarrow Nx \wedge F(x)) \text{ und}$$
$$\forall x(x \in c \leftrightarrow Nx \wedge F(x))$$

ergeben rein logisch die Folgerung

$$\forall x(x \in b \leftrightarrow x \in c).$$

Mit (Ext) folgt dann $Mb \wedge Mc \rightarrow b = c$, und das ist zu zeigen.

Mit dem Komprehensions-Schema (CA) erhält man aus dem Induktions-Axiom (IND) wieder ein Induktions-Schema:

15.1.10 Lemma Für jede Formel $F(a)$, in der x nicht auftritt, gilt in Z^2

$$F(0) \rightarrow \forall x(Nx \rightarrow F(x) \rightarrow F(Sx)) \rightarrow \forall x(Nx \rightarrow F(x)).$$

Beweis. $I(b)$ sei (IND) aus 15.1.7 und $I(F)$ sei unsere Behauptung. Wegen Axiom 1 aus 15.1.7 gilt in Z^2

$$\begin{aligned} b = \{x \in N \mid F(x)\} \quad & \rightarrow ((0 \in b \leftrightarrow F(0)) \\ & \wedge((a \in b \rightarrow Sa \in b) \leftrightarrow (Na \rightarrow F(a) \rightarrow F(Sa))) \\ & \wedge((Na \rightarrow a \in b) \leftrightarrow (Na \rightarrow F(a)))). \end{aligned}$$

Aus den linken Seiten dieser Äquivalenzen setzt sich $I(b)$ genauso zusammen, wie $I(F)$ aus den rechten Seiten. Also folgt hieraus

$$b = \{x \in N \mid F(x)\} \to I(b) \to I(F)$$

und mit einer Verteilungsregel (vgl. 5.3.3)

$$\exists y\ y = \{x \in N \mid F(x)\} \to \forall y I(y) \to I(F),$$

woraus mit (CA) und (IND) die Behauptung $I(F)$ folgt.

Hiermit haben wir das Induktionsschema (Ind) von Z in Z^2 als gültig nachgewiesen, allerdings unter Relativierung der auftretenden Allquantoren auf das Prädikatszeichen N. Nun folgt auch, dass N gegen Addition und Multiplikation abgeschlossen ist.

15.1.11 Lemma In Z^2 gilt

$$\begin{array}{l} Na \to Nb \to N(a+b) \\ Na \to Nb \to N(a \cdot b) \end{array}.$$

Beweis. Wir setzen $F \equiv Na \to N(a + *_1)$. Dann gilt $F(0)$ wegen

$$Na \to a + 0 = a$$

und $\forall x(Nx \to F(x) \to F(Sx))$ wegen

$$Na \to Nb \to a + Sb = S(a+b) \text{ und } N(a+b) \to NS(a+b).$$

Da $I(F)$ nach 15.1.10 gilt, folgt

$$\forall x(Nx \to Na \to N(a+x)),$$

und das ist äquivalent zur ersten Behauptung.

Die zweite Behauptung beweist man analog unter Rückgriff auf Axiom 4 und die erste Behauptung.

15.2 Relativierung und Übersetzung

Aus der Art, wie die zahlentheoretischen Axiome 1. bis 5. und (IND) von Z^2 aus den entsprechenden Axiomen von Z hervorgegangen sind, kann man das allgemeine Verfahren der Relativierung und Übersetzung entwickeln. Zunächst: In welchem Sinne „enthält“ jedes Modell von Z^2 ein Modell von Z?

15.2.1 Definition Ist $\mathcal{A}$ ein Modell von Z^2, so nennen wir die Struktur

$$\mathcal{B} := (N_{\mathcal{A}}; 0_{\mathcal{A}}, S_{\mathcal{A}}, +_{\mathcal{A}}, \cdot_{\mathcal{A}})$$

den *erststufigen Anteil* von $\mathcal{A}$. (Genau genommen ist für die Funktion $S_{\mathcal{A}}$ deren Beschränkung auf $N_{\mathcal{A}}$ zu nehmen, ebenso für $+_{\mathcal{A}}$ und $\cdot_{\mathcal{A}}$.)

Der erststufige Anteil $\mathcal{B}$ von $\mathcal{A}$ ist eine Struktur zur Sprache $L(Z)$, weil nach Axiom 1 aus 15.1.7 und nach 15.1.11 die Menge $N_{\mathcal{A}}$ das Element $0_{\mathcal{A}}$ enthält und abgeschlossen ist unter den Funktionen $S_{\mathcal{A}}, +_{\mathcal{A}}, \cdot_{\mathcal{A}}$. Wir wollen zeigen, dass $\mathcal{B}$ *ein Modell von* Z ist.

Man kann nicht erwarten, dass jede in $\mathcal{B}$ gültige Formel auch in $\mathcal{A}$ gilt. Z. B. ist

$$\forall x(x = 0 \vee \exists y\ x = Sy) \tag{1}$$

in $\mathcal{B}$ wahr, aber in $\mathcal{A}$ falsch, weil in $\mathcal{A}$ der Quantor $\forall x$ über Zahlen *und Mengen* läuft. Man muss die Quantoren in (1) also auf das Prädikatszeichen N beschränken oder *relativieren*, und die relativierte Formel ist dann

$$\forall x(Nx \to (x = 0 \vee \exists y(Ny \wedge x = Sy)), \tag{2}$$

die nun in $\mathcal{A}$ denselben Sachverhalt ausdrückt wie (1) in $\mathcal{B}$. Man erhält also zu einer Formel C von $L(Z)$ die *Relativierung auf* N, indem man jeden Quantor $\forall x \ldots$ durch $\forall x(Nx \to \ldots)$ und $\exists x \ldots$ durch $\exists x(Nx \wedge \ldots)$ ersetzt.

Wir behandeln den Prozess der Relativierung in allgemeinerem Rahmen.

Konvention
Im folgenden sei L' eine Erweiterung einer Sprache L, die außer den Grundzeichen von L noch mindestens ein einstelliges Prädikatszeichen U enthält.

Die Formel Ua wird benutzt, um in L' auszudrücken, dass a zum Universum (Individuenbereich) einer Struktur zu L gehört. Im Fall der Zahlentheorie ist $L' = L(Z^2)$ und $L = L(Z)$, und U ist N.

15.2.2 Rekursive Definition der *Relativierung* C^U der Formeln C aus L auf U.

1. C^U ist C für Primformeln C;
2. $(A \to B)^U$ ist $A^U \to B^U$;
3. $(\forall x F(x))^U$ ist $\forall x(Ux \to F^U(x))$, wobei $F^U(a)$ die Formel $F(a)^U$ bezeichnet.

Sind $a_1, \ldots, a_k$ die freien Variablen, die in C auftreten, so bezeichnet $C^{(U)}$ jede Formel

$$Ua_1 \to \ldots \to Ua_k \to C^U.$$

Beispiel. Für den Fall der Zahlentheorie mit $U \equiv N$ wird sich ergeben: $C^{(N)}$ hat in einem Modell $\mathcal{A}$ von Z^2 dieselbe Bedeutung wie C im erststufigen Anteil von $\mathcal{A}$.

Bemerkung. Die Relativierung des Existenzquantors ist

$$\begin{aligned}
&(\exists x F(x))^U \equiv (\neg\forall x \neg F(x))^U \\
&\equiv \neg\forall x(Ux \to \neg F^U(x)) \\
&\Leftrightarrow \neg\forall x\neg\neg(Ux \to \neg F^U(x)) \\
&\equiv \exists x(Ux \wedge F^U(x)).
\end{aligned}$$

$(\exists x F(x))^U$ stimmt also bis auf eine doppelte Negation mit $\exists x(Ux \wedge F^U(x))$ überein. Abgesehen von dieser doppelten Negation ist also (2) die Formel $(1)^U$.

15.2.3 Definition Es sei T' eine Theorie mit Sprache L'. Die Relativierung von L auf U ist eine *Übersetzung von L in T'*, wenn

$$\exists x Ux$$

und für jedes n-stellige Funktionszeichen f aus L

$$Ua_1 \to \ldots \to Ua_n \to Uf a_1 \ldots a_n$$

in T' gelten.

15.2.4 Lemma Wenn die Relativierung von L auf U eine Übersetzung von L in T' ist und $\mathcal{A}$ ein Modell von T' ist, dann ist $U_{\mathcal{A}}$ der Individuenbereich einer Unterstruktur der Beschränkung $\mathcal{A}|L$ von $\mathcal{A}$ auf L.

Beweis. Zu zeigen ist nur, dass in jedem Modell $\mathcal{A}$ von T' die Menge $U_\mathcal{A}$ nicht leer und abgeschlossen ist gegen alle Funktionen $f_{\mathcal{A}|L}$, also gegen alle Funktionen $f_\mathcal{A}$ mit f aus L. Das ist aber nach 15.2.3 der Fall.

Beispiel. Die Relativierung von $L(Z)$ auf N ist eine Übersetzung von $L(Z)$ in Z^2 wegen Axiom 1 aus 15.1.7 und 15.1.11. Ist $\mathcal{A}$ ein Modell von Z^2, so ist die Unterstruktur von $\mathcal{A}|L(Z)$ mit Individuenbereich $N_\mathcal{A}$ gerade der erststufige Anteil von $\mathcal{A}$.

15.2.5 Satz Die Relativierung von L auf U sei eine Übersetzung von L in T'. $\mathcal{A}$ sei ein Modell von T', und $\mathcal{B}$ sei die Unterstruktur von $\mathcal{A}|L$ mit Individuenbereich $U_\mathcal{A}$. Dann ist

$$\mathcal{B}(C) = \mathcal{A}(C^U)$$

für jeden Satz C aus $L(\mathcal{B})$.

Beweis durch Induktion nach dem Aufbau von C:

1. Für Primsätze P aus $L(\mathcal{B})$ ist $P^U \equiv P$, also

$$\mathcal{B}(P) = \mathcal{A}(P) = \mathcal{A}(P^U).$$

2. C ist ein Satz $A \to B$. Dann ist

$$\begin{aligned} \mathcal{B}(C) = w \quad &\Leftrightarrow \text{ aus } \mathcal{B}(A) = w \text{ folgt } \mathcal{B}(B) = w \\ &\Leftrightarrow \text{ aus } \mathcal{A}(A^U) = w \text{ folgt } \mathcal{A}(B^U) = w, \text{ nach IV} \\ &\Leftrightarrow \mathcal{A}(A^U \to B^U) = \mathcal{A}(C^U) = w. \end{aligned}$$

3. C ist ein Satz $\forall x F(x)$. Dann ist

$$\begin{aligned} \mathcal{B}(C) = w \quad &\Leftrightarrow \text{ für alle } k \in U_\mathcal{A} \text{ ist } \mathcal{B}(F(k)) = w \\ &\Leftrightarrow \text{ für alle } k \in U_\mathcal{A} \text{ ist } \mathcal{A}(F(k)^U) = \mathcal{A}(F^U(k)) = w, \text{ nach IV} \\ &\Leftrightarrow \text{ für alle } k \in |\mathcal{A}| \text{ ist } \mathcal{A}(Uk \to F^U(k)) = w \\ &\Leftrightarrow \mathcal{A}(\forall x(Ux \to F^U(x))) = \mathcal{A}(C^U) = w. \end{aligned}$$

Aus 1. bis 3. folgt mit Induktion die Behauptung.

15.2.6 Korollar Unter den Voraussetzungen von 15.2.5 gilt die Formel C aus L in $\mathcal{B}$ genau dann, wenn $C^{(U)}$ in $\mathcal{A}$ gilt.

Beweis. C sei eine Formel $F(a_1, \ldots, a_n)$ (F geschlossen), und $a_1, \ldots, a_n$ treten in C tatsächlich auf. Wir schreiben x für $x_1, \ldots, x_n$.

$$\begin{aligned}
\mathcal{B} \models C &\Leftrightarrow \mathcal{B}(\forall x_1 \ldots \forall x_n F(x)) = w \\
&\Leftrightarrow \mathcal{A}(\forall x_1(Ux_1 \to \ldots \forall x_n(Ux_n \to F^U(x)) \ldots)) = w \text{ nach 15.2.5} \\
&\Leftrightarrow \mathcal{A}(\forall x_1 \ldots \forall x_n(Ux_1 \to \ldots Ux_n \to F^U(x))) = w \\
&\Leftrightarrow \mathcal{A} \models C^{(U)}.
\end{aligned}$$

15.2.7 Definition T und T' seien Theorien mit Sprachen L bzw. L'. Die Relativierung von L auf U ist eine *Übersetzung von T in T'*, wenn sie

1. eine Übersetzung von L in T' ist und
2. von jedem Axiom $C \in Ax(T)$ die Relativierung C^U in T' gilt.

15.2.8 Beispiel Die Relativierung von $L(Z)$ auf N ist eine Übersetzung von Z in Z^2.

Denn ist C ein Allabschluss von

$$Sa = 0 \to \bot \text{ oder } Sa = Sb \to a = b$$

oder von einer Rekursionsgleichung für $+$ oder $\cdot$, so ist C^U ein Axiom 2 bis 4 aus 15.1.7; ist C ein Induktionsaxiom (Ind), so gilt C^U nach 15.1.10 in Z^2.

15.2.9 Satz Die Relativierung von L auf U sei eine Übersetzung von T in T'. Ist $\mathcal{A}$ ein Modell von T', so ist die Unterstruktur $\mathcal{B}$ von $\mathcal{A}|L$ mit Individuenbereich $U_\mathcal{A}$ ein Modell von T.

Beweis. Es sei C ein Axiom von T. Weil eine Übersetzung vorliegt, gilt dann C^U in T', insbesondere im Modell $\mathcal{A}$. Nach 15.2.5 ist dann $\mathcal{B}(C) = w$.
Also ist $\mathcal{B}$ ein Modell von T.

15.2.10 Korollar Der erststufige Anteil eines Modells von Z^2 ist ein Modell von Z.

Beweis. Dies ist wegen 15.2.8 ein Spezialfall von 15.2.9.

15.2.11 Satz Die Relativierung von L auf U sei eine Übersetzung von T in T'. Wenn C in T gilt, gilt $C^{(U)}$ in T'.

Beweis. Sei $\mathcal{A}$ ein beliebiges Modell von T' und $\mathcal{B}$ die Unterstruktur von $\mathcal{A}|L$ mit Individuenbereich $U_{\mathcal{A}}$. Nach 15.2.9 ist $\mathcal{B}$ ein Modell von T. Also gilt C in $\mathcal{B}$, wenn C in T gilt. Nach 15.2.6 gilt dann $C^{(U)}$ in $\mathcal{A}$.
Also gilt $C^{(U)}$ in T'.

Mit der Korrektheit für T und der Vollständigkeit für T' kann man die Behauptung dieses Satzes umformen zu:

$$\text{Aus } T \vdash C \text{ folgt } T' \vdash C^{(U)}.$$

In dieser Form lässt sich der Satz rein syntaktisch beweisen, nämlich durch Herleitungsinduktion in T. Dieser Beweis, der die Zulässigkeit der Schnittregel verwendet, sei dem Leser überlassen.

15.2.12 Korollar Gilt C in Z, so gilt $C^{(U)}$ in Z^2.

Beweis. Dies ist wegen 15.2.8 ein Spezialfall von 15.2.11.

Bemerkung. Wird T in T' übersetzt, so ist T' nach 15.2.11 mindestens so ausdrucksstark wie T. Quantorenfreie Sätze, die in T gelten, gelten ebenso in T'. Ist insbesondere T widerspruchsvoll, so ist T' erst recht widerspruchsvoll. Das ergibt sich schon aus 15.2.9: in jedem Modell von T' steckt ein Modell von T.

Das Umgekehrte gilt i.a. nicht. Ein Modell von T braucht keine Oberstruktur zu besitzen, die sich zu einem Modell von T' expandieren lässt.

Als Beispiel einer Übersetzung haben wir nur die von Z in Z^2 betrachtet. In ähnlicher Weise lassen sich mathematische Theorien – darunter Z und Z^2, aber auch die Theorien aus 1.2 – in die Zermelo-Fraenkelsche Mengenlehre ZF übersetzen. Gerade dies macht die große Ausdruckskraft von ZF aus, die über die Ausdruckskraft von Z^2 noch weit hinausgeht.

15.3 Modelle von Z^2

Wir beschäftigen uns jetzt mit speziellen Modellen von Z^2.

15.3.1 Definition Das *Standardmodell* von Z^2 ist die Struktur

$$\mathcal{N}^2 := (\mathbb{N} \cup \mathit{Pot}(\mathbb{N}); 0, +1, +, \cdot; \mathbb{N}, \mathit{Pot}(\mathbb{N}), \in),$$

wobei $\mathit{Pot}(\mathbb{N})$ die Potenzmenge von $\mathbb{N}$ bezeichnet.

Das Standardmodell $\mathcal{N}^2$ ist das von Z^2 intendierte Modell. Sein erststufiger Anteil ist das Standardmodell $\mathcal{N}$ von Z. Weil Z^2 eine abzählbare Theorie ist, besitzt Z^2 nach dem Satz von Löwenheim und Skolem auch abzählbare Modelle.

15.3.2 Lemma Das Modell $\mathcal{N}^2$ von Z^2 besitzt abzählbare elementare Untermodelle. Der erststufige Anteil jeder Unterstruktur von $\mathcal{N}^2$ ist $\mathcal{N}$.

Beweis. Die erste Behauptung ist ein Spezialfall des absteigenden Satzes von Löwenheim und Skolem 12.3.1. Ist $\mathcal{A}$ eine Unterstruktur von $\mathcal{N}^2$, so gelten die quantorenfreien Formeln

$$\begin{array}{ll} N0 & Na \to NSa \\ Na \to & Nb \to N(a+b) \wedge N(a \cdot b) \end{array}$$

nach 11.2.8 auch in $\mathcal{A}$. Dann ist nach 15.2.4 der erststufige Anteil $\mathcal{B}$ von $\mathcal{A}$ eine Struktur zu $L(Z)$ und deshalb eine Unterstruktur des erststufigen Anteils $\mathcal{N}$ von $\mathcal{N}^2$. Nach 14.1.10 hat aber $\mathcal{N}$ keine echten Unterstrukturen, so dass $\mathcal{B} = \mathcal{N}$ ist.

$\mathcal{N}^2$ besitzt also abzählbare Untermodelle $\mathcal{A}$ mit erststufigem Anteil $\mathcal{N}$, in denen auch $\in_{\mathcal{A}}$ die gewöhnliche $\in$-Relation ist. Der einzige, allerdings wesentliche Unterschied liegt dann im Mengen-Prädikat M, das in $\mathcal{A}$ nur durch eine abzählbare Teilmenge $M_{\mathcal{A}}$ von $Pot(\mathbb{N})$ interpretiert ist.

15.3.3 Definition Ein Modell $\mathcal{A}$ von Z^2 heißt *regulär*, wenn $\in_{\mathcal{A}}$ die gewöhnliche $\in$-Relation ist und

$$M_{\mathcal{A}} \subseteq Pot(N_{\mathcal{A}})$$

ist. Ein reguläres Modell $\mathcal{A}$ ist ein *ω-Modell*, wenn

$$N_{\mathcal{A}} = \{\mathcal{A}(k) \mid k \text{ ist Ziffer}\}$$

ist. Ein reguläres Modell $\mathcal{A}$ heißt *vollständig*, wenn

$$M_{\mathcal{A}} = Pot(N_{\mathcal{A}})$$

ist.

Beispiele. Die Untermodelle von $\mathcal{N}^2$ sind ω-Modelle. Das einzige vollständige Untermodell von $\mathcal{N}^2$ ist $\mathcal{N}^2$ selbst.

In regulären Modellen ist jedenfalls das Zeichen $\in$ der Anschauung entsprechend interpretiert. In ω-Modellen ist nach 14.1.11 darüber hinaus der erststufige Anteil isomorph zu $\mathcal{N}$. Deshalb haben wir in den Untermodellen von $\mathcal{N}^2$ im wesentlichen schon alle ω-Modelle vor uns.

15.3.4 Lemma Jedes ω-Modell ist zu einem Untermodell von $\mathcal{N}^2$ isomorph.

Beweis. Ist $\mathcal{A}$ ein ω-Modell und $\mathcal{B}$ sein erststufiger Anteil, so gibt es nach 14.1.11 einen Isomorphismus φ von $\mathcal{B}$ auf $\mathcal{N}$. Wir definieren φ auch auf $M_{\mathcal{A}}$ durch

$$\varphi(m) := \{\varphi(k) \in \mathbb{N} \mid k \in m\} \text{ für } m \in M_{\mathcal{A}}.$$

Dann ist

$$\varphi(k) \in \varphi(m) \Leftrightarrow k \in m,$$

und φ ist auch auf $M_{\mathcal{A}}$ injektiv. Also ist φ ein Isomorphismus von $\mathcal{A}$ auf eine Unterstruktur von $\mathcal{N}^2$.

15.3.5 Korollar ω-Modelle haben höchstens die Mächtigkeit $2^{\aleph_0}$.

Beweis. $\mathcal{N}^2$ hat die Mächtigkeit

$$card(\mathbb{N}) + card(Pot(\mathbb{N})) = \aleph_0 + 2^{\aleph_0} = 2^{\aleph_0}.$$

Dann haben Unterstrukturen von $\mathcal{N}^2$ und nach 15.3.4 auch alle ω-Modelle höchstens diese Mächtigkeit.

15.3.6 Korollar Es gibt keine konsistente Theorie, deren Modelle bis auf Isomorphie nur ω-Modelle sind.

Beweis. Eine solche Theorie hätte ein unendliches Modell und deshalb nach dem aufsteigenden Satz von Löwenheim und Skolem 12.2.4 auch Modelle von höherer Mächtigkeit als $2^{\aleph_0}$. Diese könnten nach 15.3.5 keine ω-Modelle sein.

Auch mit Mitteln der zweiten Stufe lässt sich die Struktur $\mathcal{N}$ der natürlichen Zahlen also nicht syntaktisch charakterisieren. Bisher haben wir nur reguläre Modelle betrachtet. Daher konnten wir – etwa im Beweis von 15.3.4 – im gewohnten mengentheoretischen Rahmen argumentieren. Die Regularität ist aber keine wesentliche Voraussetzung.

15.3.7 Satz Jedes Modell von Z^2 ist isomorph zu einem regulären Modell.

Beweis. Es sei $\mathcal{A}$ ein Modell von Z^2, und o.E. sei

$$N_{\mathcal{A}} \cap Pot(N_{\mathcal{A}}) = \emptyset.$$

Wir definieren eine Abbildung

$$\varphi : |\mathcal{A}| \to N_{\mathcal{A}} \cup Pot(N_{\mathcal{A}})$$

durch

$$\begin{array}{ll} \varphi(k) = k & \text{für } k \in N_{\mathcal{A}} \text{ und} \\ \varphi(m) = \{k \in N_{\mathcal{A}} \mid k \in_{\mathcal{A}} m\} & \text{für } m \in M_{\mathcal{A}}. \end{array}$$

φ ist wohldefiniert, weil $N_{\mathcal{A}} \cap M_{\mathcal{A}} = \emptyset$ ist. Wir definieren eine Struktur $\mathcal{B}$ zu $L(Z^2)$ wie folgt:

Der erststufige Anteil von $\mathcal{B}$ ist der erststufige Anteil von $\mathcal{A}$;

$M_{\mathcal{B}}$ ist $\{\varphi(m) \mid m \in M_{\mathcal{A}}\}$;

$\in_{\mathcal{B}}$ ist $\in$ (beschränkt auf $N_{\mathcal{A}} \times M_{\mathcal{B}}$).

Dann ist $|\mathcal{B}| = N_{\mathcal{A}} \cup M_{\mathcal{B}} \subseteq N_{\mathcal{A}} \cup Pot(N_{\mathcal{A}})$.

Wir behaupten: φ ist ein Isomorphismus von $\mathcal{A}$ auf $\mathcal{B}$. Dies ist für den erststufigen Anteil klar. φ ist per Definition surjektiv. φ ist auch injektiv, denn:

Aus $\varphi(m) = \varphi(m')$ und $m \in M_{\mathcal{A}}$ folgt $\varphi(m') \in Pot(N_{\mathcal{A}})$, wegen
$N_{\mathcal{A}} \cap Pot(N_{\mathcal{A}}) = \emptyset$ also $m' \in M_{\mathcal{A}}$ und

$$k \in_{\mathcal{A}} m \;\Leftrightarrow\; k \in_{\mathcal{A}} m' \text{ für alle } k \in N_{\mathcal{A}}.$$

Weil die Extensionalität (Ext) in $\mathcal{A}$ gilt, folgt hieraus

$$m = m'.$$

Schließlich ist

$$k \in_{\mathcal{A}} m \;\Leftrightarrow\; k \in \varphi(m) \;\Leftrightarrow\; \varphi(k) \in_{\mathcal{B}} \varphi(m).$$

Insgesamt ist φ ein Isomorphismus von $\mathcal{A}$ auf $\mathcal{B}$, und $\mathcal{B}$ ist ein reguläres Modell von Z^2.

Bemerkung. Es bedeutet also keine Einschränkung, wenn man sich auf reguläre Modelle von Z^2 beschränkt. Die Nicht-Standard-Eigenschaften dieser Modelle liegen jedenfalls nicht an der Interpretation des Prädikatszeichens $\in$.

Wieweit der erststufige Anteil eines regulären Modells $\mathcal{A}$ von $\mathcal{N}$ abweicht, hängt wesentlich von der Lage oder Größe von $M_{\mathcal{A}}$ in der vollen Potenzmenge $Pot(N_{\mathcal{A}})$ ab. Als *(volle) Semantik der zweiten Stufe* bezeichnet man die maximale Interpretation des Mengenprädikats M, den Fall $M = Pot(N_{\mathcal{A}})$, der die vollständigen Modelle charakterisiert. Wie G. Peano schon vor 1900 bemerkte, sind die natürlichen Zahlen durch diese Semantik der zweiten Stufe bis auf Isomorphie eindeutig festgelegt:

15.3.8 Satz von Peano
Jedes vollständige Modell von Z^2 ist isomorph zum Standardmodell $\mathcal{N}^2$.

Beweis. In jedem Modell $\mathcal{A}$ von Z^2 ist $\mathcal{A}[\mathbb{N}]$ ein Schnitt im erststufigen Anteil von $\mathcal{A}$ (vgl. 14.3.1). Wenn nun $\mathcal{A}$ vollständig ist, so ist $\mathcal{A}[\mathbb{N}] \in M_{\mathcal{A}}$, und das Induktionsaxiom (IND) von Z^2 gilt auch für $\mathcal{A}[\mathbb{N}]$: Der Schnitt $\mathcal{A}[\mathbb{N}]$ enthält alle Zahlen von $\mathcal{A}$, d.h. $N_{\mathcal{A}} \subseteq \mathcal{A}[\mathbb{N}]$ und damit $N_{\mathcal{A}} = \mathcal{A}[\mathbb{N}]$. Das vollständige Modell $\mathcal{A}$ ist also ein ω-Modell. Nach 15.3.4 ist dann $\mathcal{A}$ isomorph zu einem vollständigen Untermodell von $\mathcal{N}^2$ und ist damit isomorph zu $\mathcal{N}^2$ selbst.

15.3.9 Korollar Es gibt keine Theorie, deren Modelle bis auf Isomorphie genau die vollständigen Modelle von Z^2 sind.

Beweis. Eine solche Theorie wäre eine abzählbare Theorie mit dem überabzählbaren Modell $\mathcal{N}^2$. Nach dem Satz von Löwenheim und Skolem 12.3.2 hätte sie auch ein abzählbares Modell, das offenbar nicht isomorph zu $\mathcal{N}^2$ und nach 15.3.8 auch nicht vollständig ist.

Die (volle) Semantik der zweiten Stufe lässt sich hiernach nicht syntaktisch charakterisieren.

Zusammenfassung. Die Zahlentheorie der zweiten Stufe Z^2 ist eine wesentliche Erweiterung von Z. Z lässt sich in Z^2 übersetzen, aber nicht umgekehrt. Trotzdem ist Z^2 denselben „Defekten“ unterworfen wie jede Theorie der ersten Stufe: Z^2 hat Nicht-Standard-Modelle, deren erststufiger Anteil Standard sein kann, aber nicht sein muss.

Wir haben die Modelle von Z^2 haupsächlich nach zwei Merkmalen unterschieden:

1. Der erststufige Anteil ist Standard (ω-Modelle) oder nicht;
2. Das Mengenprädikat ist maximal, $M_{\mathcal{A}} = Pot(N_{\mathcal{A}})$, (vollständige Modelle) oder nicht.

Wenn wir isomorphe Modelle identifizieren, ergibt sich folgendes Diagramm:

Modelle von Z^2	vollständig	unvollständig
ω-Modell	$\mathcal{N}^2$	echte Untermodelle von $\mathcal{N}^2$
kein ω-Modell	–	übrige reguläre Modelle

In diesem Diagramm sind (Teile von) 15.3.4, 7 und 8 zusammengefasst.

15.4 Aufgaben

15.4.1 Nach Leibniz sind Objekte identisch, wenn sie dieselben Eigenschaften haben. Weisen Sie dieses Prinzip für Zahlen in Z^2 nach, indem Sie zeigen:

$$Z^2 \models Na \to (a = b \leftrightarrow \forall y(a \in y \to b \in y)).$$

Hinweis: Verwenden Sie das Singleton $\{a\} = \{x \in N \mid x = a\}$.

15.4.2 Zeigen Sie die zweite Behauptung von 15.1.11:

$$Z^2 \models Na \to Nb \to N(a \cdot b).$$

15.4.3 Weisen Sie in Z^2 das Prinzip der kleinsten Zahl nach:

$$Z^2 \models \exists x\, x \in b \to \exists x(x \in b \wedge \forall y(y < x \to \neg y \in b)).$$

Hinweis: Betrachten Sie die Kontraposition von 14.5.6 für eine geeignete (kurze) Formel $F(a)$ aus $L(Z^2)$.

15.4.4 Zeigen Sie: $Th(\mathcal{N}^2)$ hat abzählbare reguläre Modelle, die keine ω-Modelle sind.
Hinweis: Übertragen Sie das Argument von 14.2.2 auf $Th(\mathcal{N}^2)$.

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 6:
Beweistheorie der Prädikatenlogik

Autor: Justus Diller

01301-7-06-S 1

Klassische Prädikatenlogik

Kurseinheit 6: Inhalt

Studienhinweise 331

Verzeichnis der definierten Begriffe und der wichtigen Sätze 333

1. Sprache, Semantik und Syntax der Prädikatenlogik

2. Syntaktische Sätze und Regeln der Prädikatenlogik

3. Vollständigkeit

4. Modelltheorie

5. Beweistheorie der Prädikatenlogik

§16 Der Hauptsatz von Gentzen 335

16.1 r-Schnitt und r-Herleitungen 336

16.2 Das Schnitt-Lemma 338

16.3 Schnitt-Reduktion und Schnitt-Elimination 341

§17 Prämissen-abgeschlossene Klassen 345

17.1 Identitätsfreie Logik 346

17.2 Der Herbrandsche Satz 351

17.3 Interpolation 355

17.4 Vereinte Konsistenz 363

17.5 Implizite und explizite Definierbarkeit 365

17.6 Monotonie und Positivität 366

17.7 Aufgaben 368

§18 Definitorische Erweiterungen und Skolemisierung 370

18.1 Erweiterungen und Expansionen 370

18.2 Definitorische Erweiterungen um Prädikatszeichen 371

18.3 Definitorische Erweiterungen um Funktionszeichen 374

18.4 Pränexe Normalformen 379

18.5 Skolemisierung 381

18.6 Aufgaben 386

6. Automatisches Beweisen

Klassische Prädikatenlogik

Kurseinheit 6: Studienhinweise

1. Lehrziele

Im Mittelpunkt der Kurseinheit steht eine Liste von Ergebnissen der Prädikatenlogik, die überwiegend der Beweistheorie zuzuordnen sind und dem entsprechend vorwiegend mit syntaktischen Methoden bewiesen werden:

- Hauptsatz von Gentzen
- identitätsfreie Logik: Vollständigkeit und Löwenheim–Skolem–Sätze
- Satz von Herbrand
- Interpolationssätze von Craig und Lyndon
- Satz von Robinson über vereinte Konsistenz
- Definierbarkeitssatz von Beth
- Positivität monotoner Formeln
- Definitorische Erweiterungen
- Pränexe Normalformen
- Satz von Skolem, Skolemisierung

Einzelne modelltheoretische Sätze (Löwenheim–Skolem-Sätze für die identitätsfreie Logik, Satz von Robinson) sind Folgerungen oder Anwendungen der vorangehenden beweistheoretischen Ergebnisse.

Ein wesentliches Ziel der Kurseinheit ist die Erarbeitung dieser Liste von Ergebnissen. Zusätzlich soll verstanden werden, welche Methoden jeweils eingesetzt werden: Das Beweistheorie-interne Thema von § 16 wird naturgemäß rein syntaktisch behandelt. § 17 bringt wichtige Beispiele für die beweistechnischen Vorteile unseres schnittfreien Herleitungsbegriffs. In § 18 wird schließlich demonstriert, dass die Mischung von semantischen und syntaktischen Methoden in vielen Fällen kurze und durchsichtige Beweise liefern kann.

2. Eingangsvoraussetzungen

In der Kurseinheit wird ständig auf den Herleitungsbegriff aus 3.1 und seine Eigenschaften aus dem 2. Kapitel, besonders aus § 5 zurückgegriffen. In § 16 ist außerdem Kenntnis der Grundeigenschaften der Exponentialfunktion $n \mapsto 2^n$ nützlich.

In 17.1 und 17.4 werden Korrektheits- und Vollständigkeitssatz (3.2 und § 8) sowie die Löwenheim–Skolem-Sätze aus § 12 auf einen neuen Kontext übertragen bzw. für eine modelltheoretische Folgerung verwendet und dabei als bekannt vorausgesetzt.

Korrektheits- und Vollständigkeitssatz werden auch im Beweis des Expansionslemmas in 18.1 verwendet, dazu die Begriffe der Expansion und Beschränkung von Strukturen. Das Expansionslemma spielt in Beweisen in 18.2, 3 und 5 eine wichtige Rolle. Ebenfalls kommen elementare Schlussweisen der Semantik, die aus § 2 bekannt sind, vielfach zum Einsatz.

Klassische Prädikatenlogik

Kurseinheit 6: Verzeichnis der definierten Begriffe und der wichtigen Sätze

16.1.1 Rang $rg(C)$ einer Formel C

16.1.2 r-Herleitungen, Herleitungen vom Schnittrang r; $T \overset{n}{\underset{r}{\vdash}} \Gamma : \Delta$

16.2.1 **Schnitt-Lemma**

16.3.1 **Satz zur Schnitt-Reduktion**

16.3.3 **Satz zur Schnitt-Elimination**

16.3.4 **Hauptsatz von Gentzen**

17.1.1 identitätsfrei, fast identitätsfrei, $\Gamma^- : \Delta$

17.1.3 **Satz** Sind T und $\Gamma : \Delta$ identitätsfrei und ist $T \vdash \Gamma : \Delta$, so hat $\Gamma : \Delta$ eine identitätsfreie Herleitung in T ohne Gleichheitsschlüsse.

17.1.4 **Korollar** Vollständigkeit der identitätsfreien Logik

17.1.5 **Korollar** Hauptsatz von Gentzen für die identitätsfreie Logik

17.1.8 **Lemma** Ist T identitätsfrei und $\mathcal{B} \models T$, so gibt es zu jedem $\kappa \geq \text{card}(\mathcal{B})$ ein Modell $\mathcal{A} \supseteq \mathcal{B}$ von T der Mächtigkeit κ.

17.1.9 **Satz von Löwenheim, Skolem und Tarski** für identitätsfreie Theorien

17.2.1 **Satz von Herbrand**

17.2.2 $\exists$-Sequenz, Herbrand-Instanz

17.3.1 $\langle \Gamma : \Delta \rangle$, IP-Formel

17.3.2 **Interpolationssatz, Craigs Lemma**

17.3.3 positives und negatives Auftreten

17.3.5 $\langle \Gamma : \Delta \rangle^+$, $\langle \Gamma : \Delta \rangle^-$, SIP-Formel

17.3.6 **Interpolationssatz von Craig und Lyndon**

17.4.1 Vereinigung $T_1 \cup T_2$

17.4.3 **Satz von A. Robinson über vereinte Konsistenz**

17.5.2 **Definierbarkeitssatz von Beth**

17.6.1 p-positive Formeln, (in T beweisbar) monotone Nennformen

17.6.2 **Satz** Beweisbar in p monotone Formeln sind beweisbar äquivalent zu p-positiven Formeln

18.1.1 $\vdash$-Erweiterung, Herleitbarkeitserweiterung

18.1.2 **Expansions-Lemma**

18.2.1 definitorische Erweiterung um Prädikatszeichen

18.2.3 Eliminationsübersetzung $^q : C \mapsto C^q$

18.2.4 **Satz** über definitorische Erweiterungen um Prädikatszeichen

18.3.1 definitorische Erweiterung um Funktionszeichen

18.3.3 f-einfache Formeln

18.3.4 f-Vereinfachung $^1 : C \mapsto C^1$

18.3.6 Eliminationsübersetzung $^f : C \mapsto C^f$

18.3.7 **Satz** über definitorische Erweiterungen um Funktionszeichen

18.3.8 definitorische $\vdash$-Erweiterung

18.4.1 pränexe Formeln, in pränexer Form; Präfix, Matrix; pränexe Normalform

18.4.2 pränexe Umformung

18.4.4 **Satz** Jede Formel besitzt eine pränexe Normalform

18.5.1 Skolem-Erweiterung, Skolem-Funktion

18.5.3 Skolem-Normalform A^S, Skolem-Funktionszeichen, $SF(A^S)$; Skolem-Erweiterung

18.5.5 **Satz von Skolem**

18.5.6 Skolemisierung einer Theorie

18.5.7 **Satz über Skolemisierung** Jede Theorie besitzt eine offene konservative $\vdash$-Erweiterung

Kapitel 5

Beweistheorie der Prädikatenlogik

Wir unterziehen den Herleitungsbegriff aus § 3 einer genaueren Analyse. In dem kurzen § 16 geben wir ein rein syntaktisches Verfahren an, Herleitungen mit Schnitten in schnittfreie Herleitungen umzuwandeln. Das Verfahren gibt auch präzise numerische Auskunft darüber, wie stark die Ordnung einer Herleitung bei der Elimination der Schnitte anwachsen kann.

In § 17 folgen einige wichtige Sätze, die wegen der Schnittfreiheit unseres Herleitungsbegriffs besonders elementare syntaktische Beweise haben. In erster Linie handelt es sich um den Satz von Herbrand in 17.2 und den Interpolationssatz in 17.3, von dem bemerkenswerte Konsequenzen in 17.4–6 behandelt werden.

Schließlich studieren wir in § 18 einige logische Techniken, die in der mathematischen Praxis ständig – und sei es unbewusst – verwendet werden. Um in diesem Gebiet kurze und einfache Beweise zu erhalten, mischen wir syntaktische und semantische Methoden. Charakteristisch ist dafür der wiederholte Einsatz des Expansionslemmas aus 18.1.

§16 Der Hauptsatz von Gentzen

16.1 r-Schnitt und r-Herleitungen

16.2 Das Schnitt-Lemma

16.3 Schnitt-Reduktion und Schnitt-Elimination

Die Zulässigkeit der Schnittregel folgt bereits aus dem Korrektheits- und Vollständigkeitssatz, weil die Schnittregel offenbar korrekt ist. Auf dem Wege erhält man aber keine Abschätzung über das Wachstum der Herleitungsordnung, wenn man aus einer Herleitung alle Schnitte eliminiert. Gerhard Gentzen hat 1935 ein konstruktives syntaktisches Verfahren zur Schnittelimination angegeben, mit dem man in beliebigen Theorien aus Herleitungen, die die Schnittregel verwenden, durch iteriertes Umformen schnittfreie Herleitungen gewinnt. Eine scharfe Abschätzung der Herleitungsordnung bei Schnittelimination geht auf W. W. Tait 1967 zurück. Sie bezieht sich allerdings auf einen etwas anderen Kalkül. Dieses primitiv-rekursive Verfahren, angepasst an den Gentzen-Kalkül aus § 3, ist Gegenstand dieses Paragraphen.

16.1 r-Schnitt und r-Herleitungen

Die Einsicht, dass ein Schnitt zulässig ist, wird umso schwieriger, je länger die Schnittformel ist. Ein geeignetes Maß für die Länge einer Formel ist in diesem Zusammenhang ihr *Rang*.

16.1.1 Rekursive Definition des *Ranges* $rg(C)$ für jede Formel C.

1. $rg(P) = 0$ für Primformeln P.
2. $rg(A \to B) = \max(rg(A) + 1, rg(B))$.
3. $rg(\forall x F(x)) = rg(F(a)) + 1$.

Beispiele. Die Festlegung unter 2. liefert

$$rg(A_1 \to A_2 \to \ldots \to B) = \max\{rg(B), rg(A_i) + 1 | 1 \leq i \leq n\}$$

und wirkt dadurch relativ sparsam. Allerdings ist

$$rg(\neg A) = rg(A) + 1 \text{ und daher } rg(\exists x F(x)) = rg(\forall x F(x)) + 2,$$

und i. a. ist $rg(A \to B) \neq rg(B \to A)$.

Wir erweitern den Herleitungsbegriff aus 3.1 für $r \in \mathbb{N}$ um die Regel

$$(r\text{-Schnitt}) \quad \Gamma, C : \Delta \text{ und } \Gamma : C, \Delta \vdash \Gamma : \Delta, \text{ falls } rg(C) < r.$$

16.1.2 Definition Herleitungen, die neben den Grundschlussregeln aus 3.1 (u. U.) die Regel (r-Schnitt) verwenden, heißen *r-Herleitungen* oder *Herleitungen vom Schnittrang r*. $T \mid\!\!\frac{n}{r}\, \Gamma : \Delta$ besagt: In T gibt es eine r-Herleitung von $\Gamma : \Delta$ mit einer Ordnung $\leq n$ (vgl. 5.1.2).

Hiermit ist für jedes $r \in \mathbb{N}$ ein Herleitungsbegriff $\mid\!\!\frac{}{r}$ definiert. Es ist $\mid\!\!\frac{n}{0} = \mid\!\!\frac{n}{}$, weil es keine Formel mit Rang < 0 gibt, und $\mid\!\!\frac{n}{r} \subseteq \mid\!\!\frac{n}{r+1}$, aber nicht umgekehrt.

Für $r > 0$ ist $\mid\!\!\frac{}{r}$ nicht schnittfrei, sondern erlaubt Schnitte über „kurze“ Formeln, wenn man (zu gegebenem $r \in \mathbb{N}$) die Formeln vom Rang $< r$ als „kurz“ betrachtet.

16.1.3 Lemma Folgende Regeln sind schwache Schlussregeln:

$$
\begin{array}{ll}
(Subst) & \mid\!\!\frac{n}{r}\, \Gamma(b) : \Delta(b) \Rightarrow \mid\!\!\frac{n}{r}\, \Gamma(s) : \Delta(s) \quad (b \text{ nicht in } \Gamma, \Delta) \\
(Str) & \mid\!\!\frac{n}{r}\, \Gamma : \Delta \Rightarrow \mid\!\!\frac{n}{r}\, \Gamma^+ : \Delta^+, \text{ falls } \Gamma : \Delta \subset_S \Gamma^+ : \Delta^+ \\
(\rightarrow SInv) & \mid\!\!\frac{n}{r}\, \Gamma : A \rightarrow B, \Delta \Rightarrow \mid\!\!\frac{n}{r}\, \Gamma, A : B, \Delta \\
(\rightarrow AInv) & \mid\!\!\frac{n}{r}\, \Gamma, A \rightarrow B : \Delta \Rightarrow \mid\!\!\frac{n}{r}\, \Gamma : A, \Delta \text{ und } \mid\!\!\frac{n}{r}\, \Gamma, B : \Delta \\
(\forall SInv) & \mid\!\!\frac{n}{r}\, \Gamma : \forall x F(x), \Delta \Rightarrow \mid\!\!\frac{n}{r}\, \Gamma : F(s), \Delta.
\end{array}
$$

Die **Beweise** erfolgen jeweils durch Induktion nach n. Wir übernehmen die Induktionsbeweise aus 5.1 vollständig, wobei sich Rückgriffe etwa auf $(Subst)$ in den Beweisen von (Str) und $(\forall SInv)$ jetzt selbstverständlich auf die Substitutionsregel für $\mid\!\!\frac{}{r}$ beziehen, anders als in 5.1. Nachzutragen bleibt jeweils der neue Induktionsschritt

$$(r\text{-Schnitt}) \ \mid\!\!\frac{n}{r}\, \Gamma, C : \Delta \text{ und } \mid\!\!\frac{n}{r}\, \Gamma : C, \Delta \text{ ergibt } \mid\!\!\frac{n+1}{r}\, \Gamma : \Delta, \text{ falls } rg(C) < r.$$

Alle sechs Behauptungen ($(\rightarrow AInv)$ besteht aus zweien) haben bei diesem Induktionsschritt die Gestalt

$$\mid\!\!\frac{n+1}{r}\, \Gamma : \Delta \Rightarrow \mid\!\!\frac{n+1}{r}\, \Gamma^+ : \Delta^+.$$

Im Fall der Strukturschlussregel (Str) ist $\Gamma : \Delta \subset_S \Gamma^+ : \Delta^+$, also auch $\Gamma, C : \Delta \subset_S \Gamma^+, C : \Delta^+$ und $\Gamma : C, \Delta \subset_S \Gamma^+ : C, \Delta^+$. Daher ist in jedem Fall die Induktionsvoraussetzung auf die Prämissen von (r-Schnitt) anwendbar, und man erhält

$$\mid\!\!\frac{n}{r}\, \Gamma^+, C' : \Delta^+ \text{ und } \mid\!\!\frac{n}{r}\, \Gamma^+ : C', \Delta^+,$$

wobei im Fall von $(Subst)$ $C \equiv F(b)$ (b nicht in F) und $C' \equiv F(s)$ ist. In jedem anderen Fall ist $C' \equiv C$. Also ist stets $rg(C') = rg(C) < r$, und mit (r-Schnitt) folgt

$$\vdash_r^{n+1} \Gamma^+ : \Delta^+.$$

Mit Induktion nach n folgen nun die Behauptungen (für alle $r \in \mathbb{N}$).
Aus $(\to SInv)$ (mit (Str)) und $(\to AInv)$ folgt:

16.1.4 Korollar

$$\begin{array}{ll} (\neg SInv) & \vdash_r^n \Gamma : \neg A, \Delta \Rightarrow \vdash_r^n \Gamma, A : \Delta \\ (\neg AInv) & \vdash_r^n \Gamma, \neg A : \Delta \Rightarrow \vdash_r^n \Gamma : A, \Delta. \end{array}$$

16.2 Das Schnitt-Lemma

Der wesentliche Schritt zur effektiven Elimination von Schnitten ist die folgende Elimination eines einzigen $(r+1)$-Schnittes auf Kosten möglicherweise vieler r-Schnitte:

16.2.1 Schnitt-Lemma Sei $rg(C) \leq r$.

$$\text{Aus } \vdash_r^k \Gamma, C : \Delta \text{ und } \vdash_r^l \Gamma : C, \Delta \text{ folgt } \vdash_r^{k+l} \Gamma : \Delta.$$

Beweis durch Fallunterscheidung nach der Gestalt von C.

1. C ist eine Primformel P. Wir induzieren nach l.

   1.1 $\Gamma : P, \Delta$ ist ein logisches Axiom. Dann ist $\Gamma : \Delta$ selbst schon ein logisches Axiom, oder es ist $P \in \Gamma$, also $\Gamma, P = \Gamma$, und es folgt $\vdash_r^k \Gamma : \Delta$. Das ergibt wegen $k \leq k + l$ die Behauptung.

   1.2 Es ist $\vdash_r^{l+1} \Gamma : P, \Delta$ die Konklusion eines Schlusses nach einer Grundschlussregel (R) (einschließlich (r-Schnitt)). Dann haben dessen Prämissen eine Gestalt

   $$\vdash_r^l \Gamma_i : P, \Delta_i \ (i = 1 \text{ oder } i = 1, 2).$$

   Strukturschlüsse (Str), angewandt auf beide Voraussetzungen, ergeben

   $$\vdash_r^k \Gamma, \Gamma_i, P : \Delta, \Delta_i \text{ und } \vdash_r^l \Gamma, \Gamma_i : P, \Delta, \Delta_i.$$

Nach Induktionsvoraussetzung ist dann $\vdash_r^{k+l} \Gamma, \Gamma_i : \Delta, \Delta_i$, und daraus folgt mit einem Schluss nach derselben Grundschlussregel (R) $\vdash_r^{(k+l)+1} \Gamma : \Delta$. Das ist wegen $k+(l+1) = (k+l)+1$ die Behauptung.

(Im Fall $(R) = (\forall S)$ ist z. B. $i = 1, \Gamma_1 = \Gamma; \Delta = \Delta_0, \forall x F(x); \Delta_1 = \Delta_0, F(a)$ und $a \notin FV(P)$. Im Fall $(R) = (= P)$ ist $i = 1, \Delta_1 = \Delta; \Gamma = \Gamma_0, pt$ und $\Gamma_1 = \Gamma_0, ps, s = t$. Weil die Schnittformel P im Sukzedens steht, wird sie mit den Primformeln $ps, s = t$ nicht vermischt.)

2. C ist $A \to B$ oder $\forall x F(x)$. Ist $C \in \Gamma$, so ist die Behauptung trivial. Sei nun $C \notin \Gamma$. Für diesen Fall induzieren wir nach k.

   2.1 $\Gamma, C : \Delta$ ist ein logisches Axiom. Dann ist auch $\Gamma : \Delta$ ein logisches Axiom, weil C keine Primformel ist.

   2.2 Es ist $\vdash_r^{k+1} \Gamma, C : \Delta$ die Konklusion eines $(\to A)$-Schlusses mit der Hauptformel $C \equiv A \to B$. Dessen Prämissen sind dann

   $$\vdash_r^k \Gamma_0 : A, \Delta \text{ und } \vdash_r^k \Gamma_0, B : \Delta$$

   mit $\Gamma_0, C = \Gamma, C$, also $\Gamma_0 = \Gamma$ oder $\Gamma_0 = \Gamma, C$. Im zweiten Fall folgt mit $(\to AInv)$ aus 16.1.3

   $$\vdash_r^k \Gamma : A, \Delta \text{ und } \vdash_r^k \Gamma, B : \Delta,$$

   was im ersten Fall ohnehin schon klar war. Aus der zweiten Voraussetzung $\vdash_r^l \Gamma : A \to B, \Delta$ folgt mit $(\to SInv)$ $\vdash_r^l \Gamma, A : B, \Delta$. Da $rg(B) \leq r$ ist, folgt mit (Str) nach Induktionsvoraussetzung

   $$\vdash^{k+l} \Gamma, A : \Delta,$$

   und weil $rg(A) < r$ ist, folgt mit (r-Schnitt) $\vdash_r^{(k+l)+1} \Gamma : \Delta$, und das ist wegen $(k + l) + 1 = (k + 1) + l$ die Behauptung.

   2.3 Es ist $\vdash_r^{k+1} \Gamma, C : \Delta$ die Konklusion eines $(\forall A)$-Schlusses mit der Hauptformel $C \equiv \forall x F(x)$. Dessen Prämisse ist dann

   $$\vdash_r^k \Gamma_0, F(t) : \Delta$$

mit $\Gamma_0, C = \Gamma, C$. Mit (Str) folgt

$$\left|\frac{k}{r}\right. \Gamma, \forall x F(x), F(t) : \Delta.$$

Aus der zweiten Voraussetzung $\left|\frac{l}{r}\right. \Gamma : \forall x F(x), \Delta$ folgt mit (Str) $\left|\frac{l}{r}\right. \Gamma, F(t) : \forall x F(x), \Delta$, so dass wegen $rg(\forall x F(x)) \leq r$ nach Induktionsvoraussetzung

$$\left|\frac{k+l}{r}\right. \Gamma, F(t) : \Delta$$

folgt. Die zweite Voraussetzung ergibt mit $(\forall SInv)$ aber auch $\left|\frac{l}{r}\right. \Gamma : F(t), \Delta$, so dass nun wegen $rg(F(t)) < r$ mit (r-Schnitt) die Behauptung $\left|\frac{k+l+1}{r}\right. \Gamma : \Delta$ folgt.

2.4 Es ist $\left|\frac{k+1}{r}\right. \Gamma, C : \Delta$ die Konklusion eines Grundschlusses nach einer Regel (R), und C ist nicht die Hauptformel dieses Schlusses. Dessen Prämissen haben eine Gestalt

$$\left|\frac{k}{r}\right. \Gamma_i, C : \Delta_i (i = 1 \text{ oder } i = 1, 2).$$

Strukturschlüsse (Str), angewandt auf beide Voraussetzungen des Lemmas, ergeben

$$\left|\frac{k}{r}\right. \Gamma, \Gamma_i, C : \Delta, \Delta_i \text{ und } \left|\frac{l}{r}\right. \Gamma, \Gamma_i : C, \Delta, \Delta_i.$$

Nach Induktionsvoraussetzung ist dann $\left|\frac{k+l}{r}\right. \Gamma, \Gamma_i : \Delta, \Delta_i$, und daraus schließt man nach derselben Regel (R) auf die Behauptung $\left|\frac{k+l+1}{r}\right. \Gamma : \Delta$.

Damit ist das Schnitt-Lemma vollständig bewiesen.

Die Definition 16.1.1 des Ranges einer Formel ist so gewählt, dass sich das Schnitt-Lemma möglichst einfach und durchsichtig beweisen lässt: Der Beweis ist eine einfache Induktion – je nach der Gestalt der Schnittformel C – nach der Ordnung der zweiten bzw. der ersten gegebenen Herleitung, die man in jedem Fall als Induktion nach deren Summe $k + l$ lesen kann. Die kritischen Fälle sind 2.2 und 2.3, in denen die Inversionsregeln zum Zuge kommen: Mit einer Kombination von Induktionsvoraussetzung und (r-Schnitt) lässt sich jeweils auch der „unangenehme" Fall beherrschen, in dem die Schnittformel C

sowohl Hauptformel des betrachteten Schlusses ist als auch in dessen Prämissen auftritt. Dafür nimmt man die nach 16.1.1 angemerkten Symmetriemängel bei der Definition des Ranges in Kauf.

16.2.2 Alternativen Eine sparsame Alternative zu 16.1.1 erhält man durch

2.0 $rg(\neg A) = rg(A)$

und die Beschränkung des 2. Schrittes in 16.1.1 auf den Fall $A \to B$ mit $B \not\equiv \bot$. Dann wird

$$rg(A \to B) = rg(A \vee B) = rg(A \wedge B) \text{ und } rg(\forall x F(x)) = rg(\exists x F(x)),$$

aber i. a. ist immer noch $rg(A \to B) \neq rg(B \to A)$. Der Beweis des Schnitt-Lemmas bleibt technisch derselbe; die Fallunterscheidung nach der Schnittformel C geht allerdings in eine Induktion nach C über, weil im Fall $C \equiv \neg A$ auf die Gültigkeit des Lemmas für A zurückgegriffen werden muss, nicht mehr auf einen r-Schnitt. (Für diesen Fall greift man zweckmäßig auf Korollar 16.1.4 zurück.)

Im kritischen Fall 2.2 des Beweises des Schnittlemmas, in dem die Schnittformel $C \equiv A \to B$ Hauptformel des betrachteten Schlusses ist, kommt es nur darauf an, dass man auf eine der direkten Subformeln A, B die Induktionsvoraussetzung anwenden kann, auf die andere einen r-Schnitt. Diese Beobachtung führt zu der symmetrischen Alternative.

2.1 $rg(A \to B) = \min(\max(rg(A) + 1, rg(B)), \max(rg(A), rg(B) + 1))$

zur Formulierung 2. in 16.1.1. Diese Alternative kann man mit 2.0 verknüpfen, indem man 2.1 auf den Fall $A \not\equiv \bot, B \not\equiv \bot$ beschränkt und 2.0 erweitert zu

2.0.1 $rg(\neg A) = rg(\bot \to A) = rg(A)$

Beweis und Aussage des Schnitt-Lemmas werden davon nicht wesentlich beeinflusst.

16.3 Schnitt-Reduktion und Schnitt-Elimination

Man kann alle $(r+1)$-Schnitte zugunsten von r-Schnitten aus einer $(r+1)$-Herleitung eliminieren, wenn man das Schnitt-Lemma systematisch „von oben

nach unten“ auf alle $(r+1)$-Schnitte der Herleitung anwendet:

16.3.1 Satz Schnitt-Reduktion

$$\vdash^{n}_{r+1} \Gamma : \Delta \Rightarrow \vdash^{2^n}_{r} \Gamma : \Delta.$$

Beweis durch Induktion nach n.

1. $\Gamma : \Delta$ ist ein logisches Axiom. Dann ist $\Gamma : \Delta$ schnittfrei hergeleitet, und es gilt $\vdash^{2^n}_{r} \Gamma : \Delta$ für jedes n (und jedes r).

2. Es ist $\vdash^{n+1}_{r+1} \Gamma : \Delta$ die Konklusion eines Grundschlusses, der *kein* $(r+1)$-Schnitt ist. Die Prämissen dieses Schlusses sind dann

   $$\vdash^{n}_{r+1} \Gamma_i : \Delta_i \quad (i = 1 \text{ oder } i = 1, 2).$$

   Nach Induktionsvoraussetzung folgt

   $$\vdash^{2^n}_{r} \Gamma_i : \Delta_i \quad (i = 1 \text{ oder } i = 1, 2),$$

   und der betrachtete Schluss ergibt

   $$\vdash^{2^n+1}_{r} \Gamma : \Delta.$$

   Es ist $1 \leq 2^n$, also $2^n + 1 \leq 2^n + 2^n = 2^{n+1}$, so dass die Behauptung $\vdash^{2^{n+1}}_{r} \Gamma : \Delta$ folgt.

3. Der letzte Schluss ist ein $(r+1)$-Schnitt von

   $$\vdash^{n}_{r+1} \Gamma, C : \Delta \text{ und } \vdash^{n}_{r+1} \Gamma : C, \Delta \text{ auf } \vdash^{n+1}_{r+1} \Gamma : \Delta$$

   mit $rg(C) < r+1$, also $rg(C) \leq r$. Nach Induktionsvoraussetzung ist

   $$\vdash^{2^n}_{r} \Gamma, C : \Delta \text{ und } \vdash^{2^n}_{r} \Gamma : C, \Delta,$$

   so dass mit dem Schnitt-Lemma 16.2.1 folgt

   $$\vdash^{2^n+2^n}_{r} \Gamma : \Delta,$$

   und das ist wegen $2^n + 2^n = 2^{n+1}$ die Behauptung.

Mit Induktion nach n folgt nun der Schnitt-Reduktions-Satz.

Bemerkung. Gegenüber der Summe der Herleitungsordnungen, wie sie im Schnitt-Lemma auftritt, erscheint das exponentielle Wachstum der Herleitungsordnungen bei der Schnitt-Reduktion beträchtlich. Wie der Induktionsschritt 3. zeigt, lässt sich dieses Wachstum nicht wesentlich besser beschränken: Wenn die meisten Schlüsse in einer $(r+1)$-Herleitung $(r+1)$-Schnitte sind, muss man in den meisten Fällen zur Summe der für die Prämissen errechneten Herleitungsordnungen übergehen und dabei im wesentlichen auf die Rekursionsgleichung $2^n + 2^n = 2^{n+1}$ der Exponentialfunktion zurückgreifen.

Bei der Elimination aller Schnitte aus einer Herleitung verschärft sich das Wachstum der Herleitungsordnung entsprechend.

16.3.2 Rekursive Definition der *iterierten Potenz* 2^n_r.

$$2^n_0 = n \text{ und } 2^n_{r+1} = 2^{2^n_r}.$$

Beispiele. Es ist $2^1_0 = 1$, $2^1_1 = 2^1 = 2$, $2^1_2 = 2^2 = 4$, $2^1_3 = 2^4 = 16$,

$$2^1_4 = 2^{16} = 65536 \text{ und } 2^1_5 = 2^{65536}.$$

Allgemein ist $2^n_r = 2^{2^{\cdot^{\cdot^{\cdot^{2^n}}}}}$ (r-mal), wobei Klammerung nach rechts oben zu ergänzen ist. Daraus liest man ab:

$$(*) \quad 2^n_{r+1} = 2^{(2^n)}_r,$$

was man leicht durch Induktion nach r beweist.

16.3.3 Satz Schnitt-Elimination

$$\vdash^n_r \Gamma : \Delta \Rightarrow \vdash^{2^n_r}_0 \Gamma : \Delta.$$

Beweis durch Induktion nach r.

1. Für $r = 0$ stimmen wegen $2^n_0 = n$ Voraussetzung und Behauptung überein.

2. Aus $\vdash^n_{r+1} \Gamma : \Delta$ folgt mit Schnitt-Reduktion

$$\vdash^{2^n}_r \Gamma : \Delta$$

und daraus nach Induktionsvoraussetzung

$$\vdash_0^{2^{(2^n)}} \Gamma : \Delta,$$

was wegen $(*)$ schon die Behauptung ist.

Jede Herleitung H in einer Theorie T, in der die Schnittregel in beliebiger Weise verwendet wird, ist immer noch ein endlicher Sequenzenbaum. In H treten daher nur endlich viele Schnittformeln auf. Genügend große $r \in \mathbb{N}$ sind dann größer als die Ränge aller dieser Schnittformeln, so dass H eine r-Herleitung für diese r ist. Der letzte Satz lässt sich daher auch so formulieren:

16.3.4 Hauptsatz von Gentzen Jede Herleitung H in einer Theorie T mit Schnitten besitzt einen Schnittgrad r. Ist n die Ordnung dieser Herleitung, so lässt sich H umformen zu einer schnittfreien Herleitung in T mit derselben Endsequenz und einer Herleitungsordnung $\leq 2_r^n$.

§17 Prämissen-abgeschlossene Klassen

17.1 Identitätsfreie Logik

17.2 Der Herbrandsche Satz

17.3 Interpolation

17.4 Vereinte Konsistenz

17.5 Implizite und explizite Definierbarkeit

17.6 Monotonie und Positivität

17.7 Aufgaben

Wegen der Subformel-Eigenschaft unseres schnittfreien Herleitungskalküls kann man aus der Gestalt der Konklusion eines Grundschlusses Folgerungen ziehen für die Gestalt der Prämissen. In günstigen Fällen übertragen sich Eigenschaften von der Konklusion auf die Prämissen, unabhängig davon, unter welche Regel der Schluss fällt. Solche Eigenschaften übertragen sich offenbar von der Endsequenz einer Herleitung auf alle Sequenzen in der Herleitung.

17.0.1 Definition Sei T eine Theorie und $\mathcal{K}$ eine Menge von Sequenzen von $L(T)$. Wir nennen $\mathcal{K}$ eine *prämissen-abgeschlossene Klasse* in T, wenn $\mathcal{K}$ mit einer Sequenz $\Gamma : \Delta$ auch alle Prämissen jedes Grundschlusses in T enthält, dessen Konklusion $\Gamma : \Delta$ ist.

Beispiele.

1. Die Klasse aller Sequenzen aus $L(T)$ ist selbstverständlich prämissen-abgeschlossen in T. Ist aber L' eine Teilsprache von $L(T)$, so ist die Klasse aller Sequenzen aus L' i. a. nicht prämissen-abgeschlossen, auch wenn $Ax(T)$ nur aus L'-Sätzen besteht. Denn z. B. geht bei $(\forall A)$-Schlüssen ein Term t verloren, der Funktionszeichen aus $L(T)$ enthalten kann, die nicht zu L' gehören.
   Allerdings kann man eine gegebene Herleitung einer Sequenz $\Gamma : \Delta$ aus L' so abändern zu einer Herleitung H', dass H' nur noch aus Sequenzen von L' besteht. In dem Spezialfall, dass die Funktionszeichen, die nicht zu L' gehören, alle 0-stellig, also Konstanten sind, ist diese Behauptung das Lemma über neue Konstanten 8.3.2.

2. Enthält $L(T)$ überhaupt keine Funktionszeichen, also auch keine Konstanten, so ist eine Teilsprache L' von $L(T)$ genau dann prämissen-abgeschlossen in T, wenn alle Axiome von T Sätze aus L' sind.

3. Die Klasse der quantorenfreien Sequenzen von $L(T)$ ist prämissen-abgeschlossen, wenn alle Axiome von T quantorenfreie Sätze aus $L(T)$ sind.

Gerade die im 1. Beispiel angesprochene Möglichkeit, Herleitungen in einer prämissen-abgeschlossenen Klasse so abzuändern, dass sie eine gewünschte Gestalt bekommen, macht das Studium spezieller prämissen-abgeschlossener Klassen interessant. Wir untersuchen hier drei Beispiele wachsender Komplexität.

17.1 Identitätsfreie Logik

Die identitätsfreien Sequenzen bilden auch in logischen Theorien keine prämissen-abgeschlossene Klasse, denn

$$\frac{P, a = a : P}{P : P}$$

ist für jede Primformel P ein $(= I)$-Schluss und sogar eine Herleitung, in der $=$ auftritt, auch wenn P keine Gleichung ist. Die Beobachtung, dass dieser $(= I)$-Schluss „überflüssig“ ist und die $(= I)$-Regel im Grunde die einzige Grundschlussregel ist, die hier stört, führt zu folgender Definition:

17.1.1 Definition Eine Formel bzw. Formelmenge ist *identitätsfrei*, wenn in ihr bzw. in allen ihrer Formeln das Gleichheitszeichen $=$ nicht auftritt. Eine Theorie T bzw. eine Sequenz $\Gamma : \Delta$ ist *identitätsfrei*, wenn $Ax(T)$ bzw. $\Gamma \cup \Delta$ identitätsfrei ist. Eine Formelmenge Γ nennen wir *fast identitätsfrei*, wenn die Formeln aus Γ entweder identitätsfrei oder triviale Gleichungen der Gestalt $t = t$ sind. Γ^- entsteht dann aus Γ, indem man aus Γ alle diese Gleichungen $t = t$ entfernt. Eine Sequenz $\Gamma : \Delta$ nennen wir *fast identitätsfrei*, wenn Γ fast identitätsfrei und Δ (völlig) identitätsfrei ist.

Bemerkungen.

1. Identitätsfreie Sequenzen sind auch fast identitätsfrei. Ist $\Gamma : \Delta$ fast identitätsfrei, so ist $\Gamma^- : \Delta$ identitätsfrei.

2. In einer identitätsfreien Herleitung H treten keine Gleichheits-Grundschlüsse $(=I), (=F)$ oder $(=P)$ auf. Denn die Prämissen von $(=I)$- und $(=F)$-Schlüssen sind offenbar nicht identitätsfrei, und weil wir $(=F)$- und $(=P)$-Schlüsse nur auf Funktions- bzw. Prädikatszeichen positiver Stellenzahl anwenden, sind die Konklusionen dieser Schlüsse nicht identitätsfrei.

Im Gegensatz zu den identitätsfreien Sequenzen bilden die fast identitätsfreien Sequenzen eine prämissen-abgeschlossene Klasse (in einer identitätsfreien Theorie). Diese Beobachtung führt zu

17.1.2 Lemma Sei T eine identitätsfreie Theorie und $\Gamma : \Delta$ fast identitätsfrei aus $L(T)$.

Ist $T \stackrel{n}{\vdash} \Gamma : \Delta$, so ist $T \stackrel{n}{\vdash} \Gamma^- : \Delta$ mit einer Herleitung, die keine Gleichheitsschlüsse $(=I), (=F)$ oder $(=P)$ verwendet.

Beweis durch Herleitungsinduktion

1. $\Gamma : \Delta$ ist ein logisches Axiom. Dann ist auch $\Gamma^- : \Delta$ ein logisches Axiom, weil Gleichungen $t = t$ zwar in Γ, aber nicht in Δ auftreten können und deshalb nicht ursächlich für die Axiomeigenschaft von $\Gamma : \Delta$ sind.

2. Der letzte Schluss

$$T \stackrel{n}{\vdash} \Gamma_i : \Delta_i (i = 1 \text{ oder } i = 1, 2) \vdash \Gamma : \Delta$$

ist ein $(\to S)$-, $(\to A)$-, $(\forall S)$- oder $(\forall A)$-Schluss. Dessen Hauptformel ist dann keine Gleichung und mithin identitätsfrei. Daher sind auch die Prämissen $\Gamma_i : \Delta_i$ fast identitätsfrei. Nach Induktionsvoraussetzung ist dann

$$T \stackrel{n}{\vdash} \Gamma_i^- : \Delta_i \text{ ohne Gleichheitsschlüsse } (i = 1 \text{ oder } i = 1, 2),$$

und mit einem Schluss nach derselben Regel folgt $T \stackrel{n+1}{\vdash} \Gamma^- : \Delta$ ohne Gleichheitsschlüsse.

3. Der letzte Schluss ist

$$(=I) \quad T \stackrel{n}{\vdash} \Gamma, t = t : \Delta \vdash \Gamma : \Delta.$$

Mit $\Gamma : \Delta$ ist auch die Prämisse $\Gamma, t = t : \Delta$ fast identitätsfrei, und es ist $(\Gamma, t = t)^- = \Gamma^-$. Nach Induktionsvoraussetzung ist $T \overset{n}{\vdash} \Gamma^- : \Delta$ ohne Gleichheitsschlüsse, und das ist die Behauptung: der $(= I)$-Schluss entfällt.

4. Der letzte Schluss ist

$$(= F) \quad T \overset{n}{\vdash} \Gamma, fs_1 \dots s_k = ft_1 \dots t_k : \Delta \vdash \Gamma, s_1 = t_1, \dots, s_k = t_k : \Delta \text{ bzw.}$$
$$(= P) \quad T \overset{n}{\vdash} \Gamma, pt_1 \dots t_k : \Delta \vdash \Gamma, ps_1 \dots s_k, s_1 = t_1, \dots, s_k = t_k : \Delta.$$

Da die jeweilige Konklusion fast identitätsfrei ist, ist $s_i \equiv t_i$ für $i = 1, \dots, k$ und, falls im $(= P)$-Schluss p das Gleichheitszeichen ist, $k = 2$ und $s_1 \equiv s_2$. Dann ist auch

$$fs_1 \dots s_k \equiv ft_1 \dots t_k \text{ und } ps_1 \dots s_k \equiv pt_1 \dots t_k,$$

also, falls p das Gleichheitszeichen ist, auch $t_1 \equiv t_2$. In jedem Fall ist daher die Prämisse wieder fast identitätsfrei, und es ist

$$(\Gamma, fs_1 \dots s_k = ft_1 \dots t_k)^- = \Gamma^- = (\Gamma, s_1 = t_1, \dots, s_k = t_k)^- \text{ bzw.}$$
$$(\Gamma, pt_1 \dots t_k)^- = (\Gamma, ps_1 \dots s_k, s_1 = t_1, \dots, s_k = t_k)^-.$$

Nach Induktionsvoraussetzung ist dann

$$T \overset{n}{\vdash} \Gamma^- : \Delta \text{ bzw. } T \overset{n}{\vdash} (\Gamma, pt_1 \dots t_k)^- : \Delta$$

jeweils ohne Gleichheitsschlüsse, und das ist in beiden Fällen schon die Behauptung: die Gleichheitsschlüsse entfallen.

5. Der letzte Schluss ist

$$(T) \quad T \overset{n}{\vdash} B, \Gamma : \Delta \vdash \Gamma : \Delta \text{ mit } B \in Ax(T).$$

Mit T ist auch B identitätsfrei, also ist mit $\Gamma : \Delta$ auch die Prämisse fast identitätsfrei. Nach Induktionsvoraussetzung ist dann $T \overset{n}{\vdash} B, \Gamma^- : \Delta$ ohne Gleichheitsschlüsse, und mit einem (T)-Schluss folgt $T \overset{n+1}{\vdash\!\!\!-} \Gamma^- : \Delta$ ohne Gleichheitsschlüsse.

Mit Herleitungsinduktion folgt aus 1. bis 5. das Lemma. Als Korollar ergibt sich nun das Ziel dieser Betrachtung:

17.1.3 Satz Sind T und $\Gamma : \Delta$ identitätsfrei und ist $T \vdash \Gamma : \Delta$, so hat $\Gamma : \Delta$ eine identitätsfreie Herleitung in T ohne Gleichheitsschlüsse.

Beweis. Da Γ identitätsfrei ist, ist $\Gamma^- = \Gamma$. Nach Lemma 17.1.2 hat dann $\Gamma^- : \Delta$, also $\Gamma : \Delta$ eine Herleitung in T ohne Gleichheitsschlüsse. Diese Herleitung ist notwendig identitätsfrei (vgl. Bemerkung 2 nach 17.1.1).

Mit diesem Satz übertragen sich etliche Ergebnisse zur Prädikatenlogik mit Identität auf die identitätsfreie Logik.

17.1.4 Korollar Vollständigkeit der identitätsfreien Logik.
Jede identitätsfreie Sequenz, die in einer identitätsfreien Theorie T gilt, hat eine identitätsfreie Herleitung in T.

Beweis. Aus $T \models \Gamma : \Delta$ folgt $T \vdash \Gamma : \Delta$ mit dem Vollständigkeitssatz in der Fassung 8.3.8, und daraus folgt mit 17.1.3 die Behauptung.

Der Korrektheitssatz für die identitätsfreie Logik ist nur ein Spezialfall des Korrektheitssatzes 3.2.1 und unabhängig von 17.1.3. Entsprechendes gilt für den Kompaktheitssatz.

17.1.5 Korollar Hauptsatz von Gentzen für die identitätsfreie Logik.
Sind T und $\Gamma : \Delta$ identitätsfrei und ist

$$T \vdash \Gamma, B : \Delta \text{ und } T \vdash \Gamma : B, \Delta,$$

so hat $\Gamma : \Delta$ eine identitätsfreie Herleitung in T: Es gibt ein primitiv-rekursives Verfahren, aus Herleitungen von $\Gamma, B : \Delta$ und $\Gamma : B, \Delta$ in T eine identitätsfreie Herleitung von $\Gamma : \Delta$ in T zu gewinnen.

Beweis. Mit dem Verfahren, das in 16.2 und 3 zum Hauptsatz von Gentzen führte, erhält man eine Herleitung von $\Gamma : \Delta$, die notwendig fast identitätsfrei ist, und diese streicht man nach dem Beweis von 17.1.2 zu einer identitätsfreien Herleitung von $\Gamma : \Delta$ zusammen.

Für die Löwenheim–Skolem-Sätze sieht die Situation anders aus, weil jede konsistente identitätsfreie Theorie ein unendliches Modell hat, was sich für den abzählbaren Fall schon mit 9.4.3 ergibt. Wir skizzieren diesen Sachverhalt für beliebige Mächtigkeiten.

17.1.6 Lemma Ist $\varphi : \mathcal{A} \to \mathcal{B}$ ein surjektiver Homomorphismus mit

$$(*) \quad \text{aus } (\varphi(k_1), \ldots, \varphi(k_n)) \in p_{\mathcal{B}} \text{ folgt } (k_1, \ldots, k_n) \in p_{\mathcal{A}}$$

für alle $k_1, \ldots, k_n \in |\mathcal{A}|$ und nicht-logischen Prädikatszeichen p aus L, so gilt für alle identitätsfreien Sätze C aus $L(\mathcal{A})$

$$\mathcal{A}(C) = w \Leftrightarrow \mathcal{B}(\varphi(C)) = w.$$

Der Beweis durch Induktion nach dem Aufbau von C ist derselbe wie für Isomorphismen φ in 11.2.11 unter Vernachlässigung der Gleichheit.

17.1.7 Korollar Ist T eine identitätsfreie Theorie und ist $\varphi : \mathcal{A} \to \mathcal{B}$ ein surjektiver Homomorphismus mit $(*)$ zwischen Strukturen $\mathcal{A}, \mathcal{B}$ zu $L(T)$, so ist

$$\mathcal{A} \models T \Leftrightarrow \mathcal{B} \models T.$$

Denn die Behauptung besagt $\mathcal{A}(C) = w \Leftrightarrow \mathcal{B}(C) = w$ für alle $C \in Ax(T)$, und diese C sind identitätsfrei, und $\varphi(C) \equiv C$. Also ergibt das Lemma die Behauptung.

17.1.8 Lemma Ist T identitätsfrei und $\mathcal{B} \models T$, so gibt es zu jedem $\kappa \geq \text{card}(\mathcal{B})$ ein Modell $\mathcal{A} \supseteq \mathcal{B}$ von T der Mächtigkeit κ.

In der Prädikatenlogik mit Identität ist diese Aussage jedenfalls für endliche κ und auch, wenn $L(T)$ keine κ-Sprache ist, falsch.

Beweis des Lemmas. Wir setzen $B := |\mathcal{B}|$ und wählen eine zu B disjunkte Menge E, so dass $A := B \cup E$ die Mächtigkeit κ hat, ferner ein Element $b_0 \in B$. Dann sei $\varphi : A \to B$ die Surjektion, die gegeben ist durch

$$\varphi(b) = b \text{ für } b \in B, \ \varphi(e) = b_0 \text{ für } e \in E.$$

Die Struktur $\mathcal{A} = (A, (f_{\mathcal{A}})_{f \in L}, (p_{\mathcal{A}})_{p \in L})$ sei nun definiert durch

$$f_{\mathcal{A}}(a_1, \ldots, a_k) := f_{\mathcal{B}}(\varphi(a_1), \ldots, \varphi(a_k))$$
$$(a_1, \ldots, a_k) \in p_{\mathcal{A}} :\Leftrightarrow (\varphi(a_1), \ldots, \varphi(a_k)) \in p_{\mathcal{B}}$$

für alle $a_1, \ldots, a_n \in A$. Da $\varphi \restriction B = \text{id}$ ist, ist $\mathcal{B} \subseteq \mathcal{A}$, und umgekehrt ist $\varphi : \mathcal{A} \to \mathcal{B}$ ein surjektiver Homomorphismus. Nach 17.1.7 ist daher mit $\mathcal{B}$ auch $\mathcal{A}$ ein Modell von T.

Das Vergrößern von Modellen identitätsfreier Theorien ist nach diesem Lemma unproblematisch. Eine Aussage wie „Es gibt höchstens n Elemente“ ist daher identitätsfrei nicht möglich. Dagegen impliziert der identitätsfreie Satz

$$\exists x \exists y (px \wedge \neg py),$$

dass es mindestens zwei verschiedene Elemente gibt. Das Verkleinern von Modellen ist also auch hier nicht uneingeschränkt möglich.

17.1.9 Satz von Löwenheim, Skolem und Tarski für identitätsfreie Theorien. Jede konsistente identitätsfreie κ-Theorie hat ein Modell der Mächtigkeit κ.

Beweis. Ist T konsistent, so hat T ein Modell $\mathcal{B}$. Ist $\mathrm{card}(\mathcal{B}) \leq \kappa$, so folgt die Behauptung mit 17.1.8. Ist $\mathrm{card}(\mathcal{B}) > \kappa$, so ist $\mathcal{B}$ (wie κ) unendlich, und die Behauptung folgt mit dem gewöhnlichen absteigenden Satz von Löwenheim und Skolem 12.3.1.

17.2 Der Herbrandsche Satz

Wir wenden uns einer weiteren prämissen-abgeschlossenen Klasse von Sequenzen zu. Eine Theorie T ist *offen*, wenn alle ihre Axiome (reine) $\forall$-Sätze sind, d. h. Sätze der Gestalt $\forall x_1 \ldots \forall x_n F(x_1, \ldots, x_n)$ mit quantorenfreiem $F(a_1, \ldots, a_n)$.

17.2.1 Satz von Herbrand. Ist T offen und

$$T \vdash \exists x_1 \ldots \exists x_n F(x) \quad (F \text{ quantorenfrei}, x :\equiv x_1, \ldots, x_n)$$

so gibt es endlich viele Termtupel $t_1, \ldots, t_k$ der Länge n, so dass

$$T \vdash F(t_1) \vee \ldots \vee F(t_k).$$

Der Beweis dieses Satzes ist das Thema dieses Abschnitts.

Die Aussage des Herbrandschen Satzes liegt nicht auf der Hand. Dazu folgende Bemerkungen:

1. Die naheliegende Verschärfung der Aussage auf $k = 1$ ist i. A. falsch, wie das untenstehende Beispiel zeigt.

2. Die Nennform F kann freie Variablen enthalten. T beweist dann einen $\forall\exists$-Satz $\forall y \exists x F_0(y,x)$, der in jedem Modell $\mathcal{A} \models T$ die Totalität einer Funktion φ mit $\mathcal{A}(F_0(k,\varphi(k))) = w$ ausdrückt. Die Termtupel t_i müssen in Abhängigkeit von den freien Variablen von F und u. U. von weiteren freien Variablen solche Funktionen φ „stückweise“ definieren, und zwar uniform für alle Modelle von T zugleich.

3. Wenn F geschlossen ist, sind die Termtupel t_i i. A. nicht geschlossen. Sie können es gar nicht sein, wenn $L(T)$ keine Konstanten enthält. Die t_i geben also nicht einfach konkrete einzelne Instanzen „verteilt auf k Optionen“ an, sondern sie liefern in Abhängigkeit von ihren freien Variablen ganze Instanzenscharen zu dem hergeleiteten $\exists$-Satz.

Beispiel. $LO \vdash \exists x \exists y (a = b \vee x < y)$.
Es gibt sicher kein einzelnes Variablenpaar $\tilde{a}, \tilde{b}$, so dass

$$LO \vdash a = b \vee \tilde{a} < \tilde{b},$$

wohl aber bringen die zwei Variablenpaare a, b und b, a die Lösung

$$LO \vdash (a = b \vee a < b) \vee (a = b \vee b < a),$$

was sich unmittelbar aus $LO3$ ergibt.

Wenn man fragt, wie die Sequenzen in einer Herleitung einer reinen $\exists$-Formel in einer offenen Theorie aussehen, stößt man auf folgendes Konzept:

17.2.2 Definition Wir nennen $\Gamma : \Delta$ eine $\exists$-*Sequenz*, wenn Γ aus reinen $\forall$-Formeln besteht und Δ quantorenfrei ist. Dann nennen wir jede Formelmenge Γ^H, die zu jeder Formel $\forall x F(x) \in \Gamma$ ($x = x_1, \ldots, x_n, F$ quantorenfrei) endlich viele Instanzen $F(t_1), \ldots, F(t_k)$ enthält, eine *Herbrand-Instanz* von Γ. Ist $\Gamma = \{C\}$, so schreiben wir C^H für Γ^H.

Bemerkung. Γ^H ist stets quantorenfrei. Ist Γ quantorenfrei, so ist stets $\Gamma^H = \Gamma$. Sind Γ^{H_1} und Γ^{H_2} Herbrand-Instanzen von Γ, so auch $\Gamma^{H_1} \cup \Gamma^{H_2}$.

Der Herbrandsche Satz beruht wesentlich darauf, dass die $\exists$-Sequenzen in einer offenen – es genügt: in einer logischen – Theorie eine prämissen-abgeschlossene Klasse bilden.

17.2.3 Lemma Sei T logisch und $\Gamma : \Delta$ eine $\exists$-Sequenz aus $L(T)$. Wenn

$$T \vdash^n \Gamma : \Delta,$$

dann gibt es eine Herbrand-Instanz Γ^H von Γ, so dass

$$T \vdash^n \Gamma^H : \Delta.$$

Beweis durch Herleitungsinduktion.

1. $\Gamma : \Delta$ sei ein logisches Axiom. Dann setzen wir

   $$\Gamma^H = \{C \in \Gamma | C \text{ quantorenfrei}\},$$

   und dann ist auch $\Gamma^H : \Delta$ ein logisches Axiom.

2. Der letzte Schluss ist

($\rightarrow S$) $\vdash^n \Gamma, A : B, \Delta \vdash \Gamma : A \rightarrow B, \Delta$.

Dann ist $A \rightarrow B$, also auch A und B quantorenfrei, so dass die Prämisse eine $\exists$-Sequenz ist. Nach Induktionsvoraussetzung gibt es eine Herbrand-Instanz $(\Gamma, A)^H = \Gamma^H, A$, so dass

$$T \vdash^n \Gamma^H, A : B, \Delta.$$

Mit ($\rightarrow S$) folgt $T \vdash^{n+1} \Gamma^H : A \rightarrow B, \Delta$.

($\rightarrow A$) $\vdash^n \Gamma : A, \Delta$ und $\vdash^n \Gamma, B : \Delta \vdash \Gamma, A \rightarrow B : \Delta$.

Weil $A \rightarrow B$ keine Allformel ist, ist $A \rightarrow B$, also auch A und B quantorenfrei, so dass die Prämissen $\exists$-Sequenzen sind. Nach Induktionsvoraussetzung gibt es Herbrand-Instanzen Γ^{H_1} von Γ und $(\Gamma, B)^{H_2} = \Gamma^{H_2}, B$ von Γ, B, so dass

$$T \vdash^n \Gamma^{H_1} : A, \Delta \text{ und } T \vdash^n \Gamma^{H_2}, B : \Delta.$$

Setzen wir $\Gamma^H := \Gamma^{H_1} \cup \Gamma^{H_2}$, so ist auch Γ^H eine Herbrand-Instanz von Γ, und mit Strukturschlüssen folgt

$$T \vdash^n \Gamma^H : A, \Delta \text{ und } T \vdash^n \Gamma^H, B : \Delta.$$

Mit ($\rightarrow A$) folgt $T \vdash^{n+1} \Gamma^H, A \rightarrow B : \Delta$.

($\forall S$) kommt nicht vor, weil Δ quantorenfrei ist.

($\forall A$) $\vdash^n \Gamma, F(t) : \Delta \vdash \Gamma, \forall x F(x) : \Delta$.

Mit $\forall x F(x)$ ist auch $F(t)$ eine $\forall$-Formel (möglicherweise quantorenfrei), so dass die Prämisse eine $\exists$-Sequenz ist. Nach Induktionsvoraussetzung gibt es eine Herbrand-Instanz $(\Gamma, F(t))^H = \Gamma^H \cup F(t)^H$, die offenbar auch eine Herbrand-Instanz von $\Gamma, \forall x F(x)$ ist, so dass

$$T \vdash^n \Gamma^H, F(t)^H : \Delta \equiv (\Gamma, \forall x F(x))^H : \Delta,$$

und das ist bereits die Behauptung: Der ($\forall A$)-Schluss entfällt.

3. Der letzte Schluss ist ein ($= I$)-, ($= F$)- oder ($= P$)-Schluss, hat also eine Gestalt
$$\vdash^n \Gamma, P : \Delta \vdash \Gamma, \Pi : \Delta,$$
worin P eine geeignete Primformel und Π eine (im Fall ($= I$) leere) Menge von Primformeln ist, so dass die Prämisse eine $\exists$-Sequenz ist. Nach Induktionsvoraussetzung gibt es eine Herbrand-Instanz $(\Gamma, P)^H = \Gamma^H, P$ von Γ, P, so dass
$$T \vdash^n \Gamma^H, P : \Delta.$$
Mit einem Schluss nach derselben Regel folgt
$$T \vdash^{n+1} \Gamma^H, \Pi : \Delta \equiv (\Gamma, \Pi)^H : \Delta.$$

4. (T)-Schlüsse treten nicht auf, weil $Ax(T)$ leer ist.

Mit Herleitungsinduktion folgt nun die Behauptung.

Den Übergang zu einer Herbrand-Instanz kann man stückweise auch wieder rückgängig machen.

17.2.4 Lemma Ist C eine $\forall$-Formel und C^H eine Herbrand-Instanz von C, so folgt aus $T \vdash \Gamma, C^H : \Delta$ stets $T \vdash \Gamma, C : \Delta$.

Beweis. Ist $C \equiv \forall x F(x)$ ($x \equiv x_1, \ldots, x_n, F$ quantorenfrei), so ist C^H eine Menge $\{F(t_i) | t_i$ Termtupel der Länge $n, 1 \leq i \leq k\}$. Wenn man nacheinander $n \cdot k$ ($\forall A$)-Schlüsse auf $\Gamma, C^H : \Delta$ anwendet, erhält man schließlich $\Gamma, C : \Delta$.

17.2.5 Beweis des Satzes von Herbrand. Durch $(n-1)$-fache Anwendung der Stabilitäts- und Verteilungsregeln 5.2.2 bzw. 5.3.3 auf eine nach 4.1.2 herleitbare Sequenz der Gestalt $C : C$ erhält man für $x \equiv x_1, x_2, ..., x_n$

$$\vdash \exists x F(x) : \neg \forall x \neg F(x)$$

Hieraus und aus der Voraussetzung des Satzes $T \vdash \exists x F(x)$ folgt mit einem nach dem Hauptsatz von Gentzen 16.3.4 zulässigen Schnitt und anschließender $(\neg AInv)$ $T \vdash \forall x \neg F(x) : \bot$. Nach dem Deduktionstheorem gibt es Axiome $A_1, \ldots, A_m$ von T, so dass

$$(L(T), \emptyset) \vdash A_1, \ldots, A_m, \forall x \neg F(x) : \bot,$$

und das ist eine $\exists$-Sequenz, weil T offen ist. Nach 17.2.3 ist dann eine Herbrand-Instanz hiervon in $(L(T), \emptyset)$, also erst recht in T herleitbar:

$$T \vdash A_1^H, \ldots, A_m^H, \forall x \neg F(x)^H : \bot.$$

m-fache Anwendung von 17.2.4 und m (T)-Schlüsse ergeben

$$T \vdash \forall x \neg F(x)^H : \bot.$$

Sei nun $\forall x \neg F(x)^H = \{\neg F(t_1), \ldots, \neg F(t_k)\}$. Mit k $(\neg AInv)$-Schlüssen (und einem Strukturschluss, um $\bot$ im Sukzedens loszuwerden) folgt

$$T \vdash \emptyset : F(t_1), \ldots, F(t_k),$$

und $(k-1)$ $(\vee S)$-Schlüsse ergeben die Behauptung.

17.3 Interpolation

Nach dem Hauptsatz von Gentzen kann man aus Herleitungen von Sequenzen $A : B$ und $B : C$ eine Herleitung von $A : C$ konstruieren. Umgekehrt gibt es zu jeder herleitbaren Sequenz $A : C$ viele Formeln B, für die $A : B$ und $B : C$ herleitbar sind. Am einfachsten kann man für B die Formeln A und C selbst wählen.

Im Fall, dass $A : C$ logisch herleitbar ist, kann man die Formel B auch so wählen, dass sie zwischen A und C *interpoliert*, d. h. dass die Prädikatszeichen (außer $=$) und die freien Variablen aus B sowohl in A als auch in C auftreten. Wir definieren etwas allgemeiner:

17.3.1 Definition Zu einer Sequenz $\Gamma : \Delta$ aus einer Sprache L bezeichne $\langle \Gamma : \Delta \rangle$ die Menge der in $\Gamma : \Delta$ auftretenden Prädikatszeichen (außer $=$) und freien Variablen; speziell $\langle B \rangle := \langle \emptyset : \{B\} \rangle$.

Eine Formel B ist *IP-Formel* (Interpolationsformel für Prädikatszeichen) zu zwei Sequenzen $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$, wenn

1. $(L, \emptyset) \vdash \Gamma_1, B : \Delta_1$ und $(L, \emptyset) \vdash \Gamma_2 : B, \Delta_2$ und
2. $\langle B \rangle \subseteq \langle \Gamma_1 : \Delta_1 \rangle \cap \langle \Gamma_2 : \Delta_2 \rangle$.

17.3.2 Interpolationssatz, Craig's Lemma. Ist

$$(L, \emptyset) \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2,$$

so gibt es eine IP-Formel B zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$.

Oben wurde hiervon der Spezialfall $\Gamma_1 : \Delta_1 = \emptyset : C$ und $\Gamma_2 : \Delta_2 = A : \emptyset$ diskutiert. Die allgemeine Formulierung ist nötig, damit der Beweis durch Herleitungsinduktion glatt abläuft.

Beispiele.

1. Für Primformeln P, Q, R (mit verschiedenen Prädikatszeichen p, q, r) ist

   $$P, P \to Q : (Q \to R) \to R$$

   eine Tautologie, also logisch herleitbar. Weil p nicht rechts und r nicht links auftritt, darf jede IP-Formel zu $\emptyset : (Q \to R) \to R$ und $P, P \to Q : \emptyset$ nur mit q aufgebaut sein, und tatsächlich ist Q IP-Formel zu diesen beiden Sequenzen.

2. Sei p 0-stelliges, q 1-stelliges Prädikatszeichen. Dann ist

   $$\forall x(p \to qx), p : qa$$

   logisch herleitbar. Jede IP-Formel zu $\emptyset : qa$ und $\forall x(p \to qx), p : \emptyset$ darf q, aber weder p noch a enthalten. Also ist qa keine IP-Formel, wohl aber $\forall xqx$.

Wenn man sich den Beweis von Craig's Lemma durch Herleitungsinduktion näher ansieht, stellt man fest, dass er schon ein stärkeres Ergebnis beweist, was zuerst Lyndon bemerkte: Nicht nur Prädikatszeichen pauschal, sondern ihre *positiven* und *negativen* Auftreten können je für sich interpoliert werden.

17.3.3 Induktive Definition der *positiven* und *negativen* Auftreten von Prädikatszeichen in Formeln und Sequenzen.

1. $*_1 t_1 \ldots t_n$ ist ein positives Auftreten ($n \geq 0$).
2. Ist A eine Formel, $\Gamma : \Delta$ eine Sequenz und $\mathcal{F}$ ein positives (negatives) Auftreten in einer Formel, so sind $(A \to \mathcal{F})$ und $\Gamma : \mathcal{F}, \Delta$ positive (negative) und $(\mathcal{F} \to A)$ und $\Gamma, \mathcal{F} : \Delta$ negative (positive) Auftreten.
3. Mit $\mathcal{F}(*_1, a)$ ist auch $\forall x\, \mathcal{F}(*_1, x)$ ein positives (negatives) Auftreten.

Ein Prädikatszeichen p *tritt positiv (negativ) in* einer Formel C bzw. in einer Sequenz $\Gamma : \Delta$ *auf*, wenn es ein positives (negatives) Auftreten von p in C bzw. in $\Gamma : \Delta$ gibt.

17.3.4 Bemerkungen

1. Das Vorzeichen eines Auftretens ändert sich also nur beim Übergang zum Vorderglied einer Implikation und zum Antezedens einer Sequenz.
2. Jedes Auftreten von p in C und in $\Gamma : \Delta$ ist entweder positiv oder negativ.
3. Tritt p positiv (negativ) in A oder in B bzw. in $\mathcal{F}(a)$ auf, so tritt p positiv (negativ) in $A \wedge B$ und $A \vee B$ bzw. in $\exists x\, \mathcal{F}(x)$ auf.

17.3.5 Definition $\langle \Gamma : \Delta \rangle^+$ (bzw. $\langle \Gamma : \Delta \rangle^-$) bezeichnet die Menge der positiv (bzw. negativ) in $\Gamma : \Delta$ auftretenden Prädikatszeichen (außer $=$), vereinigt mit $FV(\Gamma : \Delta)$. Speziell ist $\langle B \rangle^s := \langle \emptyset : \{B\} \rangle^s$ für $s \in \{+, -\}$. Wir nennen eine Formel B eine *SIP-Formel (signierte Interpolationsformel für Prädikatszeichen)* zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$, wenn

1. $(L, \emptyset) \vdash \Gamma_1, B : \Delta_1$ und $(L, \emptyset) \vdash \Gamma_2 : B, \Delta_2$ und
2. $\langle B \rangle^{+}_{-} \subseteq \langle \Gamma_1 : \Delta_1 \rangle^{+}_{-} \cap \langle \Gamma_2 : \Delta_2 \rangle^{-}_{+}$.

Beispiele.

1. Die in den Beispielen nach 17.3.2 angegebenen IP-Formeln sind auch SIP-Formeln.

2. $p : p \to p$ ist eine Tautologie. $p \to p$ ist eine IP-Formel zu $\emptyset : p \to p$ und $p : \emptyset$, aber keine SIP-Formel, weil p in $p \to p$ sowohl positiv als auch negativ auftritt, in p aber nur positiv. SIP-Formeln sind hier p, aber auch $\top$, weil $p \to p$ eine Tautologie ist.

Der „natürliche“ Beweis von Craig's Lemma ergibt schon:

17.3.6 Interpolationssatz von Craig und Lyndon. Ist

$$(L, \emptyset) \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2,$$

so gibt es eine SIP-Formel B zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$.

Dem Beweis schicken wir ein Lemma voraus, das die Anzahl der Fälle im Beweis etwa halbiert. Der Begriff der SIP-Formel ist nicht symmetrisch in den beiden Sequenzen $\Gamma_i : \Delta_i$ $(i = 1, 2)$, aber doch fast:

17.3.7 Lemma Ist B SIP-Formel zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$, so ist $\neg B$ SIP-Formel zu $\Gamma_2 : \Delta_2$ und $\Gamma_1 : \Delta_1$ (und umgekehrt).

Beweis. Sei B SIP-Formel zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$, also:

1. $\vdash \Gamma_1, B : \Delta_1$ und $\vdash \Gamma_2 : B, \Delta_2$, was äquivalent ist zu

$$\vdash \Gamma_1 : \neg B, \Delta_1 \text{ und } \vdash \Gamma_2, \neg B : \Delta_2;$$

2. wegen $\neg B \equiv B \to \bot$ und $\langle \bot \rangle = \emptyset$ ist

$$\langle \neg B \rangle^{+}_{-} = \langle B \rangle^{-}_{+} \subseteq \langle \Gamma_1 : \Delta_1 \rangle^{-}_{+} \cap \langle \Gamma_2 : \Delta_2 \rangle^{+}_{-},$$

so dass $\neg B$ SIP-Formel zu $\Gamma_2 : \Delta_2$ und $\Gamma_1 : \Delta_1$ ist. Die Schlüsse sind umkehrbar. Die Umkehrung folgt aber auch aus dem Bewiesenen, weil mit $\neg\neg B$ auch B SIP-Formel zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$ ist.

Beweis des Interpolationssatzes 17.3.6 durch Herleitungsinduktion. Zur Abkürzung schreiben wir (in diesem Beweis) Σ für $\Gamma : \Delta$ (auch mit Indizes), und auch

$$C, \Sigma \text{ für } \Gamma : C, \Delta \text{ und } C : \Sigma \text{ für } \Gamma, C : \Delta.$$

1. Σ_1, Σ_2 ist ein logisches Axiom. Dann gibt es vier Fälle.

   1.1 Σ_1 ist logisches Axiom. Setze $B \equiv \top$ (verum). Dann ist

   $$\vdash \top : \Sigma_1 \text{ und } \vdash \top, \Sigma_2 \text{ und } \langle \top \rangle = \emptyset.$$

   1.2 Σ_2 ist logisches Axiom. Setze $B \equiv \bot$ (falsum). Dann ist

   $$\vdash \bot : \Sigma_1 \text{ und } \vdash \bot, \Sigma_2 \text{ und } \langle \bot \rangle = \emptyset.$$

   1.3 Eine Primformel P liegt in $\Delta_1 \cap \Gamma_2$. Setze $B \equiv P$. Dann ist

   $$\vdash P : \Sigma_1 \text{ und } \vdash P, \Sigma_2 \text{ und } \langle P \rangle^{+}_{-} \subseteq \langle \Sigma_1 \rangle^{+}_{-} \cap \langle \Sigma_2 \rangle^{-}_{+}.$$

   1.4 Eine Primformel P liegt in $\Gamma_1 \cap \Delta_2$. Aus dem vorigen Fall folgt mit 17.3.7, dass $B \equiv \neg P$ SIP-Formel ist.

2. Letzter Schluss ist:

$(\rightarrow S)$ $A : C, \Sigma_1, \Sigma_2 \vdash A \rightarrow C, \Sigma_1, \Sigma_2$.

1. Fall: $A \rightarrow C$ wird Σ_1 zugeschlagen, d. h. gesucht wird eine SIP-Formel B zu $A \rightarrow C, \Sigma_1$ und Σ_2. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B zu $A : C, \Sigma_1$ und Σ_2. Mit $(\rightarrow S)$ folgt

$$\vdash B : A \rightarrow C, \Sigma_1 \text{ und } \vdash B, \Sigma_2,$$

und wegen $\langle A \rightarrow C \rangle^{+}_{-} = \langle A : C \rangle^{+}_{-}$ ist B dann auch SIP-Formel zu $A \rightarrow C, \Sigma_1$ und Σ_2.

2. Fall: $A \rightarrow C$ zu Σ_2. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B zu Σ_1 und $A : C, \Sigma_2$. Nach 17.3.7 ist dann $\neg B$ SIP-Formel zu $A : C, \Sigma_2$ und Σ_1, also nach dem 1. Fall auch zu $A \rightarrow C, \Sigma_2$ und Σ_1. Nach 17.3.7 ist wieder B SIP-Formel zu Σ_1 und $A \rightarrow C, \Sigma_2$.

$(\rightarrow A)$ A, Σ_1, Σ_2 und $C : \Sigma_1, \Sigma_2 \vdash A \rightarrow C : \Sigma_1, \Sigma_2$.

1. Fall: $A \rightarrow C$ zu Σ_1. Nach Induktionsvoraussetzung gibt es SIP-Formeln B_1 zu A, Σ_1 und Σ_2 sowie B_2 zu $C : \Sigma_1$ und Σ_2. Also haben wir

(1) $\vdash B_1 : A, \Sigma_1$ und $\vdash B_2, C : \Sigma_1$ und

(2) $\vdash B_1, \Sigma_2$ und $\vdash B_2, \Sigma_2$.

Aus (1) folgt mit (Str)

$$\vdash B_1, B_2 : A, \Sigma_1 \text{ und } \vdash B_1, B_2, C : \Sigma_1$$

und weiter für $B :\equiv B_1 \wedge B_2$ mit $(\wedge A)$

$$\vdash B : A, \Sigma_1 \text{ und } \vdash B, C : \Sigma_1, \text{ und mit } (\to A) \vdash B, A \to C : \Sigma_1.$$

Aus (2) folgt mit $(\wedge S) \vdash B, \Sigma_2$. Schließlich ist

$$\begin{aligned} \langle B \rangle^{+}_{-} = \langle B_1 \rangle^{+}_{-} \cup \langle B_2 \rangle^{+}_{-} &\subseteq (\langle A, \Sigma_1 \rangle^{+}_{-} \cap \langle \Sigma_2 \rangle^{-}_{+}) \cup (\langle C : \Sigma_1 \rangle^{+}_{-} \cap \langle \Sigma_2 \rangle^{-}_{+}) \\ &= \langle A \to C : \Sigma_1 \rangle^{+}_{-} \cap \langle \Sigma_2 \rangle^{-}_{+}. \end{aligned}$$

Also ist $B \equiv B_1 \wedge B_2$ SIP-Formel zu $A \to C : \Sigma_1$ und Σ_2.

2. Fall: $A \to C$ zu Σ_2. Nach Induktionsvoraussetzung gibt es SIP-Formeln B_1 zu Σ_1 und A, Σ_2 sowie B_2 zu Σ_1 und $C : \Sigma_2$. Nach 17.3.7 sind $\neg B_1$ und $\neg B_2$ SIP-Formeln zu den vertauschten Sequenzen. Nach dem 1. Fall ist dann $\neg B_1 \wedge \neg B_2$ SIP-Formel zu $A \to C : \Sigma_2$ und Σ_1. Wieder nach 17.3.7 ist dann $\neg(\neg B_1 \wedge \neg B_2)$ und ebenso $B_1 \vee B_2$ SIP-Formel zu Σ_1 und $A \to B : \Sigma_2$.

$(\forall S)$ $F(a), \Sigma_1, \Sigma_2 \vdash \forall x F(x), \Sigma_1, \Sigma_2$ (a nicht in F, Σ_1, Σ_2).

1. Fall: $\forall x F(x)$ zu Σ_1. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B zu $F(a), \Sigma_1$ und Σ_2, also

$$\vdash B : F(a), \Sigma_1 \text{ und } \vdash B, \Sigma_2,$$

ferner $FV(B) \subseteq FV(\Sigma_2)$. Also tritt a in B nicht auf.
Mit $(\forall S)$ folgt $\vdash B : \forall x F(x), \Sigma_1$, und wegen

$$\langle \forall x F(x) \rangle^{+}_{-} = \langle F(a) \rangle^{+}_{-} - \{a\}$$

ist B auch SIP-Formel zu $\forall x F(x), \Sigma_1$ und Σ_2

2. Fall: $\forall x F(x)$ zu Σ_2. Wie beim 2. Fall von $(\to S)$ sieht man, dass jede SIP-Formel zu Σ_1 und $F(a), \Sigma_2$ auch SIP-Formel zu Σ_1 und $\forall x F(x), \Sigma_2$ ist.

$(\forall A)$ $F(t) : \Sigma_1, \Sigma_2 \vdash \forall x F(x) : \Sigma_1, \Sigma_2$.

1. Fall: $\forall x F(x)$ zu Σ_1. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B' zu $F(t) : \Sigma_1$ und Σ_2. Sei

$$\{a_1, \ldots, a_n\} := FV(t) - FV(F, \Sigma_1) \text{ und } a :\equiv a_1, \ldots, a_n.$$

Die freien Variablen a_i können in B' vorkommen, aber nicht in F oder Σ_1. Es ist $B' \equiv G(a)$ für ein G, in dem die a_i nicht mehr auftreten. Nach Induktionsvoraussetzung ist

$$\vdash F(t), G(a) : \Sigma_1 \text{ und } \vdash G(a), \Sigma_2.$$

$(\forall A)$ macht aus der linken Sequenz $\vdash \forall x F(x), G(a) : \Sigma_1$. Da die a_i nicht in F, G, Σ_1 auftreten, folgt hieraus mit n $(\exists A)$-Schlüssen und ohnehin aus der rechten Sequenz mit n $(\exists S)$-Schlüssen für $B :\equiv \exists y G(y)$

$$\vdash \forall x F(x), B : \Sigma_1 \text{ und } \vdash B, \Sigma_2, \text{ ferner}$$

$$\begin{aligned} \langle B \rangle^{\overset{+}{-}} = \langle G(a) \rangle^{\overset{+}{-}} - \{a\} &\subseteq (\langle F(t) : \Sigma_1 \rangle^{\overset{+}{-}} - \{a\}) \cap \langle \Sigma_2 \rangle^{\overset{-}{+}} \\ &= \langle \forall x F(x) : \Sigma_1 \rangle^{\overset{+}{-}} \cap \langle \Sigma_2 \rangle^{\overset{-}{+}}. \end{aligned}$$

Also ist B SIP-Formel zu $\forall x F(x) : \Sigma_1$ und Σ_2.

2. Fall: $\forall x F(x)$ zu Σ_2. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B' zu Σ_1 und $F(t) : \Sigma_2$. Ist

$$\{a_1, \ldots, a_n\} := FV(t) - FV(F, \Sigma_2), a :\equiv a_1, \ldots, a_n \text{ und } B' \equiv G(a),$$

so ergibt der 1. Fall mit 17.3.7, dass $B \equiv \forall y G(y)$ SIP-Formel zu Σ_1 und $\forall x F(x) : \Sigma_2$ ist.

$(= I)$ $t = t : \Sigma_1, \Sigma_2 \vdash \Sigma_1, \Sigma_2$.

Der Beweis verläuft wegen $\langle t = t \rangle = FV(t)$ analog zum $(\forall A)$-Schluss. Nach Induktionsvoraussetzung gibt es eine SIP-Formel B' zu $t = t : \Sigma_1$ und Σ_2. Für $\{a_1, \ldots, a_n\} := FV(t) - FV(\Sigma_1)$ und $a :\equiv a_1, \ldots, a_n$ sei wieder $B' \equiv G(a)$. Dann ist

$$\vdash t = t,\ G(a) : \Sigma_1 \text{ und } \vdash G(a), \Sigma_2.$$

$(= I)$ macht aus der linken Sequenz $\vdash G(a) : \Sigma_1$. Da die a_i nicht in G, Σ_1 auftreten, folgt wie oben für $B :\equiv \exists y G(y)$.

$$\vdash B : \Sigma_1 \text{ und } \vdash B, \Sigma_2,$$

und wie oben erweist sich B als SIP-Formel zu Σ_1 und Σ_2.

$(= F)$ $fs_1 \dots s_n = ft_1 \dots t_n : \Sigma_1, \Sigma_2 \vdash s_1 = t_1, \dots, s_n = t_n : \Sigma_1, \Sigma_2$

$(= P)$ $pt_1 \dots t_n : \Sigma_1, \Sigma_2 \vdash s_1 = t_1, \dots, s_n = t_n, ps_1 \dots s_n : \Sigma_1, \Sigma_2$.

1. Fall: Es sei $s^i = t^i$ jeweils das Teiltupel von $s_1 = t_1, \dots, s_n = t_n$, das zu Σ_i geschlagen wird ($i = 1, 2$; die beiden Teiltupel brauchen nicht disjunkt zu sein). Ferner komme bei $(= P)$ $ps_1 \dots s_n$ zu Σ_1. Nach Induktionsvoraussetzung gibt es eine SIP-Formel D zu $X : \Sigma_1$ und Σ_2, wobei $X \equiv fs_1 \dots s_n = ft_1 \dots t_n$ für $(= F)$ und $X \equiv pt_1 \dots t_n$ für $(= P)$ ist. Also ist

$$\vdash D, X : \Sigma_1 \vdash D, s^1 = t^1, s^2 = t^2, Y : \Sigma_1$$

woraus mit $(\wedge A)$-Schlüssen folgt:

$$\vdash \wedge(s^2 = t^2) \wedge D, s^1 = t^1, Y : \Sigma_1$$

worin Y fehlt bei $(= F)$, $Y \equiv ps_1 \dots s_n$ bei $(= P)$ und $\wedge(s^2 = t^2)$ die Konjunktion aller Gleichungen des Tupels $s^2 = t^2$ bezeichnet; ferner $\vdash D, \Sigma_2$, woraus mit (Str) und $(\wedge S)$-Schlüssen folgt:

$$\vdash s^2 = t^2 : \wedge(s^2 = t^2) \wedge D, \Sigma_2.$$

Für $\{a_1, \dots, a_m\} := FV(s^2 = t^2) - FV(s^1 = t^1, \Sigma_1, Y)$ und $a = a_1, \dots, a_m$ ist $\wedge(s^2 = t^2) \wedge D \equiv G(a)$ für ein G, in dem die Variablen a_i nicht mehr auftreten. Wir haben dann

$$\vdash G(a), s^1 = t^1, Y : \Sigma_1 \text{ und } \vdash s^2 = t^2 : G(a), \Sigma_2,$$

und mit $(\exists A)$- bzw. $(\exists S)$-Schlüssen folgt

$$\vdash \exists y G(y), s^1 = t^1, Y : \Sigma_1 \text{ und } \vdash s^2 = t^2 : \exists y G(y), \Sigma_2.$$

Dann ist $B :\equiv \exists y G(y)$ SIP-Formel zu $s^1 = t^1, Y : \Sigma_1$ und $s^2 = t^2 : \Sigma_2$.

Mit dem 1. Fall ist $(= F)$ vollständig behandelt.

2. Fall für $(= P) : ps_1 \ldots s_n$ zu Σ_2. Nach Induktionsvoraussetzung gibt es eine SIP-Formel D zu Σ_1 und $pt_1 \ldots t_n : \Sigma_2$. Nach 17.3.7 ist $\neg D$ SIP-Formel zu $pt_1 \ldots t_n : \Sigma_2$ und Σ_1. Nach dem 1. Fall und 17.3.7 ist dann eine Formel $\forall y \neg G(y)$ mit $G(a) \equiv \wedge(s^1 = t^1) \wedge \neg D$, also auch die Formel $\forall y F(y)$ mit $F(a) \equiv (s^1 = t^1 \rightarrow D)$ eine SIP-Formel zu $s^1 = t^1 : \Sigma_1$ und $s^2 = t^2, ps_1 \ldots s_n : \Sigma_2$.

Da nur logische Theorien betrachtet werden, ist damit die Induktion abgeschlossen und Satz 17.3.6, also auch Craig's Lemma 17.3.2 bewiesen.

Als wichtigen Spezialfall notieren wir

17.3.8 Korollar Ist in einer logischen Theorie $\vdash A : C$, dann gibt es eine Formel B, so dass

1. $\vdash A : B$ und $\vdash B : C$
2. $\langle B \rangle^{+}_{-} \subseteq \langle A \rangle^{+}_{-} \cap \langle C \rangle^{+}_{-}$, also erst recht

2'. $\langle B \rangle \subseteq \langle A \rangle \cap \langle C \rangle$.

Denn nach 17.3.6 gibt es eine SIP-Formel B zu $\emptyset : C$ und $A : \emptyset$, und für diese gilt 1. und wegen $\langle A : \emptyset \rangle^{-}_{+} = \langle A \rangle^{+}_{-}$ auch 2.

Wir ziehen noch einige Folgerungen aus Craig's Lemma und dem Satz von Craig und Lyndon.

17.4 Vereinte Konsistenz

17.4.1 Definition Unter der *Vereinigung* $T_1 \cup T_2$ von zwei Theorien T_1 und T_2 versteht man die Theorie T, deren Sprache $L(T)$ die kleinste Sprache ist, die sowohl $L(T_1)$ als auch $L(T_2)$ enthält, und deren Axiomensystem $Ax(T) = Ax(T_1) \cup Ax(T_2)$ ist.

17.4.2 Lemma T_1 und T_2 seien Theorien, deren Sprachen $L(T_1)$ und $L(T_2)$ dieselben Funktionszeichen enthalten. Ist $T_1 \cup T_2$ inkonsistent, so gibt es einen Satz B aus $L(T_1) \cap L(T_2)$, so dass

$$T_1 \vdash B \text{ und } T_2 \vdash \neg B.$$

Beweis. Wenn $T_1 \cup T_2 \vdash \square$, dann gibt es nach dem verallgemeinerten Deduktionstheorem je eine endliche Menge Γ_i von Axiomen von T_i $(i = 1, 2)$, so dass

$$T_0 := (L(T_1 \cup T_2), \emptyset) \vdash \Gamma_1, \Gamma_2 : \emptyset.$$

Nach Craig's Lemma 17.3.2 gibt es dann eine IP-Formel B zu $\Gamma_2 : \emptyset$ und $\Gamma_1 : \emptyset$, also

1. $T_0 \vdash \Gamma_2, B : \emptyset$ und $T_0 \vdash \Gamma_1 : B$
2. $\langle B \rangle \subseteq \langle \Gamma_2 : \emptyset \rangle \cap \langle \Gamma_1 : \emptyset \rangle$.

Zu 2: Da $\Gamma_i \subseteq Ax(T_i)$ ist, ist B ein Satz aus $L(T_1) \cap L(T_2)$, weil beide Sprachen dieselben Funktionszeichen enthalten.

Zu 1: Wieder wegen $\Gamma_i \subseteq Ax(T_i)$ folgt

$$T_0 + Ax(T_1) \vdash B \text{ und } T_0 + Ax(T_2) \vdash \neg B.$$

Nach dem Subformelprinzip 3.3 können in einer (schnittfreien) Herleitung nur Prädikatszeichen (außer $=$) auftreten, die in den verwendeten Axiomen Γ_i oder in der Endformel B bzw. $\neg B$ auftreten. Also ist schon

$$T_1 \vdash B \text{ und } T_2 \vdash \neg B.$$

Durch einfache Kontraposition erhält man hieraus ein wichtiges Ergebnis der Modelltheorie:

17.4.3 Satz von A. Robinson über vereinte Konsistenz
T_1 und T_2 seien Theorien, deren Sprachen dieselben Funktionszeichen enthalten. Wenn es zu jedem Satz B von $L(T_1) \cap L(T_2)$ ein Modell $\mathcal{A}_1$ von T_1 gibt, in dem B nicht gilt, oder ein Modell $\mathcal{A}_2$ von T_2 gibt, in dem B gilt, dann hat $T_1 \cup T_2$ ein Modell.

Beweis. Angenommen, $T_1 \cup T_2$ hat kein Modell. Nach dem Vollständigkeitssatz ist dann $T_1 \cup T_2$ inkonsistent. Nach 17.4.2 gibt es dann einen Satz B aus $L(T_1) \cap L(T_2)$ mit $T_1 \vdash B$ und $T_2 \vdash \neg B$. Dieser Satz B gilt also in jedem Modell von T_1 und ist falsch in jedem Modell von T_2, und das ist die Negation unserer Voraussetzung.

Kontraposition ergibt nun die Behauptung.

17.5 Implizite und explizite Definierbarkeit

17.5.1 Lemma Sei $\Gamma : \Delta$ eine Sequenzen-Nennform aus $L(T)$ und q ein neues Prädikatszeichen, d. h. $q \notin L(T)$, von derselben Stellenzahl wie $p \in L(T)$.

$$\text{Aus } T + \{q\} \vdash \Gamma(q) : \Delta(q) \text{ folgt dann } T \vdash \Gamma(p) : \Delta(p).$$

Der **Beweis** durch Herleitungsinduktion ist dem des Lemmas über neue Konstanten analog. Da q in der gegebenen Herleitung nicht in Axiomen von T auftritt, kann man q überall durch p ersetzen (aber nicht umgekehrt), ohne Axiome oder Grundschlüsse zu verletzen. q verhält sich quasi wie eine freie Prädikatsvariable – auch wenn es die in der Logik 1. Stufe nicht gibt.

Hiermit folgt aus Craig's Lemma:

17.5.2 Definierbarkeitssatz von Beth. Es sei $F(p)$ ein Satz aus $L(T)$, das n-stellige nicht-logische Prädikatszeichen p trete in F nicht auf, und q sei ein n-stelliges neues Prädikatszeichen. Wenn

(1) $T + \{q\} \vdash F(p) \wedge F(q) \to pa_1 \dots a_n \to qa_1 \dots a_n$,

dann gibt es eine Formel B aus $L(T)$ mit $FV(B) \subseteq \{a_1, \dots, a_n\}$, in der auch p nicht auftritt, so dass

(2) $T \vdash F(p) \to (pa_1 \dots a_n \leftrightarrow B)$:

Jedes implizit in T definierbare Prädikat ist auch explizit definierbar.

Beweis. Nach einigen Inversionsschlüssen folgt mit dem verallgemeinerten Deduktionstheorem aus (1), dass es eine endliche Menge Γ von Axiomen von T gibt, so dass

$$(L(T) + \{q\}, \emptyset) \vdash \Gamma, F(p), pa_1 \dots a_n, F(q) : qa_1 \dots a_n.$$

Nun können wir Craig's Lemma auf die Teilsequenzen $F(q) : qa_1 \dots a_n$ und $\Gamma, F(p), pa_1 \dots a_n : \emptyset$ anwenden: Es gibt eine Formel B mit $FV(B) \subseteq \{a_1, \dots, a_n\}$, in der weder p noch q auftritt, so dass

(3) $(L(T) + \{q\}, \emptyset) \vdash F(q), B : qa_1 \dots a_n$ und

(4) $(L(T) + \{q\}, \emptyset) \vdash \Gamma, F(p), pa_1 \dots a_n : B$.

Mit 17.5.1 (und dem Deduktionstheorem) folgt hieraus:

(3') $T \vdash F(p), B : pa_1 \dots a_n$ und

(4') $T \vdash F(p), pa_1 \dots a_n : B$.

Hieraus schließt man mit $(\to S)$ und $(\wedge S)$ auf die Behauptung (2).

Bemerkung. Wegen (1) *definiert* (in T) der Satz $F(p)$ *das Prädikat p implizit.* Tatsächlich folgt in jedem Modell $\mathcal{A} \models T$ aus $\mathcal{A} \models F(p)$, dass das Prädikat $p_\mathcal{A}$ in *jeder* Teilmenge $X \subseteq |\mathcal{A}|^n$ enthalten ist, für die $\mathcal{A} \models F(X)$, also

$$p_\mathcal{A} = \bigcap \{X \subseteq |\mathcal{A}|^n | \mathcal{A} \models F(X)\}.$$

(Hier steht X für die Interpretation von q, die wegen $q \notin L(T)$ nicht in $\mathcal{A}$, sondern erst in geeigneten Expansionen von $\mathcal{A}$ festgelegt ist.) Dies ist eine zweitstufige Definition von $p_\mathcal{A}$, und es ist von vornherein nicht klar, dass es in $L(T) - \{p\}$ eine Formel $B \equiv G(a_1, \dots, a_n)$ gibt, die das Prädikat p in $T + \{F(p)\}$ explizit gemäß (2) definiert: Uniform für alle Modelle $\mathcal{A} \models T + F(p)$ ist

$$p_\mathcal{A} = \{(k_1, \dots, k_n) \in |\mathcal{A}|^n | \mathcal{A} \models G(k_1, \dots, k_n)\}.$$

Beispiel. Ist $\forall x \forall y(px \wedge py \to x = y)$ ein Axiom von T und ist $F(p) \equiv pc$ (c eine Konstante aus $L(T)$), dann folgt (1), weil in $T + \{F(p)\}$ gilt, dass c das einzige Element ist, auf das p zutrifft, und wegen $F(q)$ das neue Prädikat q mindestens auf c zutrifft. Die gesuchte explizite Definition B ist dann natürlich $a = c$.

17.6 Monotonie und Positivität

17.6.1 Definition Seien p und q n-stellige Prädikatszeichen. Man nennt eine Formel B *p-positiv*, wenn p in B nur positiv auftritt, d. h. wenn $p \notin \langle B \rangle^-$ ist. Seien nun $p, q \notin L(T)$. Eine Nennform F aus $L(T)$ heißt (in T beweisbar) *monoton*, wenn für $x :\equiv x_1 \dots x_n$

(1) $T + \{p, q\} \vdash \forall x(px \to qx) \to F(p) \to F(q)$.

In dieser Situation nennt man auch die Formel $F(p)$ (in $T + \{p\}$ beweisbar) *monoton in p.*

Man überlegt sich leicht, dass p-positive Formeln auch (logisch beweisbar) monoton in p sind. Die Umkehrung hiervon folgt aus dem Interpolationssatz von Craig und Lyndon, sogar in etwas allgemeinerer Form.

17.6.2 Satz Seien p und q n-stellige nicht-logische Prädikatszeichen, $p \in L(T)$, $q \notin L(T)$ und sei F eine Nennform, in der p und q nicht auftreten. Wenn mit $x :\equiv x_1 \dots x_n$

(2) $T + \{q\} \vdash \forall x(px \to qx) \to F(p) \to F(q)$,

dann gibt es eine p-positive Formel B aus $L(T)$, so dass

(3) $T \vdash F(p) \leftrightarrow B$.

Insbesondere:
Beweisbar in p monotone Formeln sind beweisbar äquivalent zu p-positiven Formeln.

(Da p in $Ax(T)$ auftreten darf, ist die Voraussetzung (2) allgemeiner als die beweisbare Monotonie von F, die in (1) formuliert ist.

Beweis. Mit $(\to SInv)$ und dem Deduktionstheorem folgt aus (2), dass es eine endliche Menge Γ von Axiomen von T gibt, so dass

$$(L(T) + \{q\}, \emptyset) \vdash \Gamma, F(p), \forall x(px \to qx) : F(q).$$

Nach dem Interpolationssatz 17.3.6 gibt es dann eine SIP-Formel B zu $\forall x(px \to qx) : F(q)$ und $\Gamma, F(p) : \emptyset$. Das bedeutet:

(4) $(L(T) + \{q\}, \emptyset) \vdash B, \forall x(px \to qx) : F(q)$

(5) $(L(T) + \{q\}, \emptyset) \vdash \Gamma, F(p) : B$.

Weil q nicht in $Ax(T)$, also nicht in Γ und auch nicht in $F(p)$ auftritt, tritt q auch nicht in B auf. Weil p nur positiv in $\forall x(px \to qx) : F(q)$ auftritt, tritt p auch nur positiv in B auf, B ist p-positiv. Aus (5) folgt

(5') $T + \{q\} \vdash F(p) : B$.

Ersetzt man in (4) (und (5')) q durch p, so folgt mit 17.5.1

(4') $T \vdash B, \forall x(px \to px) : F(p)$,

also mit einem Schnitt, da $T \vdash \forall x(px \to px)$,

(4*) $T \vdash B : F(p)$

(5*) $T \vdash F(p) : B$,

woraus sich mit $(\to S)$, $(\wedge S)$ die Behauptung (3) ergibt.

17.7 Aufgaben

17.7.1 Sei f ein Funktionszeichen $\notin L(T)$ und $\Gamma : \Delta$ eine Sequenz aus $L(T)$, die in $T + \{f\}$ herleitbar ist. Zeigen Sie:

a. Es gibt eine Herleitung von $\Gamma : \Delta$ in $T + \{f\}$, in der f auftritt.

b. $T \vdash \Gamma : \Delta$.

(Hinweis zu b.: Ersetzen Sie in der gegebenen Herleitung von $\Gamma : \Delta$ jeden Term, der mit f anfängt, durch eine freie Variable, die in der Herleitung nicht auftritt.)

17.7.2 Geben Sie (in Abgrenzung zu 17.1.8) zu jedem $k > 0$ eine endlich axiomatisierte Theorie T_k an, die Modelle jeder Mächtigkeit $\leq k$ und $\geq \aleph_0$, aber keine endlichen Modelle mit mehr als k Elementen besitzt.

17.7.3 Zeigen Sie, dass in einfachen, konsistenten, offenen Erweiterungen von LO der Satz $\exists x \exists y\, x < y$ nicht herleitbar ist.

17.7.4 Sei f ein einstelliges Funktionszeichen, $f^k t$ sei $\underbrace{f \ldots f}_{k\text{-mal}} t$, und für $k > 1$ sei T_k die Theorie

$$T_k := LO + \{f\} + \{\forall x fx \neq x, \forall x f^k x = x\}.$$

a. Zeigen Sie: $T_k \vdash \exists x fx < x$

b. Geben Sie eine in T_k herleitbare Herbrand-Instanz $(\exists x fx < x)^H$ minimaler Länge an.

17.7.5 Zeigen Sie für ein einstelliges Prädikatszeichen p :

a. $\vdash a = b, pa : pb$

b. Jede IP-Formel B zu $pa : pb$ und $a = b : \emptyset$ enthält das Gleichheitszeichen: „$=$ lässt sich nicht interpolieren.“

17.7.6 Zeigen Sie durch Auswertung des Beweises von 17.3.6: Ist eine identitätsfreie Sequenz $\Gamma_1, \Gamma_2 : \Delta_1, \Delta_2$ logisch herleitbar, so gibt es eine identitätsfreie SIP-Formel zu $\Gamma_1 : \Delta_1$ und $\Gamma_2 : \Delta_2$.

17.7.7 Konstruieren Sie aus der nächstliegenden logischen Herleitung der Sequenz $\forall xpx : pt$ gemäß dem Beweis von 17.3.6 eine SIP-Formel zu $\emptyset : pt$ und $\forall xpx : \emptyset$.

17.7.8 Zeigen Sie: Ist B p-positiv, so ist B (logisch beweisbar) monoton in p.

§18 Definitorische Erweiterungen und Skolemisierung

18.1 Erweiterungen und Expansionen

18.2 Definitorische Erweiterungen um Prädikatszeichen

18.3 Definitorische Erweiterungen um Funktionszeichen

18.4 Pränexe Normalformen

18.5 Skolemisierung

18.6 Aufgaben

Es gibt eine Reihe von Techniken in der Prädikatenlogik, die beim Arbeiten mit mathematischen Theorien nützlich sind und in der Mathematik häufig ohne weitere Begründung eingesetzt werden. Einige dieser Techniken wollen wir hier studieren. Wegen ihres syntaktischen Charakters gehören sie zur Beweistheorie der Prädikatenlogik, aber zu ihrer Begründung trägt die Schnittfreiheit des Herleitungsbegriffs nicht bei. Manchmal sind es gerade semantische Methoden, die die Beweise verkürzen. In diesem Paragraphen verwenden wir deshalb syntaktische und semantische Methoden nebeneinander. Kürze und Durchsichtigkeit erhalten Vorrang vor der Forderung nach methodischer Reinheit.

18.1 Erweiterungen und Expansionen

Als ersten Komplex untersuchen wir Erweiterungen von Theorien um Definitionen.

Im 3. Kapitel haben wir den Begriff der Erweiterung sehr eng gefasst, nämlich als *Axiom-Erweiterung*:

> Ist $T' \succ T$, so ist jedes Axiom von T auch Axiom von T'.

Das war zum Beweis des Vollständigkeitssatzes notwendig, solange die Zulässigkeit der Schnittregel nicht bewiesen war. Wir lockern den Begriff jetzt etwas auf:

18.1.1 Definition T' ist $\vdash$-*Erweiterung (Herleitbarkeitserweiterung)* von T, wir schreiben $T' \vdash T$, wenn

1. $L(T) \subseteq L(T')$ ist und
2. $T' \vdash B$ für jedes $B \in Ax(T)$.

Ferner nennen wir T und T' $\vdash$-*äquivalent*, wenn $T' \vdash T$ und $T \vdash T'$.

Ist $T' \vdash T$, so ist jede in T herleitbare Sequenz auch in T' herleitbar. Ist dagegen $T' \succ T$, so ist sogar jede Herleitung in T auch eine Herleitung in T'. Also:

$$\text{Aus } T' \succ T \text{ folgt } T' \vdash T,$$

aber nicht umgekehrt.

18.1.2 Expansions-Lemma Es sei $T' \vdash T$. Wenn sich jedes Modell von T zu einem Modell von T' expandieren lässt, ist T' konservativ über T.

Beweis. Sei $\Gamma : \Delta$ eine Sequenz aus $L(T)$ und $T' \vdash \Gamma : \Delta$. Zu zeigen ist $T \vdash \Gamma : \Delta$. Sei $\mathcal{A} \models T$ und $\mathcal{B}$ eine Expansion von $\mathcal{A}$, die Modell von T' ist. Ein solches $\mathcal{B}$ gibt es nach Voraussetzung.

Aus $T' \vdash \Gamma : \Delta$ folgt $\mathcal{B} \models \Gamma : \Delta$ nach dem Korrektheitssatz. Weil $\Gamma : \Delta$ aus $L(T)$ ist, folgt wie in 8.2.4 auch $\mathcal{A} = \mathcal{B}|L(T) \models \Gamma : \Delta$. Da dies für jedes Modell $\mathcal{A} \models T$ gilt, folgt mit dem Vollständigkeitssatz $T \vdash \Gamma : \Delta$.

Die Umkehrung dieses Lemmas ist falsch, wie man sich leicht überlegt (vgl. hierzu die Aufgaben 18.6.1 und 2).

18.2 Definitorische Erweiterungen um Prädikatszeichen

Ein Beispiel. In 14.1 haben wir die Kleiner-Relation $<$ in der gewöhnlichen Zahlentheorie Z definiert durch die Äquivalenz

$$a < b \leftrightarrow \exists x(a + Sx = b).$$

Wir haben dort die linke Seite $a < b$ als (externe) Abkürzung für die rechte Seite aufgefasst.

Die elegantere, logik-interne Lösung ist, das Zeichen $<$ als neues Prädikatszeichen zu $L(Z)$ hinzuzunehmen. Dann muss man als erstes sicherstellen, dass

$$Z + \{<\} + \{a < b \leftrightarrow \exists x(a + Sx = b)\}$$

eine konservative Erweiterung von Z ist.

Wir behandeln dieses Problem allgemein.

18.2.1 Definition Das n-stellige Prädikatszeichen q sei nicht in $L(T)$, und G sei eine geschlossene Nennform aus $L(T)$. Dann heißt $T' := T + \{q\} + \{\forall(1)\}$ mit

(1) $qa_1 \dots a_n \leftrightarrow G(a_1, \dots, a_n)$

definitorische Erweiterung von T *um die Definition* (1) von q.

18.2.2 Lemma Die definitorische Erweiterung T' von T um die Definition (1) von q ist eine konservative Erweiterung von T.

Beweis. Sei $\mathcal{A} \models T$. Wir setzen

$$q_{\mathcal{B}} := \{(k_1, \dots, k_n) \in |\mathcal{A}|^n | \mathcal{A}(G(k_1, \dots, k_n)) = w\},$$

ferner $\mathcal{B}|L(T) := \mathcal{A}$. Dann ist $\mathcal{B}$ eine Expansion von $\mathcal{A}$ zu einer Struktur zu $L(T) + \{q\}$, in der (1) gilt. Also ist $\mathcal{B} \models T'$. Mit 18.1.2 folgt die Behauptung.

18.2.3 Rekursive Definition *der Eliminationsübersetzung* $^q : C \mapsto C^q$ von $L(T')$ in $L(T)$. Sei T' die definitorische Erweiterung von T um die Definition (1) von q.

1. $(qt_1 \dots t_n)^q :\equiv G(t_1, \dots, t_n)$
2. $P^q :\equiv P$ für andere Primformeln P
3. $(A \to B)^q :\equiv A^q \to B^q$
4. $(\forall x F(x))^q :\equiv \forall y F^q(y)$ mit y nicht in $F^q(a) :\equiv F(a)^q$.

Hiernach geht C^q aus C hervor, indem man jede Primformel qt durch $G(t)$ ersetzt und, soweit nötig, gebunden umbenennt. In 4. wird man $y \equiv x$ wählen, solange x nicht in F^q auftritt.

18.2.4 Satz Sei T' definitorische Erweiterung von T um die Definition (1) von q.

1) Für Formeln $C \in L(T')$ ist $C^q \in L(T)$. Für $C \in L(T)$ kann $C^q \equiv C$ gewählt werden; dann ist stets $C^{qq} \equiv C^q$.

2) $T' \vdash C \leftrightarrow C^q$.

3) $T' \vdash C \Leftrightarrow T \vdash C^q$.

Beweis. 1) ist klar – jedenfalls nach der Bemerkung über sparsame gebundene Umbenennung im Fall 4. von 18.2.3.

Beweis von 2) durch Induktion nach dem Aufbau von C.

1. Ist $C \equiv qt_1 \dots t_n$, so ist $T' \vdash C \leftrightarrow C^q$, weil dies ein Einsetzungsfall des Axioms (1) von T' ist.

2. Ist C eine andere Primformel, so ist $C^q \equiv C$, und $C \leftrightarrow C^q$ ist eine Tautologie.

3. Ist $C \equiv A \to B$, so ist nach Induktionsvoraussetzung

$$T' \vdash A \leftrightarrow A^q \text{ und } T' \vdash B \leftrightarrow B^q,$$

und da $(A \leftrightarrow A^q) \to (B \leftrightarrow B^q) \to (C \leftrightarrow C^q)$ eine Tautologie ist, folgt hieraus (mit Schnitten) die Behauptung.

4. Ist $C \equiv \forall x F(x)$, so ist nach Induktionsvoraussetzung

$$T' \vdash F(a) \leftrightarrow F^q(a) \quad (a \text{ nicht in } F, F^q),$$

woraus mit Quantorenverteilung die Behauptung folgt.

3) Ist $T' \vdash C$, so ist wegen 2) auch $T' \vdash C^q$. Da $C^q \in L(T)$ ist, folgt $T \vdash C^q$ aus 18.2.2. Ist umgekehrt $T \vdash C^q$, so ist erst recht $T' \vdash C^q$, also wegen 2) auch $T' \vdash C$.

Die Definition (1) von q lässt sich also aus der Theorie T' mittels der Übersetzung ^{q} wieder eliminieren: Die Theorie T und ihre definitorische Erweiterung T' sind gleich ausdrucksstark.

18.3 Definitorische Erweiterungen um Funktionszeichen

Ergebnisse, die denen von 18.2 entsprechen, lassen sich auch für Funktionen beweisen. Allerdings ist der Aufwand dafür etwas größer. Während man jede Formel $G(a)$ für die Definition einer neuen Relation heranziehen kann (vgl. 18.2.1), kann man neue Funktionen nur durch solche Formeln $G(a, b)$ definieren, die in T erkennbar den Graphen einer Funktion beschreiben. Ferner wird die Eliminations-Übersetzung komplizierter, weil Funktionszeichen (in Termen) geschachtelt auftreten können, Prädikatszeichen (in Primformeln) aber nicht.

18.3.1 Definition Das n-stellige Funktionszeichen f sei nicht in $L(T)$, und G sei eine geschlossene Nennform aus $L(T)$. Wenn

$$T \vdash \exists! y G(a_1, \ldots, a_n, y),$$

dann heißt $T' := T + \{f\} + \{\forall(2)\}$ mit

$$(2)\ f a_1 \ldots a_n = b \leftrightarrow G(a_1, \ldots, a_n, b)$$

definitorische Erweiterung von T *um die Definition* (2) von f.

Die Definition (2) von f besagt nur, dass man *das* in T eindeutig bestimmte y, für das $G(a_1, \ldots, a_n, y)$ – in Abhängigkeit von $a_1, \ldots, a_n$ – gilt, mit $f a_1 \ldots a_n$ bezeichnet. Von daher ist zu erwarten:

18.3.2 Lemma Die definitorische Erweiterung T' von T um die Definition (2) von f ist eine konservative Erweiterung von T.

Beweis. Sei $\mathcal{A} \models T$. Für $k_1, \ldots, k_n \in |\mathcal{A}|$ ist dann nach der Voraussetzung über T $\mathcal{A}(\exists! y G(k_1, \ldots, k_n, y)) = w$, d. h. es gibt genau ein $l \in |\mathcal{A}|$, so dass

$$\mathcal{A}(G(k_1, \ldots, k_n, l) = w)$$

ist. Wir setzen $f_{\mathcal{B}}(k_1, \ldots, k_n)$ gleich diesem l, ferner $\mathcal{B}|L(T) := \mathcal{A}$. Dann ist $\mathcal{B}$ eine Expansion von $\mathcal{A}$ zu einer Struktur zu $L(T) + \{f\}$, in der (2) gilt. Also ist $\mathcal{B} \models T'$. Mit 18.1.2 folgt die Behauptung.

18.3.3 Definition Sei f ein n-stelliges Funktionszeichen. Eine Formel C ist *f-einfach*, wenn f in C höchstens in Gleichungen $fs_1 \dots s_n = t$ und darin auch nur einmal auftritt (d. h. f tritt nicht mehr in $s_1, \dots, s_n, t$ auf).

Wir geben ein Verfahren an, mit dem man eine beliebige Formel in eine f-einfache Formel äquivalent umformt.

18.3.4 Rekursive Definition der *f-Vereinfachung* $^1 : C \mapsto C^1$.

1. Ist P eine Primformel, in der f nicht auftritt, so sei $P^1 :\equiv P$.

2.1 Tritt f in $s_1, \dots, s_n, t$ nicht auf, so sei

$$(fs_1 \dots s_n = t)^1 :\equiv fs_1 \dots s_n = t.$$

2.2 Tritt f nicht in t, aber in mindestens einem s_i auf, so sei

$$(gs_1 \dots s_m = t)^1 :\equiv \exists y_1 \dots \exists y_m((s_1 = y_1)^1 \wedge \dots \wedge (s_m = y_m)^1 \wedge gy_1 \dots y_m = t).$$

2.3 Tritt f in t auf, so sei

$$(s = t)^1 :\equiv \exists y((s = y)^1 \wedge (t = y)^1).$$

3. Ist p nicht das Gleichheitszeichen und tritt f in mindestens einem t_i auf, so sei

$$(pt_1 \dots t_m)^1 :\equiv \exists y_1 \dots \exists y_m((t_1 = y_1)^1 \wedge \dots \wedge (t_m = y_m)^1 \wedge py_1 \dots y_m).$$

4. $(A \to B)^1 :\equiv A^1 \to B^1$.

5. $(\forall x F(x))^1 :\equiv \forall y F^1(y)$ mit y nicht in $F^1(a) :\equiv F(a)^1$.

18.3.5 Lemma

1) C^1 ist stets f-einfach. Ist C f-einfach, so kann $C^1 \equiv C$ gewählt werden; dann ist stets $C^{11} \equiv C^1$.

2) Logisch gilt $C \leftrightarrow C^1$.

Beweis. 1) Zunächst überzeugt man sich, dass C^1 durch 18.3.4 für jede Formel C aus $L(T)+\{f\}$ definiert ist. Dazu induziert man sowohl nach der Länge der Terme als auch nach der Länge der Formeln (wobei wegen 2.3 allerdings ein Term rechts vom Gleichheitszeichen mehr zählt als links davon). Indem man die Fälle der Definition 18.3.4 durchgeht, erkennt man dann, dass C^1 stets f-einfach ist, wobei man ab dem Fall 2.2 auf die Induktionsvoraussetzung zurückgreift.

Ist C selbst schon f-einfach, so kommen die Fälle 2.2, 2.3 und 3 nicht zum Zuge: In C^1 werden keine zusätzlichen Quantoren $\exists y$ eingeführt, so dass in 5 $y \equiv x$ die natürliche Wahl ist, und es wird $C^1 \equiv C$.

Da in jedem Fall C^1 f-einfach ist, ist danach $C^{11} \equiv C^1$ für jede Formel C.

2) Man überlegt sich, dass

$$F(t_1, \ldots, t_n) \leftrightarrow \exists y_1 \ldots \exists y_n (t_1 = y_1 \wedge \ldots \wedge t_n = y_n \wedge F(y_1, \ldots, y_n))$$

logisch gilt. Indem man wieder die Fälle von 18.3.4 durchgeht, folgt hieraus und etwa dem Äquivalenzsatz 6.2.3 mit Induktion, dass $C \leftrightarrow C^1$ für jede Formel C logisch gilt.

Es sind die f-einfachen Formeln, die eine Eliminationsübersetzung analog zu 18.2.3 gestatten, weil in ihnen f nicht geschachtelt auftritt. Wegen 18.3.5 ist das offenbar keine wirkliche Einschränkung.

18.3.6 Rekursive Definition *der Eliminationsübersetzung* ${}^f : C \mapsto C^f$ von $L(T')$ in $L(T)$. Dabei sei T' die definitorische Erweiterung von T um die Definition (2) von f. Zunächst sei C f-einfach.

1. $(f s_1 \ldots s_n = t)^f :\equiv G(s_1, \ldots, s_n, t)$, falls f nicht in $s_1, \ldots, s_n, t$ auftritt.
2. $P^f :\equiv P$ für Primformeln P, in denen f nicht auftritt.
3. $(A \to B)^f :\equiv A^f \to B^f$.
4. $(\forall x F(x))^f :\equiv \forall y F^f(y)$ mit y nicht in $F^f(a) :\equiv F(a)^f$.
5. Ist C nicht f-einfach, so sei $C^f :\equiv C^{1f}$.

Hiermit ist nach 18.3.5, 1) C^f für alle Formeln C aus $L(T')$ definiert, und bis auf gebundene Umbenennung ist stets $C^f \equiv C^{1f}$. Insgesamt geht C^f aus

C hervor, indem man, soweit nötig, C zunächst f-einfach macht und in dem f-einfachen C^1 jede Gleichung $fs = t$ durch $G(s,t)$ ersetzt. In 4. wird man wieder $y \equiv x$ wählen, solange x nicht in F^f auftritt.

18.3.7 Satz Sei T' die definitorische Erweiterung von T um die Definition (2) von f.

1) Für Formeln $C \in L(T')$ ist $C^f \in L(T)$. Für $C \in L(T)$ kann $C^f \equiv C$ gewählt werden; dann ist stets $C^{ff} \equiv C^f$.

2) $T' \vdash C \leftrightarrow C^f$.

3) $T' \vdash C \Leftrightarrow T \vdash C^f$.

Beweis. 1) ergibt sich aus 18.3.5, 1), weswegen auch C^f stets definiert ist, und wieder bei sparsamer gebundener Umbenennung.

Beweis von 2) durch Induktion nach dem Aufbau von C, zunächst für f-einfaches C.

1. Ist C eine Gleichung $fs_1 \dots s_n = t$, so ist $T' \vdash C \leftrightarrow C^f$, weil das ein Einsetzungsfall des Axioms (2) von T' ist.

2. bis 4. kann man (mit ^{f} anstelle von ^{q}) aus dem Beweis von 18.2.4 übernehmen.

5. Ist C nicht f-einfach, so gilt $C \leftrightarrow C^1$ logisch, $T' \vdash C^1 \leftrightarrow C^{1f}$ nach 1. bis 4., und wegen $C^f \equiv C^{1f}$ folgt $T' \vdash C \leftrightarrow C^f$.

Mit Induktion folgt 2).

3) Ist $T' \vdash C$, so ist wegen 2) auch $T' \vdash C^f$. Da $C^f \in L(T)$ ist, folgt $T \vdash C^f$ aus 18.3.2. Ist umgekehrt $T \vdash C^f$, so ist erst recht $T' \vdash C^f$, also wegen 2) auch $T' \vdash C$.

Analog zu 18.2 lässt sich auch die Definition (2) aus der Theorie T' mittels ^{f} eliminieren: Die Theorie T und ihre definitorische Erweiterung T' sind gleich ausdrucksstark.

18.3.8 Definition Wir nennen T' eine definitorische $\vdash$-Erweiterung von T, wenn es Theorien $T_0, \ldots, T_n$ gibt, so dass $T_0 = T$ ist, für $i < n$ T_{i+1} eine definitorische Erweiterung von T_i um eine Definition (1) oder (2) ist und schließlich T_n und T' $\vdash$-äquivalent sind.

Beispiel. In der Algebra werden die Gruppen oft als die Modelle einer Theorie T_0 eingeführt, deren Sprache allein durch ein zweistelliges Funktionszeichen $\circ$ gegeben ist und die die Axiome

$$a \circ (b \circ c) = (a \circ b) \circ c$$
$$\exists z \forall x \exists y (x \circ z = x \wedge x \circ y = z)$$

besitzt. In T_0 ist herleitbar

$$\exists! z \forall x \exists y (x \circ z = x \wedge x \circ y = z).$$

Dann ist $T_0 + \{e\} + \{\forall(2.e)\}$ mit

$$(2.e)\ e = b \leftrightarrow \forall x \exists y (x \circ b = x \wedge x \circ y = b)$$

definitorische Erweiterung T_1 von T_0 um die Definition $(2.e)$ der Konstanten e – die offenbar als das neutrale Element fungiert.

In dieser Theorie T_1 ist herleitbar

$$\exists! y (a \circ y = e).$$

Dann ist $T_1 + \{^{-1}\} + \{\forall(2.^{-1})\}$ mit

$$(2.^{-1}) \quad a^{-1} = b \leftrightarrow a \circ b = e$$

definitorische Erweiterung T_2 von T_1 um die Definition $(2.^{-1})$ der Inversenfunktion $^{-1}$. T_2 hat dieselbe Sprache wie die Gruppentheorie T_G, und in beiden Theorien sind die sämtlichen Axiome der jeweils anderen Theorie herleitbar. Damit ist T_G eine definitorische $\vdash$-Erweiterung von T_0: Die Modelle von T_0 sind genau die Beschränkungen von Gruppen auf die Sprache von T_0.

18.4 Pränexe Normalformen

Wir stellen ein Verfahren dar, das die Quantoren $\forall$ und $\exists$ in einer Formel durch äquivalente prädikatenlogische Umformungen alle an ihren Anfang bringt.

18.4.1 Definition Eine Formel B der Gestalt

$$Q_1 x_1 \dots Q_n x_n F(x_1, \dots, x_n) \quad (n \geq 0)$$

heißt *pränex* oder *in pränexer Form*, wenn jedes Q_i ein Quantor $\forall$ oder $\exists$ ist und die Formel $F(a_1, \dots, a_n)$ quantorenfrei ist. $Q_1 x_1 \dots Q_n x_n$ heißt dann das *Präfix* von B, $F(a_1, \dots, a_n)$ heißt die *Matrix* von B. Eine pränexe Formel B heißt *pränexe Normalform* einer Formel A, wenn $A \leftrightarrow B$ logisch gilt.

Wir behandeln hier den Existenzquantor $\exists$ wie ein Grundzeichen, ebenso wie $\forall$. Dadurch hat die pränexe Normalform $\exists x \neg px$ von $\neg \forall x px$ dieselbe Länge wie $\neg \forall x px$, während sie als $\neg \forall x \neg \neg px$ um eine doppelte Negation länger wäre.

Man nähert eine beliebige Formel A durch pränexe Umformung von Teilformeln schrittweise einer pränexen Formel an. Dabei bleibt die Zahl der Quantoren $\forall$ und $\exists$ insgesamt, die Zahl der Implikationen und die Zahl der Falsum unverändert.

Beispiel. Die Formel $A \equiv \exists x px \to \forall x (px \to \exists y qxy)$ können wir schrittweise äquivalent umformen zu

$$\begin{aligned} &\exists x px \to \forall x \exists y (px \to qxy) \\ &\forall z (pz \to \forall x \exists y (px \to qxy)) \\ &\forall z \forall x \exists y (pz \to px \to qxy). \end{aligned}$$

18.4.2 Definition Eine Formel D heißt *pränexe Umformung* einer Formel C in den folgenden Fällen:

i) C hat eine Gestalt $B \to Qx F(x)$, und D ist eine Formel $Qy(B \to F(y))$, wobei y nicht in B, F auftritt;

ii) C hat eine Gestalt $Qx F(x) \to B$, und D ist eine Formel $\overline{Q} y(F(y) \to B)$, wobei y nicht in F, B auftritt und $\overline{Q}$ der von Q verschiedene Quantor ist ($\overline{\forall} \equiv \exists, \overline{\exists} \equiv \forall$).

Im obigen Beispiel werden insgesamt vier pränexe Umformungen innerhalb längerer Formeln vorgenommen, und zwar nach i), ii), i), i). Bei der zweiten wird eine gebundene Variable umbenannt, von x in z; die dritte und vierte Umformung sind in einer Zeile zusammengefasst.

18.4.3 Lemma Ist D pränexe Umformung von C, so ist $C \leftrightarrow D$ logisch gültig.

Beweis. Es sei $\mathcal{A}$ eine beliebige Struktur zur gegebenen Sprache und $'$ eine $\mathcal{A}$-Belegung.

1. C und D seien nach 18.4.2, i) gegeben.

   1.1 $\mathcal{A}(B') = f$. Dann ist $\mathcal{A}(C') = w = \mathcal{A}(D')$.

   1.2 $\mathcal{A}(B') = w$. Dann ist $\mathcal{A}(C') = \mathcal{A}(QxF'(x)) = \mathcal{A}(D')$.

2. C und D seien nach 18.4.2, ii) gegeben.

   2.1 $\mathcal{A}(B') = w$. Dann ist $\mathcal{A}(C') = w = \mathcal{A}(D')$.

   2.2 $\mathcal{A}(B') = f$. Dann ist

   $$\mathcal{A}(C') = \mathcal{A}(\neg QxF'(x)) = \mathcal{A}(\overline{Q}y\neg F'(y)) = \mathcal{A}(D')$$

In jedem Fall ist also $\mathcal{A}(C') = \mathcal{A}(D')$, und damit ist $C \leftrightarrow D$ logisch gültig.

Iterierte Anwendung dieses Lemmas ergibt:

18.4.4 Satz Jede Formel A besitzt eine pränexe Normalform, die ebenso viele Quantoren $\forall, \exists$ insgesamt und ebenso viele Implikationszeichen $\rightarrow$ enthält wie A.

Beweis durch Induktion nach der Anzahl der Auftreten von $\forall, \exists$ und $\rightarrow$ insgesamt in A, die wir hier als die Länge von A bezeichnen.

1. A ist quantorenfrei. Dann ist A pränex (mit leerem Präfix), und $A \leftrightarrow A$ ist eine Tautologie. Also ist A selbst eine pränexe Normalform von A.

2. A ist $B \rightarrow C$, und C enthält mindestens einen Quantor. Nach Induktionsvoraussetzung hat C eine pränexe Normalform $QxF(x)$, und ebenso hat $B \rightarrow F(a)$ eine pränexe Normalform $G(a)$, jeweils von derselben Länge. Dann sind nach Induktionsvoraussetzung bzw. 18.4.3 logisch äquivalent:

   $$A, B \rightarrow QxF(x),\ Qy(B \rightarrow F(y)),\ QyG(y).$$

Dabei sei y so gewählt, dass es in B, F, G nicht auftritt. Also gilt $A \leftrightarrow QyG(y)$ logisch, und $QyG(y)$ ist eine pränexe Normalform von A, von derselben Länge wie A.

3. A ist $B \to C$, und B enthält mindestens einen Quantor. Nach Induktionsvoraussetzung hat B eine pränexe Normalform $QxF(x)$, und $F(a) \to C$ hat eine pränexe Normalform $G(a)$, jeweils von derselben Länge. Dann sind nach Induktionsvoraussetzung bzw. 18.4.3 logisch äquivalent:

$$A, QxF(x) \to C,\ \overline{Q}y(F(y) \to C),\ \overline{Q}yG(y).$$

   Wie unter 2. ist $\overline{Q}yG(y)$ die gesuchte pränexe Normalform von A.

4. A ist $QxF(x)$. Nach Induktionsvoraussetzung hat $F(a)$ eine pränexe Normalform $G(a)$ von gleicher Länge. Dann gilt $A \leftrightarrow QyG(y)$ (mit einem y, das nicht in G auftritt) logisch, und $QyG(y)$ ist die gesuchte pränexe Normalform von A.

Durch Induktion nach der Länge von A folgt der Satz.

18.5 Skolemisierung

Wenn eine Nennform G den Graphen einer Funktion beschreibt und dies in einer Theorie T herleitbar ist, so führt die Hinzunahme dieser Funktion und ihrer Definition (2) (vgl. 18.3.1) zu T zu einer definitorischen, daher konservativen und wieder eliminierbaren Erweiterung von T. Aber auch wenn G nur eine in T erkennbar rechts-existente Relation beschreibt, erhält man noch eine zu 18.3.2 analoge Konservativitätsaussage, auch wenn von einer definitorischen Erweiterung keine Rede mehr sein kann. Dies geht auf den norwegischen Mathematiker und Logiker Thoralf Skolem zurück.

18.5.1 Definition Das n-stellige Funktionszeichen f sei nicht in $L(T)$, und G sei eine geschlossene Nennform aus $L(T)$. Wenn

$$T \vdash \exists y G(a_1, \ldots, a_n, y),$$

dann heißt $T' := T + \{f\} + \{\forall(3)\}$ mit

(3) $G(a_1, \ldots, a_n, fa_1 \ldots a_n)$

Skolem-Erweiterung von T um die *Skolem-Funktion* f mit (3).

Nun lässt sich 18.3.2 verallgemeinern zu:

18.5.2 Lemma von Skolem Die Skolem-Erweiterung T' von T um f mit (3) ist eine konservative Erweiterung von T.

Beweis. Sei $\mathcal{A} \models T$. Dann gibt es zu $k_1, \ldots, k_n \in |\mathcal{A}|$ stets mindestens ein $l \in |\mathcal{A}|$, so dass

$$\mathcal{A}(G(k_1, \ldots, k_n, l)) = w$$

ist. Mit dem Auswahlaxiom folgt, dass es eine Auswahlfunktion $\varphi : |\mathcal{A}|^n \to |\mathcal{A}|$ gibt, die $k_1, \ldots, k_n \in |\mathcal{A}|$ stets ein solches l zuordnet. Dann ist stets

$$\mathcal{A}(G(k_1, \ldots, k_n, \varphi(k_1, \ldots, k_n))) = w.$$

Wir setzen $f_{\mathcal{B}} := \varphi$, ferner $\mathcal{B}|L(T) := \mathcal{A}$. Dann ist $\mathcal{B}$ eine Expansion von $\mathcal{A}$ zu einer Struktur zu $L(T) + \{f\}$, in der (3) gilt. Also ist $\mathcal{B} \models T'$. Mit 18.1.2 folgt die Behauptung.

In diesem Beweis ist die Expansion $\mathcal{B}$ durch das Modell $\mathcal{A}$ von T i. a. nicht eindeutig festgelegt. $\mathcal{B}$ hängt zusätzlich noch von der Auswahlfunktion φ ab, anders als in 18.3.2.
Der Spezialfall $n = 0$ von 18.5.2 ist uns schon bekannt: Ist in T ein Satz $\exists y G(y)$ herleitbar, so ist $T + \{\varepsilon y G(y)\} + \{G(\varepsilon y G(y))\}$ konservativ über T. Mehrfache Anwendung dieser Aussage ergibt 8.3.3.

Beispiel. *DLO* sei die Theorie der dichten linearen Ordnung aus 1.2.5 und §13. Wegen des Dichteaxioms *LO*4 ist

$$DLO \vdash \exists y(a < b \to a < y \wedge y < b).$$

Nach dem Lemma von Skolem ist

$$DLO + \{f\} + \{\forall x \forall y(x < y \to x < fxy \wedge fxy < y)\}$$

eine konservative Erweiterung von *DLO*, obwohl Sprache und Axiome von *DLO* offenbar keine Möglichkeit bieten, ein solches f zu definieren. In dem Modell $(\mathbb{Q}, <)$ von *DLO* kann man zwar extern

$$f_{(\mathbb{Q},<)}(k, l) := \frac{1}{2} \cdot (k + l)$$

setzen, aber jede andere Wahl eines Zwischenwertes täte es auch.

Mit dem Lemma von Skolem kann man in einer herleitbaren Existenzformel den führenden Existenzquantor eliminieren und die durch ihn gebundene Variable durch einen Term ersetzen, der mit einem neuen Funktionszeichen beginnt. Durch mehrfache Anwendung dieses Lemmas kann man schließlich alle Existenzquantoren im Präfix einer herleitbaren Formel in analoger Weise eliminieren.

18.5.3 Rekursive Definition der *Skolem-Normalform* A^S einer Formel A. Zunächst sei A eine pränexe Formel.

1. Ist A quantorenfrei, so sei $A^S :\equiv A$.
2. $(\forall x G(x))^S$ sei $G(a)^S$, wobei a nicht in G auftritt.
3. $(\exists y G(y))^S$ sei $G(f a_1 \dots a_n)^S$, wobei $FV(G) = \{a_1, \dots, a_n\}$ (mit paarweise verschiedenen a_i) und f ein n-stelliges Funktionszeichen ist, das in $G(b)^S$ nicht auftritt.
4. Ist A nicht pränex, so sei B eine pränexe Normalform von A gemäß 18.4.4, und A^S sei B^S.

Die neuen Funktionszeichen f, die durch 3. in A^S, aber nicht in A auftreten, sind die *Skolem-Funktionszeichen* von A^S. Die Menge dieser Funktionszeichen bezeichnen wir mit $SF(A^S)$.
$T + SF(A^S) + \{\forall(A^S)\}$ heißt *Skolem-Erweiterung* von T.

Beispiel. Sei T_0 die nach 18.3.8 betrachtete Theorie und A deren Axiom

$$\exists z \forall x \exists y (x \circ z = x \wedge x \circ y = z).$$

Wählt man wie dort e und $^{-1}$ als neue Funktionszeichen, so ist $SF(A^S) = \{e, ^{-1}\}$ und

$$A^S \equiv (a \circ e = a \wedge a \circ a^{-1} = e).$$

Dann ist $T_0 + SF(A^S) + \{\forall A^S\}$ eine (zweifache) Skolem-Erweiterung von T_0 und nach 18.5.2 konservativ über T_0. Ohne die zusätzlichen Überlegungen aus 18.3.8 ist damit noch nicht klar, dass dies auch eine definitorische Erweiterung von T_0 ist.

Bemerkungen. A^S ist stets quantorenfrei. Ist A pränex, so enthält A^S zu jedem nicht-überflüssigen Existenzquantor im Präfix von A genau ein Skolem-Funktionszeichen.
A^S ist durch A nicht eindeutig festgelegt. Willkürlich bleiben: die Wahl der neuen freien Variablen a in 2. aus 18.5.3, die Wahl des Skolem-Funktionszeichens f in 3., die Reihenfolge der freien Variablen $a_1, \ldots, a_n$ hinter f in 3. und schließlich die Wahl der pränexen Normalform B in 4.

18.5.4 Lemma $\forall(A^S) \to A$ ist logisch gültig.

Beweis durch Induktion nach A gemäß 18.5.3.

1. Ist A quantorenfrei, so ist die Behauptung klar.

2. Ist $A \equiv \forall x G(x)$, so gilt $\forall A^S \to G(a)$ nach Induktionsvoraussetzung, und die Behauptung folgt etwa mit $(\forall S)$.

3. Ist $A \equiv \exists y G(y)$, so gilt $\forall A^S \to G(f a_1 \ldots a_n)$ nach Induktionsvoraussetzung, und die Behauptung folgt mit $(\exists S)$.

4. Ist A nicht pränex, so gilt $\forall A^S \to B$ nach Induktionsvoraussetzung für eine pränexe Normalform B von A. Nach 18.4.4 gilt $B \to A$, und ein Kettenschluss ergibt die Behauptung.

18.5.5 Satz von Skolem Ist $T \vdash A$ und sind die Skolem-Funktionszeichen von A^S alle nicht in $L(T)$, so ist die Skolem-Erweiterung $T+ SF(A^S) + \{\forall A^S\}$ eine konservative Erweiterung von T.

Beweis. T^A bezeichne die angegebene Skolem-Erweiterung. Offenbar ist T^A eine Erweiterung von T. Zu zeigen bleibt die Konservativität von T^A über T. Wir induzieren wieder nach dem Aufbau von A gemäß 18.5.3.

1. Ist A quantorenfrei, so ist $T^A = T + \{\forall A\}$ eine Erweiterung von T um einen herleitbaren Satz und somit konservativ über T.

2. Ist $A \equiv \forall x G(x)$, so ist $T^A = T^{G(a)}$ und daher nach Induktionsvoraussetzung konservativ über T.

3. Ist $A \equiv \exists y G(y)$, so ist nach dem Lemma von Skolem

$$T \prec T + \{f\} + \{\forall G(f a_1, \ldots, a_n)\}$$

konservativ. Da (mit $a :\equiv a_1 \dots a_n$) $G(fa)$ pränex und kürzer als A ist, ist nach Induktionsvoraussetzung auch

$$T + \{f\} + \{\forall G(fa)\} \prec (T + \{f\} + \{\forall G(fa)\})^{G(fa)}$$

konservativ. Diese konservative Erweiterung ist, ausführlich hingeschrieben, die Theorie

$$T + \{f\} \cup SF(G(fa)^S) + \{\forall G(fa), \forall G(fa)^S\}.$$

Hierin ist $G(fa)^S \equiv A^S$ und $\{f\} \cup SF(G(fa)^S) = SF(A^S)$, und nach 18.5.4 gilt $\forall G(fa)^S \to \forall G(fa)$. Also ist diese Theorie $\vdash$-äquivalent zu

$$T + SF(A^S) + \{\forall A^S\} = T^A.$$

Dann ist T^A auch konservativ über T.

4. Ist A nicht pränex, so ist $T^A = T^B$ für eine pränexe Normalform B von A. Nach 18.4.4 ist auch $T \vdash B$, so dass mit der Induktionsvoraussetzung die Behauptung folgt.

Damit ist der Satz bewiesen.

Den Satz von Skolem kann man offenbar simultan auf alle Axiome einer Theorie anwenden.

18.5.6 Definition Sei T eine Theorie. Wir betrachten hier ${}^S : A \mapsto A^S$ als eine Abbildung, die (zumindest) jedem Axiom A von T genau eine ihrer Skolem-Normalformen zuordnet. Dabei seien $SF(A^S)$ und $SF(B^S)$ disjunkt, wenn $A \not\equiv B$ ist, und auch disjunkt zu $L(T)$. Als *Skolemisierung von* T bezeichnet man die Theorie

$$T^S := (L(T) \cup \bigcup\{SF(A^S) | A \in Ax(T)\}, \{\forall A^S | A \in Ax(T)\}).$$

Bemerkung. Da A^S stets quantorenfrei ist, ist T^S stets eine offene Theorie. Also ist die Klasse der Modelle von T^S nach 11.2.10 stets gegen Unterstrukturen abgeschlossen – was für T nicht zu gelten braucht. Die Skolemisierung von T hat i. a. eine reichere Sprache als T, ist aber beweistechnisch nicht stärker als T:

18.5.7 Satz über Skolemisierung.
Die Skolemisierung T^S einer Theorie T ist eine offene konservative $\vdash$-Erweiterung von T: *Jede Theorie besitzt eine offene konservative $\vdash$-Erweiterung.*

Beweis. Jedes Axiom A von T ist in T^S allein aus dem entsprechenden Axiom $\forall A^S$ von T^S nach 18.5.4 herleitbar. Also ist $T^S \vdash T$.
Sei nun C aus $L(T)$ und $T^S \vdash C$. Wegen der Endlichkeit der Herleitungen gibt es dann eine kleinste Zahl n und verschiedene Axiome A_i von T $(i < n)$, so dass

$$T_n := T + \bigcup\{SF(A_i^S)|i < n\} + \{\forall A_i^S|i < n\} \vdash C.$$

Ist $n = 0$, so ist $T_n = T \vdash C$. Ist $n > 0$, so ist T_n eine Skolem-Erweiterung von

$$T_{n-1} := T + \bigcup\{SF(A_i^S)|i < n-1\} + \{\forall A_i^S|i < n-1\}.$$

Es ist $T_{n-1} \vdash A_{n-1}$, weil $A_{n-1} \in Ax(T)$ ist, und die Skolem-Funktionszeichen von A_{n-1}^S gehören alle nicht zu $L(T_{n-1})$. Dann ist aber nach dem Satz 18.5.5 von Skolem $T_n \succ T_{n-1}$ konservativ, und es folgt $T_{n-1} \vdash C$. Also war n nicht minimal bezüglich der Eigenschaft $T_n \vdash C$.
Daher ist notwendig $n = 0$, und $T \vdash C$: T^S ist konservativ über T.

18.6 Aufgaben

18.6.1 Widerlegen Sie die Umkehrung des Expansionslemmas, indem Sie zeigen: Die Theorie DLO besitzt eine überabzählbare konservative Erweiterung T', die nur überabzählbare Modelle besitzt, so dass sich $(\mathbb{Q}, <)$ nicht zu einem Modell von T' expandieren lässt.

18.6.2 Zeigen Sie: Es ist $T' \vdash T$ konservativ genau dann, wenn es zu jedem Modell von T ein elementar äquivalentes Modell gibt, das sich zu einem Modell von T' expandieren lässt.

18.6.3 Zeigen Sie: Die Theorie $Z + \{<, |, prim\} + \{\forall(<), \forall(|), \forall(prim)\}$ mit

$(<)$ $\quad a < b \leftrightarrow \exists x a + \operatorname{suc} x = b$

$(|)$ $\quad a|b \leftrightarrow \exists x a \cdot x = b$

$(prim)$ $\quad prim\, a \leftrightarrow a \neq 1 \wedge \forall x(x|a \rightarrow x = 1 \vee x = a)$

ist eine konservative Erweiterung von Z.
Schreiben Sie die Aussage „Es gibt unendlich viele Primzahlen“

a. als Satz von $Z + \{<, |, prim\}$

b. als Satz von Z.

18.6.4 Zeigen Sie die logische Gültigkeit von

$$F(t_1, \ldots, t_n) \leftrightarrow \exists y_1 \ldots \exists y_n (t_1 = y_1 \wedge \ldots \wedge t_n = y_n \wedge F(y_1, \ldots, y_n)).$$

18.6.5 Geben Sie verschiedene Theorien T und T' an mit

$$T \prec T' \text{ und } T' \vdash T.$$

Kann dann auch $T' \prec T$ sein?

18.6.6 Zeigen Sie: Theorien T und T' sind genau dann $\vdash$-äquivalent, wenn sie dieselben Modelle haben.

18.6.7 Zeigen Sie: Ist T' eine definitorische $\vdash$-Erweiterung von T, so ist T' konservativ über T, und die Modelle von T sind genau die Beschränkungen der Modelle von T' auf $L(T)$.

18.6.8 Sei $F(a_1, \ldots, a_n, b)$ eine beliebige Formel (F geschlossen) von $L(Z^2)$ (vgl. 15.1). Zeigen Sie: $Z^2 + \{\forall(*)\}$ mit

$(*)$ $f a_1 \ldots a_n = \{x \in N | F(a_1, \ldots, a_n, x)\}$

ist eine definitorische $\vdash$-Erweiterung von Z^2.

18.6.9 Eine Formel B ist *in Negationsnormalform (in NNF)*, wenn B aus Primformeln und negierten Primformeln allein mit $\wedge, \vee, \forall, \exists$ aufgebaut ist. B ist *Negationsnormalform* einer Formel A *(NNF von A)*, wenn B in NNF ist und $A \leftrightarrow B$ logisch gilt.

a. Geben Sie eine induktive Definition der Formeln in NNF an.

b. Zeigen Sie: Jede Formel A besitzt eine NNF.

18.6.10 Bilden Sie eine Skolemisierung $(T_K)^S$ der Körpertheorie T_K aus 1.2.3. Ist $(T_K)^S$ eine definitorische $\vdash$-Erweiterung von T_K?

FernUniversität in Hagen

Klassische Prädikatenlogik

Kurseinheit 7:
Automatisches Beweisen

Autor: Justus Diller

Klassische Prädikatenlogik

Kurseinheit 7: Inhalt

Studienhinweise 391

Verzeichnis der definierten Begriffe und der wichtigen Sätze 393

1. Sprache, Semantik und Syntax der Prädikatenlogik

2. Syntaktische Sätze und Regeln der Prädikatenlogik

3. Vollständigkeit

4. Modelltheorie

5. Beweistheorie der Prädikatenlogik

6. Automatisches Beweisen

§19 Gentzen-Theorien und Widerlegungsvollständigkeit 396

19.1 Herbrand-Strukturen 396
19.2 Gentzen-Theorien 398
19.3 Substitutionen 400
19.4 Widerlegungsvollständige Regelsysteme 405
19.5 Aufgaben 412

§20 Unifikation 413

20.1 Unifikatoren 413
20.2 Berechnung allgemeinster Unifikatoren 417
20.3 Zum Aufwand der Unifikation 425
20.4 Aufgaben 429

§21 Resolution 431

21.1 Einfache Resolution und Faktorisierung 431
21.2 Volle Resolution 435
21.3 Aufgaben 442

Klassische Prädikatenlogik

Kurseinheit 7: Studienhinweise

1. Lehrziele

Die Kurseinheit gibt eine Einführung in das Automatische Beweisen. Ziel ist die Erarbeitung zweier Resolutionskalküle, insbesondere ihrer Formulierung (in 21.1), ihrer Korrektheit und ihrer Widerlegungsvollständigkeit (21.2.6). Während ihre Korrektheit trivial ist, ist ihre Widerlegungsvollständigkeit (zur Definition vgl. 19.4.4) das wichtigste Ergebnis der Kurseinheit.

Die entscheidende neue Technik, die die Resolutionskalküle schon zu ihrer Formulierung verwenden, ist die Unifikation von Termen, die in §20 erarbeitet wird. Sie ist ein völlig elementarer neuer Ansatz, der nur auf der Definition der Terme und der Substitution aus §1 basiert. Allerdings muss man in die elementar-kombinatorischen Eigenschaften von (Term-) Substitutionen tiefer eindringen, als das vorher nötig war. Sie werden in 19.3 und 20.1 recht präzise erarbeitet. Die zentralen Ergebnisse über Unifikatoren sind die Eindeutigkeit und Berechenbarkeit von allgemeinsten Unifikatoren (Sätze 20.1.12 und 20.2.10).

Die Resolutionskalküle werden nur auf Gentzen-Theorien angewandt, die nach ihrer Definition 19.2.1 eine sehr spezielle Klasse bilden. Dieses Bild ändert sich, wenn man statt der formalen Definition einer Theorie ihre Beweiskraft in den Vordergrund rückt: Nach Satz 19.2.5 und der anschließenden Bemerkung sind Gentzen-Theorien im wesentlichen die Skolemisierungen beliebiger abzählbarer Theorien und damit von (mindestens) gleicher Beweis- und Ausdruckskraft wie diese. Satz 19.2.5 ist deshalb wesentlich für das Verständnis des allgemeinen prädikatenlogischen Rahmens, auf den sich Widerlegungsvollständigkeit, Resolution etc. beziehen.

2. Eingangsvoraussetzungen

Da für das Hauptthema der Kurseinheit, die Korrektheit und Vollständigkeit von Resolutionskalkülen, ab 19.3 ein neuer Ansatz gemacht wird, sind die erforderlichen Vorkenntnisse von diesem Punkt ab sehr gering. Im wesentlichen werden der Termbegriff, die Eindeutigkeit des Termaufbaus, daneben der Begriff der mathematischen Theorie, alles Begriffe aus §1, verwendet. Zusätzlich sollte aus den Anfängen der Semantik aus §2 die Korrektheit von Schlüssen

und Regeln, aus §3 die Aussage des Korrektheitssatzes vertraut sein.

Anders verhält es sich mit dem Einstieg in die Kurseinheit. Neben den Grundbegriffen der Semantik für den Vergleich von beliebigen Modellen mit Term-(Herbrand-) Modellen (Satz 19.1.4) wird auf konjunktive Normalformen nach 4.4.7 und auf den Satz über Skolemisierung 18.5.7 zurückgegriffen.

Klassische Prädikatenlogik

Kurseinheit 7: Verzeichnis der definierten Begriffe und der wichtigen Sätze

19.1.1 Herbrand-Universum $CT(L)$, Herbrand-Struktur, Termstruktur

19.1.3 Herbrand-Modell

19.1.4 **Satz** Eine offene identitätsfreie Theorie (mit einer Konstanten) hat ein Modell genau dann, wenn sie ein Herbrand-Modell hat.

19.2.1 prime Sequenz, Gentzen-Theorie

19.2.5 **Satz** Jede abzählbare identitätsfreie Theorie besitzt eine konservative $\vdash$-Erweiterung, die Gentzen-Theorie ist.

19.3.1 Substitution σ, Definitionsbereich $dom\sigma$, (erweiterte) Variablenmenge $var\sigma$, $var^{+}\sigma$, Grundsubstitution

19.3.2 Instanz $A\sigma$, Grundinstanz

19.3.3 $[a_1/t_1, \ldots, a_n/t_n]$

19.3.4 **Homomorphie-Lemma**

19.3.6 Komposition $(\sigma\tau)$

19.3.8 Permutation, Umbenennung, Variante, $\sigma \sim \tau$

19.4.1 (prime) Regel, R-Schluss, Regelsystem, $(Subst), (Mix), (sMix)$

19.4.2 $\mathcal{R}$-Herleitung, Startsequenz

19.4.3 $\mathcal{R}$-herleitbar, $T \vdash_{\mathcal{R}} \Gamma : \Delta$

19.4.4 widerlegungsvollständig

19.4.7 **Konsistenzlemma**

19.4.8 **Satz** $\{(sMix)\}$ ist widerlegungsvollständig für quantorenfreie Gentzen-Theorien.

19.4.10 quantorenfreie Version, Grundversion

19.4.13 **Lifting-Satz**

19.4.14 **Satz** $\{(Mix), (Subst)\}$, sogar $\{(sMix), (Subst)\}$ ist widerlegungsvollständig.

20.1.1 Unifikator, unifizieren

20.1.2 σ allgemeiner als τ, σ subsumiert τ, $\sigma \leq \tau$, äquivalent, $\sigma \approx \tau$

20.1.9 **Satz** $\sigma \approx \tau$ ist äquivalent zu $\sigma \sim \tau$.

20.1.10 allgemeinster Unifikator, $mgu(M)$

20.1.12 **Satz** Eindeutigkeit des allgemeinsten Unifikators

20.2.2 σ löst, ist Lösung von Γ

20.2.4 **Lemma** Ist $dom\tau \cap var\tau = \emptyset$, so ist τ allgemeinste Lösung von $\{a = a\tau \mid a \in dom\tau\}$.

20.2.5 Martelli-Montanari-Regeln

20.2.6 MM-Herleitung, MM-herleitbar, Martelli-Montanari-Algorithmus (MM), Start-Gleichungsmenge, terminierend, (MM) terminiert für Γ_0

20.2.9 **Satz** Für jede Start-Gleichungsmenge terminiert (MM).

20.2.10 **Satz von Martelli-Montanari**

20.3.1 $l(M)$, $l(\Gamma)$

20.3.3 **Satz** Für ein geeignetes $M = \{t, t'\}$ und $\sigma = mgu(M)$ wächst $l(\sigma[M])$ gegenüber $l(M)$ exponentiell.

21.1.2 Einfache Resolutionsregel (Res)

21.1.4 Faktorisierungsregeln (Fak)

21.1.9 **Satz** zur Iteration von (Fak)

21.2.1 Volle Resolutionsregel (RES)

21.2.2 **Satz** $(RES) \subseteq \{(Res), (Fak)\}$

21.2.4 **Transformationssatz**

21.2.6 **Satz, Widerlegungsvollständigkeit der Resolutionskalküle**

Kapitel 6

Automatisches Beweisen

Der Sequenzenkalkül aus §3 regelt den logischen Aufbau der Formeln in den Herleitungen systematisch und so dicht an der Wahrheitsdefinition entlang, wie es sich unabhängig vom einzelnen Modell einrichten lässt. Das sieht man besonders deutlich im Beweis des Vollständigkeitssatzes nach Schütte in §9. Dort zeigt sich aber auch der eine Punkt, in dem die Suche nach einer Herleitung nicht kanonisch ist: Der Term, der bei einem $(\forall A)$-Schluss verloren geht, hat i. a. keinen Bezug zu der Formel, in die er eingesetzt wird, oder zu der Herleitung, in der er steht. Im Suchbaum wird er einer von außen gegebenen, beliebigen Aufzählung aller Terme entnommen, und es liegt auf der Hand, dass das Verfahren das Auffinden einer Herleitung enorm verzögern kann.

Das Automatische Beweisen hat nun das Ziel, die Suche nach den passenden Termen grundlegend effizienter anzulegen. Das Aufbauen längerer Formeln in den Herleitungen, das für das Beweisen mathematisch interessanter Sätze zentral ist, tritt in den Hintergrund; man arbeitet nur noch mit Sequenzen, die ausschließlich aus Primformeln bestehen. Man sucht praktischerweise auch nicht nach korrekten Beweisen solcher Sequenzen, sondern man prüft die Vermutung, ob eine Menge von primen Sequenzen bei geeigneten Termsubstitutionen falsch wird. Wenn das der Fall ist, so muss sich aus der Menge von primen Sequenzen mit einem widerlegungsvollständigen Regelsystem auch ein Widerspruch, die leere Sequenz $\square$, herleiten lassen. Dieses Programm wird in §19 ausgearbeitet.

Wie findet man nun möglichst effizient die Terme, die man in den primen Sequenzen substituiert, um zu einem Widerspruch zu gelangen? Hier bietet die Technik der Unifikation einen neuen Ansatz, der mit den Mitteln von §1 zu

behandeln ist, der aber erst im Kontext des Automatischen Beweisens seine Bedeutung gewinnt. Unifikation und die effiziente Berechnung möglichst allgemeiner Unifikatoren sind der Gegenstand von §20.

Wenn man die Unifikation mit der Schnitt-Regel (in der Fassung der Mix-Regel) kombiniert, so erhält man die sogenannte Resolutions-Regel. Sie bildet bereits für sich allein ein widerlegungsvollständiges Regelsystem. Der Resolutionskalkül wird in §21 behandelt.

§19 Gentzen-Theorien und Widerlegungsvollständigkeit

19.1 Herbrand-Strukturen

19.2 Gentzen-Theorien

19.3 Substitutionen

19.4 Widerlegungsvollständige Regelsysteme

19.5 Aufgaben

19.1 Herbrand-Strukturen

19.1.1 Definition Mit $CT(L)$ bezeichnen wir die Menge der geschlossenen Terme (closed terms) einer Sprache L. Ist $CT(L)$ nicht leer (d. h. enthält L mindestens eine Konstante), so wird $CT(L)$ gelegentlich auch das *Herbrand-Universum* von L genannt.
Eine Struktur $\mathcal{A}$ zu L heißt *Herbrand-Struktur* oder *Term-Struktur* zu L, wenn

1. $|\mathcal{A}| = CT(L) \neq \emptyset$ ist und
2. $f_{\mathcal{A}}(t_1, \ldots, t_n) := ft_1 \ldots t_n$ ist für alle n-stelligen Funktionszeichen f von L ($n \in \mathbb{N}$).

Beispiel $(\mathbb{N}; 0, S_{\mathbb{N}})$ mit $S_{\mathbb{N}}(k) = Sk$ ist die einzige Herbrand-Struktur zur Sprache, die durch 0 und S gegeben ist. Für jede Relation $R \subseteq \mathbb{N} \times \mathbb{N}$ ist $(\mathbb{N}; 0, S_{\mathbb{N}}; R)$ eine Herbrand-Struktur zur Sprache, die durch $0, S$ und $<$ gegeben ist.

19.1.2 Lemma Ist $\mathcal{A}$ Herbrand-Struktur zu L, so ist

$$\mathcal{A}(t) = t \text{ für alle } t \in CT(L).$$

Beweis durch Induktion nach dem Aufbau von t.

Ist $t \equiv ft_1 \dots t_n$ $(n \geq 0)$, so ist

$$\begin{aligned} \mathcal{A}(t) &= f_{\mathcal{A}}(\mathcal{A}(t_1), \dots, \mathcal{A}(t_n)) \\ &= f_{\mathcal{A}}(t_1, \dots, t_n) \quad \text{nach IV} \\ &= ft_1 \dots t_n \equiv t, \end{aligned}$$

weil $\mathcal{A}$ Herbrand-Struktur ist.

19.1.3 Definition Eine Struktur $\mathcal{A}$ ist *Herbrand-Modell* einer Theorie T, wenn $\mathcal{A} \models T$ und $\mathcal{A}$ eine Herbrand-Struktur zu $L(T)$ ist.

19.1.4 Satz Es sei T eine offene identitätsfreie Theorie, und $CT(L(T))$ sei nicht leer.
T hat ein Modell genau dann, wenn T ein Herbrand-Modell hat.

Beweis. Die Richtung von rechts nach links ist trivial. Für die Richtung von links nach rechts sei zu $\mathcal{B} \models T \quad \mathcal{A}$ die Herbrand-Struktur, die durch

$$(t_1, \dots, t_n) \in p_{\mathcal{A}} :\Leftrightarrow (\mathcal{B}(t_1), \dots, \mathcal{B}(t_n)) \in p_{\mathcal{B}}$$

für $t_i \in CT(L(T))$ und nicht-logische Prädikatszeichen p aus $L(T)$ definiert ist.

Wir zeigen zunächst durch Induktion nach C:

(1) Ist C ein identitäts- und quantorenfreier Satz aus $L(T)$, so ist
$\mathcal{A}(C) = \mathcal{B}(C)$.

1. Sei $C \equiv pt_1 \dots t_n$, p nicht das Gleichheitszeichen.

$$\begin{aligned} \mathcal{A}(C) = w &\Leftrightarrow (\mathcal{A}(t_1), \dots, \mathcal{A}(t_n)) \in p_{\mathcal{A}} \\ &\Leftrightarrow (t_1, \dots, t_n) \in p_{\mathcal{A}} \quad \text{nach 19.1.2} \\ &\Leftrightarrow (\mathcal{B}(t_1), \dots, \mathcal{B}(t_n)) \in p_{\mathcal{B}} \\ &\Leftrightarrow \mathcal{B}(C) = w. \end{aligned}$$

2. $\mathcal{A}(\bot) = \mathcal{B}(\bot) = f$.

3. $\mathcal{A}(C_1 \to C_2) = w \quad \Leftrightarrow$ Aus $\mathcal{A}(C_1) = w$ folgt $\mathcal{A}(C_2) = w$
$\Leftrightarrow$ Aus $\mathcal{B}(C_1) = w$ folgt $\mathcal{B}(C_2) = w$ nach IV
$\Leftrightarrow \mathcal{B}(C_1 \to C_2) = w.$

Mit Induktion folgt (1).

Sei nun $C \in Ax(T)$. Dann ist C ein identitätsfreier reiner $\forall$-Satz $\forall x_1 \ldots \forall x_n \; F(x_1, \ldots, x_n)$ und $\mathcal{B}(C) = w$. Daraus folgt für alle $t_1, \ldots, t_n \in CT(L(T))$

$$\mathcal{B}(\mathcal{F}(t_1, \ldots, t_n)) = \mathcal{B}(\mathcal{F}(\mathcal{B}(t_1), \ldots, \mathcal{B}(t_n))) = w.$$

Wegen (1) ist dann auch $\mathcal{A}(\mathcal{F}(t_1, \ldots, t_n)) = w$ für alle $t_1, \ldots, t_n \in CT(L(T)) = |\mathcal{A}|$. Also ist $\mathcal{A}(C) = w$.
Insgesamt folgt $\mathcal{A} \models T$. Damit ist $\mathcal{A}$ Herbrand-Modell von T.

19.2 Gentzen-Theorien

19.2.1 Definition Wir nennen eine Sequenz $\Gamma : \Delta$ *prim*, wenn Γ, Δ nur aus Primformeln bestehen, die keine Gleichungen sind.

Eine Theorie T ist eine *Gentzen-Theorie*, wenn

1. $L(T)$ eine abzählbare Sprache ist, die mindestens eine Konstante und ein nicht-logisches Prädikatszeichen enthält,

2. $Ax(T)$ aus Sätzen $\forall(\Gamma \to \Delta)$ besteht, die primen Sequenzen $\Gamma : \Delta$ zugeordnet sind (vgl. 6.3.4 und 7.1.3).

Gentzen-Theorien sind also insbesondere abzählbar, offen und identitätsfrei. Formal sind sie offenbar recht spezielle Theorien. Man könnte denken, dass sie auch unter den abzählbaren Theorien hinsichtlich ihrer Ausdrucks- oder Beweiskraft nur eine kleine Klasse bilden. Das ist aber nicht der Fall, im wesentlichen wegen des Satzes über Skolemisierung. Wir behandeln hier den identitätsfreien Fall.

19.2.2 Definition Primformeln und negierte Primformeln nennt man auch *Literale*.

19.2.3 Lemma Zu jeder Disjunktion D von Literalen gibt es eine prime Sequenz $\Gamma : \Delta$, so dass

$$D \leftrightarrow (\Gamma \to \Delta)$$

eine Tautologie ist. Mit D ist auch $\Gamma : \Delta$ identitätsfrei.

Beweis. Sind $P_1, \ldots, P_m$ die Primformeln und $\neg Q_1, \ldots, \neg Q_n$ die negierten Primformeln, die Disjunktionsglieder von D sind, so ist

$$\begin{aligned} D &\leftrightarrow \neg Q_1 \vee \ldots \vee \neg Q_n \vee P_1 \vee \ldots \vee P_m \\ &\leftrightarrow Q_1 \to \ldots \to Q_n \to P_1 \vee \ldots \vee P_m \\ &\leftrightarrow (\Gamma \to \Delta) \end{aligned}$$

mit $\Gamma = \{Q_1, \ldots, Q_n\}$ und $\Delta = \{P_1, \ldots, P_m\}$ eine Tautologie. Mit den Konventionen aus 6.3.4 stimmt das auch für $m = 0$ oder $n = 0$.

19.2.4 Lemma Jede offene Theorie T ist $\vdash$-äquivalent zu einer Theorie T', deren Axiome alle eine Gestalt $\forall(\Gamma \to \Delta)$ mit primen Sequenzen $\Gamma : \Delta$ haben. Mit T ist auch T' identitätsfrei.

Beweis. Sei $C \in Ax(T)$. Dann ist $C \equiv \forall B$ mit quantorenfreiem B, weil T offen ist. B besitzt nach 4.4.7 eine konjunktive Normalform

$$K \equiv D_1 \wedge \ldots \wedge D_l,$$

wobei die D_i Disjunktionen von Literalen sind. Wegen 19.2.3 kann man o. E. $D_i \equiv \Gamma_i \to \Delta_i$ mit primen Sequenzen $\Gamma_i : \Delta_i$ annehmen $(i = 1, \ldots, l)$. Da sich Allquantoren über Konjunktionen verteilen, ist

$$\forall B \leftrightarrow \forall K \leftrightarrow \forall(\Gamma_1 \to \Delta_1) \wedge \ldots \wedge \forall(\Gamma_l \to \Delta_l)$$

logisch gültig. Ersetzt man also jedes Axiom $C \in Ax(T)$ durch die l Sätze $\forall(\Gamma_i \to \Delta_i)$ $(i = 1, \ldots, l)$, so erhält man die gesuchte Theorie T'.

19.2.5 Satz Jede abzählbare identitätsfreie Theorie T besitzt eine konservative $\vdash$-Erweiterung T', die eine Gentzen-Theorie ist.

Beweis. Wir können o. E. voraussetzen, dass $L(T)$ eine Konstante und ein nicht-logisches Prädikatszeichen enthält, weil $T + \{c, p\} \succ T$ konservativ ist für jede neue Konstante c und jedes neue Aussagezeichen p. Die Skolemisierung T^S von T ist dann nach 18.5.7 eine offene konservative $\vdash$-Erweiterung

von T, die offenbar auch abzählbar und identitätsfrei ist. Wenn man T^S nach 19.2.4 $\vdash$-äquivalent zu T' umformt, so ist T' eine Gentzen-Theorie und eine konservative $\vdash$-Erweiterung von T.

Ein ähnliches Ergebnis erhält man auch für abzählbare Theorien T mit Identität. Dafür legt man zuerst ein neues zweistelliges Prädikatszeichen $\sim$ fest, fügt Allabschlüsse von

$(\sim I)$ $a \sim a$

$(\sim F)$ $a_1 \sim b_1 \to \ldots \to a_n \sim b_n \to fa_1 \ldots a_n \sim fb_1 \ldots b_n$
für jedes n-stellige Funktionszeichen f aus $L(T)$ und

$(\sim P)$ $a_1 \sim b_1 \to \ldots \to a_n \sim b_n \to pa_1 \ldots a_n \to pb_1 \ldots b_n$
für jedes n-stellige Prädikatszeichen p aus $L(T) \cup \{\sim\}$

zu $Ax(T)$ hinzu und ersetzt $=$ in den Axiomen von T überall durch $\sim$. So erhält man eine abzählbare identitätsfreie Theorie $T^\sim$, auf die 19.2.5 anwendbar ist und die dasselbe Ausdrucksvermögen hat wie T. Ferner ist $T^\sim$ offen, wenn T offen ist, und die neuen Axiome sind primen Sequenzen zugeordnet.

19.3 Substitutionen

Substitutionen wurden in §1 in allgemeiner Form eingeführt: Sie entstehen, indem man in beliebigen Nennformen Nennzeichen durch Nennformen ersetzt. Im Folgenden interessiert uns nur die spezielle Situation, dass man in Termen oder Formeln für freie Variablen Terme einsetzt. Diese Situation soll hier mathematisch etwas weiter analysiert werden, als es in §1 nötig war. Wir führen deshalb den Substitutionsbegriff noch einmal ein, zwar mit dem alten Namen, aber in der Form, in der er beim Automatischen Beweisen üblich ist.

19.3.1 Definition Es sei L eine Sprache, FV die Menge der freien Variablen und $T(L)$ die Menge der Terme von L. Eine Abbildung

$$\sigma : FV \to T(L)$$

ist eine *Substitution*, wenn die Menge

$$dom\sigma := \{a \in FV \mid a \not\equiv \sigma(a)\}$$

endlich ist. Weiter sei

$$var\sigma := \bigcup\{FV(\sigma(a)) \mid a \in dom\sigma\} \text{ und}$$
$$var^+\sigma := \bigcup\{FV(\sigma(a)) \mid a \in FV\}.$$

Man nennt $dom\sigma$ den *Definitionsbereich* von σ (in Abweichung von üblicher mathematischer Terminologie, vgl. 11.1, passender wäre die Bezeichnung *Support* von σ), $var\sigma$ die *Variablenmenge von* σ und $var^+\sigma$ die *erweiterte Variablenmenge von* σ. Ist $var\sigma = \emptyset$, also $\sigma(a) \in CT(L)$ für alle $a \in dom\sigma$, so heißt σ *Grundsubstitution*.

Für Substitutionen σ ist $dom\sigma$ per Definition stets eine endliche Menge (von freien Variablen). Daher ist auch $var\sigma$ eine endliche Menge als endliche Vereinigung endlicher Mengen. Dagegen ist $var^+\sigma$ stets eine unendliche Menge, weil $var^+\sigma$ jedenfalls das gesamte Komplement von $dom\sigma$ enthält. Tatsächlich ist

$$var^+\sigma = var\sigma \cup (FV - dom\sigma).$$

Beispiele.

1. Die Identität id (auf FV) ist eine Substitution mit $dom\ id = var\ id = \emptyset$, also eine Grundsubstitution.

2. Ist $dom\sigma = \{a, b\}, \sigma(a) = b$ und $\sigma(b) = (b + c)$, so ist σ eine Substitution mit $var\sigma = \{b, c\}$ und $var^+\sigma = FV - \{a\}$.

3. Ist 0 eine Konstante und $\sigma(a) = 0$ für $a \in dom\sigma$, so ist σ eine Grundsubstitution.

Substitutionen dienen dazu, die freien Variablen a in Termen und Formeln durch die Terme $\sigma(a)$ zu ersetzen. Daher ist klar, wie σ von FV auf Terme und Formeln fortzusetzen ist.

19.3.2 Rekursive Definition der *Instanz* $A\sigma$ von dem Term, der Formel bzw. der Sequenz A von L unter der Substitution σ.

1.1 $a\sigma :\equiv \sigma(a)$ für (freie) Variablen a

1.2 $(ft_1 \ldots t_n)\sigma :\equiv f(t_1\sigma) \ldots (t_n\sigma)$

2.1 $(pt_1 \ldots t_n)\sigma :\equiv p(t_1\sigma) \ldots (t_n\sigma)$ (auch für $p \equiv =$)

2.2 $\bot\sigma :\equiv \bot$

2.3 $(A \to B)\sigma :\equiv A\sigma \to B\sigma$

2.4 $\forall x F(x)\sigma :\equiv \forall x F\sigma(x)$, wobei $F\sigma(a) :\equiv F(a)\sigma$, falls $a \notin dom\sigma$ ist.

3.1 $\Gamma\sigma :\equiv \{C\sigma \mid C \in \Gamma\}$

3.2 $(\Gamma : \Delta)\sigma :\equiv \Gamma\sigma : \Delta\sigma$.

Ist $A\sigma$ geschlossen, so heißt $A\sigma$ *Grundinstanz* von dem Term, der Formel bzw. der Sequenz A.

19.3.3 Schreibweisen Wir schreiben $A\sigma \equiv B$, wenn die Zeichenreihe (Term oder Formel) $A\sigma$ die Zeichenreihe B ist. (Dagegen bezeichnet $t\sigma = t'$ die Gleichung zwischen den Termen $t\sigma$ und t', also eine Formel.)

Für die Substitution σ mit $dom\sigma = \{a_1, \ldots, a_n\}$ und $\sigma(a_i) \equiv t_i$ für $i = 1, \ldots, n$ (wobei die a_i paarweise verschiedene Variablen bezeichnen) schreibt man auch

$$[a_1/t_1, \ldots, a_n/t_n].$$

Insbesondere steht $[a/t]$ für die Substitution τ mit $dom\tau = \{a\}$ und $a\tau \equiv t$.

Um die (allgemeine) Nennform-Schreibweise aus 1.1 mit der Substitutions-Schreibweise bequem vergleichen zu können, setzen wir in 19.3.2 für Nennzeichen noch

$$*_i\sigma \equiv *_i.$$

Wir behandeln also Nennzeichen wie freie Variablen, die nicht zu $dom\sigma$ gehören. Dann ist $F\sigma$ definiert für alle Term- und Formelnennformen F.

Terme werden durch Substitutionen nicht kürzer. Das findet seinen wesentlichen Ausdruck in folgendem Analogon zum Homomorphie-Prinzip aus §2.3:

19.3.4 Homomorphie-Lemma

1. Ist t ein Term und F eine Term- oder Formelnennform, so ist

$$F(t)\sigma \equiv F\sigma(t\sigma).$$

2. Ist F n-stellig und geschlossen, so ist

$$F(a_1, \ldots, a_n)\sigma \equiv F(a_1\sigma, \ldots, a_n\sigma).$$

3. Ist speziell t ein Term, aber keine Variable, so beginnen t und $t\sigma$ mit demselben Funktionszeichen. Ist $t\sigma$ eine freie Variable, so auch t.

Analog zur entsprechenden Aussage über Belegungen haben wir hier:

19.3.5 Lemma Ist $a\sigma \equiv a\tau$ für alle $a \in FV(B)$, so ist

$$B\sigma \equiv B\tau,$$

gleichermaßen für Terme, Formeln und Sequenzen B.

Die einfachen Beweise überlegt man sich selbst.

Beispiele.

1. Es ist stets $Aid \equiv A$.

2. Ist $t(a)$ ein Term, so dass a in der Nennform t nicht auftritt, so ist

$$t(a)[a/s] \equiv t(s).$$

3. Es ist $(a + b)[b/a] \equiv (a + a)$ und $(a + b)[a/b, b/a] \equiv (b + a)$.

Durch die Fortsetzung 19.3.2 wird jede Substitution σ zu einer Abbildung, die Termen Terme und Formeln Formeln zuordnet. Die Komposition wird selbstverständlich für diese Fortsetzung definiert:

19.3.6 Definition Sind σ, τ Substitutionen, so ist $(\sigma\tau)$ (erst σ, dann τ) die Substitution, die durch

$$a(\sigma\tau) := (a\sigma)\tau \text{ für } a \in FV$$

gegeben ist.

Es ist also $a(\sigma\tau) \equiv (\sigma(a))\tau$, auch wenn $\sigma(a)$ keine Variable und daher $\tau(\sigma(a))$ gemäß 19.3.1 nicht definiert ist.

19.3.7 Lemma Für Substitutionen ρ, σ, τ gilt:

1. $\sigma id = id\sigma = \sigma$
2. $(\rho\sigma)\tau = \rho(\sigma\tau)$
3. $dom(\sigma\tau) \subseteq dom\sigma \cup dom\tau$
4. $var(\sigma\tau) \subseteq var\sigma \cup var\tau$
5. $var^+(\sigma\tau) \subseteq var^+\tau$.

Beweis.

1. und 2. sind trivial.
3. Ist $a \notin dom\sigma \cup dom\tau$, so ist $a(\sigma\tau) \equiv (a\sigma)\tau \equiv a\tau \equiv a$, und es ist $a \notin dom(\sigma\tau)$.
4. Sei $c \in var\sigma\tau$. Dann ist $c \in FV(a\sigma\tau)$ für ein $a \in dom\sigma\tau$.
    1. Fall. $a \notin dom\sigma$. Dann ist $a\sigma \equiv a$ und wegen 3. $a \in dom\tau$. Also ist $a\sigma\tau \equiv a\tau$ und $c \in var\tau$.
    2. Fall. $a \in dom\sigma$. Dann gibt es ein $b \in FV(a\sigma)$ mit $c \in FV(b\tau)$. Ist $b \in dom\tau$, so ist $c \in var\tau$. Sonst ist $c \equiv b\tau \equiv b \in var\sigma$.
5. Sei $c \in var^+\sigma\tau$. Dann ist $c \in FV(a\sigma\tau)$ für eine Variable a. Es gibt also ein $b \in FV(a\sigma)$, so dass $c \in FV(b\tau)$ ist. Es folgt $c \in var^+\tau$.

Bijektive Substitutionen $\pi : FV \to FV$ spielen eine besondere Rolle.

19.3.8 Definition *Permutationen* sind (bekanntlich) Bijektionen einer Menge auf sich selbst. Substitutionen, die zugleich Permutationen von FV sind, nennt man *Umbenennungen*. Unterscheiden sich zwei Substitutionen σ, τ nur um eine Umbenennung, d. h. gibt es zu σ, τ eine Umbenennung π mit $\sigma\pi = \tau$, so ist τ eine *Variante* von σ, man schreibt $\sigma \sim \tau$.

19.3.9 Lemma
1. Ist die Substitution $\pi : FV \to FV$ injektiv, so ist π bereits eine Umbenennung.
2. Die Umbenennungen bilden bezüglich der Komposition $\circ$ eine Gruppe.

3. Ist τ Variante von σ, so ist σ Variante von τ: Aus $\sigma \sim \tau$ folgt $\tau \sim \sigma$.

Beweis von 1. π ist außerhalb von $dom\pi$ die Identität. Da π injektiv ist, ist dann auch $\pi \restriction dom\pi : dom\pi \to dom\pi$ injektiv. Da $dom\pi$ endlich ist, ist deshalb $\pi \restriction dom\pi$ bijektiv, also auch π.

Die Beweise von 2. und 3. überlegt man sich leicht selbst.

19.4 Widerlegungsvollständige Regelsysteme

Uns interessieren Regeln, nach denen man aus primen Sequenzen auf andere prime Sequenzen und eventuell auf die leere Sequenz $\square$ schließen kann. Formal kann man diesen Regelbegriff im Rahmen einer festen Sprache L so definieren:

19.4.1 Definition Eine *(prime) Regel* R mit n Prämissen ist eine Menge von $(n+1)$-Tupeln von (primen) Sequenzen der Sprache L. Jedes Element $(\Gamma_0 : \Delta_0, \ldots, \Gamma_n : \Delta_n)$ von R ist ein *R-Schluss* mit den n *Prämissen* $\Gamma_0 : \Delta_0, \ldots, \Gamma_{n-1} : \Delta_{n-1}$ und der *Konklusion* $\Gamma_n : \Delta_n$ und wird meistens

$$(R)\ \ \Gamma_0 : \Delta_0, \ldots, \Gamma_{n-1} : \Delta_{n-1} \vdash \Gamma_n : \Delta_n$$

geschrieben.

Ein *Regelsystem* $\mathcal{R}$ ist eine endliche Menge von primen Regeln.

Typische Beispiele solcher primen Regeln sind die Substitutionsregel

$$(Subst)\quad \Gamma : \Delta \vdash \Gamma\sigma : \Delta\sigma$$

(für Substitutionen σ) und die Mix-Regel als formale Verallgemeinerung der Schnittregel (für Primformeln)

$$(Mix)\quad \Gamma_1, P : \Delta_1 \text{ und } \Gamma_2 : P, \Delta_2 \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2$$

Neben der (Mix)-Regel interessiert uns auch der *strikte Mix*

$$(sMix)\quad \Gamma_1, P : \Delta_1 \text{ und } \Gamma_2 : P, \Delta_2 \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2,$$
$$\text{falls } P \notin \Gamma_1 \cup \Delta_2 \text{ ist.}$$

Der Regelbegriff ist uns seit §3 bekannt. Während es dort darum ging, ein bestimmtes, für die gesamte klassische Prädikatenlogik vollständiges System einzuführen und Herleitungen auf seiner Basis eingehend zu studieren, wollen wir hier verschiedene Regelsysteme unter dem Aspekt der Widerlegungsvollständigkeit miteinander vergleichen. Da hier Varianten der Schnittregel in jedem betrachteten Regelsystem eine tragende Rolle spielen, werden auch die Theorie-Axiome in Herleitungen anders verwendet als in §3.

19.4.2 Induktive Definition der $\mathcal{R}$-Herleitungen in einer Gentzen-Theorie T für ein primes Regelsystem $\mathcal{R}$.

1. Ist $\forall(\Gamma \to \Delta) \in Ax(T)$, so nennen wir $\Gamma : \Delta$ eine *Startsequenz* von T. $\Gamma : \Delta$ ist dann eine $\mathcal{R}$-Herleitung von $\Gamma : \Delta$ in T. (Im Fall $\Delta = \{\bot\}$ sei $\Gamma : \emptyset$ Startsequenz von T (statt $\Gamma : \bot$).)

2. Sind H_i $\mathcal{R}$-Herleitungen von $\Gamma_i : \Delta_i$ in T $(i < n)$ und ist

(R) $\Gamma_0 : \Delta_0, \ldots, \Gamma_{n-1} : \Delta_{n-1} \vdash \Gamma : \Delta$
ein R-Schluss für ein $R \in \mathcal{R}$, so ist

$$\frac{H_1 \ldots H_n}{\Gamma : \Delta}$$

eine $\mathcal{R}$-Herleitung von $\Gamma : \Delta$ in T.

Wir unterscheiden freie und gebundene Variablen. Dadurch ist zu einem Axiom $B \equiv \forall(\Gamma \to \Delta)$ die Startsequenz $\Gamma : \Delta$ i. a. nicht eindeutig festgelegt: Für jede Umbenennung π ist auch $\Gamma\pi : \Delta\pi$ eine Startsequenz zum Axiom B, weil B auch ein Allabschluss von $\Gamma\pi \to \Delta\pi$ ist. Und umgekehrt: Alle Startsequenzen zu $\forall(\Gamma \to \Delta)$ sind Varianten voneinander, insbesondere von $\Gamma : \Delta$.

19.4.3 Definition Gibt es eine $\mathcal{R}$-Herleitung von $\Gamma : \Delta$ in T, so ist $\Gamma : \Delta$ *$\mathcal{R}$-herleitbar in T*, und wir schreiben $T \mid\!\!\!\frac{}{\mathcal{R}}\, \Gamma : \Delta$.

19.4.4 Definition Ein Regelsystem $\mathcal{R}$ ist *widerlegungsvollständig*, wenn für alle Gentzen-Theorien T gilt:

$$\text{Hat } T \text{ kein Modell, so ist } T \mid\!\!\!\frac{}{\mathcal{R}}\, \square.$$

Dies ist offenbar eine Vollständigkeitsaussage, denn es ist die Kontraposition von:

Jede $\mathcal{R}$-konsistente Gentzen-Theorie hat ein Modell.

Man halte sich allerdings vor Augen, dass hieraus für viele Regelsysteme $\mathcal{R}$ keineswegs die allgemeine Vollständigkeit folgt in der Form:

$$\textit{Wenn } \Gamma : \Delta \textit{ in } T \textit{ gilt, ist } T \mid\!\!\!\frac{}{\mathcal{R}}\, \Gamma : \Delta.$$

Ziel dieses Abschnitts ist der Nachweis, dass das Regelsystem $\{(Mix), (Subst)\}$ widerlegungsvollständig ist. Als erstes zeigen wir einige Eigenschaften der strikten Mix-Regel $(sMix)$, die nicht für die (Mix)-Regel gelten. Im folgenden sei T stets eine Gentzen-Theorie.

19.4.5 Lemma Es sei P ein Primsatz und $\Gamma : \Delta \not\equiv \emptyset : P$.
Ist $T + \{P\} \mid\!\!\!\frac{}{sMix}\, \Gamma : \Delta$, dann ist $T \mid\!\!\!\frac{}{sMix}\, \Gamma, P : \Delta$ oder $T \mid\!\!\!\frac{}{sMix}\, \Gamma : \Delta$.

Beweis durch Herleitungsinduktion.

1. $\Gamma : \Delta$ ist Startsequenz von $T + \{P\}$. Da $\Gamma : \Delta \not\equiv \emptyset : P$ ist, ist $\Gamma : \Delta$ auch Startsequenz von T, und es ist $T \mid\!\!\!\frac{}{sMix}\, \Gamma : \Delta$.

2. $\Gamma : \Delta \equiv \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2$ ist $(sMix)$-hergeleitet in $T + \{P\}$, und der letzte Schluss ist

   $$(sMix) \qquad \Gamma_1, Q : \Delta_1 \text{ und } \Gamma_2 : Q, \Delta_2 \vdash \Gamma : \Delta$$

   Wir schreiben $\Gamma, (P) : \Delta$ für $\Gamma, P : \Delta$ oder $\Gamma : \Delta$.

   2.1 Es ist $\Gamma_2 : Q, \Delta_2 \not\equiv \emptyset : P$. Nach Induktionsvoraussetzung ist

   $$T \mid\!\!\!\frac{}{sMix}\, \Gamma_1, Q, (P) : \Delta_1 \text{ und } T \mid\!\!\!\frac{}{sMix}\, \Gamma_2, (P) : Q, \Delta_2.$$

   Mit $(sMix)$ folgt $T \mid\!\!\!\frac{}{sMix}\, \Gamma, (P) : \Delta$.

   2.2 Es ist $\Gamma_2 : Q, \Delta_2 \equiv \emptyset : P$, also $\Gamma_2 = \Delta_2 = \emptyset$ und $P \equiv Q$.
   Nach Induktionsvoraussetzung ist

   $$T \mid\!\!\!\frac{}{sMix}\, \Gamma_1, P, (P) : \Delta_1 \equiv \Gamma, P : \Delta.$$

Mit Induktion folgt die Behauptung.

19.4.6 Lemma Es sei P ein Primsatz und $\Gamma : \Delta \not\equiv P : \emptyset$.
Ist $T + \{\neg P\} \mid\!\!\!\frac{}{sMix}\, \Gamma : \Delta$, dann ist $T \mid\!\!\!\frac{}{sMix}\, \Gamma : P, \Delta$ oder $T \mid\!\!\!\frac{}{sMix}\, \Gamma : \Delta$.

Der Beweis ist analog zu dem vorigen. Man überlegt ihn sich leicht selbst.

19.4.7 Konsistenzlemma Aus $T + \{P\} \overline{\vdash}_{sMix} \square$ und $T + \{\neg P\} \overline{\vdash}_{sMix} \square$ folgt $T \overline{\vdash}_{sMix} \square$.

Beweis. Aus den beiden Voraussetzungen folgt nach den beiden vorigen Lemmata:

$$T \overline{\vdash}_{sMix} (P) : \emptyset \text{ und } T \overline{\vdash}_{sMix} \emptyset : (P).$$

Dann steht die Behauptung $T \overline{\vdash}_{sMix} \square$ bereits da, oder sie folgt mit dem Schluss

$$(sMix) \qquad\qquad P : \emptyset \text{ und } \emptyset : P \vdash \square.$$

19.4.8 Satz Sei T eine quantorenfreie Gentzen-Theorie, d. h. alle Startsequenzen von T sind geschlossen. Hat T kein Modell, so ist $T \overline{\vdash}_{sMix} \square$:

$\{(sMix)\}$ *ist widerlegungsvollständig für quantorenfreie Gentzen-Theorien.*

Beweis. Wir nehmen an, dass nicht $T \overline{\vdash}_{sMix} \square$, und konstruieren daraus ein Herbrand-Modell von T. Es sei $(P_n)_{n \in \mathbb{N}}$ eine Aufzählung aller Primsätze (ohne Gleichungen) von $L(T)$. Wir definieren rekursiv eine Folge $(T_n)_{n \in \mathbb{N}}$ von Gentzen-Theorien wie folgt:

$$T_0 := T$$
$$T_{n+1} := \begin{cases} T_n + \{P_n\}, \text{ falls } T_n + \{P_n\} \not\overline{\vdash}_{sMix} \square \\ T_n + \{\neg P_n\} \text{ sonst.} \end{cases}$$

Wir zeigen $T_n \not\overline{\vdash}_{sMix} \square$ durch Induktion nach n.

1. $T_0 \not\overline{\vdash}_{sMix} \square$ nach Annahme.

2. Wenn $T_{n+1} = T_n + \{P_n\}$ ist, ist $T_{n+1} \not\overline{\vdash}_{sMix} \square$ nach Konstruktion von T_{n+1}. Sei nun $T_{n+1} = T_n + \{\neg P_n\}$. Dann ist $T_n + \{P_n\} \overline{\vdash}_{sMix} \square$ nach Konstruktion von T_{n+1}. Also folgt aus $T_{n+1} \overline{\vdash}_{sMix} \square$ mit dem Konsistenzlemma 19.4.7 auch $T_n \overline{\vdash}_{sMix} \square$, im Widerspruch zur Induktionsvoraussetzung.
   Also ist stets $T_{n+1} \not\overline{\vdash}_{sMix} \square$.

Mit Induktion nach n folgt $T_n \not\vdash_{sMix} \square$ für alle n.

Wir setzen

$$T^+ := \bigcup\{T_n \mid n \in \mathbb{N}\} := (L(T), \bigcup\{Ax(T_n) \mid n \in \mathbb{N}\}).$$

Dann ist auch $T^+ \not\vdash_{sMix} \square$ wegen der Endlichkeit der $(sMix)$-Herleitungen, und für jeden Primsatz P aus $L(T)$, der keine Gleichung ist, ist $P \in Ax(T^+)$ oder $\neg P \in Ax(T^+)$. Sei nun $\mathcal{A}$ die Herbrand-Struktur zu $L(T)$ mit

$$(t_1, \ldots, t_n) \in p_{\mathcal{A}} \Leftrightarrow pt_1 \ldots t_n \in Ax(T^+),$$

also $\mathcal{A}(P) = w \Leftrightarrow P \in Ax(T^+) \Leftrightarrow \neg P \notin Ax(T^+)$.
Wir behaupten, dass $\mathcal{A}$ ein Modell von T^+ ist. Angenommen, das wäre nicht so. Dann gibt es einen Satz $\Gamma \to \Delta \in Ax(T^+)$, so dass $\mathcal{A} \not\models \Gamma : \Delta$. Dann ist für alle $P \in \Gamma$:

$$\mathcal{A}(P) = w, \text{ also } P \in Ax(T^+) \text{ und } T^+ \vdash \emptyset : P$$

und für alle $Q \in \Delta$:

$$\mathcal{A}(Q) = f, \text{ also } \neg Q \in Ax(T^+) \text{ und } T^+ \vdash Q : \emptyset.$$

Dann schließt man in T^+ aus den Startsequenzen $Q : \emptyset$ für $Q \in \Delta$ und $\Gamma : \Delta$ mit $card(\Delta)$ vielen $(sMix)$-Schlüssen auf $\Gamma : \emptyset$ und hieraus und den Startsequenzen $\emptyset : P$ für $P \in \Gamma$ mit $card(\Gamma)$ vielen (Mix)-Schlüssen auf $\square$.
Es folgt $T^+ \vdash_{sMix} \square$, im Widerspruch zum oben Bewiesenen.

Also ist $\mathcal{A}$ ein Modell von T^+ und deshalb auch von T, weil T^+ eine einfache Erweiterung von T ist. Damit ist der Satz bewiesen.

19.4.9 Korollar $\{(Mix)\}$ ist widerlegungsvollständig für quantorenfreie Gentzen-Theorien.

Beweis. Jeder $(sMix)$-Schluss ist offenbar auch ein (Mix)-Schluss. Also ist jede $\{(sMix)\}$-Herleitung auch eine $\{(Mix)\}$-Herleitung, und aus $T \vdash_{sMix} \square$ folgt $T \vdash_{Mix} \square$. Daher liefert der Satz die Behauptung.
Dies ist ein Vollständigkeitssatz relativ zu einer einzigen Regel, allerdings für recht spezielle Theorien, und diese einzige Regel ist ausgerechnet eine Fassung

der Schnittregel, die wir in Kapitel 3 sorgfältig vermieden haben, weil wir dort mit dem Vollständigkeitssatz die Zulässigkeit der Schnittregel überhaupt erst beweisen wollten. Der Beweis verwendet wieder ein Henkin-Verfahren. Er ist allerdings sehr viel einfacher als in Kapitel 3, weil hier die Axiome quantorenfrei sind und deshalb schon jede einfache maximal-konsistente Erweiterung T^+ ein kanonisches Modell besitzt und damit ein Modell von T liefert.

Noch zu tun bleibt, dieses Ergebnis von quantorenfreien auf beliebige Gentzen-Theorien hochzuheben.

19.4.10 Definition Zu einer Gentzen-Theorie T sei die *quantorenfreie* oder *Grundversion* von T die Theorie

$$\begin{aligned}\mathrm{grund}(T) := (L(T), \{(\Gamma \to \Delta)\sigma | \forall(\Gamma \to \Delta) \in Ax(T),\\ (\Gamma \to \Delta)\sigma \text{ Grundinstanz von } \Gamma \to \Delta\}).\end{aligned}$$

Die Theorien T und $\mathrm{grund}(T)$ haben also dieselbe Sprache $L(T)$, aber die Startsequenzen von $\mathrm{grund}(T)$ sind gerade die sämtlichen Grundsequenzen, die durch Substitution aus den Startsequenzen von T hervorgehen. T ist offenbar eine $\vdash$-Erweiterung von $\mathrm{grund}(T)$.

19.4.11 Lemma Eine Gentzen-Theorie T hat genau dann ein Modell, wenn ihre Grundversion $\mathrm{grund}(T)$ ein Modell hat.

Beweis. Jedes Modell von T ist offenbar auch ein Modell von $\mathrm{grund}(T)$.
Hat nun $\mathrm{grund}(T)$ ein Modell, so hat $\mathrm{grund}(T)$ nach Satz 19.1.4 ein Herbrand-Modell $\mathcal{A}$. Da $|\mathcal{A}|$ gerade aus den geschlossenen Termen von $L(T)$ besteht, ist $\mathcal{A}(\forall B) = w$ äquivalent zu

$$\mathcal{A}(B\sigma) = w \text{ für alle Grundinstanzen } B\sigma \text{ von } B.$$

Für $B \equiv \Gamma \to \Delta$ ist aber $\forall B \in Ax(T)$ genau dann, wenn alle Grundinstanzen $B\sigma$ von B Axiome von $\mathrm{grund}(T)$ sind. Also ist $\mathcal{A}$ auch ein Modell von T.

19.4.12 Lemma Sei $\mathcal{R}$ ein Regelsystem.

$$\text{Aus } \mathrm{grund}(T) \mid\!\!\!\frac{}{\mathcal{R}} \Gamma : \Delta \text{ folgt } T \mid\!\!\!\frac{}{\mathcal{R}, Subst} \Gamma : \Delta.$$

Beweis durch $\mathcal{R}$-Herleitungsinduktion.

1. $\Gamma : \Delta$ ist Startsequenz von grund(T), d. h. $\Gamma \to \Delta \in Ax(\text{grund}(T))$. Dann ist $\Gamma \to \Delta$ eine Grundinstanz $(\Gamma_0 \to \Delta_0)\sigma$ von einer Formel $\Gamma_0 \to \Delta_0$, deren Allabschluss in $Ax(T)$ ist. Also ist $\Gamma_0 : \Delta_0$ Startsequenz von T, und mit dem Substitutionsschluss

$$(Subst) \quad \Gamma_0 : \Delta_0 \vdash \Gamma_0\sigma : \Delta_0\sigma \equiv \Gamma : \Delta$$

   folgt $T \mid\!\!\overline{}_{Subst}\, \Gamma : \Delta$.

2. $\Gamma : \Delta$ ist in grund(T) mit einem R-Schluss (für $R \in \mathcal{R}$) als letztem Schluss hergeleitet. Dann ist dieser Schluss auch ein R-Schluss in T. Nach Induktionsvoraussetzung sind die Prämissen des Schlusses in T $\mathcal{R} \cup \{(Subst)\}$-herleitbar. Also ist auch $\Gamma : \Delta$ in T $\mathcal{R} \cup \{(Subst)\}$-herleitbar.

Mit Induktion folgt die Behauptung.

19.4.13 Lifting-Satz Ist $\mathcal{R}$ für quantorenfreie Gentzen-Theorien widerlegungsvollständig, so ist $\mathcal{R} \cup \{(Subst)\}$ (für alle Gentzen-Theorien) widerlegungsvollständig.

Beweis. Sei T eine Gentzen-Theorie, die kein Modell besitzt. Dann hat nach 19.4.11 ihre Grundversion grund(T) auch kein Modell. Da grund(T) eine quantorenfreie Gentzen-Theorie ist, folgt aus der Voraussetzung des Satzes

$$\text{grund}(T) \mid\!\!\overline{}_{\mathcal{R}}\, \square.$$

Mit 19.4.12 folgt $T \mid\!\!\overline{}_{\mathcal{R},Subst}\, \square$.

Also ist $\mathcal{R} \cup \{(Subst)\}$ widerlegungsvollständig.

19.4.14 Satz $\{(Mix), (Subst)\}$ und sogar $\{(sMix), (Subst)\}$ ist widerlegungsvollständig (für alle Gentzen-Theorien).

Beweis. Nach 19.4.8 und 19.4.9 sind $\{(sMix)\}$ und $\{(Mix)\}$ widerlegungsvollständig für quantorenfreie Gentzen-Theorien. Mit Lifting gemäß 19.4.13 folgt dann die Behauptung.

19.5 Aufgaben

19.5.1 T sei eine identitätsfreie Theorie, deren Sprache durch ein 2-stelliges Prädikatszeichen $<$ und die endlich vielen Konstanten $a_0, \ldots, a_n (n \geq 0)$ gegeben ist, mit den Axiomen

$$\forall x \neg x < x, \forall x \forall y \forall z (x < y \to y < z \to x < z), \forall x \exists y\ x < y.$$

Zeigen Sie:

1. T hat ein Modell, aber kein Herbrand-Modell.
2. Ist f ein 1-stelliges Funktionszeichen, so hat $T + \{f\}$ auch ein Herbrand-Modell.

19.5.2 Man gebe an

a. eine Grundsubstitution σ und einen Term t, so dass $t\sigma$ kein Grundterm ist;

b. einen Grundterm $t\sigma$, wobei t kein Grundterm und σ keine Grundsubstitution ist;

c. ein (einfaches) Kriterium dafür, dass $t\sigma$ ein Grundterm ist.

19.5.3 Zeigen Sie

$$F(t_1, \ldots, t_n)\sigma \equiv F\sigma(t_1\sigma, \ldots, t_n\sigma)$$

für Term- und Formelnennformen F und folgern Sie daraus 19.3.4, 1 und 2.

19.5.4 Zeigen Sie: $\sigma\sigma = \sigma \Leftrightarrow dom\sigma \cap var\sigma = \emptyset$.

19.5.5 Zeigen Sie 19.3.9, 2 und folgern Sie daraus 19.3.9, 3.

19.5.6 Zeigen Sie für die Regelsysteme $\mathcal{R} = \{(Mix)\}$ und $\mathcal{R} = \{(sMix)\}$ sowie für alle Regelsysteme $\mathcal{R}$, die $(Subst)$ enthalten:

Aus $T \mid\!\!\frac{}{\mathcal{R}}\ \Gamma : \Delta$ folgt $T \mid\!\!\frac{}{\mathcal{R}}\ \Gamma\pi : \Delta\pi$ für jede Umbenennung π.

19.5.7 $L(T)$ sei durch das Aussagezeichen p gegeben, und $\neg p$ sei einziges Axiom von T (so dass $p : \emptyset$ einzige Startsequenz von T ist). Zeigen Sie:

a. $T + \{p\} \mid\!\!\frac{}{Mix}\ p : p$, aber $T + \{p\} \not\mid\!\!\frac{}{sMix}\ p : p$.

b. Lemma 19.4.5 wird falsch, wenn man dort $\mid\!\!\frac{}{sMix}$ durch $\mid\!\!\frac{}{Mix}$ ersetzt.

§20 Unifikation

20.1 Unifikatoren

20.2 Berechnung allgemeinster Unifikatoren

20.3 Zum Aufwand der Unifikation

20.4 Aufgaben

20.1 Unifikatoren

20.1.1 Definition Eine Substitution σ ist *Unifikator* einer endlichen nichtleeren Menge M von Termen, σ *unifiziert* M, wenn

$$s\sigma \equiv t\sigma$$

für alle $s, t \in M$ ist, wenn also

$$\sigma[M] := \{t\sigma \mid t \in M\}$$

aus einem einzigen Term besteht.

M ist *unifizierbar*, wenn es eine Substitution σ gibt, die M unifiziert.

Zwei Terme s, t sind *unifizierbar*, wenn die Menge $\{s, t\}$ unifizierbar ist.

Beispiele.

1. Sind $a \cdot (a + 0)$ und $a \cdot (b \cdot c + b)$ unifizierbar? Für jeden Unifikator σ von $\{a \cdot (a + 0), a \cdot (b \cdot c + b)\}$ muss jedenfalls

   $$a\sigma \equiv (b \cdot c)\sigma \text{ und } b\sigma \equiv 0\sigma \equiv 0$$

   sein. Für $dom\sigma = \{a, b\}$ und

   $$\begin{array}{l} a\sigma \equiv 0 \cdot c \text{ und } b\sigma \equiv 0 \quad \text{folgt} \\ a \cdot (a + 0)\sigma \equiv 0 \cdot c \cdot (0 \cdot c + 0) \equiv a \cdot (b \cdot c + b)\sigma, \end{array}$$

   und σ unifiziert $a \cdot (a + 0)$ und $a \cdot (b \cdot c + b)$.

2. Sind $a + (b + c)$ und $(a + b) + c$ unifizierbar? Jeder Unifikator σ von $\{a + (b + c), (a + b) + c\}$ müsste erfüllen:

$$a\sigma \equiv (a + b)\sigma \equiv (a\sigma + b\sigma) \text{ und } c\sigma \equiv (b + c)\sigma \equiv (b\sigma + c\sigma),$$

und das geht nicht, weil z. B. $+b\sigma)$ und $(b\sigma+$ nicht leer sein können. Auch in der klammerfreien Schreibweise $+a + bc$ und $++abc$ sind die Terme nicht unifizierbar. Dagegen sind $+a+bc$ und $++a'bc'$ mit verschiedenen Variablen a, a', b, c, c' unifizierbar.

20.1.2 Definition Seien σ, τ Substitutionen. σ ist *allgemeiner* als τ, σ *subsumiert* τ, man schreibt $\sigma \leq \tau$, wenn es eine Substitution ρ gibt, so dass

$$\sigma\rho = \tau$$

ist. σ und τ sind *äquivalent*, man schreibt $\sigma \approx \tau$, wenn

$$\sigma \leq \tau \text{ und } \tau \leq \sigma$$

ist. Man sieht sofort:

20.1.3 Lemma Die Relation $\leq$ (zwischen Substitutionen) ist reflexiv und transitiv. Die Relation $\approx$ der Äquivalenz von Substitutionen ist eine Äquivalenzrelation.

Beweis. Wegen $\sigma id = \sigma$ ist $\sigma \leq \sigma$. Ist $\rho\sigma' = \sigma$ und $\sigma\tau' = \tau$, so ist $\rho(\sigma'\tau') = \sigma\tau' = \tau$, also $\rho \leq \tau$, und $\leq$ ist transitiv. Hieraus folgt (in allgemeiner Form), dass $\approx$ eine Äquivalenzrelation ist.

Beispiele.

1. Wegen $id\tau = \tau$ ist $id \leq \tau$ für jede Substitution τ.
2. Es ist stets $\sigma \leq \sigma^2 = \sigma\sigma \leq \sigma^3 \leq \ldots$.

Für die Unifikation von Termmengen erhält man:

20.1.4 Lemma Ist σ Unifikator von M und ist $\sigma \leq \tau$, so ist auch τ Unifikator von M.

Beweis. Wenn $\{t\sigma \mid t \in M\}$ aus einem einzigen Term s besteht, besteht $\{t\sigma\tau \mid t \in M\} = \{s\tau\}$ aus dem einzigen Term $s\tau$.

20.1.5 Korollar Ist $\sigma \approx \tau$, so unifizieren σ und τ dieselben nicht-leeren Mengen von Termen.

Äquivalente Substitutionen sind bereits Varianten voneinander. Um das zu beweisen, zeigen wir zunächst:

20.1.6 Lemma Ist $\sigma\tau = \sigma$, so ist $\tau \restriction var^+\sigma = id \restriction var^+\sigma$.

Beweis. Sei $b \in var^+\sigma$. Dann gibt es eine Variable a, so dass $b \in FV(a\sigma)$. Wegen $a\sigma \equiv (a\sigma)\tau$ ist dann $b\tau \equiv b$. Also ist $\tau \restriction var^+\sigma$ die Identität auf $var^+\sigma$. Man überlegt sich, dass τ außerhalb von $var^+\sigma$ von der Identität abweichen kann.

20.1.7 Lemma Seien π, ρ, σ Substitutionen. Aus $\pi \restriction var^+\sigma = \rho \restriction var^+\sigma$ folgt stets $\sigma\pi = \sigma\rho$.

Beweis. Für jede Variable a ist $FV(a\sigma) \subseteq var^+\sigma$. Nach der Voraussetzung operieren also π und ρ auf $FV(a\sigma)$ und damit nach 19.3.5 auch auf $a\sigma$ gleich: es ist $a\sigma\pi = a\sigma\rho$.

Außerdem brauchen wir eine allgemeine kombinatorische Aussage über endliche Abbildungen:

20.1.8 Lemma Es sei X_0 eine endliche Teilmenge einer Menge Y. Dann lässt sich jede injektive Abbildung $\varphi_0 : X_0 \to Y$ fortsetzen zu einer Permutation φ der endlichen Teilmenge $X_0 \cup im(\varphi_0)$ von Y.

Beweis. Sei $\varphi_0 : X_0 \to Y$ injektiv. Dann ist

$$\varphi_0 : X_0 \to im(\varphi_0)$$

bijektiv. Nun ist X_0 endlich. Wenn X_0 n Elemente hat, hat auch $im(\varphi_0)$ n Elemente, und auch $X := X_0 \cup im(\varphi_0)$ ist endlich, etwa mit k Elementen, wobei $n \leq k \leq 2n$ ist. Dann haben $X - X_0$ und $X - im(\varphi_0)$ beide $k - n$ Elemente, sind also gleichmächtig, so dass es eine Bijektion

$$\varphi_1 : X - X_0 \to X - im(\varphi_0)$$

gibt. Dann ist offenbar $\varphi := \varphi_0 \cup \varphi_1 : X \to X$ eine Bijektion, mithin eine Permutation von X, und es ist $\varphi \restriction X_0 = \varphi_0$, so dass φ eine Fortsetzung von φ_0 ist.

Nun folgt endlich:

20.1.9 Satz $\sigma \approx \tau$ ist äquivalent zu $\sigma \sim \tau$.

Beweis. Seien zunächst σ, τ Varianten voneinander, $\sigma \sim \tau$. Dann gibt es nach 19.3.8 eine Umbenennung π, für die $\sigma\pi = \tau$, also $\sigma \leq \tau$, und $\tau\pi^{-1} = \sigma$, also $\tau \leq \sigma$ ist. Es folgt $\sigma \approx \tau$.

Sei umgekehrt $\sigma \approx \tau$, also $\sigma \leq \tau$ und $\tau \leq \sigma$. Es gibt also Substitutionen ρ, ρ', für die $\sigma\rho = \tau$ und $\tau\rho' = \sigma$ ist. Dann ist $\sigma\rho\rho' = \tau\rho' = \sigma$, und nach 20.1.6 ist $\rho\rho' \restriction var^+\sigma$ die Identität auf $var^+\sigma$. Für $b, c \in var^+\sigma$ folgt also aus $b\rho \equiv c\rho$ stets

$$b \equiv b\rho\rho' \equiv c\rho\rho' \equiv c :$$
$$\rho \restriction var^+\sigma : var^+\sigma \to FV$$

ist injektiv. Damit ist für die endliche Menge $X_0 := dom\rho \cap var^+\sigma$ und $Y := dom\rho$ die Situation von 20.1.8 hergestellt, weil ρ außerhalb $dom\rho$ ohnehin die Identität ist. Also gibt es eine Permutation φ von $X := X_0 \cup \rho[X_0]$, die $\rho \restriction X_0$ fortsetzt. Setzt man

$$\pi(a) = \varphi(a) \text{ für } a \in X$$
$$\pi(a) = a \text{ für } a \in FV - X,$$

so ist π eine Umbenennung, die $\rho \restriction var^+\sigma$ fortsetzt. Nach 20.1.7 ist dann $\sigma\pi = \sigma\rho = \tau$, und τ ist eine Variante von σ.

20.1.10 Definition Sei M eine nicht-leere Menge von Termen. Eine Substitution σ ist ein *allgemeinster Unifikator von* M, σ ist $mgu(M)$ ($m$ost $g$eneral $u$nifier von M), man schreibt $\sigma = mgu(M)$, wenn

1. σ Unifikator von M ist und
2. für jeden Unifikator τ von M $\sigma \leq \tau$ ist.

Die Gleichung $\sigma = mgu(M)$ ist offenbar nicht im strikten Sinne zu lesen, denn die $mgu(M)$ sind gegen Äquivalenz abgeschlossen:

20.1.11 Lemma Ist $\sigma = mgu(M)$, so ist jedes zu σ äquivalente τ ebenfalls $mgu(M)$.

Beweis. Ist $\sigma \approx \tau$, so ist mit σ nach 20.1.5 auch τ Unifikator von M. Ist σ' ein beliebiger Unifikator von M, so ist $\tau \leq \sigma \leq \sigma'$, also nach 20.1.3 auch $\tau \leq \sigma'$, und τ ist $mgu(M)$.

Wichtiger ist die Umkehrung:

20.1.12 Satz *Eindeutigkeit des allgemeinsten Unifikators.* Sind σ und τ allgemeinste Unifikatoren einer nicht-leeren Menge M von Termen, so ist $\sigma \sim \tau$.

Beweis. Sind σ und τ $mgu(M)$, so ist $\sigma \leq \tau$ und $\tau \leq \sigma$, also $\sigma \approx \tau$. Nach 20.1.9 ist dann $\sigma \sim \tau$.

Damit ist der mgu einer Menge M modulo $\sim$ eindeutig bestimmt, sofern er existiert. Unter welchen Bedingungen M überhaupt einen mgu besitzt und wie man ihn berechnet, wird im nächsten Abschnitt untersucht.

20.2 Berechnung allgemeinster Unifikatoren

Es ist klar, dass zwei Terme i. a. nicht unifizierbar sind.

Beispiele.

1. Zwei Terme, die mit verschiedenen Funktionszeichen beginnen, sind nach 19.3.4, 3 nicht unifizierbar.

2. Ist $a \in FV(t)$, aber $a \not\equiv t$, so sind a und t nicht unifizierbar. Denn da a echter Subterm von t ist, ist nach 19.3.4, 1 auch $a\sigma$ stets echter Subterm von $t\sigma$.

Anders liegt der Fall, wenn a in t nicht auftritt.

20.2.1 Lemma Ist $a \notin FV(t)$, so ist $[a/t]$ ein mgu von a und t. Genauer: Für jeden Unifikator σ von a und t ist

$$[a/t]\sigma = \sigma.$$

Beweis. Wegen $a \notin FV(t)$ ist

$$a[a/t] \equiv t \equiv t[a/t].$$

Also ist $[a/t]$ Unifikator von a und t.
Ist nun σ irgendein Unifikator von a und t, so ist

$$a[a/t]\sigma \equiv t\sigma \equiv a\sigma,$$

und für alle anderen Variablen b ist

$$b[a/t]\sigma \equiv b\sigma.$$

Also ist insgesamt $[a/t]\sigma = \sigma$ und damit $[a/t] \leq \sigma : [a/t]$ ist $mgu\{a, t\}$.

Eine naheliegende Verallgemeinerung des Unifikators von zwei Termen ist der Begriff der *Lösung* eines Gleichungssystems.

20.2.2 Definition Eine Substitution σ *löst* eine endliche Menge Γ von Gleichungen, σ ist *Lösung* von Γ, wenn $s\sigma \equiv t\sigma$ ist für alle Gleichungen $s = t$ aus Γ.

Im folgenden stehen Γ und Δ (auch mit Indizes) für endliche Mengen von Gleichungen. $\Gamma, s = t$ steht für $\Gamma \cup \{s = t\}$, und $\Gamma\tau$ steht für die Gleichungsmenge

$$\{s\tau = t\tau \mid s = t \in \Gamma\}.$$

$\Gamma[a/t]$ geht nach 19.3.3 also aus Γ hervor, indem man jedes Auftreten von a in Γ durch t ersetzt.

Unmittelbar aus der Definition ergibt sich:

20.2.3 Lemma Für Substitutionen σ, τ gilt:

(1) σ löst $\Gamma \cup \Delta \Leftrightarrow \sigma$ löst Γ und σ löst Δ

(2) σ löst $\Gamma\tau \Leftrightarrow \tau\sigma$ löst Γ.

Beweis von (2). Beide Seiten besagen, dass $s\tau\sigma \equiv t\tau\sigma$ ist für alle Gleichungen $s = t \in \Gamma$.

Völlig analog zu 20.2.1 zeigt man:

20.2.4 Lemma Eine Substitution τ mit

$$dom\tau \cap var\tau = \emptyset$$

löst die Gleichungsmenge

$$\Gamma = \{a = a\tau \mid a \in dom\tau\}.$$

τ ist allgemeinste Lösung von Γ in dem Sinne, dass für jede Lösung σ von Γ sogar $\tau\sigma = \sigma$ ist.

Es gibt verschiedene Algorithmen, die entscheiden, ob eine Gleichungsmenge eine Lösung besitzt, und die dabei zugleich allgemeinste Lösungen suchen und auch finden, falls sie überhaupt lösbar ist. Einen solchen Algorithmus betrachten wir hier näher.

20.2.5 Definition der Martelli-Montanari-Regeln

$$(MM1) \quad \Gamma, fs_1 \dots s_n = ft_1 \dots t_n \vdash \Gamma, s_1 = t_1, \dots, s_n = t_n, \quad \text{falls } fs_1 \dots s_n = ft_1 \dots t_n \text{ nicht in } \Gamma \text{ ist.}$$

$$(MM2) \quad \Gamma, a = t \vdash \Gamma[a/t], a = t, \text{ falls } a = t \notin \Gamma, \quad a \notin FV(t), \text{ aber } a \in FV(\Gamma) \text{ ist.}$$

$$(MM3) \quad \Gamma, t = a \vdash \Gamma, a = t, \text{ falls } t = a \notin \Gamma, t \notin FV \text{ ist.}$$

$$(MM4) \quad \Gamma, a = a \vdash \Gamma, \text{ falls } a = a \notin \Gamma \text{ ist.}$$

Durch die Martelli-Montanari-Regeln ist offenbar ein Herleitungsbegriff gegeben, der endliche Gleichungsmengen aus endlichen Gleichungsmengen herleitet.

20.2.6 Definition Eine *MM-Herleitung* (Martelli-Montanari-Herleitung) ist eine endliche Folge

$$\Gamma_0, \Gamma_1, \dots, \Gamma_n$$

von endlichen Gleichungsmengen, in der

$$\Gamma_i \vdash \Gamma_{i+1}$$

für alle $i < n$ Schlüsse nach $(MM1)$ bis $(MM4)$ sind. Δ ist aus Γ_0 *MM-herleitbar* (nach dem *Martelli-Montanari-Algorithmus* (MM)), wir schreiben $\Gamma_0 \vdash_{MM} \Delta$, wenn es eine MM-Herleitung $\Gamma_0, \dots, \Gamma_n$ gibt, in der $\Gamma_n = \Delta$ ist. Dabei heißt Γ_0 auch *Start-Gleichungsmenge*; Δ heißt *terminierend*, wenn auf Δ keine der Regeln $(MM1)$ bis $(MM4)$ mehr anwendbar ist. Man sagt, (MM) *terminiert für* Γ_0, wenn es ein terminierendes Δ gibt mit $\Gamma_0 \vdash_{MM} \Delta$.

Wenn man in MM-Schlüssen Prämisse und Konklusion jeweils als Konjunktion liest und in Herbrand-Strukturen interpretiert, dann sind die vier Regeln auch logisch korrekt, sogar in beiden Richtungen. Also lösen dieselben Substitutionen Prämisse und Konklusion eines MM-Schlusses und damit auch die sämtlichen Gleichungsmengen einer MM-Herleitung.

Beispiel. Wir unterwerfen das erste Beispiel nach 20.1.1 den Martelli-Montanari-Regeln

$$\begin{array}{ll}
\dfrac{a \cdot (a+0) = a \cdot (b \cdot c + b)}{a = a, \quad a + 0 = b \cdot c + b} & \text{nach } MM1 \\[2ex]
\dfrac{a + 0 = b \cdot c + b}{a = b \cdot c, 0 = b} & \text{nach } MM4 \\[2ex]
\dfrac{a = b \cdot c, 0 = b}{a = b \cdot c, b = 0} & \text{nach } MM1 \\[2ex]
\dfrac{a = b \cdot c, b = 0}{a = 0 \cdot c, b = 0} & \text{nach } MM3 \\[2ex]
a = 0 \cdot c, b = 0 & \text{nach } MM2.
\end{array}$$

Die letzte Zeile löst offenbar die Substitution

$$\tau = [a/0 \cdot c, b/0],$$

und sie ist eine (allgemeinste) Lösung aller Gleichungsmengen der Herleitung. Insbesondere ist

$$\tau = mgu\{a \cdot (a+0), a \cdot (b \cdot c + b)\}.$$

Die oben angekündigte Korrektheit der MM-Regeln in Herbrand-Strukturen kann man so formulieren:

20.2.7 Lemma Ist $\Gamma \vdash \Delta$ ein Schluss nach einer Martelli-Montanari-Regel, so haben die Gleichungsmengen Γ und Δ dieselben Lösungen.

Beweis. Für die Regeln $(MM1)$, $(MM3)$ und $(MM4)$ ist die Behauptung unmittelbar einsichtig.

Zu $(MM2)$. Wenn σ die Prämisse oder Konklusion von

$$\Gamma, a = t \vdash \Gamma[a/t], a = t$$

löst, löst σ in jedem Fall auch $a = t$, wobei a in t nicht auftritt. Nach 20.2.1 ist also

$$[a/t]\sigma = \sigma.$$

Dann folgt mit 20.2.3:

$$[a/t]\sigma = \sigma \text{ löst } \Gamma, a = t \;\Leftrightarrow\; \sigma \text{ löst } \Gamma[a/t], a = t.$$

Durch Induktion nach der Länge der MM-Herleitungen folgt hieraus unmittelbar:

20.2.8 Korollar Ist $\Gamma_0 \mathrel{\vert\!\overline{\underline{\;\;}}_{MM}} \Delta$, so haben Γ_0 und Δ dieselben Lösungen.

Die MM-Regeln sind so formuliert, dass Prämisse und Konklusion eines Schlusses niemals identisch sind. Man kann eventuell mehrere MM-Regeln auf eine Gleichungsmenge anwenden, aber in jedem Fall ist die Konklusion eines MM-Schlusses in einem bestimmten Sinne einfacher als die Prämisse:

20.2.9 Satz Für jede Start-Gleichungsmenge terminiert (MM).

Beweis. Zunächst ordnen wir jeder Gleichungsmenge Γ ein Tripel (k, l, n) von natürlichen Zahlen zu. Dabei sei

- k die Anzahl der Variablen in Γ, aber ohne die Variablen a, zu denen es eine Gleichung $a = t \in \Gamma$ gibt, wobei a weder in t noch in $\Gamma - \{a = t\}$ auftritt;
- l die Summe der Längen aller Termauftreten s, t mit $s = t \in \Gamma$;
- n die Anzahl der Gleichungen der Gestalt $t = a \in \Gamma$, in denen t keine Variable ist.

Man beachte: k und n sind Anzahlen von unmittelbaren (formalen) Objekten der Theorie, nämlich von gewissen Variablen bzw. Gleichungen. l ist dagegen eine Anzahl von Auftreten von Grundzeichen. Man schreibe die Gleichungen aus Γ ohne Wiederholungen nebeneinander und lösche die Gleichheitszeichen; die Länge der so entstehenden Zeichenreihe ist l.

Die Menge aller dieser Tripel wird durch die lexikographische Ordnung von $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$, die wir wieder mit $<$ bezeichnen, linear geordnet, sogar wohlgeordnet (vgl. 12.1.8).

Sei nun $\Gamma \vdash \Delta$ ein MM-Schluss und sei $(k_\Gamma, l_\Gamma, n_\Gamma)$ der Menge Γ, $(k_\Delta, l_\Delta, n_\Delta)$ der Menge Δ zugeordnet. Wir zeigen durch Fallunterscheidung:

(3) $$(k_\Delta, l_\Delta, n_\Delta) < (k_\Gamma, l_\Gamma, n_\Gamma).$$

($MM1$): In Γ und Δ treten dieselben Variablen auf, es geht keine Gleichung $a = t$ verloren, wohl aber können Gleichungen $a_i = t$ hinzukommen, wenn $s_i \equiv a_i$ ist. Also ist $k_\Delta \leq k_\Gamma$. Weiter gehen zwei Auftreten von f verloren; im übrigen können noch Auftreten von Gleichungen $s_i = t_i$ verschwinden, wenn sie schon in der Prämisse auftraten. Dagegen werden die u. U. hinzutretenden Auftreten von $=$ nicht gezählt. Also ist stets $l_\Delta + 2 \leq l_\Gamma$ und daher $l_\Delta < l_\Gamma$. Insgesamt folgt (3).

($MM2$): Die Variable a trägt nicht zu k_Δ bei, wohl aber zu k_Γ. Trägt ein $b \in FV(\Gamma)$ nicht zu k_Γ bei, so tritt b nur einmal in Γ in einer Gleichung $b = s$ auf, insbesondere nicht in t. Dann tritt b auch in Δ nur einmal auf, in $b = s[a/t]$, und b trägt nicht zu k_Δ bei. Also ist $k_\Delta < k_\Gamma$, und es folgt (3).

($MM3$): Die Gleichung $a = t$ kann so in Δ auftreten, dass a zu k_Δ nicht beiträgt. Dann ist $k_\Delta < k_\Gamma$, sonst $k_\Delta = k_\Gamma$, jedenfalls $k_\Delta \leq k_\Gamma$. Ist $a = t \in \Gamma$, so ist $l_\Delta < l_\Gamma$, sonst $l_\Delta = l_\Gamma$, jedenfalls $l_\Delta \leq l_\Gamma$. In jedem Fall ist aber $n_\Delta < n_\Gamma$, es folgt (3).

($MM4$): a trägt in jedem Fall zu k_Γ bei. Daher ist $k_\Delta \leq k_\Gamma$. Ferner ist $l_\Delta + 2 = l_\Gamma$, also $l_\Delta < l_\Gamma$, und es folgt (3).

Damit ist (3) für alle MM-Schlüsse bewiesen. Also ist jeder MM-Herleitung eine streng monoton fallende Folge von Tripeln zugeordnet. Weil die lexikographische Ordnung $<$ eine Wohlordnung ist, lässt sich diese Folge und damit auch die gegebene MM-Herleitung nicht beliebig ins Unendliche hinein fortsetzen: Nach endlich vielen MM-Schlüssen ist eine terminierende Menge erreicht: (MM) terminiert, und das war zu zeigen.

Beispiel. Die obige MM-Herleitung (nach 20.2.6) verwendet 5 MM-Schlüsse, besteht also aus 6 Gleichungsmengen. Ihnen sind nacheinander folgende Tripel zugeordnet (wenn man die Terme klammerfrei schreibt):

(3, 12, 0)
(3, 10, 0)
(3, 8, 0)
(2, 6, 1)
(2, 6, 0)
(1, 6, 0)

In der lexikographischen Ordnung von $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ ist dies eine streng monoton fallende Folge, auch wenn an einer Stelle der n-Wert wächst. Der Martelli-Montanari-Algorithmus löst auch das allgemeine Unifikationsproblem.

20.2.10 Satz von Martelli-Montanari
Sei M eine endliche, nicht leere Menge von Termen, gegeben in einer Aufzählung $t_0, \ldots, t_n$. Sei

$$\Gamma_0 = \{t_i = t_{i+1} \mid i < n\}$$

die zugehörige Start-Gleichungsmenge. Sei Δ eine terminierende Gleichungsmenge, die aus Γ_0 MM-herleitbar ist. Dann besteht folgende Alternative:

(i) Δ ist *vollständig gelöst*, d. h. es ist

$$\Delta = \{a_1 = s_1, \ldots, a_m = s_m\},$$

wobei $m \geq 0$ ist, die a_i paarweise verschieden sind und

$$\{a_1, \ldots, a_m\} \cap FV(\{s_1, \ldots, s_m\}) = \emptyset$$

ist. Dann ist M unifizierbar, und

$$\tau = [a_1/s_1, \ldots, a_m/s_m]$$

ist $mgu(M)$.

(ii) Sonst ist Γ_0 nicht lösbar und M nicht unifizierbar.

Beweis. Gegeben seien M und – nach Festlegung einer Aufzählung von $M - \Gamma_0$ als Start-Gleichungsmenge. Die Lösungen von Γ_0 sind genau die Substitutionen σ mit

$$t_0\sigma \equiv t_1\sigma \equiv \ldots \equiv t_n\sigma,$$

also gerade die Unifikatoren von M. Nach dem vorigen Satz terminiert (MM) für Γ_0. Es gibt also ein terminierendes Δ, das aus Γ_0 MM-hergeleitet ist. Nach 20.2.8 hat diese Menge Δ dieselben Lösungen wie Γ_0.

Sei zunächst die Alternative (i) gegeben. Wir setzen

$$\tau := [a_1/s_1, \ldots, a_m/s_m].$$

Dann ist $dom\tau \cap var\tau = \emptyset$, und nach 20.2.4 ist τ allgemeinste Lösung von Δ, also auch von Γ_0 und damit auch $mgu(M)$. Damit ist (i) bewiesen.

Im Fall (ii) ist die Alternative (i) verletzt, und dem gegebenen terminierenden Δ fehlt eine der unter (i) aufgeführten Eigenschaften. Wir gehen die verschiedenen Möglichkeiten hierfür durch.

1. Δ hat nicht die angegebene Gestalt $\{a_i = s_i \mid 1 \leq i \leq m\}$, unabhängig von den Variablenbedingungen. Dann gibt es eine Gleichung $r = s_i \in \Delta$, in der r keine Variable ist.

   1.1 s_i kann keine Variable sein, weil dann $(MM3)$ auf Δ anwendbar wäre.

   1.2 r und s_i können nicht mit demselben Funktionszeichen anfangen, weil dann $(MM1)$ auf Δ anwendbar wäre.

   1.3 Also fangen r und s_i mit verschiedenen Funktionszeichen an. Dann sind r und s_i nicht unifizierbar (nach 19.3.4, 3), und Δ ist nicht lösbar.

2. Δ hat die angegebene Gestalt, aber eine der Variablenbedingungen ist verletzt.

   2.1 Für kein i kann $a_i \equiv s_i$ sein, weil dann $(MM4)$ auf Δ anwendbar wäre.

   2.2 Ist $a_i \not\equiv s_i$, aber $a_i \in FV(s_i)$ für ein i, so sind a_i und s_i nicht unifizierbar (nach 19.3.4, 1), und Δ ist nicht lösbar.

   2.3 Für keine zwei verschiedenen Indizes i, j kann $a_i \in FV(a_j = s_j)$ und zugleich $a_i \notin FV(s_i)$ sein, weil dann $(MM2)$ auf Δ anwendbar wäre.

In den beiden einzigen möglichen Fällen 1.3 und 2.2 ist das gegebene terminierende Δ nicht lösbar. Also ist nach 20.2.8 auch Γ_0 nicht lösbar, und M ist nicht unifizierbar.

Damit ist der Satz bewiesen.

Wir brauchten in diesem Beweis nicht zu untersuchen, ob oder in welchem Maße die terminierende Menge Δ und gegebenenfalls die vollständige Lösung

τ durch die Startmenge Γ_0 eindeutig festgelegt ist. Nach 20.1.12 sind die allgemeinsten Lösungen von Γ_0 jedenfalls Varianten voneinander, und die nach dem MM-Algorithmus berechneten Lösungen können sich offenbar auch nur um Umbenennungen unterscheiden, die Variablen aus M bzw. Γ_0 permutieren. Das ist auch tatsächlich möglich.

Beispiel. M sei $\{fab, fba\}$, Γ_0 sei $\{fab = fba\}$, mit verschiedenen Variablen a, b.
Dann gibt es die beiden MM-Herleitungen

$$\dfrac{\dfrac{\dfrac{fab = fba}{a = b,\ b = a}}{a = a,\ b = a}}{b = a} \qquad \text{und} \qquad \dfrac{\dfrac{\dfrac{fab = fba}{a = b,\ b = a}}{a = b,\ b = b}}{a = b.}$$

Die terminierenden Mengen $\Delta_1 = \{b = a\}$ und $\Delta_2 = \{a = b\}$ liefern die beiden Lösungen $\tau_1 = [b/a]$ und $\tau_2 = [a/b]$, die beide a und b identifizieren. Sie sind nicht identisch, sondern unterscheiden sich um die Umbenennung $[a/b, b/a]$.

20.3 Zum Aufwand der Unifikation

Zum Schluss untersuchen wir den Aufwand, der mit dem Prozess der Unifikation verbunden ist.

Offenbar gibt es zu einer unifizierbaren Menge M beliebig aufwendige Unifikatoren, weil mit σ immer auch $\sigma\tau$ Unifikator von M ist. Allerdings enthält $\sigma[M]$ per Definition stets genau einen Term. Die Anzahl der Elemente von $\sigma[M]$ ist also gewiss kein Maß für den Aufwand der Unifikation von M.

Dagegen können $dom\sigma$ und $var\sigma$ beliebig große endliche Variablenmengen sein, selbst wenn $\sigma = mgu(M)$ ist. Denn für jede Umbenennung ρ ist mit σ auch $\sigma\rho$ ein $mgu(M)$, und dabei kann $dom\rho = var\rho$ von beliebiger endlicher Größe sein.

Ist σ ein $mgu(M)$, wie er sich aus dem MM-Algorithmus ergibt, also möglichst sparsam gewählt, dann ist offenbar

$$dom\sigma \subseteq FV(M) \text{ und } var\sigma \subseteq FV(M),$$

und man erkennt, dass auch $dom\sigma$ und $var\sigma$ kein gutes Maß für den Aufwand der Unifikation sind.

Mehr Information liefert die Gesamtlänge, der l-Wert einer Term- oder Gleichungsmenge, wie er im Beweis von Satz 20.2.9 aufgetreten ist.

20.3.1 Definition Ist t ein Term, so sei $l(t)$ die Länge von t, also die Anzahl von Auftreten von Grundzeichen in t. Ferner sei für eine endliche Termmenge M

$$l(M) = \sum\{l(t) \mid t \in M\}$$

und für eine endliche Gleichungsmenge Γ

$$l(\Gamma) = \sum\{l(s) + l(t) \mid s = t \in \Gamma\}.$$

Offenbar ist $l(\Gamma)$ der in 20.2.9 eingeführte l-Wert von Γ. Ist $M_\Gamma = \{s, t \mid s = t \in \Gamma\}$, so ist $l(M_\Gamma) \leq l(\Gamma)$, und es kann durchaus $l(M_\Gamma) < l(\Gamma)$ sein.

Beispiel. Ist $\Gamma = \{a_i = a_j \mid i, j < n\}$, so ist $l(\Gamma) = 2 \cdot n^2$, aber das zugehörige $M_\Gamma = \{a_i, a_j \mid i, j < n\}$ hat l-Wert $l(M_\Gamma) = n$.

Die triviale Aussage, dass Terme bei Substitutionen nicht kürzer werden, spiegelt sich in der Abschätzung

$$l(t) \leq l(t\sigma).$$

Dagegen ist selbst für $\sigma = mgu(M)$ sowohl $l(M) \leq l(\sigma[M])$ als auch $l(M) \geq l(\sigma[M])$ möglich.

Beispiel. Ist $M = \{a_i \mid i < n\}$, so ist $\sigma[M] = \{a_i\}$ für ein $i < n$. Dann ist $l(M) = n$, aber $l(\sigma[M]) = 1$.

Interessant und hier nicht vollständig zu behandeln ist die andere Abschätzung: Um wieviel kann $l(\sigma[M])$ stärker wachsen als $l(M)$? Wir studieren zunächst eine Schar von Gleichungsmengen.

20.3.2 Lemma Zu $n > 0$ sei

$$\Gamma = \{a_i = f a_{i+1} a_{i+1} \mid i < n\}.$$

Es ist $l(\Gamma) = 4n$, aber für die allgemeinste Lösung σ von Γ ist

$$l(a_{n-i}\sigma) = 2^{i+1} - 1 \text{ für } i \leq n, \text{ speziell } l(a_0\sigma) = 2^{n+1} - 1 :$$

das Wachstum von $l(\Gamma\sigma)$ gegenüber $l(\Gamma)$ ist exponentiell.

Beweis. Γ besteht aus n Gleichungen, jede Gleichung hat den l-Wert 4, also ist $l(\Gamma) = 4n$. Man betrachte die „Translation“

$$\tau = [a_0/fa_1a_1, \ldots, a_i/fa_{i+1}a_{i+1}, \ldots, a_{n-1}/fa_na_n].$$

Dann ist

$$\begin{aligned}
&a_{n-1}\tau \equiv fa_na_n\\
&a_{n-2}\tau^2 \equiv f(a_{n-1}\tau)(a_{n-1}\tau) \equiv f(fa_na_n)(fa_na_n)\\
&a_{n-3}\tau^3 \equiv fa_{n-2}\tau^2a_{n-2}\tau^2 \equiv f(f(fa_na_n)(fa_na_n))(f(fa_na_n)(fa_na_n)).
\end{aligned}$$

Allgemein ist (für $i \leq n$) $a_{n-i}\tau^i$ ein Term, in dem nur noch a_n frei auftritt, und zwar 2^i mal, so dass

$$l(a_{n-i}\tau^i) = 2^{i+1} - 1$$

ist, wie man mit Induktion nach i nachrechnet. Wegen $a_n \notin dom\tau$ ist $a_{n-i}\tau^i \equiv a_{n-i}\tau^n$, so dass $\sigma := \tau^n$ eine Lösung von Γ ist. Insbesondere $a_0\sigma \equiv a_0\tau^n$ kommt auch erst nach n $(MM2)$-Schlüssen zustande, so dass σ eine allgemeinste Lösung von Γ ist.

20.3.3 Satz Ist zusätzlich g ein n-stelliges Funktionszeichen und

$$M = \{ga_0 \ldots a_{n-1}, g(fa_1a_1) \ldots (fa_na_n)\},$$

so ist $l(M) = 4n + 2$, aber für $\sigma = mgu(M)$ ist

$$l(\sigma[M]) = 2^{n+2} - (n + 3) :$$

$l(\sigma[M])$ wächst gegenüber $l(M)$ exponentiell.

Beweis. M besteht aus zwei Termen t, t'. Ein $(MM1)$-Schluss, angewandt auf $\Gamma_0 = \{t = t'\}$, ergibt die Menge Γ aus 20.3.2. Also ist die allgemeinste Lösung σ von Γ aus 20.3.2 auch $mgu(M)$. Es ist

$$\sigma[M] = \{g(a_0\sigma) \ldots (a_{n-1}\sigma)\}$$

und mit 20.3.2 folgt

$$l(\sigma[M]) = 1 + \sum_{0<i\leq n} (2^{i+1} - 1) = \sum_{i\leq n+1} (2^i - 1) = 2^{n+2} - (n + 3).$$

Es wird nicht behauptet, dass $l(\sigma[M])$ etwa wie $2^{l(M)}$ wächst. Vielmehr wächst $l(\sigma[M])$ wie 2^n, wobei die Anzahl der Variablen $n = \frac{1}{4}(l(M) - 2)$ ist. Also wächst $l(\sigma[M])$ wie $(\sqrt[4]{2})^{l(M)} \approx 1,19^{l(M)}$.

Wenn man das Funktionszeichen f in 20.3.2 k-stellig statt 2-stellig wählt ($k \geq 2$) und

$$\Gamma = \{a_i = fa_{i+1} \dots a_{i+1} \mid i < n\}$$

setzt, ändert sich am Wachstum nicht viel und insbesondere kaum nach oben. Für k-stelliges f wächst $l(\sigma[M])$ wie $\sqrt[k+2]{k}^{l(M)}$. Die Basis $\sqrt[k+2]{k}$ konvergiert bekanntlich gegen 1, so dass für große k das Wachstum von $l(\sigma[M])$ gegen $l(M)$ immer schwächer wird. Das Maximum für ganzzahlige k erreicht $\sqrt[k+2]{k}$ bei $k = 4$, und es ist $\sqrt[6]{4} \approx 1,26$.

Exponentielles Wachstum zur Basis 1,26 ist also das stärkste Wachstum, das man mit diesen Beispielen erreichen kann.

Die Vermutung drängt sich auf, dass man mit diesen Beispielen schon in die Nähe der ungünstigsten Fälle gekommen ist und dass jedenfalls $l(\sigma[M])$ höchstens exponentiell gegenüber $l(M)$ wachsen kann. Ein Beweis dafür ist uns aber nicht bekannt. Die denkbar gröbste Abschätzung liefert:

20.3.4 Lemma Ist σ allgemeinste Lösung von Γ und $l(\Gamma) = l$, so ist

$$l(\Gamma\sigma) \leq l^{2^l}.$$

Beweis. Gegeben sei eine MM-Herleitung

$$\Gamma = \Gamma_0, \Gamma_1, \dots, \Gamma_n = \Gamma\sigma$$

zur Berechnung einer allgemeinsten Lösung σ von Γ. In dieser Herleitung wächst der l-Wert allenfalls bei $MM2$-Schlüssen. Sei nun

$$\Delta, a = t \vdash \Delta[a/t], a = t$$

ein solcher $MM2$-Schluss und $l(\Delta, a = t) = m$. Dann ist $l(\Delta) < m, l(t) < m$, und a tritt $p < m$-mal in Δ auf. Also ist

$$l(\Delta[a/t], a = t) \leq l(\Delta) + p \cdot (m - 2) + l(a = t) \leq m^2.$$

In der obigen MM-Herleitung treten höchstens l verschiedene Variablen auf, weil sie alle schon in Γ auftreten müssen und $l(\Gamma) = l$ ist. Dann enthält die

MM-Herleitung auch höchstens l $(MM2)$-Schlüsse, weil jede dieser Variablen a nur einmal die substituierte Variable sein kann (wegen $a \notin FV(t), a \in FV(\Delta)$). Ist nun $\Gamma_j \vdash \Gamma_{j+1}$ der i-te $(MM2)$-Schluss in der MM-Herleitung und ist $l(\Gamma_j) \leq l^{2^{i-1}}$, so ist nach unserer Rechnung

$$l(\Gamma_{j+1}) \leq (l^{2^{i-1}})^2 = l^{2^{i-1} \cdot 2} = l^{2^i}.$$

Da $l(\Gamma) = l = l^{2^0}$ ist, ist nach maximal l $(MM2)$-Schlüssen $l(\Gamma\sigma) \leq l^{2^l}$.

Diese Abschätzung überträgt sich offenbar auf die l-Werte von Termmengen M und $\sigma[M]$, wenn σ ein $mgu(M)$ ist.

Die Abschätzung ist so schlecht, dass sie sich leicht verbessern lässt. Die Frage ist, ob sie sich entscheidend, nämlich auf exponentielles Wachstum verbessern lässt. Dieser Frage können wir nicht weiter nachgehen.

20.4 Aufgaben

20.4.1 Geben Sie einen mgu von $+a + bc$ und $+ + a'bc'$ an.

20.4.2 Sei $\leq$ eine reflexive und transitive Relation auf einer Menge $X \neq \emptyset$. Ferner sei für $x, y \in X$:

$$xRy \Leftrightarrow x \leq y \text{ und } y \leq x.$$

Zeigen Sie: R ist eine Äquivalenzrelation.

20.4.3 Geben Sie Substitutionen σ, τ an, für die $\sigma\tau = \sigma$, aber τ keine Umbenennung ist.

20.4.4 Gilt die Umkehrung von 20.1.7:
Aus $\sigma\pi = \sigma\rho$ folgt $\pi \restriction var^+\sigma = \rho \restriction var^+\sigma$?
Begründen Sie Ihre Antwort.

20.4.5 Es sei $X_0 = \{n \in \mathbb{N} \mid n < 10\}$ und

$$\varphi_0 : X_0 \to \mathbb{N},\ n \mapsto 2n.$$

Geben Sie eine Permutation φ von $X_0 \cup im(\varphi_0)$ an, die φ_0 fortsetzt.

20.4.6 Zeigen Sie an einem Beispiel, dass Lemma 20.1.8 falsch ist für unendliche Mengen X_0.

20.4.7 Die Sprache L enthalte mindestens ein zweistelliges Funktionszeichen. Zeigen Sie:
Zu jeder endlichen Gleichungsmenge Γ gibt es Terme s, t, für die stets gilt:

$$\sigma \text{ unifiziert } s \text{ und } t \Leftrightarrow \sigma \text{ löst } \Gamma.$$

20.4.8 Beweisen Sie 20.2.4 (ohne Rückgriff auf 20.2.1).

20.4.9 Entscheiden Sie mit dem Martelli-Montanari-Algorithmus, ob folgende Gleichungen lösbar sind, und geben Sie gegebenenfalls eine allgemeinste Lösung an ($\cdot$ bindet stärker als $+$):

a. $a + b \cdot b = c \cdot c + a$

b. $a + b \cdot b = c \cdot a + c$

c. $a + b \cdot c = c \cdot b + a$

§21 Resolution

21.1 Einfache Resolution und Faktorisierung

21.2 Volle Resolution

21.3 Aufgaben

Wir bauen nun die Technik der Unifikation in das widerlegungsvollständige Regelsystem $\{(Mix), (Subst)\}$ aus 19.4 ein, um weitere widerlegungsvollständige Regelsysteme zu erhalten, sogenannte Resolutionskalküle. Sie sind dadurch besonders effizient, dass sie statt beliebiger Substitutionen nur noch allgemeinste Unifikatoren verwenden und so besonders zielgerichtet auf die angestrebte Widerlegung einer gegebenen Theorie hinarbeiten.

21.1 Einfache Resolution und Faktorisierung

Im folgenden sei T stets eine Gentzen-Theorie und $\mathcal{R}$ ein Regelsystem. Es verbessert die Übersicht über $\mathcal{R}$-Herleitungen, wenn man sicherstellen kann, dass gewisse Variablen in vorliegenden primen Sequenzen nicht auftreten. Das lässt sich für viele $\mathcal{R}$ erreichen.

21.1.1 Lemma *zur Variablentrennung.* Zu jeder endlichen Menge V von Variablen und jeder (primen) Sequenz $\Gamma : \Delta$ gibt es eine Umbenennung π, so dass

(1) $FV(\Gamma\pi : \Delta\pi) \cap V = \emptyset$.

Beweis. Weil V endlich ist, ist $FV - V$ unendlich, und es gibt eine Injektion π_0 von $FV(\Gamma : \Delta)$ in $FV - V$. π_0 lässt sich nach 20.1.8 fortsetzen zu einer Umbenennung π mit (1).

Dieses einfache Lemma ist auf viele Regelsysteme anwendbar. Es gilt

(2) Aus $T \vdash_{\mathcal{R}} \Gamma : \Delta$ folgt $T \vdash_{\mathcal{R}} \Gamma\pi : \Delta\pi$ für Umbenennungen π

einerseits, wenn $(Subst) \in \mathcal{R}$ ist, weil Umbennungen Substitutionen sind. Andererseits gilt (2), wenn alle Regeln $R \in \mathcal{R}$ gegen Umbenennungen abgeschlossen sind (was etwa für $(Mix), (sMix)$ und auch $(Subst)$ der Fall ist), weil mit jeder Startsequenz $\Gamma : \Delta$ von T auch $\Gamma\pi : \Delta\pi$ eine Startsequenz von T ist.

Auch die folgende Resolutionsregel, eine spezielle Kombination von Umbenennungen, mgu und (Mix), ist gegen Umbenennung abgeschlossen, wie man sich leicht überlegt:

21.1.2 Definition der *einfachen Resolutionsregel*

$$(Res) \qquad \Gamma_1, P : \Delta_1 \text{ und } \Gamma_2 : Q, \Delta_2 \vdash (\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu,$$

falls

1. π eine Umbenennung mit

$$FV(\Gamma_1, P : \Delta_1) \cap FV(\Gamma_2\pi : Q\pi, \Delta_2\pi) = \emptyset,$$

2. $\mu = mgu\{P, Q\pi\}$ ist.

Einen einzelnen (Res)-Schluss nennt man auch eine *(einfache) Resolution*.

Durch die Variablentrennung, die π bewirkt, ist ein $mgu\{P, Q\pi\}$ in mehr Fällen zu bestimmen als $mgu\{P, Q\}$. Dadurch kann man (Res) flexibler einsetzen als (Mix). Da Umbenennungen und Unifikatoren Substitutionen sind, ergibt die Definition unmittelbar:

21.1.3 Lemma Jeder (Res)-Schluss lässt sich aus $(Subst)$- und (Mix)-Schlüssen zusammensetzen: Die einfache Resolutionsregel ist korrekt.

Beweis. Die Korrektheit von (Res) ergibt sich wegen der ersten Teilaussage aus der Korrektheit von $(Subst)$ und (Mix), die im Prinzip seit §2, für die Schnittregel und damit auch für (Mix) jedenfalls seit 8.4.1 bekannt ist.

Will man andererseits einen (Mix)-Schluss durch eine Resolution darstellen, muss man die Wirkung der Umbenennung π auf $\Gamma_2 : \Delta_2$ zum Schluss wieder kompensieren, und zwar durch eine geeignete Substitution: Nicht jeder (Mix)-Schluss ist ein (Res)-Schluss.

Beispiel. p sei 0-stellig, q sei 1-stellig.

$$qa, p : \emptyset \text{ und } qa : p \vdash qa : \emptyset$$

ist ein (Mix)-Schluss. Um (Res) auf die beiden Prämissen anzuwenden, brauchen wir als erstes eine Umbenennung π mit $\pi(a) \not\equiv a$. Es ist $p\pi \equiv p$, und id

ist $mgu\{p, p\pi\}$, und nach 20.1.12 modulo Umbenennungen auch der einzige. (Res) ergibt also $qa, qa\pi : \emptyset$ mit zwei verschiedenen Variablen a und $a\pi$, und erst eine Substitution σ, die $a\pi$ mit a, also $qa\pi$ mit qa identifiziert, liefert die Konklusion $qa : \emptyset$.

Ein (Mix)-Schluss

$$P : \emptyset \text{ und } \emptyset : P \vdash \square,$$

der zum Widerspruch $\square$ führt, lässt sich allerdings auch als Resolution auffassen. Trotzdem ist (Res) für sich allein nicht widerlegungsvollständig.

Beispiel. p sei 1-stellig, $a \not\equiv b$. Von den drei Sequenzen

$$pa, pb : \emptyset \text{ und } \emptyset : pa, pb \text{ und } pa : pb$$

sind die beiden ersten zusammen inkonsistent. Mit $(Subst)$ und (Mix) leitet man aus ihnen $\square$ leicht her. Um (Res) auf zwei dieser drei Sequenzen anzuwenden, trennt man zuerst die Variablen mit einer Umbenennung π, so dass $\pi(a), \pi(b) \not\equiv a, b$ ist. Dann sucht man einen mgu für eine Antezedenz- und eine Sukzedenzformel und erhält modulo Umbenennung

$$pa : pb \text{ bzw. } \emptyset : pa, pb \text{ bzw. } pa, pb : \emptyset,$$

je nachdem man Resolution auf die beiden ersten, die beiden letzten bzw. die beiden äußeren Sequenzen anwendet: Die einfache Resolution führt (modulo Umbenennung) nicht aus dieser Menge von drei Sequenzen heraus; insbesondere führt sie nicht zum Widerspruch $\square$.

Um doch dahin zu kommen, wird man vor der Resolution die beiden Primformeln der ersten oder zweiten Sequenz unifizieren, man wird *faktorisieren*:

21.1.4 Definition der *Faktorisierungsregeln*

$$(Fak) \quad \Gamma, P, Q : \Delta \vdash (\Gamma, P : \Delta)\mu \quad \text{und} \quad \Gamma : P, Q, \Delta \vdash (\Gamma : P, \Delta)\mu,$$
$$\text{falls } \mu = mgu\{P, Q\} \text{ ist.}$$

Da mgu's Substitutionen sind, ist klar:

21.1.5 Lemma Jeder (Fak)-Schluss ist ein $(Subst)$-Schluss: (Fak) ist korrekt.

Aus den Lemmata 21.1.3 und 5 folgt sofort:

21.1.6 Lemma Jede $\{(Res),(Fak)\}$-Herleitung in einer Theorie T lässt sich auch als $\{(Mix),(Subst)\}$-Herleitung darstellen: Das Regelsystem $\{(Res),(Fak)\}$ ist korrekt.

Wichtiger ist natürlich die Frage, ob dieses Regelsystem auch widerlegungsvollständig ist. Einen positiven Hinweis geben unsere beiden Beispiele. Die Substitution, die man im ersten Beispiel zur Kompensation der Umbenennung anhängt, entfällt, wenn $\square$ die Konklusion der Resolution ist. Die Substitution, die man im zweiten Beispiel zur Unifikation vorschalten muss, ist eine Faktorisierung. Man kann das zweite Beispiel so aufblähen, dass mehrere Faktorisierungen vor einer Resolution nötig werden.

21.1.7 Lemma Äquivalent sind:

(1) μ unifiziert $\Gamma \cup \Delta$:

(2) Es ist $\mu = \sigma\tau$ für ein $\sigma = mgu(\Gamma)$ und einen Unifikator τ von $\Gamma\sigma \cup \Delta\sigma$.

Beweis. Gilt (1), so unifiziert μ insbesondere Γ, und nach 20.2.10 gibt es einen mgu σ von Γ. Dann ist $\sigma \leq \mu$, also $\mu = \sigma\tau$ für ein τ. Jedes solche τ unifiziert $(\Gamma \cup \Delta)\sigma$ gemäß 20.2.3 genau dann, wenn $\mu = \sigma\tau$ $\quad \Gamma \cup \Delta$ unifiziert.

21.1.8 Lemma Äquivalent sind:

(1') $\mu = mgu(\Gamma \cup \Delta)$

(2') $\mu = \sigma\tau$ für ein $\sigma = mgu(\Gamma)$ und ein $\tau = mgu(\Gamma\sigma \cup \Delta\sigma)$.

Beweis. Gelte (2'). Nach 21.1.7 unifiziert μ $\Gamma \cup \Delta$. Sei μ' ein weiterer Unifikator von $\Gamma \cup \Delta$. Wieder nach 21.1.7 gibt es einen Unifikator τ' von $\Gamma\sigma \cup \Delta\sigma$ mit $\mu' = \sigma\tau'$. Da $\tau = mgu(\Gamma\sigma \cup \Delta\sigma)$ ist, folgt $\tau \leq \tau'$, also

$$\mu = \sigma\tau \leq \sigma\tau' = \mu',$$

und μ ist $mgu(\Gamma \cup \Delta)$.

Gelte (1'). Nach 21.1.7 und 20.2.10 gibt es dann ein $\sigma = mgu(\Gamma)$ und ein $\tau' = mgu(\Gamma\sigma \cup \Delta\sigma)$. Nach dem Bewiesenen ist $\sigma\tau' = mgu(\Gamma \cup \Delta)$. Also gibt es nach 20.1.12 eine Umbenennung π, für die

$$\mu = \sigma\tau'\pi$$

ist. Dann folgt (2') für $\tau := \tau'\pi$.

21.1.9 Satz Durch Iteration der Faktorisierung erhält man:

$$\Gamma, P_0, \ldots, P_n : \Delta \underset{Fak}{\vdash} \Gamma\mu, P_0\mu : \Delta\mu \quad \text{und}$$
$$\Gamma : P_0, \ldots, P_n, \Delta \underset{Fak}{\vdash} \Gamma\mu : P_0\mu, \Delta\mu,$$

falls $\mu = mgu\{P_0, \ldots, P_n\}$ ist.

Beweis durch Induktion nach n. Für $n = 0$ ist $\mu = id$, und Prämisse und Konklusion stimmen überein.

Sei $n > 0$ und $\mu = mgu\{P_0, \ldots, P_n\}$. Nach 21.1.8 ist $\mu = \sigma\tau$ für ein $\sigma = mgu\{P_0, \ldots, P_{n-1}\}$ und $\tau = mgu\{P_0\sigma, P_n\sigma\}$. Nach Induktionsvoraussetzung folgt

$$\Gamma, P_0, \ldots, P_n : \Delta \underset{Fak}{\vdash} \Gamma\sigma, P_0\sigma, P_n\sigma : \Delta\sigma \underset{Fak}{\vdash} \Gamma\sigma\tau, P_0\sigma\tau : \Delta\sigma\tau,$$

und das ist bereits die erste Behauptung. Die zweite folgt analog.

21.2 Volle Resolution

Die iterierte Faktorisierung nach 21.1.9 und die einfache Resolution kombinieren wir zu einer neuen Regel:

21.2.1 Definition der *vollen Resolutionsregel.*

$$(RES) \quad \Gamma_1, P_0, \ldots, P_n : \Delta_1 \text{ und } \Gamma_2 : Q_0, \ldots, Q_m, \Delta_2 \vdash (\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu,$$

falls

1. π eine Umbenennung ist mit

$$FV(\Gamma_1, P_0, \ldots, P_n : \Delta_1) \cap FV((\Gamma_2 : Q_0, \ldots, Q_m, \Delta_2)\pi) = \emptyset,$$

2. $\mu = mgu\{P_0, \ldots, P_n, Q_0\pi, \ldots, Q_m\pi\}$ ist.

Jeder Schluss nach dieser Regel heißt auch eine (volle) Resolution.

Die einfache Resolutionsregel (Res) ist offenbar der Spezialfall $n = m = 0$ von (RES).

Das Beispiel vor 21.1.4 ist mit der vollen Resolutionsregel ohne weiteres zu behandeln. Denn

$$pa, pb : \emptyset \text{ und } \emptyset : pa, pb \vdash \square$$

ist ein einziger (RES)-Schluss, nämlich die volle Resolution mit

$\Gamma_i = \Delta_i = \emptyset \quad (i = 1, 2)$
$P_0 \equiv Q_0 \equiv pa$ und $P_1 \equiv Q_1 \equiv pb \quad (n = m = 1)$
$\pi = [a/c, b/d, c/a, d/b]$ mit verschiedenen Variablen a, b, c, d und
$\mu = [b/a, c/a, d/a]$.

Dann ist

$$\{P_0, P_1, Q_0\pi, Q_1\pi\}\mu = \{pa, pb, pc, pd\}\mu = \{pa\},$$

und μ ist Unifikator und offenbar auch mgu dieser vier Primformeln.

An diesem einfachen Beispiel zeigt sich schon, dass die genaue Angabe eines einzelnen (RES)-Schlusses recht aufwendig werden kann. Dafür ist die volle Resolutionsregel aber auch von großer Leistungsfähigkeit. Zunächst zu ihrer Korrektheit.

21.2.2 Satz $(RES) \subseteq \{(Res), (Fak)\}$: Jede volle Resolution lässt sich darstellen als eine Folge von Faktorisierungen, gefolgt von einer einfachen Resolution.

Beweis. Gegeben sei ein (RES)-Schluss wie in 21.2.1. Gegeben ist uns also

1. eine Umbenennung π, die die Variablen der beiden Prämissen vollständig voneinander trennt, und

2. ein mgu μ von $\Gamma \cup \Delta$, wobei

$$\Gamma := \{P_0, \ldots, P_n\} \text{ und } \Delta := \{Q_0\pi, \ldots, Q_m\pi\} \text{ ist.}$$

Nach 21.1.8 ist $\mu = \sigma\tau = \sigma'\tau'$ für einen mgu σ von Γ und einen mgu σ' von Δ. Dann ist $\pi\sigma' = mgu\{Q_0, \ldots, Q_m\}$. Wegen der Eindeutigkeit 20.1.12 der mgu's können wir σ und σ' auch nach dem Martelli-Montanari-Algorithmus 20.2.6 konstruieren, so dass

$$dom\sigma \cup var\sigma \subseteq FV(\Gamma) \text{ und } dom\sigma' \cup var\sigma' \subseteq FV(\Delta)$$

und wegen der Variablentrennung durch π allgemein $\sigma\sigma' = \sigma'\sigma$ ist, σ' die erste Prämisse und σ das π-Bild der zweiten Prämisse nicht verändert.

Weil τ ein *mgu* von $(\Gamma \cup \Delta)\sigma = \{P_0\sigma\} \cup \Delta\sigma$ ist, ist wieder nach 21.1.8 $\tau = \sigma'\mu_0$, also $\mu = \sigma\sigma'\mu_0$ für einen *mgu* μ_0 von $(\Gamma \cup \Delta)\sigma\sigma' = \{P_0\sigma, Q_0\pi\sigma'\}$:

$$(\Gamma \cup \Delta)\mu = \{P_0\sigma\mu_0\} = \{Q_0\pi\sigma'\mu_0\}.$$

Mit iterierter Faktorisierung nach 21.1.9 folgt

$$\Gamma_1, P_0, \ldots, P_n : \Delta_1 \vdash_{\mathit{Fak}} \Gamma_1\sigma, P_0\sigma : \Delta_1\sigma \quad \text{und}$$
$$\Gamma_2 : Q_0, \ldots, Q_m, \Delta_2 \vdash_{\mathit{Fak}} \Gamma_2\pi\sigma' : Q_0\pi\sigma', \Delta_2\pi\sigma'.$$

Die Variablen dieser beiden Konklusionen sind bereits getrennt, so dass mit einer einfachen Resolution für den *mgu* μ_0 folgt:

$$\Gamma_1\sigma\mu_0, \Gamma_2\pi\sigma'\mu_0 : \Delta_1\sigma\mu_0, \Delta_2\pi\sigma'\mu_0,$$

und das ist die angestrebte Sequenz $(\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu$.

21.2.3 Korollar $\{(RES)\}$ ist korrekt.

Denn nach 21.1.6 ist $\{(Res), (Fak)\}$ korrekt, also hiernach erst recht $\{(RES)\}$.

Die Korrektheit von (RES) ist ohnehin klar, weil sich jede volle (wie jede einfache) Resolution per Definition aus speziellen Substitutionen und einem (Mix)-Schluss zusammensetzt. Nicht ganz trivial ist dagegen Satz 21.2.2, dass sich jede volle Resolution auch aus Faktorisierungen und einer einfachen Resolution zusammensetzt.

Betrachten wir noch einmal das Regelsystem $\mathcal{R} = \{(Subst), (Mix)\}$. In jedem Faden einer $\mathcal{R}$-Herleitung kann man aufeinanderfolgende Substitutionen offenbar zu einer einzigen Substitution zusammenfassen, und je zwei aufeinanderfolgende (Mix)-Schlüsse kann man durch eine triviale (identische) Substitution trennen. Man kann also $\mathcal{R}$-Herleitungen so normieren, dass sich in ihren Fäden $(Subst)$- und (Mix)-Schlüsse jeweils abwechseln. Solche Herleitungsteile – (Mix)-Schlüsse, deren beide Prämissen Konklusionen von $(Subst)$-Schlüssen sind – lassen sich nun durch volle Resolutionen, *gefolgt* von einer Substitution, darstellen:

21.2.4 Transformationssatz Gegeben sei ein (Mix)-Schluss, dessen beide Prämissen durch Substitutionen entstanden sind, d. h. ein Ausschnitt aus einer $\{(Subst), (Mix)\}$-Herleitung der Form

$$\dfrac{\dfrac{\Gamma_1, \Gamma : \Delta_1}{\Gamma_1\sigma, P : \Delta_1\sigma}\,(Subst) \qquad \dfrac{\Gamma_2 : \Delta, \Delta_2}{\Gamma_2\tau : P, \Delta_2\tau}\,(Subst)}{\Gamma_1\sigma, \Gamma_2\tau : \Delta_1\sigma, \Delta_2\tau}\,(Mix)$$

mit $\Gamma\sigma = \Delta\tau = \{P\}$. Dann gibt es eine volle Resolution mit variablentrennender Umbenennung π und $\mu = mgu(\Gamma \cup \Delta\pi)$ und eine Substitution, so dass

$$\dfrac{\dfrac{\dfrac{\Gamma_1, \Gamma : \Delta_1 \qquad \Gamma_2 : \Delta, \Delta_2}{(\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu}\,(RES)}{\Gamma_1\sigma, \Gamma_2\tau : \Delta_1\sigma, \Delta_2\tau}\,(Subst)$$

Jede $(Subst)$-(Mix)-Kombination der angegebenen Art lässt sich durch eine (RES)-$(Subst)$-Kombination darstellen.

Beweis. Zunächst wählen wir eine Umbenennung π mit

$$FV(\Gamma_1, \Gamma : \Delta_1) \cap FV(\Gamma_2\pi : \Delta\pi, \Delta_2\pi) = \emptyset.$$

Dann können wir eine Substitution μ' definieren durch

$$\mu'(a) = \begin{cases} \sigma(a) & \text{für } a \in FV(\Gamma_1, \Gamma : \Delta_1) \\ \tau \circ \pi^{-1}(a) & \text{für } a \in FV(\Gamma_2\pi : \Delta\pi, \Delta_2\pi) \\ a & \text{sonst.} \end{cases}$$

Es ist $\Gamma\mu' = \Gamma\sigma = \{P\}$ und

$$(\Delta\pi)\mu' = \Delta\pi\pi^{-1}\tau = \Delta\tau = \{P\}.$$

Also ist μ' Unifikator von $\Gamma \cup \Delta\pi$. Nach 20.2.10 gibt es einen mgu μ von $\Gamma \cup \Delta\pi$. Mithin ist $\mu \leq \mu'$, es ist $\mu\rho = \mu'$ für eine Substitution ρ. Dann gibt es folgende (RES)-$(Subst)$-Kombination

$$\dfrac{\dfrac{\Gamma_1, \Gamma : \Delta_1 \qquad \Gamma_2 : \Delta, \Delta_2}{(\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu}}{(\Gamma_1, \Gamma_2\pi : \Delta_1, \Delta_2\pi)\mu\rho\,.}$$

Wegen $\mu\rho = \mu'$ ist aber

$$\Gamma_1\mu\rho = \Gamma_1\mu' = \Gamma_1\sigma, \text{ analog } \Delta_1\mu\rho = \Delta_1\sigma,$$

ferner

$$\Gamma_2\pi\mu\rho = \Gamma_2\pi\mu' = \Gamma_2\pi\pi^{-1}\tau = \Gamma_2\tau, \text{ analog } \Delta_2\pi\mu\rho = \Delta_2\tau.$$

Damit stimmt die Endsequenz mit $\Gamma_1\sigma, \Gamma_2\tau : \Delta_1\sigma, \Delta_2\tau$ überein, und der Satz ist bewiesen.

Kürzt man die beiden Ausgangssequenzen mit S_1, S_2 ab, das Ergebnis der vollen Resolution mit R und die Endsequenz mit E, so kann man die Aussage des Satzes so symbolisieren:

$$\begin{array}{ccccc} S_1 & S_2 & \xrightarrow{\sigma,\tau} & S_1\sigma & S_2\tau \\ & | & & & \\ RES & \pi, \mu & & & \downarrow Mix \\ & \downarrow & & & \\ & R & \xrightarrow{\rho} & R\rho \equiv E. & \end{array}$$

Durch iterierte Anwendung dieses Satzes ergibt sich:

21.2.5 Satz Aus $T \mid\!\!\!\overline{}_{Mix,Subst}\, \Gamma : \Delta$ folgt

$$T \mid\!\!\!\overline{}_{RES}\, \Gamma_0 : \Delta_0$$

für eine Sequenz $\Gamma_0 : \Delta_0$, aus der $\Gamma : \Delta$ durch eine Substitution ρ hervorgeht.

Beweis durch $\{(Mix), (Subst)\}$-Herleitungsinduktion.

1. Ist $\Gamma : \Delta$ eine Startsequenz, so ist $T \mid\!\!\!\overline{}_{RES}\, \Gamma : \Delta$, und es ist $\rho = id$.

2. Der letzte Schluss in der Herleitung von $\Gamma : \Delta$ ist eine Substitution

   $$\Gamma_1 : \Delta_1 \vdash \Gamma_1\sigma : \Delta_1\sigma \equiv \Gamma : \Delta.$$

   Nach Induktionsvoraussetzung ist $T \mid\!\!\!\overline{}_{RES}\, \Gamma_0 : \Delta_0$, so dass $\Gamma_0\rho_1 : \Delta_0\rho_1 \equiv \Gamma_1 : \Delta_1$ für ein ρ_1 ist. Wir setzen $\rho = \rho_1\sigma$ und erhalten

   $$\Gamma_0\rho : \Delta_0\rho \equiv \Gamma_0\rho_1\sigma : \Delta_0\rho_1\sigma \equiv \Gamma_1\sigma : \Delta_1\sigma \equiv \Gamma : \Delta.$$

3. Der letzte Schluss in der Herleitung von $\Gamma : \Delta$ ist ein (Mix)-Schluss

$$\Gamma_1, P : \Delta_1 \text{ und } \Gamma_2 : P, \Delta_2 \vdash \Gamma_1, \Gamma_2 : \Delta_1, \Delta_2 \equiv \Gamma : \Delta.$$

Nach Induktionsvoraussetzung ist

$$T \mid\!\!\overline{}_{RES}\, \Gamma_{10}, \Gamma_0 : \Delta_{10} \text{ und } T \mid\!\!\overline{}_{RES}\, \Gamma_{20} : \Delta_0, \Delta_{20},$$

und es gibt Substitutionen σ und τ mit

$$\Gamma_{10}\sigma = \Gamma_1, \Gamma_0\sigma = \{P\}, \Delta_{10}\sigma = \Delta_1, \Gamma_{20}\tau = \Gamma_2, \Delta_0\tau = \{P\}, \Delta_{20}\tau = \Delta_2.$$

Damit sind die Voraussetzungen des Transformationssatzes erfüllt, so dass mit einem weiteren (RES)-Schluss für geeignete π und μ folgt:

$$T \mid\!\!\overline{}_{RES}\, (\Gamma_{10}, \Gamma_{20}\pi : \Delta_{10}, \Delta_{20}\pi)\mu,$$

und $\Gamma : \Delta$ geht aus dieser Sequenz durch eine Substitution ρ hervor.

Mit $\{(Mix), (Subst)\}$-Herleitungsinduktion folgt nun der Satz.

Als Korollare erhält man nun:

21.2.6 Satz Widerlegungsvollständigkeit der Resolutionskalküle

1. $\{(RES)\}$ ist widerlegungsvollständig.
2. $\{(Res), (Fak)\}$ ist widerlegungsvollständig.

Beweis von 1.
Sei T eine Gentzen-Theorie ohne Modell. Nach 19.4.14 ist dann $T \mid\!\!\overline{}_{Mix,Subst}\, \square$. Nach dem letzten Satz 21.2.5 folgt $T \mid\!\!\overline{}_{RES}\, \Gamma_0 : \Delta_0$ für eine Sequenz $\Gamma_0 : \Delta_0$, aus der die leere Sequenz $\square$ durch eine Substitution hervorgeht. Dann ist $\Gamma_0 : \Delta_0 \equiv \square$ und $T \mid\!\!\overline{}_{RES}\, \square$. Es folgt 1.

Beweis von 2.
Ist $T \mid\!\!\overline{}_{RES}\, \square$, so ist nach 21.2.2 auch $T \mid\!\!\overline{}_{Res,Fak}\, \square$. Also folgt 2. aus 1.

Zur Illustration betrachten wir eine konkrete (RES)-Widerlegung.

Beispiel. $L(T)$ sei gegeben durch drei Prädikatszeichen p (2-stellig), q und r (1-stellig) und zwei einstellige Funktionszeichen f und g. Startsequenzen von T seien:

(1) $pafb, qc, rc : \emptyset$

(2) $pafb, qfc : rfc, rfgfb$

(3) $pab : qfgb, qfc$

(4) $\emptyset : pafb, pgcc, pgfbd.$

Ist $T \underset{RES}{\vdash} \Box$?

Wir wenden zuerst (RES) auf (1) und (2) an. Da $pafb$ in beiden Antezedenten auftritt, brauchen wir nur c in (2) durch $\pi = [c/c', c'/c]$ umzubenennen. Offenbar ist

$$mgu\{rc, rfc', rfgfb\} = [c/fgfb, c'/gfb].$$

Dann ist

$$(1) \text{ und } (2) \vdash pafb, qfgfb : \emptyset$$

eine volle Resolution.

Nun wenden wir auf diese Sequenz und (3) (RES) an. Da a in beiden Sequenzen an derselben Stelle auftritt, können wir zur Variablentrennung $\pi = [b/b', b'/b]$ wählen. Dann ergibt sich

$$mgu\{qfgfb, qfgb', qfc\} = [b'/fb, c/gfb].$$

Damit ist auch

$$pafb, qfgfb : \emptyset \text{ und } (3) \vdash pafb : \emptyset$$

wieder eine volle Resolution. Nun können wir abkürzend auf die Variablentrennung verzichten, also $\pi = id$ setzen. Mit

$$mgu\{pafb, pafb, pgcc, pgfbd\} = [a/gfb, c/fb, d/fb]$$

erhalten wir die volle Resolution

$$pafb : \emptyset \text{ und } (4) \vdash \Box.$$

Damit ist T durch volle Resolution widerlegt.

Per Hand mag es lästig sein, alle erforderlichen mgu's in einer $\{(RES)\}$-Herleitung auszurechnen. Es handelt sich dabei aber um eine endliche Prüfaufgabe, die ein Computer effizient erledigen kann. Tatsächlich wurden die Resolutionskalküle für das automatische, programmierte Beweisen bzw. Widerlegen entworfen. Dagegen erfordert das Herleiten in einem vollständigen Logikkalkül (wie in 3.1 oder 6.3) einen Einblick in das jeweilige zur Lösung anstehende Problem, eine mathematische oder kombinatorische Intuition. Und darin ist der logisch geschulte Mathematiker und Informatiker den Computern bis heute noch weit überlegen.

21.3 Aufgaben

21.3.1 Geben Sie eine einfache Resolution an, bei der die trennende Umbenennung π nicht die Identität sein kann.

21.3.2 Geben Sie Formelmengen Γ und Δ mit $\sigma = mgu(\Gamma)$ und $\tau = mgu(\Delta)$ an, so dass $\Gamma\sigma \cup \Delta\tau$ unifizierbar ist, $\Gamma \cup \Delta$ aber nicht.

21.3.3 Zeigen Sie: Zu jeder Substitution ρ gibt es eine Anwendung des Transformationssatzes 21.2.4 (sogar mit $\tau = \pi = \mu = id$), die ρ zur Darstellung einer gegebenen $(Subst)$-(Mix)-Kombination nach einer geeigneten vollen Resolution verwendet.

21.3.4 Es sei $\Gamma\sigma = \Delta$ und π eine Umbenennung.

a. Geben Sie eine Substitution τ an mit $\Gamma\pi\tau = \Delta\pi$.

b. Zeigen Sie für diese τ:

$$\sigma = mgu(\Gamma) \;\Leftrightarrow\; \tau = mgu(\Gamma\pi).$$

21.3.5 Zeigen Sie: $\{(Res)\}$ ist widerlegungsvollständig für quantorenfreie Gentzen-Theorien.

21.3.6 $L(T)$ sei gegeben durch die 1-stelligen Prädikatszeichen p, q, r, das 1-stellige Funktionszeichen f und die Konstante 0. a, b seien verschiedene Variablen. Grundsequenzen von T seien:

$$pa, qfa : \emptyset; \quad pa : qb, qf0; \quad r0 : p0, pb \quad \text{und} \quad \emptyset : p0, pb, ra.$$

Zeigen Sie $T \vdash_{\mathcal{R}} \square$

a. für $\mathcal{R} = \{(Mix), (Subst)\}$

b. für $\mathcal{R} = \{(RES)\}$.