Logische Systeme der Informatik

Hubert Wagner
Universität Dortmund
FB Informatik LS1
D-44221 Dortmund

Wintersemester 2000/2001

Inhaltsverzeichnis

Vorwort

Dieses Skript ist als Ergänzung und als Begleitmaterial zur Vorlesung „Logische Systeme der Informatik" gedacht. In seiner im Augenblick vorliegenden Form kann eine alleinige Lektüre den Besuch der Vorlesung nicht ersetzen, da insbesondere Erläuterungen, Motivationen und auch Beispiele manchmal nicht oder nur in gekürzter Form im Skript enthalten sind. Darüberhinaus aber enthält das Skript alle Begriffe, Definitionen und Sätze (inklusive Beweise) usw., die in der Vorlesung behandelt werden. Natürlich (und zugleich auch bedauerlicherweise) bleibt es nicht aus, daß sich beim Schreiben eines solchen Skripts Fehler einschleichen, inhaltliche Fehler sollten aber ausgeschlossen sein. Sollten Sie Fehler – eventuell doch noch inhaltlicher Art – finden, so wäre ich dafür dankbar, wenn Sie mir diese mitteilen würden (email: wagner@ls1.informatik.uni-dortmund.de).

Einleitung

Die Entwicklung der modernen Logik als „Wissenschaft vom richtigen Denken” geht in ihren Wurzeln auf die alten Griechen zurück. Zwei Schulen waren prägend, die beide etwa zur selben Zeit (ca. 300 vor Christus) entstanden sind und die beide unterschiedliche Ansätze verfolgten. Die von Zenon aus Kition (Zypern) begründete Schule der Stoiker betrieb die Analyse von Sätzen und Schlussfolgerungen in einer Weise, die als Vorläufer der Wahrheitswerttabellenmethode der Aussagenlogik angesehen werden kann. Die von Aristoteles entwickelte und von seinen Schülern verbreitete Syllogistik beinhaltete Schlussregeln für Aussagen der Form „alle S sind (nicht) P” und „einige S sind (nicht) P”. Als „Lehre vom folgerichtigen Schließen” war Aristoteles' Syllogistik noch bis Anfang des neunzehnten Jahrhunderts von zentraler Bedeutung. Erst Mitte bzw. Ende des neunzehnten Jahrhunderts nahm die moderne Logik durch die Arbeiten von Boole zu seiner „Algebra der Logik” und von G. Peano, G. Frege und C.S. Peirce ihre heutige Form an. Noch in den 20er Jahren unseres Jahrhunderts wurde jedoch Logik eher als philosophischer Gegenstand behandelt. Zusehends bediente man sich jedoch mathematischer Methoden zur Formalisierung der Semantik und zur Analyse von Schlüssen und Beweissystemen. Formale Logik wurde zu einer mathematischen Disziplin, in deren weiteren Entwicklung sich als fast schon eigenständige Bereiche Mengenlehre, Modelltheorie, Beweistheorie und Rekursionstheorie herauskristallisierten, die ihrerseits selbst wiederum wichtige Beiträge zur Mathematik lieferten.

Die Entwicklung der Informatik führte zu einem Schub in der Logikforschung, in deren Folge weitere bedeutende Forschungsbereiche der Logik oder zumindest solche mit Logikbestandteilen entstanden sind: hierzu zählen Programmverifikation und -spezifikation im Bereich der sequentiellen und nebenläufigen Systeme, logische und funktionale Programmierung, deduktive Datenbanksysteme, automatisches Beweisen, Beschreibungssprachen und die endliche Modelltheorie. Viele Problemstellungen der Informatik führten zum Aufbau neuer nichtklassischer Logiken und damit zu einer Vielzahl neuartiger logischer Fragestellungen. Im Wechselspiel befruchten sich somit Informatik und Logik gegenseitig.

In der Veranstaltung „Logische Systeme der Informatik” sollen die wesentlichen Grundlagen aus den Bereichen der klassischen Aussagen- und Prädikatenlogik, der Logischen Programmierung und der Modallogiken behandelt werden, wobei wir hinsichtlich der Beweiskalküle solchen Systemen den Vorzug gegeben haben, die gut zu erlernen, leicht anwendbar bzw. die Grundlage von automatischen Theorembeweisern sind.

⟨noch zu ergänzen⟩

Ergänzende Literatur:

(Es werden zu den einzelnen Schwerpunktthemen jeweils einige ergänzende und vertiefende Lehrbücher aufgeführt.)

1. Für den gesamten Stoff der Vorlesung
   (a) A. Nerode, R.A. Shore:
   Logic for Applications, Springer 1997
   (b) B. Heinemann, K. Weihrauch:
   Logik für Informatiker, Teubner Verlag 1992
   (c) M. R.A. Huth, M.D. Ryan:
   Logic in Computer Science, Cambridge University Press 2000
2. Aussagen- und/oder Prädikatenlogik, Logische Programmierung
   (a) M. Fitting:
   First-Order Logic and Automated Theorem Proving, Springer 1996

(b) H. Thiele:
Logische Systeme der Informatik, Vorlesungsskript 1994

(c) H.-D. Ebbinghaus, J. Flum, W. Thomas:
Einführung in die mathematische Logik, Spektrum Verlag 1996

(d) U. Schöning:
Logik für Informatiker, Spektrum Verlag 1995

(e) D. van Dalen:
Logic and Structure, Springer Verlag 1997

(f) W. Rautenberg:
Einführung in die mathematische Logik, Vieweg Verlag 1996

3. Logische Programmierung

(a) A. Leitsch:
The Resolution Calculus, Springer Verlag 1997

(b) J.W. Lloyd:
Foundations of Logic Programming, Springer Verlag 1993

(c) H.J. Goltz, H. Herre:
Grundlagen der logischen Programmierung, Akademie-Verlag 1990

4. Modal- und Temporallogik

(a) R. Goldblatt:
Logics of Time and Computation, CSLI 1992

(b) B.F. Chellas:
Modal Logic, Cambridge University Press 1984

(c) F. Kröger:
Temporal Logic of Programs, Springer Verlag 1987

(d) L. Bolc, A. Szalas (Hrg.):
Time and Logic: a computational approach, UCL Press 1995

(e) Z. Manna, A. Pnueli:
Temporal Verification of Reactive Systems: Safety, Springer Verlag 1995

Kapitel 1

Aussagenlogik

In der Einleitung hatten wir bereits erläutert, dass Logik in gewissem Sinne als eine Theorie des Schließens bzw. Beweisens angesehen werden kann. Gegenstände von Schlussweisen sind im intuitiven Sinne natürlich Aussagen, d.h. sprachliche Gebilde, die uns zur Mitteilung von Sachverhalten dienen. Beschreibt eine solche Aussage in zutreffender Weise einen Sachverhalt, so werden wir diese Aussage als wahr ansehen, andernfalls als falsch. Wir nehmen dabei den Standpunkt ein, dass eine Aussage immer entweder wahr oder falsch ist, auch dann, wenn wir (prinzipiell) keine Möglichkeit haben, zu entscheiden, welcher der beiden Fälle zutrifft. Dieser Standpunkt ist nicht unumstritten. Wir werden später darauf noch kurz eingehen.

Nun gibt es viele Beispiele, insbesondere in idealisierten Gegenstandsbereichen wie solche der Mathematik, in denen der Wahrheitswert einer Aussage in unstrittiger und eindeutiger Weise gegeben ist. Unsere Umgangssprache ist jedoch voll von Beispielen, in denen die Wahrheit einer Aussage abhängt z.B. vom Kontext der Äußerung (von sogenannten Indikatoren) oder von der Intention, mit der diese Aussage getroffen wird. Während mit der mathematischen Aussage „239 ist eine Primzahl" die Vorstellung verbunden ist, dass diese Aussage zu allen Zeiten und an allen Orten gilt, trifft dies auf eine Äußerung der Art „Heute regnet es in Köln" nicht mehr zu. Abhängig von dem Indikator „heute" wird diese Aussage einmal wahr und ein andermal falsch sein. Im Rahmen der Aussagenlogik wollen wir nur solche Aussagen betrachten, die keine Indikatoren beinhalten. Wir gehen daher davon aus, dass jeder Aussage ein fester Wahrheitswert zugeordnet ist.

Einige dieser Aussagen wie z.B. „239 ist eine Primzahl" sind minimal in dem Sinne, dass sie keine echten Teilaussagen enthalten. Andere wiederum sind über sprachliche Verknüpfungen gebildet, z.B. „$\sqrt{2}^{\sqrt{2}}$ ist rational oder irrational". Unsere Sprache bietet aber eine Vielzahl von Möglichkeiten, von der Intention her gleiche sprachliche Verknüpfungen in unterschiedlicher Weise sprachlich widerzugeben.

Die Verneinung einer Aussage kann z.B. in der Form

„Der Buchstabe 'a' ist in 'Wort' *nicht* enthalten"

erfolgen. Wie das Beispiel

„*Keine* ungerade Zahl ist durch 2 teilbar"

zeigt, muss aber keineswegs eine Verneinung mittels „nicht" gebildet sein.

Konjunktive Verknüpfungen sind in der Umgangssprache häufig durch „und" sowie durch „aber" formuliert:

„4 ist größer als 3 *und* 4 ist keine Primzahl"

„4 ist größer als 3, *aber* kleiner als 5".

Disjunktive Verknüpfungen werden üblicherweise mittels „oder" ausgedrückt:

„Das Programm terminiert *oder* es ist nicht korrekt". Meistens ist in der Umgangssprache dieses „oder" in einem ausschließenden Sinne zu verstehen.

Implikative Verknüpfungen bilden wir in einer der nachfolgenden Formen:

„$a^2 > 0$ für eine natürliche Zahl a *impliziert* $a > 0$"

„Aus $a^2 > 0$ für eine natürliche Zahl a *folgt* $a > 0$"

„*wenn* $a^2 > 0$, *dann* $a > 0$",

wobei im allgemeinen eine *inhaltliche Abhängigkeit* der beiden Teilaussagen unterstellt ist.

Im folgenden wollen wir nun eine formale Sprache aufbauen, die diese Vagheiten und Vieldeutigkeiten in der Bildung von Aussagenverknüpfungen vermeidet, aber dennoch die erforderlichen Ausdrucksmöglichkeiten zur Verfügung stellt.

1.1 Formalisierung der Aussagenlogik

Wir werden zunächst ein fixiertes Alphabet zugrundelegen, mit dem Zeichenreihen gebildet werden können, die einfache („atomare") Aussagen repräsentieren. Sogenannte logische „Funktoren" oder „Konnektoren" dienen dann zur Bildung neuer Zeichenreihen, die verknüpfte Aussagen repräsentieren. Diese logischen Konnektoren entsprechen in ihrer Verwendung weitestgehend dem umgangssprachlichen Gebrauch, werden aber insbesondere eine präzise definierte Semantik haben.

1.1.1 Syntax der Aussagenlogik

Im folgenden führen wir nun eine formale Sprache ein, deren Elemente Aussagen in ihrer abstrakten Form darstellen. Wir legen hierzu ein Alphabet Σ_{AL} zugrunde, das die nachfolgend aufgeführten Elemente enthält:

- Symbole zur Bildung der Aussagenvariablen: $A, |$
- Konstanten: $\perp$ (bottom oder Falsum) und $\top$ (top oder Verum)
- Zeichen für die logischen Funktoren oder Konnektoren:

  $\neg$ (einstelliger Konnektor, entspricht umgangssprachlich dem „nicht")

  $\wedge$, $\vee$, $\rightarrow$ (2-stellige Konnektoren, entsprechen umgangssprachlich „und" bzw. „oder" bzw. „wenn - dann")

- Klammersymbole als technische Hilfszeichen: (,).

Definition 1.1 *Die folgenden Zeichenreihen über dem Alphabet Σ_{AL} heißen* Aussagenvariablen (*oder aussagenlogische Variablen*)*:* $A, A|, A||, \ldots$.
Schreibweise: statt $A|\ldots|$ ($n+1$ *Striche*)*:* A_{n+1}
Die Menge dieser Aussagenvariablen sei mit Var_{AL} *bezeichnet.*

Als metasprachliche Symbole für Aussagenvariablen verwenden wir im folgenden auch die Symbole $A, A_1, A_2, \ldots$, $B, B_1, B_2, \ldots$, $C, C_1, C_2, \ldots$.
Ausgehend von den Aussagenvariablen bilden wir mittels der logischen Konnektoren Formeln:

Definition 1.2 (Formeln des Aussagenkalküls) *1. Jede Konstante und jede Aussagenvariable ist eine aussagenlogische Formel* (atomare Formel).

2. Falls α und β aussagenlogische Formeln sind, dann auch $\neg\alpha$, $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$ und $(\alpha \rightarrow \beta)$.

3. Nur Zeichenreihen, die nach (1) *oder* (2) *erhalten werden, sind aussagenlogische Formeln.*
Die Menge aller aussagenlogischer Formeln sei mit Fm_{AL} *bezeichnet.*

Bemerkung 1.3 *Es existiert eine eindeutige kontextfreie Grammatik über dem Alphabet Σ_{AL}, die die Menge der Formeln des Aussagenkalküls erzeugt.*

Der nachfolgend definierte Begriff des Formelbaumes ist mehr aus technischer Sicht von Bedeutung. Er dient im Wesentlichen dazu, Begriffe wie Teilformeln auf einfache Weise einzuführen.

Definition 1.4 *Ein* Formelbaum T *ist ein endlicher, mit aussagenlogischen Formeln markierter Binärbaum, der den folgenden Bedingungen genügt:*

1. *Die Blätter sind mit Aussagenvariablen markiert.*
2. *Ist ein Knoten σ von T mit $(\alpha \circ \beta)$ für einen logischen Konnektor $\circ \in \{\wedge, \vee, \rightarrow\}$ markiert, so ist der linke Sohn von σ mit α, und der rechte Sohn von σ mit β markiert.*
3. *Ist ein Knoten σ mit $\neg\alpha$ markiert, so hat σ genau einen Sohn. Dieser ist mit α markiert.*

Definition 1.5 *T ist ein* Formelbaum zur Formel α *gdw. gilt*

1. *T ist ein Formelbaum und*
2. *die Wurzel von T ist mit α markiert.*

Nach dem, was Sie in der Grundvorlesung GTI gehört haben, sollte die folgende Aussage klar sein.

Folgerung 1.6 *Der Formelbaum einer aussagenlogischen Formel ist eindeutig.*

Definition 1.7 *Eine Formel β ist* Teilformel *einer Formel α genau dann, wenn im Formelbaum von α ein Knoten mit β markiert ist.*
Ist β verschieden von α, so heißt β echte Teilformel *von α.*
Ist im Formelbaum von α ein Sohn des Wurzelknotens mit β markiert, so ist β direkte Teilformel *von α.*

Zur Vereinfachung der Schreibweise treffen wir die folgenden **Klammerungskonventionen**, von denen wir des öfteren Gebrauch machen.

- Das äußerste Klammerpaar muss nicht geschrieben werden. Wir dürfen also $A \vee B$ statt $(A \vee B)$ schreiben.
- Wir vereinbaren, dass in der Reihenfolge
$$\neg,\ \wedge,\ \vee,\ \rightarrow$$
jeder logische Konnektor stärker trennt als die vorhergehenden, so dass also z.B. die Formel $\neg A \rightarrow B \wedge \neg A \vee C$ eine kürzere Schreibweise für $(\neg A \rightarrow ((B \wedge \neg A) \vee C))$ ist.

Notation 1.8 $Var_{AL}(\alpha)$ *bezeichne die Menge der in α vorkommenden Aussagenvariablen.*

Um für Formelmengen bestimmte Eigenschaften nachzuweisen, werden wir oft auf das Beweisprinzip der strukturellen Induktion zurückgreifen:

Theorem 1.9 (Strukturelle Induktion oder auch Induktion über den Aufbau der Formeln) *E sei eine Eigenschaft von Formeln, für die gilt:*

1. *Alle atomaren Formeln haben die Eigenschaft E.*
2. *Haben α und β die Eigenschaft E, so auch $\neg\alpha$ und $(\alpha \circ \beta)$ für einen logischen Konnektor $\circ \in \{\wedge, \vee, \rightarrow\}$.*

Dann haben alle aussagenlogischen Formeln die Eigenschaft E.

Beweis. (Informell) Es sei $M(E)$ die Menge aller aussagenlogischer Formeln, die die Eigenschaft E haben.
Nach (1) sind alle atomaren Formeln in $M(E)$ enthalten.
Wegen (2) sind mit α und β auch $\neg\alpha$ und $\alpha \circ \beta$ für $\circ \in \{\wedge, \vee, \rightarrow\}$ in $M(E)$ enthalten. Nach Definition von Fm_{AL} folgt somit
$$Fm_{AL} \subseteq M(E) \subseteq Fm_{AL}.$$
■

Dem Beweisprinzip der strukturellen Induktion korrespondiert als Definitonsprinzip für Funktionen auf der Menge der aussagenlogischen Formeln die *Strukturelle Rekursion.*

Theorem 1.10 (Strukturelle Rekursion) *Es existiert auf der Menge der aussagenlogischen Formeln genau eine Funktion f,*

- *die auf der Menge der atomaren Formeln explizit festgelegt ist, und*
- *für die der Wert für $\neg\alpha$ über den Wert für α und der Wert für $(\alpha \circ \beta)$, $\circ \in \{\wedge, \vee, \rightarrow\}$, über die Werte für α und β bestimmt ist.*

Auf den nicht wesentlich schwierigeren, aber etwas aufwendigeren Beweis dieses Theorems verzichten wir.

1.1.2 Semantik

Mit dem Aufbau einer formalen Sprache liegen uns mit den Formeln rein syntaktische Gebilde vor. Für diese Formeln wollen wir eine Beziehung zu inhaltlichen Aussagen herstellen: zunächst dienen uns Aussagenvariablen zur Repräsentation von Aussagen. Durch Interpretation der Aussagenvariablen durch diese Aussagen und der Konnektoren durch die entsprechende sprachliche Verknüpfung geht eine Formel in eine Aussage über. Vom logischen Standpunkt interessiert uns aber nur der Wahrheitswert einer solchen Aussage. Die Konsequenz dieser Auffasung ist, dass wir Aussagenvariablen nur als Variablen für die Wahrheitswerte $True$ und $False$ (hier mit T und F abgekürzt) betrachten. Ergibt sich dann der Wahrheitswert einer solchen zusammengesetzten Aussage aus den Wahrheitswerten der Teilaussagen, und wenn ja, in welcher Weise? Wir wollen hierzu entsprechende Festlegungen treffen, die weitestgehend unserem umgangssprachlichen Gebrauch von sprachlichen Verknüpfungen entlehnt sind. Insbesondere aber stellen wir uns auf einen extensionalen Standpunkt, nach dem nur der Wahrheitswert der einzelnen Teilaussagen für den Wahrheitswert der Gesamtaussage von Relevanz ist.

Definition 1.11 *Eine* Belegung (*oder Interpretation*) *ist eine Abbildung* $I : Var_{AL} \to \{T, F\}$.

Definition 1.12 *Der* Wert α^I *einer aussagenlogischen Formel* α *unter der Belegung* I *ist wie folgt definiert:*
$A^I := I(A)$
$\top^I := T$
$\bot^I := F$
$[\neg\alpha]^I := \begin{cases} T & \textit{falls } \alpha^I = F \\ F & \textit{sonst} \end{cases}$
$[\alpha \wedge \beta]^I := \begin{cases} T & \textit{falls } \alpha^I = T \textit{ und } \beta^I = T \\ F & \textit{sonst} \end{cases}$
$[\alpha \vee \beta]^I := \begin{cases} T & \textit{falls } \alpha^I = T \textit{ oder } \beta^I = T \\ F & \textit{sonst} \end{cases}$
$[\alpha \to \beta]^I := \begin{cases} T & \textit{falls } \alpha^I = F \textit{ oder } \beta^I = T \\ F & \textit{sonst} \end{cases}$

Abkürzung: $(A \leftrightarrow B)$ für $((A \to B) \wedge (B \to A))$

Es gilt dann:

$$[\alpha \leftrightarrow \beta]^I := \begin{cases} T & \text{falls } \alpha^I = \beta^I = T \text{ oder } \alpha^I = \beta^I = F \\ F & \text{sonst} \end{cases}$$

Bemerkung 1.13 *Der Konnektor $\to$ steht für das umgangssprachliche „wenn - dann", der als Abkürzung eingeführte Konnektor $\leftrightarrow$ für das umgangssprachliche „genau dann - wenn" bzw. „ist äquivalent zu". Im Unterschied zur umgangssprachlichen Verwendung einer solchen „wenn - dann"-Aussage, für die, um wahr sein zu können, i.a. ein inhaltlicher Zusammenhang zwischen den beiden Teilaussagen bestehen muss, ist hier der Wahrheitswert nur noch von den Wahrheitswerten der beiden Teilaussagen abhängig.*

Von der Intuition her würde man erwarten, dass der Wahrheitswert einer Formel unter einer Belegung nur von den Wahrheitswerten der darin vorkommenden Aussagenvariablen abhängig ist. Das folgende Theorem zeigt, dass dies tatsächlich so ist.

Theorem 1.14 (Koinzidenztheorem) *I_1 und I_2 seien zwei Belegungen, die auf der Menge der in einer Formel α vorkommenden Aussagenvariablen übereinstimmen.*
Es gilt dann:

$$\alpha^{I_1} = \alpha^{I_2}.$$

Beweis. Beweis mittels struktureller Induktion.
E sei die folgendermaßen definierte Eigenschaft von Formeln:
E trifft auf α zu gdw. gilt
stimmen I_1 und I_2 auf $Var_{AL}(\alpha)$ überein, so ist $\alpha^{I_1} = \alpha^{I_2}$.
$M(E)$ sei dann die Menge der Formeln in Fm_{AL}, auf die E zutrifft.
Offensichtlich gilt $\top, \bot \in M(E)$. Sei A eine beliebige Aussagenvariable. Aus $I_1(A) = I_2(A)$ folgt $A^{I_1} = A^{I_2}$, also $A \in M(E)$. Damit sind alle atomaren Formeln in $M(E)$ enthalten.
Sei nun $\alpha \in M(E)$. Betrachte $\neg\alpha$. Stimmen I_1 und I_2 auf $Var_{AL}(\neg\alpha)$ überein, so wegen $Var_{AL}(\neg\alpha) = Var_{AL}(\alpha)$ auch auf $Var_{AL}(\alpha)$. Da $\alpha \in M(E)$, gilt daher $\alpha^{I_1} = \alpha^{I_2}$ und damit auch $[\neg\alpha]^{I_1} = [\neg\alpha]^{I_2}$.
Analog wird die Eigenschaft für $\alpha \circ \beta$ mit $\circ \in \{\wedge, \vee, \rightarrow\}$ gezeigt. Es folgt somit $Fm_{AL} \subseteq M(E) \subseteq Fm_{AL}$, d.h. für jede aussagenlogische Formel α gilt $\alpha^{I_1} = \alpha^{I_2}$, wenn I_1 und I_2 auf $Var_{AL}(\alpha)$ übereinstimmen. ■

Eine Folge von Theorem 1.14 ist, dass wir, um den Wert einer Formel unter einer Interpretation zu ermitteln, nur die Belegungen der aussagenlogischen Variablen, die in der Formel vorkommen, berücksichtigen müssen, d.h. nur endlich viele aussagenlogische Variablen.

Definition 1.15 1. *Eine Belegung I* erfüllt *eine Formel α genau dann, wenn $\alpha^I = T$.*

2. *I erfüllt eine Menge Σ von Formeln genau dann, wenn I jede Formel $\alpha \in \Sigma$ erfüllt. I heißt dann* Modell *von Σ.*

   $\mathcal{M}(\Sigma)$ bezeichne die Menge aller Modelle von Σ.

3. *Eine Formel α heißt* erfüllbar *genau dann, wenn eine Belegung I existiert, die α erfüllt.*

   α heißt allgemeingültig (*oder* gültig *bzw.* Tautologie) *genau dann, wenn jede Belegung I die Formel α erfüllt. Taut sei die Menge der Tautologien.*

Bemerkung 1.16 *An dieser Stelle sollte noch einmal auf das bereits aus GTI bekannte Erfüllbarkeitsproblem der Aussagenlogik hingewiesen werden:*
Gegeben*: eine aussagenlogische Formel α.*
Frage*: Ist α erfüllbar?*
Dieses Erfüllbarkeitsproblem ist NP-vollständig (bzgl. Polynomialzeitreduktion wie auch bzgl. logspace-Reduktion) und bildet den Ausgangspunkt für viele Reduktionen zum Nachweis der NP-Härte.

Aufgrund des Koinzidenztheorems können wir die Frage nach der Erfüllbarkeit oder Allgemeingültigkeit einer Formel sehr einfach mit der „Methode der Wahrheitswerttabellen" beantworten: in einer Tabelle werden sämtlich mögliche Belegungskombinationen der in der Formel vorkommenden aussagenlogischen Variablen aufgelistet und für jede solche Belegungskombination wird der Wahrheitswert der Formel dadurch ermittelt, dass wir im Formelbaum von den Blättern zur Wurzel gehend für jede Teilformel den Wahrheitswert bestimmen. Für Teilformeln, die mehrfach auftreten, braucht der Wahrheitswert natürlich nur einmal ermittelt werden.

Beispiel 1.17 *Wir wollen feststellen, ob die Formel $A \rightarrow (B \rightarrow A)$ allgemeingültig ist. Die Wurzel des Formelbaums ist dann mit $A \rightarrow (B \rightarrow A)$ markiert. Dieser Knoten hat zwei Söhne, die mit A bzw. $B \rightarrow A$ markiert sind. $B \rightarrow A$ hat ebenfalls zwei Söhne, die mit A bzw. B markiert sind.*
Wir stellen dann eine Tabelle in der nachfolgenden Form auf. Dieser Tabelle entnehmen wir, dass $A \rightarrow (B \rightarrow A)$ allgemeingültig ist.

A	B	$B \rightarrow A$	$A \rightarrow (B \rightarrow A)$
T	T	T	T
T	F	T	T
F	T	F	T
F	F	T	T

Der Folgerungsbegriff

Da Tautologien unter allen möglichen Belegungen den Wahrheitswert *True* erhalten, liefert eine Tautologie uns keine Informationen. Wohl aber, und dies wird später im Zusammenhang mit formalen Kalkülen noch klar werden, können mit ihnen aus vorhandenen Aussagen neue Aussagen gefolgert werden. Wir wollen daher als Nächstes den intuitiven Folgerungsbegriff formal fassen.

Definition 1.18 *Sei* Σ *eine Menge von Formeln und* α *eine Formel.*
Aus Σ folgt α (*notiert durch* $\Sigma \models_{AL} \alpha$) *genau dann, wenn jede Belegung* I*, die Modell von* Σ *ist,* α *erfüllt.*
(*andere Sprechweise:* α *ist logische Konsequenz von* Σ*.*)

Definition 1.19 $Cn(\Sigma) := \{\alpha \in Fm_{AL} \mid \Sigma \models_{AL} \alpha\}$ *ist die* Konsequenzenmenge *von* Σ*.*

Proposition 1.20 Σ, Σ_1, Σ_2 *seien beliebige Mengen von aussagenlogischen Formeln.*
Es gilt dann:

1. $\Sigma_1 \subseteq \Sigma_2$ *impliziert* $Cn(\Sigma_1) \subseteq Cn(\Sigma_2)$*.*
2. $\Sigma \subseteq Cn(\Sigma)$*.*
3. $Taut \subseteq Cn(\Sigma)$ *für jede Menge* Σ*.*
4. $Cn(Cn(\Sigma)) \subseteq Cn(\Sigma)$*.*

Beweis.

1. Sei $\alpha \in Cn(\Sigma_1)$ und sei I ein Modell von Σ_2. Dann ist I insbesondere ein Modell von Σ_1. Da $\alpha \in Cn(\Sigma_1)$, erfüllt I somit α, d.h. $\alpha \in Cn(\Sigma_2)$.
2. Offensichtlich erfüllt.
3. $Taut = \{\alpha \in Fm_{AL} \mid \emptyset \models_{AL} \alpha\}$.
   Da $\emptyset \subseteq \Sigma$ folgt die Behauptung nach (1).
4. Sei $\alpha \in Cn(Cn(\Sigma))$, d.h. $Cn(\Sigma) \models_{AL} \alpha$.
   Sei ferner I ein Modell von Σ. Dann erfüllt I jede aussagenlogische Formel β mit $\Sigma \models_{AL} \beta$., also ist I Modell von $Cn(\Sigma)$. Wegen $Cn(\Sigma) \models_{AL} \alpha$ erfüllt I auch α. Insgesamt erhalten wir somit $\alpha \in Cn(\Sigma)$. Dies war zu zeigen.

■

Einfache Eigenschaften

Wir wollen als Nächstes einige einfache Eigenschaften von Formeln festhalten, die wir im Folgenden meist stillschweigend verwenden werden.

Lemma 1.21 *Die folgenden Formeln sind Tautologien:*

1. $(A \vee (B \vee C)) \leftrightarrow ((A \vee B) \vee C)$
   $(A \wedge (B \wedge C)) \leftrightarrow ((A \wedge B) \wedge C)$
2. $(A \vee B) \leftrightarrow (B \vee A)$
   $(A \wedge B) \leftrightarrow (B \wedge A)$
3. $(A \vee (B \wedge C)) \leftrightarrow ((A \vee B) \wedge (A \vee C))$
   $(A \wedge (B \vee C)) \leftrightarrow ((A \wedge B) \vee (A \wedge C))$

4. $(A \leftrightarrow \neg\neg A)$

5. $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$
   $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$
   $\neg(A \rightarrow B) \leftrightarrow (A \wedge \neg B)$

6. $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$

Beweis. Wir wollen exemplarisch einen Fall, nämlich 3. beweisen. Gewissermaßen zur Übung werden wir dies noch einmal sehr ausführlich und ohne Rückgriff auf die Methode der Wahrheitswerttabellen zeigen.
Sei dazu I eine beliebige Belegung. Es gilt
$[(A \vee (B \wedge C))]^I = T$
$\Leftrightarrow$ $A^I = T$ oder $[(B \wedge C)]^I = T$
$\Leftrightarrow$ $A^I = T$ oder $\big(B^I = T$ und $C^I = T\big)$
$\Leftrightarrow$ $\big(A^I = T$ oder $B^I = T\big)$ und $\big(A^I = T$ oder $C^I = T\big)$
$\Leftrightarrow$ $[(A \vee B)]^I = T$ und $[(A \vee C)]^I = T$
$\Leftrightarrow$ $[((A \vee B) \wedge (A \vee C))]^I = T$.
Damit gilt weiter
$[((A \vee (B \wedge C)) \leftrightarrow ((A \vee B) \wedge (A \vee C)))]^I = T$.
Da dies für eine beliebige Belegung gilt, ist $((A \vee (B \wedge C)) \leftrightarrow ((A \vee B) \wedge (A \vee C)))$ somit eine Tautologie. ■

Definition 1.22 *Zwei Formeln α und β heißen (aussagenlogisch)* semantisch äquivalent *(notiert als $\alpha \equiv_{AL} \beta$) genau dann, wenn $\alpha \leftrightarrow \beta$ eine Tautologie ist.*

Aussagenlogische Formeln sind somit genau dann semantisch äquivalent, wenn sie unter beliebigen Belegungen immer den gleichen Wahrheitswert erhalten.

Lemma 1.23 *Für beliebige Formeln α und β sowie Formelmengen Σ gilt:*

1. *Wenn $\Sigma \models_{AL} \alpha \rightarrow \beta$ und $\Sigma \models_{AL} \alpha$, so $\Sigma \models_{AL} \beta$*
   (Modus ponens)

2. *Wenn $\Sigma \models_{AL} \alpha \rightarrow \beta$, so $\Sigma \cup \{\alpha\} \models_{AL} \beta$*
   (Ableitungstheorem)

3. *Wenn $\Sigma \cup \{\alpha\} \models_{AL} \beta$, so $\Sigma \models_{AL} \alpha \rightarrow \beta$*
   (Deduktionstheorem)

Beweis. Der Beweis des Lemmas ist sehr einfach. Exemplarisch zeigen wir den Fall 3.
Sei I eine Belegung, die ein Modell von Σ ist. Wir müssen zeigen, dass $[\alpha \rightarrow \beta]^I = T$ gilt, d.h. falls $\alpha^I = T$, dann muss $\beta^I = T$ gelten.
Nach Voraussetzung gilt nun: ist I ein Modell von $\Sigma \cup \{\alpha\}$, d.h. ist I ein Modell von Σ und erfüllt I α, so erfüllt I β. Genau dies war zu zeigen. ■

Das Ersetzungstheorem

Es ist leicht zu sehen, dass die semantische Äquivalenz eine Äquivalenzrelation ist, d.h. dass für beliebige aussagenlogische Formeln α und β gilt

$\alpha \equiv_{AL} \alpha$ (Reflexivität)

$\alpha \equiv_{AL} \beta \Rightarrow \beta \equiv_{AL} \alpha$ (Symmetrie)

$\alpha_1 \equiv_{AL} \alpha_2$ und $\alpha_2 \equiv_{AL} \alpha_3 \Rightarrow \alpha_1 \equiv_{AL} \alpha_3$ (Transitivität).

Ist darüber hinaus die semantische Äquivalenz auch eine Kongruenzrelation bezgüglich der logischen Konnektoren, gilt also z.B.

$$\alpha \equiv_{AL} \alpha_1 \text{ und } \beta \equiv_{AL} \beta_1 \;\Rightarrow\; (\alpha \to \beta) \equiv_{AL} (\alpha_1 \to \beta_1) \text{ ?}$$

Das nachfolgende Ersetzungstheorem wird uns hierfür eine positive Antwort liefern und uns zugleich eine Methode an die Hand geben, um aus vorhandenen Tautologien neue zu erzeugen.

Notation 1.24 *Für eine Formel β bezeichne $\alpha\left(\frac{A}{\beta}\right)$ diejenige Formel, die aus α dadurch entsteht, dass jedes Vorkommen von A in α* simultan *durch β ersetzt wird.*
(*Achtung: A muss in α nicht vorkommen!*)

Beispiel 1.25 $\alpha = (A \to B) \vee \neg A$
$\beta = (A \vee C)$
$\Rightarrow$: $\quad \alpha\left(\frac{A}{\beta}\right) = ((A \vee C) \to B) \vee \neg\,(A \vee C)$

Zur Schreibweise:
$\Rightarrow$ drückt das meta- bzw. umgangssprachliche „impliziert" aus.
Analog soll $\Leftrightarrow$ das metasprachliche „ist äquivalent" bzw. „genau dann, wenn" bezeichnen.
Für die definitorische Äquivalenz verwenden wir auch das Symbol $\Leftrightarrow_{def}$.

Theorem 1.26 (Ersetzungstheorem) *α, β_1 und β_2 seien beliebige Formeln, A eine Aussagenvariable.*

1. *I sei eine Belegung mit $[\beta_1]^I = [\beta_2]^I$.*
   Dann gilt:
   $$\left[\alpha\left(\frac{A}{\beta_1}\right)\right]^I = \left[\alpha\left(\frac{A}{\beta_2}\right)\right]^I.$$
2. $$\beta_1 \equiv_{AL} \beta_2 \quad\Rightarrow\quad \alpha\left(\frac{A}{\beta_1}\right) \equiv_{AL} \alpha\left(\frac{A}{\beta_2}\right).$$

Beweis.

1. Wir führen den Beweis durch eine strukturelle Induktion.
   Eine Formel α werde als „gut" bezeichnet (dies ist also die Eigenschaft E), falls folgendes gilt:
   $[\beta_1]^I = [\beta_2]^I$ impliziert $\left[\alpha\left(\frac{A}{\beta_1}\right)\right]^I = \left[\alpha\left(\frac{A}{\beta_2}\right)\right]^I$.
   Über eine strukturelle Induktion zeigen wir, dass alle Formeln gut sind, d.h. die Eigenschaft E haben.
   Es gelte $[\beta_1]^I = [\beta_2]^I$.
   Induktionsbasis:
   Es ist $\left[\bot\left(\frac{A}{\beta_1}\right)\right]^I = \bot^I = \left[\bot\left(\frac{A}{\beta_2}\right)\right]^I$. $\bot$ ist also gut. Analog wird gezeigt, dass $\top$ gut ist.
   Sei nun B eine Aussagenvariable.
   Fall 1: $B = A$ (d.h. B ist identisch mit A als Zeichenreihe).
   Dann gilt:
   $\left[B\left(\frac{A}{\beta_1}\right)\right]^I = \left[A\left(\frac{A}{\beta_1}\right)\right]^I = [\beta_1]^I = [\beta_2]^I = \left[A\left(\frac{A}{\beta_2}\right)\right]^I = \left[B\left(\frac{A}{\beta_2}\right)\right]^I.$
   Fall 2: $B \neq A$:
   $\left[B\left(\frac{A}{\beta_1}\right)\right]^I = B^I = \left[B\left(\frac{A}{\beta_2}\right)\right]^I.$

In beiden Fällen ist B also gut.

Damit sind alle atomaren Formeln als gut nachgewiesen, und die Induktionsbasis ist gesichert.

Induktionsschritt:

Wir müssen zeigen, dass die Menge der guten Formeln abgeschlossen unter $\neg$ und $\circ$ für $\circ \in \{\wedge, \vee, \rightarrow\}$ ist. Wir betrachten exemplarisch den Konnektor $\rightarrow$.

Seien α und β gut. Es ist

$\left[(\alpha \rightarrow \beta)\left(\frac{A}{\beta_1}\right)\right]^I = T$

genau dann, wenn $\left[\alpha\left(\frac{A}{\beta_1}\right)\right]^I = F$ oder $\left[\beta\left(\frac{A}{\beta_1}\right)\right]^I = T$

genau dann, wenn $\left[\alpha\left(\frac{A}{\beta_2}\right)\right]^I = F$ oder $\left[\beta\left(\frac{A}{\beta_2}\right)\right]^I = T$

genau dann, wenn $\left[(\alpha \rightarrow \beta)\left(\frac{A}{\beta_2}\right)\right]^I = T$.

Also ist $\alpha \rightarrow \beta$ gut

Damit ist der Induktionsschritt bewiesen.

Mit dem Theorem über strukturelle Induktion folgt dann, dass alle Formeln gut sind. Dies war zu zeigen.

2. Sei $\beta_1 \leftrightarrow \beta_2$ eine Tautologie. Dann gilt für eine beliebige Belegung I $[\beta_1]^I = [\beta_2]^I$.

   Damit gilt nach (1) für jede Belegung I $\left[\alpha\left(\frac{A}{\beta_1}\right)\right]^I = \left[\alpha\left(\frac{A}{\beta_2}\right)\right]^I$, also $\left[\alpha\left(\frac{A}{\beta_1}\right) \leftrightarrow \alpha\left(\frac{A}{\beta_2}\right)\right]^I = T$. Daher ist $\alpha\left(\frac{A}{\beta_1}\right) \leftrightarrow \alpha\left(\frac{A}{\beta_2}\right)$ eine Tautologie.

■

Beispiel 1.27 *Es gilt* $(\neg A \vee B) \equiv_{AL} (A \rightarrow B)$.
Betrachte $\alpha = (\neg A \vee B) \wedge (B \rightarrow C)$.
Dann sind semantisch äquivalent

$$\alpha\left(\frac{C}{(\neg A \vee B)}\right) = (\neg A \vee B) \wedge (B \rightarrow (\neg A \vee B))$$

und

$$\alpha\left(\frac{C}{(A \rightarrow B)}\right) = (\neg A \vee B) \wedge (B \rightarrow (A \rightarrow B)).$$

1.1.3 α- und β-Formeln

Mit der Einführung von α- und β-Formeln klassifizieren wir Formeln einer bestimmten formalen Struktur. Die hierdurch eingeführte Terminologie erweist sich als besonders nützlich für die Formulierung und Darstellung von formalen Beweiskalkülen
Wir teilen Formeln vom Typ $\alpha \circ \beta$ und $\neg(\alpha \circ \beta)$ in zwei Kategorien ein:
 −α-Formeln, die sich wie Konjunktionen verhalten, und
 −β-Formeln, die sich wie Disjunktionen verhalten.
α-Formeln sind Formeln der Gestalt $\alpha \wedge \beta$, $\neg(\alpha \vee \beta)$ und $\neg(\alpha \rightarrow \beta)$. β-Formeln sind Formeln der Gestalt $\alpha \vee \beta$, $\neg(\alpha \wedge \beta)$ und $(\alpha \rightarrow \beta)$. α-Formeln und β-Formeln (im Folgenden immer mit $\boldsymbol{\alpha}$ bzw. $\boldsymbol{\beta}$ bezeichnet) ordnen wir jeweils Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ bzw. $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$ in der folgenden Weise zu:

$\boldsymbol{\alpha}$	$\boldsymbol{\alpha}_1$	$\boldsymbol{\alpha}_2$
$\alpha \wedge \beta$	α	β
$\neg(\alpha \vee \beta)$	$\neg\alpha$	$\neg\beta$
$\neg(\alpha \rightarrow \beta)$	α	$\neg\beta$

(α-Formeln mit Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$)

$\boldsymbol{\beta}$	$\boldsymbol{\beta}_1$	$\boldsymbol{\beta}_2$
$\neg(\alpha \wedge \beta)$	$\neg\alpha$	$\neg\beta$
$\alpha \vee \beta$	α	β
$\alpha \rightarrow \beta$	$\neg\alpha$	β

(β-Formeln und ihre Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$)

Theorem 1.28 *Für alle α-Formeln und β-Formeln und für alle Belegungen I gilt:*

$$\boldsymbol{\alpha}^I = [\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2]^I$$
$$\boldsymbol{\beta}^I = [\boldsymbol{\beta}_1 \vee \boldsymbol{\beta}_2]^I$$

Auf den sehr einfachen Beweis des Theorems verzichten wir.

Über die α- und β-Formeln erhalten wir ein sehr nützliches, weil einfach anwendbares weiteres Beweisprinzip sowie ein zugehöriges Definitionsprinzip.

Theorem 1.29 (*Strukturelle Induktion über α-, β-Formeln*)
E sei eine Eigenschaft von aussagenlogischen Formeln, für die gilt:

1. *E trifft auf jede atomare Formel und deren Negation zu* (*d.h. auch auf $\neg\alpha$, falls α eine atomare Formel ist*).

   (*Induktionsbasis*)

2. *Hat α die Eigenschaft E, so auch $\neg\neg\alpha$.*

   Haben die Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ einer α-Formel die Eigenschaft E, so auch die α-Formel $\boldsymbol{\alpha}$.

   Haben die Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$ einer β-Formel die Eigenschaft E, so auch die β-Formel $\boldsymbol{\beta}$.

   (*Induktionsschritt*)

Dann trifft E auf alle aussagenlogischen Formeln zu.

Beweis. Wir führen den Beweis mittels einer strukturellen Induktion (im Folgenden zur Unterscheidung auch als gewöhnliche strukturelle Induktion bezeichnet).
Sei nach Voraussetzung E eine Eigenschaft, für die (1) und (2) gilt.
Wir bezeichnen eine Formel α als „gut" genau dann, wenn die Eigenschaft E sowohl auf die Formel α als auch auf die Formel $\neg\alpha$ zutrifft. (Die Eigenschaft einer Formel, gut zu sein, ist nun die Eigenschaft E', die wir für die gewöhnliche strukturelle Induktion zugrundelegen).
Wir zeigen mittels gewöhnlicher struktureller Induktion:

$$\text{alle aussagenlogischen Formeln sind gut.} \tag{1.1}$$

Daraus folgt dann insbesondere, dass die Eigenschaft E auf alle aussagenlogischen Formeln zutrifft, d.h. die Behauptung unseres Theorems.
Nachweis von (1.1): Wir müssen dazu zeigen, dass E' die Bedingungen (1) (Induktionsbasis) und (2) (Induktionsschritt) von Theorem 1 erfüllt.

1. Nach Voraussetzung unseres Theorems gilt E für alle atomaren Formeln und deren Negationen. Also sind alle atomaren Formeln gut. Damit ist die Induktionsbasis gezeigt.

2. Für den Induktionsschritt müssen wir die Abschluss eigenschaften nachweisen, d.h. wir müssen zeigen, dass mit α und β auch $\neg\alpha$ und $\alpha \circ \beta$ für $\circ \in \{\wedge, \vee, \rightarrow\}$ gut sind.

   Seien α und β gut. Wir betrachten exemplarisch $\alpha \wedge \beta$. Die anderen Fälle werden analog bewiesen.

   Wir wollen zeigen, dass $\alpha \wedge \beta$ gut ist, d.h. dass E auf $\alpha \wedge \beta$ sowie auf $\neg(\alpha \wedge \beta)$ zutrifft.

   $\alpha \wedge \beta$ ist eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1 = \alpha$ und $\boldsymbol{\alpha}_2 = \beta$, die nach Voraussetzung gut sind. Also trifft E auf α und β zu, damit nach Bedingung (2) unseres Theorems auch auf $\boldsymbol{\alpha}$ selbst, d.h. auf $\alpha \wedge \beta$.

$\neg(\alpha \wedge \beta)$ ist eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1 = \neg\alpha$ und $\boldsymbol{\beta}_2 = \neg\beta$. Nach Voraussetzung sind α und β gut, damit trifft E auf $\neg\alpha$ und auf $\neg\beta$ zu, wegen Bedingung (2) unseres Theorems dann auch auf $\neg(\alpha \wedge \beta)$.

E trifft also auf $\alpha \wedge \beta$ und $\neg(\alpha \wedge \beta)$ zu. Somit ist $\alpha \wedge \beta$ gut.

■

Wie im Falle der gewöhnlichen strukturellen Induktion existiert ein zugehöriges Definitionsprinzip.

Theorem 1.30 (Strukturelle Rekursion über α-, β-Formeln)
Es existiert auf der Menge der aussagenlogischen Formeln genau eine Funktion f, für die gilt:

- *Der Wert von f ist explizit für atomare Formeln und deren Negationen festgelegt.*

  (Rekursionsanfang)

- *Der Wert von f für $\neg\neg\alpha$ ist eindeutig durch den Wert von f für α festgelegt.*

  Der Wert von f für eine α-Formel (β-Formel) ist eindeutig durch den Wert von f für die Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ (bzw. $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$) festgelegt.

  (Rekursionsschritt)

1.1.4 Normalformen

Die im Folgenden eingeführten Normalformen sind Ihnen bereits aus der Veranstaltung Rechnerstrukturen bekannt. Ihre Bedeutung liegt darin, dass die maschinelle Verarbeitung etwas erleichtert wird und der Test auf Erfüllbarkeit bzw. Allgemeingültigkeit einfacher wird. Mittels α- und β-Formeln ergibt sich ein einfaches Verfahren zur Erzeugung dieser Normalformen.
Zunächst führen wir jedoch verallgemeinerte Disjunktionen und Konjunktionen ein.

Definition 1.31 (Verallgemeinerte Disjunktion und Konjunktion)
$\alpha_1, \alpha_2, \ldots$ seien Formeln.

Verallg. Disjunktion	*Verallg. Konjunktion*
$\bigvee_{i=1}^{0} \alpha_i := \bot$	$\bigwedge_{i=1}^{0} \alpha_i := \top$
$\bigvee_{i=1}^{1} \alpha_i := \alpha_1$	$\bigwedge_{i=1}^{1} \alpha_i := \alpha_1$
$\bigvee_{i=1}^{n+1} \alpha_i := \bigvee_{i=1}^{n} \alpha_i \vee \alpha_{n+1}$	$\bigwedge_{i=1}^{n+1} \alpha_i := \bigwedge_{i=1}^{n} \alpha_i \wedge \alpha_{n+1}$

für $n \geq 1$.

Zur Erleichterung des Lesens von Formeln führen wir die folgende Notation ein.

Notation 1.32 *Für $\bigvee_{i=1}^{n} \alpha_i$ schreiben wir auch $[\alpha_1, \ldots, \alpha_n]$. Insbesondere ist also $\bigvee_{i=1}^{0} \alpha_i = [\,]$). Analog schreiben wir für $\bigwedge_{i=1}^{n} \alpha_i$ auch $\langle \alpha_1, \ldots, \alpha_n \rangle$. Es gilt dann $\bigwedge_{i=1}^{0} \alpha_i = \langle\,\rangle$.*

Bemerkung 1.33 *In dem Ausdruck $[\alpha_1, \ldots, \alpha_n]$ (bzw. $\langle \alpha_1, \ldots, \alpha_n \rangle$) kann $[\,]$ (bzw. $\langle\,\rangle$) als ein n-stelliger Konnektor angewandt auf die Formeln $\alpha_1, \ldots, \alpha_n$ aufgefasst werden. Wir haben die Einführung weiterer mehrstelliger Konnektoren vermieden und haben stattdessen die verallgemeinerten Disjunktionen und Konjunktionen als Abkürzungen eingeführt, um an dieser Stelle nicht auf den Aspekt der Definierbarkeit von logischen Konnektoren eingehen zu müssen.*

Folgerung 1.34 *Für Formeln $\alpha_1, \ldots, \alpha_n$ und eine Belegung I gilt:*

1. $[\alpha_1, \ldots, \alpha_n]^I = T$ *genau dann, wenn in der Liste* $\alpha_1, \ldots, \alpha_n$ *ein* α_i *existiert mit* $[\alpha_i]^I = T$. *Speziell gilt* $[\,]^I = F$.
2. $\langle \alpha_1, \ldots, \alpha_n \rangle^I = T$ *genau dann, wenn für jede Formel* α_i *in der Liste* $\alpha_1, \ldots, \alpha_n$ *gilt* $[\alpha_i]^I = T$. *Speziell gilt* $\langle\,\rangle^I = T$.
3. *Sei* $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ *eine Permutation. Dann gilt*
   $[\alpha_1, \ldots, \alpha_n]^I = \left[\alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(n)}\right]^I$
   $\langle \alpha_1, \ldots, \alpha_n \rangle^I = \left\langle \alpha_{\sigma(1)}, \ldots, \alpha_{\sigma(n)} \right\rangle^I$.

Definition 1.35 (Literal, Klausel, duale Klausel)

1. *Ein* Literal *ist eine atomare Formel oder die Negation einer Aussagenvariablen.*
2. *Eine* Klausel *ist eine (verallgemeinerte) Disjunktion* $[\alpha_1, \ldots, \alpha_n]$ *von Literalen* $\alpha_1, \ldots, \alpha_n$.
3. *Eine* duale Klausel *ist eine (verallgemeinerte) Konjunktion* $\langle \alpha_1, \ldots, \alpha_n \rangle$ *von Literalen* $\alpha_1, \ldots, \alpha_n$.

Definition 1.36 (Konjunktive und disjunktive Normalform)
Eine aussagenlogische Formel α *ist in*

1. konjunktiver Normalform *(KNF, auch Klauselmenge oder Klauselform) genau dann, wenn* α *eine Konjunktion* $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle$ *mit Klauseln* $\alpha_1, \ldots, \alpha_n$ *ist,*
2. disjunktiver Normalform*(DNF, auch duale Klauselmenge oder duale Klauselform) genau dann, wenn* α *eine Disjunktion* $\alpha = [\alpha_1, \ldots, \alpha_n]$ *mit dualen Klauseln* $\alpha_1, \ldots, \alpha_n$ *ist.*

Theorem 1.37 *Zu jeder aussagenlogischen Formel* α *existiert*

1. *eine Formel* α_1 *in konjunktiver Normalform mit* $\alpha \equiv_{AL} \alpha_1$,
2. *eine Formel* α_2 *in disjunktiver Normalform mit* $\alpha \equiv_{AL} \alpha_2$.

Beweis. (Mittels einer strukturellen Induktion über α- und β-Formeln. Übungsaufgabe!) ■

Wir wollen uns mit einer bloßen Existenzaussage nicht begnügen, sondern darüberhinaus auch jeweils informell beschriebene, nichtdeterministische Algorithmen angeben, die zu α eine konjunktive resp. eine disjunktive Normalform bestimmen.

1. **Klauselformalgorithmus** zur Erzeugung einer konjunktiven Normalform:
   $S := \langle [\alpha] \rangle$
   While S enthält Nichtklausel als Konjunktionselement *do*
   wähle eine solche Nichtklausel D aus
   wähle in D ein Nichtliteral N aus
   if $N = \neg\top$ *then* ersetze in D das Disjunktionsglied N durch $\bot$;
   die resultierende Disjunktion sei D_1; (notiert als Ersetzungsregel: $\frac{\neg\top}{\bot}$)
   else if $N = \neg\bot$ *then* ersetze in D das Disjunktionsglied N durch $\top$;
   die resultierende Disjunktion sei D_1; (notiert als Ersetzungsregel: $\frac{\neg\bot}{\top}$)
   else if $N = \neg\neg\beta$ *then* ersetze in D das Disjunktionsglied N durch β;
   die resultierende Disjunktion sei D_1; (notiert als Ersetzungsregel: $\frac{\neg\neg\beta}{\beta}$)
   else if N eine β-Formel $\boldsymbol{\beta}$ *then* ersetze in D das Disjunktionsglied N durch 2 Disjunktionsglieder

$\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$;

die resultierende Disjunktion sei D_1; (notiert als Ersetzungsregel: $\dfrac{\boldsymbol{\beta}}{\begin{matrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{matrix}}$)

else (d.h. wenn N eine α-Formel $\boldsymbol{\alpha}$) ersetze in D das Disjunktionsglied N

einmal durch $\boldsymbol{\alpha}_1$ mit dem Resultat D_1 und einmal durch $\boldsymbol{\alpha}_2$ mit dem Resultat D_2

(notiert als Ersetzungsregel: $\dfrac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2}$)

end

ersetze in S das Konjunktionsglied D durch das Konjunktionsglied D_1 bzw.

durch die Konjunktionsglieder D_1 und D_2.

Nenne das Resultat S.

end.

Für diesen Algorithmus ist zweierlei zu zeigen:

(a) Der Algorithmus terminiert.
Nicht ganz so einfach (Übungsaufgabe: geben Sie eine obere Schranke für die Anzahl der Schleifendurchläufe an!)

(b) Das Ergebnis S ist eine zu α semantisch äquivalente konjunktive Normalform.
Hierzu zeigt man, dass die folgende Aussage eine Schleifeninvariante ist:

S ist eine Konjunktion von Disjunktionen und $S \equiv_{AL} \alpha$.

Beim Eintritt in die Schleife gilt $S = \langle[\alpha]\rangle = \alpha$, also $S \equiv_{AL} \alpha$ und S ist eine Konjunktion von Disjunktionen.

Sei $S = S_n$ das Resultat nach n Schleifendurchläufen, und es gelte $S_n \equiv_{AL} \alpha$. Ferner sei S_n eine Konjunktion von Disjunktionen. Es gebe einen weiteren Schleifendurchlauf mit dem Resultat S_{n+1}. Aufgrund der Ersetzunsgregeln ist dann S_{n+1} wiederum eine Konjunktion von Disjunktionen.
Es gilt

$$\begin{array}{ccc}
\neg\top & \equiv_{AL} & \bot \\
\neg\bot & \equiv_{AL} & \top \\
\neg\neg\beta & \equiv_{AL} & \beta \\
\boldsymbol{\alpha} & \equiv_{AL} & \boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2 \\
\boldsymbol{\beta} & \equiv_{AL} & \boldsymbol{\beta}_1 \vee \boldsymbol{\beta}_2 \ ,
\end{array}$$

weiter gilt

$[\alpha_1, \ldots, \alpha_n, \beta \vee \gamma, \beta_1, \ldots, \beta_m] \equiv_{AL} [\alpha_1, \ldots, \alpha_n, \beta, \gamma, \beta_1, \ldots, \beta_m]$,

sowie

$[\alpha_1, \ldots, \alpha_n, \beta \wedge \gamma, \beta_1, \ldots, \beta_m] \equiv_{AL} [\alpha_1, \ldots, \alpha_n, \beta, \beta_1, \ldots, \beta_m] \wedge [\alpha_1, \ldots, \alpha_n, \gamma, \beta_1, \ldots, \beta_m]$.

Mehrfache Anwendung des Ersetzungstheorems liefert somit $S_n \equiv_{AL} S_{n+1}$ und damit $\alpha \equiv_{AL} S_{n+1}$.

Insbesondere folgt, dass die nach Terminierung erzeugte Konjunktion von Disjunktionen, d.h. S, semantisch äquivalent zu α ist.

2. **Algorithmus** zur Erzeugung einer disjunktiven Normalform:

Im Falle, dass eine disjunktive Normalform erzeugt werden soll, starten wir den oben aufgeführten Algorithmus mit $S = [\langle\alpha\rangle]$. Ferner sind die Rollen von Konjunktion und Disjunktion sowie von α- und β-Formeln zu vertauschen , d.h. statt der Ersetzungsregeln

$$\frac{\boldsymbol{\beta}}{\begin{matrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{matrix}}$$

und

$$\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2}$$

haben wir die Ersetzungsregeln

$$\frac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2}$$

und

$$\frac{\boldsymbol{\alpha}}{\begin{array}{c}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{array}}.$$

Der Vollständigkeit halber listen wir den Algorithmus (dualer Klauselformalgorithmus) hier noch auf. Der Korrektheitsbeweis hierzu verläuft vollkommen analog.

Dualer Klauselformalgorithmus:

(a) $S := [\langle\alpha\rangle]$
While S enthält eine nicht duale Klausel als Disjunktionsselement *do*
wähle eine solche nicht duale Klausel K aus
wähle in K ein Nichtliteral N aus
if $N = \neg\top$ *then*
ersetze in K das Konjunktionsglied N durch $\bot$; die resultierende Konjunktion sei K_1
(notiert als Ersetzungsregel: $\frac{\neg\top}{\bot}$)
else if $N = \neg\bot$ *then*
ersetze in K das Konjunktionsglied N durch $\top$; die resultierende Konjunktion sei K_1
(notiert als Ersetzungsregel: $\frac{\neg\bot}{\top}$)
else if $N = \neg\neg\beta$ *then*
ersetze in K das Konjunktionsglied N durch β; die resultierende Konjunktion sei K_1
(notiert als Ersetzungsregel: $\frac{\neg\neg\beta}{\beta}$)
else if N eine α-Formel *then* ersetze in K das Konjunktionsglied N durch 2
Konjunktionsglieder $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$; die resultierende Konjunktion sei K_1
(notiert als Ersetzungsregel: $\frac{\boldsymbol{\alpha}}{\begin{array}{c}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{array}}$)
else (d.h. wenn N eine β-Formel) ersetze in K das Konjunktionsglied N einmal
durch $\boldsymbol{\beta}_1$ mit dem Resultat K_1 und einmal durch $\boldsymbol{\beta}_2$ mit dem Resultat K_2
(notiert als Ersetzungsregel: $\frac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2}$)
end
ersetze in S das Disjunktionsglied K durch das Disjunktionsglied K_1 bzw.
durch die Disjunktionsglieder K_1 und K_2.
Nenne das Resultat S.
end.

■

Zusammengefasst sehen die Ersetzungsregeln zur Ezeugung der konjunktiven Normalform wie folgt aus:

$$\frac{\neg\top}{\bot} \qquad \frac{\neg\bot}{\top} \qquad \frac{\neg\neg\beta}{\beta} \qquad \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2} \qquad \frac{\boldsymbol{\beta}}{\begin{array}{c}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}}$$

Die hierzu dualen Regeln zur Bildung der disjunktiven Normalform:

$$\frac{\neg\top}{\bot} \qquad \frac{\neg\bot}{\top} \qquad \frac{\neg\neg\beta}{\beta} \qquad \frac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2} \qquad \frac{\boldsymbol{\alpha}}{\begin{array}{c}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\end{array}}$$

Beispiel 1.38 *zur Erzeugung einer konjunktiven Normalform für die Formel*

$$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$$

1. $\langle[(\boldsymbol{A} \to (\boldsymbol{B} \to \boldsymbol{C})) \to ((\boldsymbol{A} \to \boldsymbol{B}) \to (\boldsymbol{A} \to \boldsymbol{C}))]\rangle$
2. $\langle[\neg\,(\boldsymbol{A} \to (\boldsymbol{B} \to \boldsymbol{C}))\,,\ (A \to B) \to (A \to C)]\rangle$
3. $\langle[A,\ (A \to B) \to (A \to C)]\,,\ [\neg\,(\boldsymbol{B} \to \boldsymbol{C}),\ (A \to B) \to (A \to C)]\rangle$
4. $\langle[A,\ (\boldsymbol{A} \to \boldsymbol{B}) \to (\boldsymbol{A} \to \boldsymbol{C})]\,,\ [B,\ (A \to B) \to (A \to C)]\,,\ [\neg C,\ (A \to B) \to (A \to C)]\rangle$
5. $\langle[A,\ \neg\,(A \to B)\,,\ A \to C]\,,[B,\ (A \to B) \to (A \to C)]\,,\ [\neg C,\ (A \to B) \to (A \to C)]\rangle$

   $\vdots$ (*mehrere Schritte*)

6. $\langle[A, A, \neg A, C]\,,\ [A, \neg B, \neg A, C]\,,\ [B, A, \neg A, C]\,,\ [B, \neg B, \neg A, C]\,,\ [\neg C, A, \neg A, C]\,,\ [\neg C, \neg B, \neg A, C]\rangle$ ■

1.2 Aussagenlogische Beweiskalküle

Mit der präzisen Definition der Semantik haben wir ein „im Prinzip" taugliches Instrument, um festzustellen, ob eine aussagenlogische Formel allgemeingültig ist oder aus einer gegebenen Menge von aussagenlogischen Formeln folgt. Mit der Methode der Wahrheitswerttabellen haben wir ein einfaches, wenngleich auch aufwendiges Verfahren, um die Allgemeingültigkeit zu entscheiden. Schwieriger wird es bereits, wenn die Frage nach der logischen Konsequenz z.B. aus einer unendlichen Formelmenge zu beantworten ist. Hierfür kann es zwar nachweislich kein algorithmisches Verfahren geben, ein allein semantisch orientiertes Vorgehen müsste jedoch bereits überabzählbar viele Belegungen in Betracht ziehen.

Nun war gerade einer der zentralen Beweggründe in der Entwicklung der modernen Logik die Vorstellung, das logische Schließen so formalisieren zu können, dass die Schlussweisen wie auch die Richtigkeit von Schlüssen rein „mechanisch", also algorithmisch überprüfbar werden. Inhaltliches Schließen soll durch rein formale Anwendungen von Regeln ersetzt werden. Von einem solch formalen Regelkalkül (oder wie wir im Folgenden sagen wollen: Beweiskalkül) erwarten wir natürlich, dass er nichts Falsches beweist, d.h. der Beweiskalkül soll „korrekt" sein. Darüberhinaus hätten wir gerne, dass in diesem Beweiskalkül auch jede Tautologie bewiesen werden kann, d.h. der Beweiskalkül soll „vollständig" sein.

Im folgenden werden wir uns einige interessante Beweiskalküle der Aussagenlogik ansehen. Diese Beweiskalküle können nach ihrem grundsätzlichem Ansatz klassifiziert werden in zum einen sogenannte axiomatische Systeme, die ausgehend von Axiomen und endlich vielen Regeln über endlich viele Regelanwendungen eine Formel ableiten, und in sogenannte Widerlegungsverfahren, die, um die Allgemeingültigkeit einer Formel zu zeigen, so vorgehen, dass sie annehmen, die Negation dieser Formel sei erfüllbar, und dann versuchen, über endlich viele Regelanwendungen daraus einen Widerspruch herzuleiten. Wir wollen zunächst zwei solcher Widerlegungsverfahren betrachten: den Tableau-Kalkül und den Resolutionskalkül.

1.2.1 Der Tableau-Kalkül

Der Tableau-Kalkül – ursprünglich als Methode der semantischen Tafeln bezeichnet – hat seinen Ursprung in der Methode der Wahrheitswerttabellen. Eine aussagenlogische Formel α ist nicht allgemeingültig genau dann, wenn sie unter einer Belegung den Wert F erhält, also wenn $\neg\alpha$ erfüllbar ist. Nun gilt, dass eine Belegung z.B. eine Formel $\beta_1 \wedge \beta_2$ erfüllt genau dann, wenn sie β_1 und β_2 erfüllt. In diesem Sinne wird die Frage nach der Erfüllbarkeit auf die Frage nach der Erfüllbarkeit von „einfacheren" Formeln zurückgeführt.

Wenn dieses Verfahren so gestaltet werden kann, dass es nach endlich vielen Schritten abbricht, dann liefert es also eine Entscheidung darüber, ob $\neg\alpha$ erfüllbar und damit zusammenhängend, ob α eine Tautologie ist.

Beispiel: (informell)
Wir wollen feststellen, ob die Formel $\alpha \rightarrow ((\alpha \rightarrow \beta) \rightarrow \beta)$ eine Tautologie ist. $\alpha \rightarrow ((\alpha \rightarrow \beta) \rightarrow \beta)$ ist genau dann keine Tautologie, wenn es eine Belegung I gibt, unter der $\neg(\alpha \rightarrow ((\alpha \rightarrow \beta) \rightarrow \beta))$ den Wert T erhält Es gilt

wenn		I erfüllt $\neg(\alpha \rightarrow ((\alpha \rightarrow \beta) \rightarrow \beta))$	
so		I erfüllt α (1)	
und		I erfüllt $\neg((\alpha \rightarrow \beta) \rightarrow \beta)$ (2)	
nach (2) : so		I erfüllt $\alpha \rightarrow \beta$ (3)	
und		I erfüllt $\neg\beta$ (4)	
nach (3) : so	I erfüllt $\neg\alpha$ (5)	oder	I erfüllt β (6)
	Widerspruch wegen (1) und (5)		Widerspruch wegen (4) und (6)

Lassen wir in diesem Beispiel die ergänzenden Bemerkungen weg, so erhalten wir den folgenden, mit aussagenlogischen Formeln markierten Binärbaum, den wir auch als Tableau bezeichnen:

$$\begin{array}{ccc} & \neg(\alpha \rightarrow ((\alpha \rightarrow \beta) \rightarrow \beta)) & \\ & | & \\ & \alpha & \\ & | & \\ & \neg((\alpha \rightarrow \beta) \rightarrow \beta) & \\ & | & \\ & \alpha \rightarrow \beta & \\ & | & \\ & \neg\beta & \\ / & & \backslash \\ \neg\alpha & & \beta \end{array}$$

Informell beschrieben sind Tableaus endliche, mit aussagenlogischen Formeln markierte Binärbäume, wobei diese Markierung noch näher anzugebenden Markierungsprinzipien genügen muss. Die folgende induktive Definition macht dies explizit.

Definition 1.39 (Tableau)
Sei $\{\alpha_0, \ldots, \alpha_n\}$, $n \geq 0$, *eine endliche Menge aussagenlogischer Formeln.*

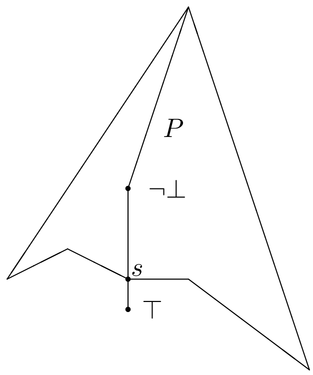
1. *Der mit* $\alpha_0, \ldots, \alpha_n$ *markierte Binärbaum*

$$\begin{array}{c} \alpha_0 \\ | \\ \alpha_1 \\ | \\ \vdots \\ | \\ \alpha_n \end{array}$$

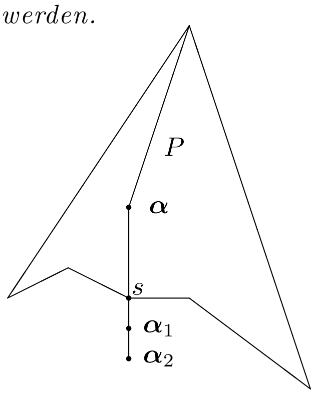
   ist ein Tableau (das Starttableau*) für* $\{\alpha_0, \ldots, \alpha_n\}$.

2. *Ist τ ein Tableau für $\{\alpha_0, \ldots, \alpha_n\}$ und P ein Zweig*[1] *in τ mit Blatt s, ist ferner ein Knoten in P mit einem Nichtliteral N markiert, dann ist der in Abhängigkeit von N nachfolgend definierte, markierte Binärbaum τ' ebenfalls ein Tableau für $\{\alpha_0, \ldots, \alpha_n\}$:*

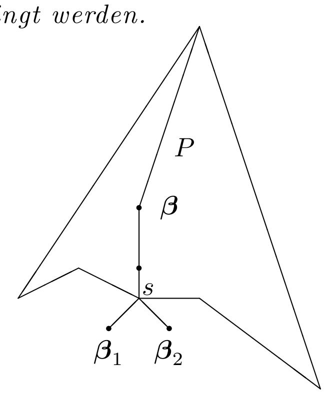
   (a) $N = \neg\bot \quad (\neg\top$ *resp.* $\neg\neg\beta)$:
   τ' resultiert aus τ, indem an das Blatt s ein mit $\top$ ($\bot$ resp. β) markierter Knoten als Sohn angehängt wird.



   (b) *N ist eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$:*
   τ' resultiert aus τ dadurch, dass das Blatt s von P einen Sohn und einen Enkel erhält, die mit $\boldsymbol{\alpha}_1$ bzw. $\boldsymbol{\alpha}_2$ markiert werden.



   (c) *N ist eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$:*
   τ' wird aus τ dadurch erhalten, dass an das Blatt s von P zwei Söhne, die mit $\boldsymbol{\beta}_1$ bzw. $\boldsymbol{\beta}_2$ markiert werden, angehängt werden.



3. *Nur markierte Binärbäume, die nach 1. und 2. erhalten werden, sind Tableaus für $\{\alpha_0, \ldots, \alpha_n\}$.*

Wir führen für das in Punkt 2. definierte Tableau τ' die folgende Sprechweise ein:
Wir sagen, das Tableau τ' geht aus dem Tableau τ durch Anwendung der in der folgenden Tabelle angegebenen Tableau-Regel hervor:

[1] Unter einem Zweig verstehen wir immer einen Wurzelpfad, d.h. einen Pfad von der Wurzel zu einem Blatt.

$$\begin{cases} \dfrac{\neg\bot}{\top} & \text{falls } N = \neg\bot \\[2ex] \dfrac{\neg\top}{\bot} & \text{falls } N = \neg\top \\[2ex] \dfrac{\neg\neg\beta}{\beta} & \text{falls } N = \neg\neg\beta \\[2ex] \dfrac{\boldsymbol{\alpha}}{\begin{matrix}\boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2\end{matrix}} & \text{falls } N \text{ eine } \alpha - Formel \\[3ex] \dfrac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2} & \text{falls } N \text{ eine } \beta - Formel \end{cases}$$

Definition 1.40 *Ein* Zweig P *eines Tableaus heißt* (*atomar*) geschlossen *genau dann, wenn*

1. *es Knoten* s, t *in* P *sowie eine* (*atomare*) *aussagenlogische Formel* α *gibt, so dass* s *mit* α *markiert ist und* t *mit* $\neg\alpha$, *oder*
2. *ein mit* $\bot$ *markierter Knoten in* P *existiert.*

Ein Tableau *heißt* (*atomar*) geschlossen, *wenn jeder Zweig des Tableaus* (*atomar*) *geschlossen ist.*

Definition 1.41 *Ein* Tableau-Beweis *für eine aussagenlogische Formel* α *ist ein geschlossenes Tableau für* $\{\neg\alpha\}$.
α *heißt* Theorem (*oder auch beweisbare Formel, notiert durch* $\vdash_{AL-T} \alpha$) *des Tableau-Kalküls genau dann, wenn* α *einen Tableau-Beweis hat.*

Bemerkung 1.42 *Ein Tableau-Beweis für eine aussagenlogische Formel* α *liegt somit genau dann vor, wenn eine Folge*

$$\tau_0, \ldots, \tau_n$$

von Tableaus gegeben ist, in der τ_0 *das Starttableau für* $\{\neg\alpha\}$ *ist,* τ_n *ein geschlossenes Tableau ist und für jedes* i *mit* $0 \leq i < n$ *das Tableau* τ_{i+1} *aus* τ_i *durch Anwendung einer Tableau-Regel hervorgeht.*

Unsere Definition von Tableaus hat die Möglichkeit offengelassen, dass eine Tableauregel in einem Zweig (bzw. in Verlängerungen davon) mehrfach auf dieselbe Formel angewandt wird. Dies scheint überflüssig zu sein und ist es in der Tat auch. In der praktischen Anwendung und insbesondere in Implementierungen wird man eine solche mehrfache Bearbeitung ausschließen. Tableaus, die in diesem Sinne ohne Mehrfachbearbeitung in einem noch zu präzisierenden Sinne systematisch aufgebaut werden, werden als strikte Tableaus bezeichnet. Im Falle der Aussagenlogik führt die Verwendung strikter Tableaus dazu, dass entweder nach endlich vielen Schritten ein geschlossenes Tableau erhalten worden ist, oder aber keine weitere Regelanwendung ohne Veletzung der Striktheitsbedingung mehr möglich ist. Die Verwendung nichtstrikter Tableaus hat hier ihren Grund einzig und allein darin, dass wir mit ihnen den Vollständigkeitsbeweis in einer Weise führen können, der den ungleich komplizierteren Vollständigkeitsbeweis der Prädikatenlogik in einfacher Weise schon vorbereitet.

Korrektheit des Tableau-Kalküls

Ein Tableau-Beweis stellt nun ein rein formales Verfahren dar, mit dem aussagenlogische Formeln bewiesen werden können. Ziel war es ja, mit einem solchen Beweisverfahren genau die Tautologien der Aussagenlogik zu erfassen. Wir verlangen daher insbesondere, dass mit unserem formalen Beweisverfahren keine falschen Formeln bewiesen werden können. Ein formales Beweissystem der Aussagenlogik heißt *korrekt* (engl. sound), falls nur Tautologien bewiesen werden können. Wir wollen im Folgenden die Korrektheit des Tableau-Kalküls zeigen.

Notation 1.43 *Für einen Zweig P eines Binärbaumes sei $Fm(P)$ die Menge der Formeln, mit denen Knoten des Zweiges P markiert sind.*

Definition 1.44 *Ein* Zweig *P eines Tableaus heißt* erfüllbar *genau dann, wenn $Fm(P)$ erfüllbar ist. Ein* Tableau *heißt* erfüllbar *genau dann, wenn es einen erfüllbaren Zweig in dem Tableau gibt.*

Lemma 1.45 *Sei τ ein erfüllbares Tableau. τ' sei das Ergebnis einer einmaligen Anwendung einer Tableau-Regel auf τ. τ' ist dann ebenfalls erfüllbar.*

Beweis. τ' ist durch die Anwendung einer Tableau-Regel für eine Formel N eines Zweiges P im Tableaus τ hervorgegangen.

Fall 1: P ist kein erfüllbarer Zweig des Tableaus τ.

Da τ erfüllbar ist, gibt es daher einen erfüllbaren Zweig $P' \neq P$ in τ. P' ist aber ebenfalls ein Zweig in τ', da nur der Zweig P erweitert wurde. Also ist τ' erfüllbar.

Fall 2: P ist ein erfüllbarer Zweig des Tableaus τ.

Wir müssen die verschiedenen Fälle betrachten. Exemplarisch werden wir den Fall der β-Formeln behandeln.

Sei also N eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$. P_1 und P_2 seien die Zweige in τ', die aus P durch die Anwendung der Tableau-Regel

$$\frac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2}$$

erhalten worden sind.

Nach Voraussetzung existiert ein Modell I von $Fm(P)$. Da $\boldsymbol{\beta} \in Fm(P)$, erfüllt I $\boldsymbol{\beta}$. Wegen $\boldsymbol{\beta} \equiv_{AL} \boldsymbol{\beta}_1 \vee \boldsymbol{\beta}_2$ erfüllt I dann $\boldsymbol{\beta}_1$ oder $\boldsymbol{\beta}_2$. Damit ist I Modell von $Fm(P) \cup \{\boldsymbol{\beta}_1\} = Fm(P_1)$ oder von $Fm(P) \cup \{\boldsymbol{\beta}_2\} = Fm(P_2)$. Daher ist P_1 oder P_2 ein erfüllbarer Zweig in τ'. Folglich ist τ' erfüllbar.

■

Mit Hilfe dieses Lemmas ist das nachfolgende Korrektheitstheorem leicht zu beweisen.

Theorem 1.46 (Korrektheit des Tableau-Kalküls)
Wenn eine aussagenlogische Formel einen Tableau-Beweis hat, ist sie eine Tautologie.
(In Zeichen: $\vdash_{AL-T} \alpha \ \Rightarrow \models_{AL} \alpha$)

Beweis. Hat α einen Tableau-Beweis, so gibt es eine Folge $\tau_0, \ldots, \tau_n$ von Tableaus, so dass τ_0 das Starttableau für $\{\neg\alpha\}$ ist, τ_n geschlossen ist, und τ_{i+1} aus τ_i für $0 \leq i < n$ durch Anwendung einer Tableau-Regel hervorgeht.
Annahme: α ist keine Tautologie.
Dann ist $\neg\alpha$ und somit τ_0 erfüllbar. Unter Verwendung von Lemma 1.45 folgt, dass alle τ_i, $0 \leq i \leq n$, erfüllbar sind. Insbesondere ist τ_n erfüllbar. τ_n ist aber ein geschlossenes Tableau und kann daher nicht erfüllbar sein (Übungsaufgabe). Daher folgt ein Widerspruch.
Unsere Annahme ist somit falsch. α is also eine Tautologie. ■

Hintikka-Mengen

Wir wollen nun umgekehrt zeigen, dass auch jede Tautologie im Tableau-Kalkül beweisbar ist, d.h., dass der Tableau-Kalkül auch vollständig ist. Wir wollen hierzu die nach dem finnischen Logiker und Philosophen Hintikka benannten Hintikka-Mengen verwenden.

Definition 1.47 *Eine aussagenlogische Formelmenge $\Sigma \subseteq Fm_{AL}$ heißt aussagenlogische* Hintikka-Menge *genau dann, wenn gilt:*

1. *Keine Aussagenvariable ist zugleich mit ihrer Negation in* Σ *enthalten, d.h.* $A \in \Sigma \Rightarrow \neg A \notin \Sigma$.
2. $\bot \notin \Sigma$, $\neg\top \notin \Sigma$.
3. $\neg\neg\beta \in \Sigma \Rightarrow \beta \in \Sigma$.
4. $\boldsymbol{\alpha} \in \Sigma \Rightarrow \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Sigma$ *(für* α*-Formeln)*
5. $\boldsymbol{\beta} \in \Sigma \Rightarrow \boldsymbol{\beta}_1 \in \Sigma$ *oder* $\boldsymbol{\beta}_2 \in \Sigma$ *(für* β*-Formeln)*

Hintikka-Mengen werden auch als abwärts-saturierte Mengen bezeichnet.

Die schöne Eigenschaft von aussagenlogischen Hintikka-Mengen ist in dem folgenden Lemma ausgedrückt:

Lemma 1.48 (Hintikka) *Jede aussagenlogische Hintikka-Menge ist erfüllbar.*

Beweis. Sei Σ eine aussagenlogische Hintikka-Menge. Wir definieren eine Belegung I und zeigen mittels struktureller Induktion über α- und β-Formeln, dass diese Belegung ein Modell von Σ ist.
$I(A) := \begin{cases} T & \text{falls } A \in \Sigma \\ F & \text{sonst.} \end{cases}$
E sei eine Eigenschaft von aussagenlogischen Formeln, die auf eine Formel α genau dann zutrifft, wenn gilt:

$$\alpha \in \Sigma \ \Rightarrow \ [\alpha]^I = T.$$

Induktionsbasis:
E trifft nach Definition auf alle Aussagenvariablen sowie auf $\top$, $\bot$ zu ($\bot \notin \Sigma$!).
Sei $\neg A \in \Sigma$. Dann gilt $A \notin \Sigma$, also ist $I(A) = F$, daher $[\neg A]^I = T$.
Für $\neg\top$ und $\neg\bot$ ist die Behauptung offensichtlich richtig.
Damit ist die Induktionsbasis gezeigt.

Induktionsschritt:
Exemplarisch wird der Fall der α-Formeln betrachtet. Sei also $\boldsymbol{\alpha}$ eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$, und es gelte nach Induktionsvoraussetzung, dass E auf $\boldsymbol{\alpha}_1$ und auf $\boldsymbol{\alpha}_2$ zutrifft.
Sei $\boldsymbol{\alpha} \in \Sigma$, und damit $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Sigma$, da Σ eine Hintikka-Menge ist. Nach Voraussetzung trifft E auf $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ zu, also gilt $[\boldsymbol{\alpha}_1]^I = [\boldsymbol{\alpha}_2]^I = T$, daher erhalten wir $\boldsymbol{\alpha}^I = [\boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2]^I = T$. ∎

Das Modellexistenztheorem

Eine zentrale Rolle im Beweis der Vollständigkeit des Tableau-Kalküls wird das Modellexistenztheorem spielen, das ausdrückt, dass Formelmengen, die bestimmte Eigenschaften besitzen, erfüllbar sind. Darüber hinaus werden wir noch weitere schöne Anwendungen dieses Theorems kennenlernen.
Wir wollen zunächst Vollständigkeit äquivalent charakterisieren.
Der Tableau-Kalkül ist vollständig genau dann, wenn jede aussagenlogische Tautologie α einen Tableau-Beweis hat, d.h. wenn gilt:

$$\models_{AL} \alpha \ \Rightarrow \vdash_{AL-T} \alpha.$$

Dies ist äquivalent dazu, dass gilt:

$$\text{nicht} \vdash_{AL-T} \alpha \ \Rightarrow \ \text{nicht} \models_{AL} \alpha.$$

Dies wiederum ist äquivalent zu:

$$\text{es existiert kein geschlossenes Tableau für } \{\neg\alpha\} \Rightarrow \neg\alpha \text{ ist erfüllbar.}$$

Um also die Vollständigkeit des Tableau-Kalküls zu zeigen, genügt es, zu zeigen, dass eine endliche aussagenlogische Formelmenge, die kein geschlossenes Tableau besitzt, erfüllbar ist.

Definition 1.49 *Sei* $\Sigma \subseteq Fm_{AL}$ *endlich.*
Σ *heißt* tableau-konsistent *genau dann, wenn kein geschlossenes Tableau für* Σ *existiert.*

Die Vollständigkeitsaussage ist somit äquivalent zu der Aussage:

ist $\{\beta\}$ tableau-konsistent, so ist β erfüllbar.

Wollen wir diese Aussage beweisen, so müssen wir also allein aufgrund der Tatsache, dass kein geschlossenes Tableau existiert, eine erfüllende Belegung angeben können. Nun wissen wir, dass Hintikka-Mengen ein Modell besitzen. Wir werden daher so vorgehen, dass wir für Formelmengen $Fm\,(P)$ für Zweige P des nicht geschlossenen Tableaus bestimmte Abschlusseigenschaften sicherstellen, die zu einer Hintikka-Menge führen.

Definition 1.50

1. *$\mathcal{C} \subseteq \mathcal{P}\,(Fm_{AL})$ (in Worten: eine Kollektion $\mathcal{C}$ von Formelmengen) heißt* aussagenlogische Konsistenzeigenschaft *genau dann, wenn für alle $\Sigma \in \mathcal{C}$ gilt:*
   (a) $A \in Var_{AL}$ *und* $A \in \Sigma \;\Rightarrow \neg A \notin \Sigma$.
   (b) $\perp \notin \Sigma\,,\; \neg\top \notin \Sigma$.
   (c) $\neg\neg\beta \in \Sigma \;\Rightarrow\; \Sigma \cup \{\beta\} \in \mathcal{C}$.
   (d) $\boldsymbol{\alpha} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$ *(für α-Formeln)*
   (e) $\boldsymbol{\beta} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\beta}_1\} \in \mathcal{C}$ *oder* $\Sigma \cup \{\boldsymbol{\beta}_2\} \in \mathcal{C}$ *(für β-Formeln).*

2. *Die aussagenlogische Konsistenzeigenschaft $\mathcal{C}$ heißt* teilmengenabgeschlossen *genau dann, wenn aus $\Sigma \in \mathcal{C}$ und $\Sigma_1 \subseteq \Sigma$ folgt $\Sigma_1 \in \mathcal{C}$.*

3. *$\mathcal{C}$ heißt* von endlichem Charakter *genau dann, wenn für $\Sigma \subseteq Fm_{AL}$ gilt:*
   $\Sigma \in \mathcal{C} \;\Leftrightarrow$ *jede endliche Teilmenge von Σ ist in $\mathcal{C}$.*

Lemma 1.51

1. *Jede aussagenlogische Konsistenzeigenschaft $\mathcal{C}$ kann zu einer aussagenlogischen Konsistenzeigenschaft, die teilmengenabgeschlossen ist, erweitert werden.*

2. *Jede aussagenlogische Konsistenzeigenschaft von endlichem Charakter ist teilmengenabgeschlossen.*

3. *Jede aussagenlogische Konsistenzeigenschaft, die teilmengenabgeschlossen ist, kann zu einer von endlichem Charakter erweitert werden.*

4. *$\mathcal{C}$ sei eine aussagenlogische Konsistenzeigenschaft von endlichem Charakter.*
   Sind $\Sigma_0,\ \Sigma_1,\ \Sigma_2,\ \ldots \in \mathcal{C}$ und gilt $\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots$, so gilt $\bigcup_{i\in\mathbb{N}} \Sigma_i \in \mathcal{C}$.

Beweis. Übungsaufgabe. ■

Theorem 1.52 (Aussagenlogisches Modellexistenztheorem)
$\mathcal{C}$ sei eine aussagenlogische Konsistenzeigenschaft.
Ist $\Sigma \in \mathcal{C}$, so ist Σ erfüllbar.

Beweis. Nach Lemma 1.51 (1) und (3) kann angenommen werden, dass $\mathcal{C}$ von endlichem Charakter und teilmengenabgeschlossen ist.
Fm_{AL} ist abzählbar. Es existiert daher eine Aufzählung $\alpha_0, \alpha_1, \alpha_2, \ldots$ aller aussagenlogischer Formeln. Wir definieren ausgehend von Σ eine Folge von Mengen, die alle in $\mathcal{C}$ liegen:
$\Sigma_0 := \Sigma \in \mathcal{C}$
$$\Sigma_{n+1} := \begin{cases} \Sigma_n \cup \{\alpha_n\} & \text{falls } \Sigma_n \cup \{\alpha_n\} \in \mathcal{C} \\ \Sigma_n & \text{sonst.} \end{cases}$$
Durch Induktion nach n folgt sofort $\Sigma_n \in \mathcal{C}$ für alle $n \in \mathbb{N}$ sowie $\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots$.
Nach Lemm 1.51 (4) gilt damit $H := \bigcup_{n\in\mathbb{N}} \Sigma_n \in \mathcal{C}$. Für H gilt dann:

1. H ist bezüglich $\subseteq$ maximales Element in $\mathcal{C}$:

   Angenommen, es existiert $\Sigma' \in \mathcal{C}$ mit $H \subsetneq \Sigma'$. Es gibt dann in der Aufzählung $\alpha_0, \alpha_1, \alpha_2, \ldots$ von Fm_{AL} eine Formel α_n mit $\alpha_n \in \Sigma' \setminus H$. Aus $\alpha_n \notin H$ folgt $\alpha_n \notin \Sigma_{n+1}$, also $\Sigma_n \cup \{\alpha_n\} \notin \mathcal{C}$. Es gilt aber $\Sigma_n \cup \{\alpha_n\} \subseteq \Sigma' \in \mathcal{C}$. $\mathcal{C}$ ist aber teilmengenabgeschlossen. Es gilt somit $\Sigma_n \cup \{\alpha_n\} \in \mathcal{C}$, im Widerspruch zur zuvor gefolgerten Eigenschaft $\Sigma_n \cup \{\alpha_n\} \notin \mathcal{C}$. Unsere Annahme, dass H nicht maximal ist, ist somit falsch.

2. H ist eine Hintikka-Menge:

   Exemplarisch wird der Fall einer α-Formel betrachtet.

   Sei $\boldsymbol{\alpha} \in H$. Da $\mathcal{C}$ eine aussagenlogische Konsistenzeigenschaft ist, gilt $H \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$.

   Da $H \subseteq H \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$ und da H maximal in $\mathcal{C}$ bezüglich $\subseteq$ ist, gilt $H = H \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$, also $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in H$.

Damit existiert also eine Hintikka-Menge H mit $\Sigma \subseteq H$. H ist erfüllbar, somit auch Σ. ■

Lemma 1.53 *$\mathcal{C}$ sei die Menge aller tableaukonsistenten Mengen von aussagenlogischen Formeln, d.h.*

$$\mathcal{C} := \{\Sigma \subseteq Fm_{AL} \mid \Sigma \text{ tableaukonsistent}\}.$$

Dann gilt:
$\mathcal{C}$ ist eine aussagenlogische Konsistenzeigenschaft.

Beweis. Wir zeigen exemplarisch die Eigenschaft (5) in der Definition von Konsistenzeigenschaft.
$\boldsymbol{\beta} \in \Sigma$ und $\Sigma \in \mathcal{C} \Rightarrow \Sigma \cup \{\boldsymbol{\beta}_1\} \in \mathcal{C}$ oder $\Sigma \cup \{\boldsymbol{\beta}_2\} \in \mathcal{C}$.
Wir zeigen die Kontraposition:
$\Sigma \cup \{\boldsymbol{\beta}_1\} \notin \mathcal{C}$, $\Sigma \cup \{\boldsymbol{\beta}_2\} \notin \mathcal{C}$ und $\boldsymbol{\beta} \in \Sigma \Rightarrow \Sigma \notin \mathcal{C}$.

Seien also $\Sigma \cup \{\boldsymbol{\beta}_1\}$ und $\Sigma \cup \{\boldsymbol{\beta}_2\}$ nicht tableaukonsistent. (Σ ist dabei endlich!)
Es gibt dann geschlossene Tableaus sowohl für $\Sigma \cup \{\boldsymbol{\beta}_1\}$ als auch für $\Sigma \cup \{\boldsymbol{\beta}_2\}$. Sei $\Sigma = \{\alpha_1, \ldots, \alpha_n, \boldsymbol{\beta}\}$.
Dann haben wir also geschlossene Tableaus der Form

$$\begin{array}{c} \alpha_1 \\ | \\ \vdots \\ | \\ \alpha_n \\ | \\ \boldsymbol{\beta} \\ | \\ \boldsymbol{\beta}_1 \\ \diagup\, T_1 \,\diagdown \end{array} \qquad \text{bzw.} \qquad \begin{array}{c} \alpha_1 \\ | \\ \vdots \\ | \\ \alpha_n \\ | \\ \boldsymbol{\beta} \\ | \\ \boldsymbol{\beta}_2 \\ \diagup\, T_2 \,\diagdown \end{array}$$

Betrachte das folgende Tableau für Σ :

$$\begin{array}{ccc} & \alpha_1 & \\ & | & \\ & \vdots & \\ & | & \\ & \alpha_n & \\ & | & \\ & \boldsymbol{\beta} & \\ \diagup & & \diagdown \\ \boldsymbol{\beta}_1 & & \boldsymbol{\beta}_2 \\ \diagup\, T_1 \,\diagdown & & \diagup\, T_2 \,\diagdown \end{array}$$

Dies ist offensichtlich ein geschlossenes Tableau für Σ. Folglich ist Σ nicht tableaukonsistent, d.h. $\Sigma \notin \mathcal{C}$. Die anderen Fälle werden analog gezeigt. ■

Der Beweis der Vollständigkeit des aussagenlogischen Tableau-Kalküls ist nun ganz einfach.

Theorem 1.54 (Vollständigkeit des Tableau-Kalküls)
Ist die aussagenlogische Formel α eine Tautologie, so ist α tableaubeweisbar.
$(\models_{AL} \alpha \ \Rightarrow \vdash_{AL-T} \alpha)$

Beweis. Wir zeigen die Kontraposition. Sei α nicht tableaubeweisbar. Dann gibt es kein geschlossenes Tableau für $\{\neg\alpha\}$. Also ist $\{\neg\alpha\}$ tableaukonsistent. Nach dem aussagenlogischen Modellexistenztheorem ist $\{\neg\alpha\}$ erfüllbar, damit ist α keine Tautologie. ■

Der folgende Kompaktheitssatz lässt sich sehr einfach aus dem Vollständigkeitssatz folgern. Wir wollen hier einen ebenso einfachen Beweis mit dem Modellexistenztheorem führen.

Theorem 1.55 (Kompaktheits- oder Endlichkeitssatz)
Sei $\Sigma \subseteq Fm_{AL}$.
Dann gilt:
Σ ist erfüllbar genau dann, wenn jede endliche Teilmenge von Σ erfüllbar ist.

Beweis.
„$\Rightarrow$": Ist Σ erfüllbar, so ist offensichtlich jede endliche Teilmenge erfüllbar.

„$\Leftarrow$":
Für die umgekehrte Beweisrichtung betrachten wir die Kollektion

$$\mathcal{C} := \{\Sigma' \subseteq Fm_{AL} \mid \text{ jede endliche Teilmenge von } \Sigma' \text{ erfüllbar}\}\,.$$

Nach Voraussetzung gilt $\Sigma \in \mathcal{C}$. Wir zeigen: $\mathcal{C}$ ist aussagenlogische Konsistenzeigenschaft. Wir betrachten hierzu exemplarisch den Fall einer α-Formel.

Sei $\Sigma' \in \mathcal{C}$ und $\boldsymbol{\alpha} \in \Sigma'$. Sei weiter Σ_1 eine beliebige endliche Teilmenge von $\Sigma' \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$.

Fall 1: $\Sigma_1 \subseteq \Sigma'$.

Dann gilt wegen $\Sigma' \in \mathcal{C}$, dass jede endliche Teilmenge von Σ', insbesondere also Σ_1, erfüllbar ist.

Fall 2: Seien $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Sigma_1$.

Sei $\Sigma_2 := \Sigma_1 \setminus \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$. Dann ist $\Sigma_2 \cup \{\boldsymbol{\alpha}\}$ eine endliche Teilmenge von Σ', also erfüllbar. Wegen $\boldsymbol{\alpha} \equiv_{AL} \boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2$ ist dann auch $\Sigma_2 \cup \{\boldsymbol{\alpha}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ und damit insbesondere $\Sigma_1 \subseteq \Sigma_2 \cup \{\boldsymbol{\alpha}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ erfüllbar.

Fall 3: Nur $\boldsymbol{\alpha}_1$ oder $\boldsymbol{\alpha}_2 \in \Sigma_1$.

Dieser Fall wird analog zum Fall 2 behandelt.

In allen Fällen folgt somit, dass Σ_1 erfüllbar ist. Da Σ_1 eine beliebige endliche Teilmenge von $\Sigma' \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ ist, folgt somit $\Sigma' \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$. Dies war in diesem Fall zu zeigen.
Damit folgt, dass $\mathcal{C}$ eine aussagenlogische Konsistenzeigenschaft ist. Aufgrund des Modellexistenztheorems ist $\Sigma \in \mathcal{C}$ somit erfüllbar. ■

Wir wollen uns eine einfache Anwendung des Kompaktheitssatzes ansehen.

Beispiel 1.56 *Nach dem Vierfarbentheorem von Appel und Haken (1977) ist jeder endliche, ungerichtete planare Graph mit 4 Farben färbbar, d.h. für jeden endlichen, ungerichteten planaren Graphen $\mathcal{G} = (V, E)$ existiert eine Partition der Knotenmenge V in 4 Mengen R, B, G, Y (R für rot, B für blau, G für grün und Y für gelb), so dass keine zwei benachbarten, d.h. durch eine Kante verbundenen Knoten in derselben Menge liegen.*

Wir zeigen unter Verwendung des Kompaktheitssatzes, dass diese Eigenschaft auch für ungerichtete planare Graphen mit abzählbar unendlich vielen Knoten gilt.

Sei $\mathcal{G} = (V, E)$ ein ungerichteter planarer Graph mit abzählbar unendlicher Knotenmenge $V = \{c_i \mid i \in \mathbb{N}\}$.
Unsere Variablenmenge Var_{AL} enthalte die folgenden aussagenlogischen Variablen:

$$A_{i,C} \text{ für } i \in \mathbb{N} \text{ und } C \in \{R, B, G, Y\}\,.$$

Die informelle Interpretation für die Variable $A_{i,C}$ ist: der Knoten c_i ist mit der Farbe C gefärbt. Betrachte die folgenden aussagenlogischen Formelmengen:

1. $\Sigma_1 := \{A_{i,R} \vee A_{i,B} \vee A_{i,G} \vee A_{i,Y} \mid i \in \mathbb{N}\}$
   Informell drückt diese Formelmenge aus:
   „Jeder Knoten ist mit einer der Farben rot, blau, grün oder gelb gefärbt".
2. $\Sigma_2 := \{\neg (A_{i,C} \wedge A_{i,C'}) \mid i \in \mathbb{N} \text{ und } C, C' \in \{R, B, G, Y\}\,,\ C \neq C'\}$
   Informell:
   „Jeder Knoten ist höchstens mit einer Farbe gefärbt".
3. $\Sigma_3 := \{A_{i,C} \to \neg A_{j,C} \mid i, j \in \mathbb{N},\ \{c_i, c_j\} \in E\}$
   Informell:
   „Benachbarte Knoten erhalten nicht die gleiche Farbe"

Es sei $\Sigma := \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$. Ist die Formelmenge Σ erfüllbar, so gibt es also eine Belegung I, die Σ erfüllt. Ausgehend von dieser Belegung I können wir sofort eine 4-Färbung des Graphen $\mathcal{G}$ angeben, d.h. eine Abbildung $f : V \to \{R, B, G, Y\}$, die die Eigenschaft hat, dass keine zwei benachbarten Knoten dieselbe Farbe erhalten:

$$I\,(A_{i,C}) = T \text{ für } C \in \{R, B, G, Y\} \quad \Rightarrow \quad f\,(c_i) := C.$$

Dass f eine 4-Färbung mit den gewünschten Eigenschaften ist, ist leicht zu verifizieren.
Sei nun Σ' eine beliebige endliche Teilmenge von Σ.
Sei

$$E' := \{\{c_i, c_j\} \mid A_{i,C} \to \neg A_{j,C} \in \Sigma' \cap \Sigma_3 \text{ für ein } C \in \{R, B, G, Y\}\}$$

und

$$V' := \bigcup_{e \in E'} e.$$

($\{c_i, c_j\}$ bezeichnet hierbei eine ungerichtete Kante.)
$\mathcal{G}' := (V', E')$ ist dann der kleinste ungerichtete Teilgraph von $\mathcal{G}$, der alle Kanten enthält, die durch die in Σ' vorkommenden Formeln beschrieben sind. Da (V', E') ebenfalls planar und auch endlich ist, existiert nach dem Vierfarbentheorem eine 4-Färbung $f' : V' \to \{R, B, G, Y\}$. Definiere eine Belegung I' durch $(i \in \mathbb{N},\ C \in \{R, B, G, Y\})$:

$$I'\,(A_{i,C}) := \begin{cases} T & \text{falls } c_i \in V' \text{ und } f'\,(c_i) = C \text{ oder falls } c_i \notin V' \text{ und } C = R \\ F & \text{sonst.} \end{cases}$$

(Die Bedingung $C = R$ ist dabei willkürlich gewählt.)
Aufgrund der Existenz der 4-Färbung ist leicht zu verifizieren, dass diese Belegung alle Formeln in Σ' erfüllt, in denen nur Variablen $A_{i,C}$ mit $c_i \in V'$ vorkommen. Falls $c_i \notin V'$, dann kommt $A_{i,C}$ nicht in einer Formel von $\Sigma' \cap \Sigma_3$ vor. Nach Definition von I erhält nur $A_{i,R}$ den Wert T, so dass also verbleibende Formeln, in denen $A_{i,C}$ vorkommt (also solche aus $\Sigma' \cap \Sigma_1$ oder aus $\Sigma' \cap \Sigma_2$) ebenfalls den Wert T erhalten.
Damit ist also Σ' erfüllbar. Da Σ' beliebig gewählt war, ist also jede endliche Teilmenge von Σ erfüllbar. Nach dem Kompaktheitssatz ist somit Σ selbst erfüllbar. Nach dem, was wir oben ausgeführt haben, folgt dann, dass der Graph $\mathcal{G}$ 4-färbbar ist.

Semantische Folgerung und formale Ableitbarkeit

Mit dem Vollständigkeitssatz haben wir die gewünschte Beziehung zwischen Allgemeingültigkeit und formaler Beweisbarkeit im Tableau-Kalkül hergestellt:

$$\models_{AL} \alpha \text{ genau dann, wenn } \vdash_{AL-T} \alpha.$$

Wir sind jedoch an Folgerungen aus vorhandenem Wissen bzw. aus Theorien interessiert. Demgemäß werden wir versuchen, das semantische Folgern ebenfalls formal zu erfassen und eine entsprechende Beziehung zwischen semantischem Folgern und formalem Ableiten aufzustellen.
Zunächst wollen wir jedoch eine wichtige Eigenschaft des Folgerungsbegriffs zeigen.

Theorem 1.57 (Endlichkeitseigenschaft des Folgerungsbegriffs)
$\Sigma \models_{AL} \alpha$ genau dann, wenn eine endliche Teilmenge $\Sigma' \subseteq \Sigma$ existiert mit $\Sigma' \models_{AL} \alpha$.

Beweis. „$\Leftarrow$": Offensichtlich.
„$\Rightarrow$":
Für beliebige aussagenlogische Formeln β und Formelmengen Σ' gilt

$$\Sigma' \models_{AL} \beta \;\Leftrightarrow\; \Sigma' \cup \{\neg\beta\} \text{ nicht erfüllbar.} \tag{1.2}$$

Es gelte $\Sigma \models_{AL} \alpha$. Dann ist nach (1.2) $\Sigma \cup \{\neg\alpha\}$ nicht erfüllbar. Nach dem Kompaktheitssatz gibt es somit eine endliche Teilmenge Σ_1 von $\Sigma \cup \{\neg\alpha\}$, die nicht erfüllbar ist. Mit Σ_1 ist auch $\Sigma_1 \cup \{\neg\alpha\}$ nicht erfüllbar. Wiederum nach (1.2) gilt dann $\Sigma_1 \models_{AL} \alpha$. Damit ist die Behauptung des Theorems gezeigt. ■

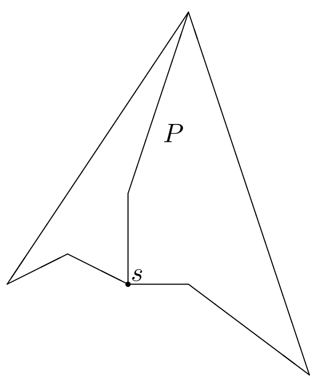
Um nun den Zusammenhang zwischen semantischem Folgern und formalen Ableiten herzustellen, verallgemeinern wir den Begriff „Tableau-Beweis" zu „Σ-Tableau-Beweis" (oder auch „Tableau-Beweis mit Prämissenmenge Σ")

Definition 1.58 *Sei $\Sigma \subseteq Fm_{AL}$, $\alpha \in Fm_{AL}$.*

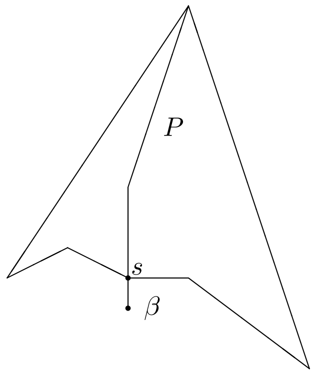
1. Σ-Tableau *für* $\{\alpha_0, \ldots, \alpha_n\}$:

   Wir übernehmen die Bedingungen 1. und 2. aus der Definition 1.39 eines Tableaus für $\{\alpha_0, \ldots, \alpha_n\}$ (wobei das Resultat dann jeweils ein Σ-Tableau ist) erweitert um den folgenden zusätzlichen Fall:

   Ist τ ein Σ-Tableau



   und P ein Zweig in τ mit Blatt s, dann ist für ein $\beta \in \Sigma$ auch der folgende markierte Binärbaum ein Σ-Tableau:



(Die entsprechende Σ-Tableau-Regel wird notiert durch: $\frac{}{\beta}$*)*

2. *Ein* Zweig P *eines Σ-Tableaus heißt* Σ-erfüllbar *genau dann, wenn ein Modell für* $\Sigma \cup Fm(P)$ *existiert.*

   Ein Σ-Tableau *heißt* Σ-erfüllbar *genau dann, wenn im Tableau ein Zweig, der Σ-erfüllbar ist, existiert.*

   Ein Σ-Tableau-Beweis für α *ist ein geschlossenes Σ-Tableau für* $\{\neg\alpha\}$ *(notiert als :* $\Sigma \vdash_{AL-T} \alpha$*).*

3. Σ *heißt* α-tableaukonsistent *genau dann, wenn nicht gilt* $\Sigma \vdash_{AL-T} \alpha$.

   Ist Σ *nicht* α*-tableaukonsistent, so heißt* Σ α-tableauinkonsistent.

Theorem 1.59 (Strenge Korrektheit und Vollständigkeit des Tableau-Kalküls)
Sei $\Sigma \subseteq Fm_{AL}$, $\alpha \in Fm_{AL}$. *Dann gilt:*
$\Sigma \models_{AL} \alpha$ *genau dann, wenn* $\Sigma \vdash_{AL-T} \alpha$.

Beweis. Skizze.
Die Vorgehensweise ist im Wesentlichen analog zum Korrektheits- und Vollständigkeitsbeweis des Tableau-Kalküls.

Strenge Korrektheit: $\Sigma \vdash_{AL-T} \alpha \ \Rightarrow\ \Sigma \models_{AL} \alpha$.

Man zeigt:

1. Ist τ ein Σ-erfüllbares Σ-Tableau und resultiert τ' aus τ mittels einer Tableau-Regel für Σ-Tableaus, so ist τ' ein Σ-erfüllbares Σ-Tableau.
2. Existiert ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$, so ist das Σ-Starttableau für $\{\neg\alpha\}$ nicht erfüllbar (unter Verwendung von (1))

Aus (2) folgt: $\Sigma \cup \{\neg\alpha\}$ ist nicht erfüllbar, also gilt $\Sigma \models_{AL} \alpha$.

Strenge Vollständigkeit: $\Sigma \models_{AL} \alpha \ \Rightarrow\ \Sigma \vdash_{AL-T} \alpha$.

Man zeigt:

1. Für jede Formel $\alpha \in Fm_{AL}$ ist
$$\mathcal{C}_\alpha := \{\Sigma \subseteq Fm_{AL} \mid \Sigma\ \alpha\text{-tableaukonsistent}\}$$
eine aussagenlogische Konsistenzeigenschaft.
2. Ist Σ α-tableaukonsistent, so auch $\Sigma \cup \{\neg\alpha\}$.

Damit gilt dann weiter: Gilt nicht $\Sigma \vdash_{AL-T} \alpha$, so ist Σ α-tableaukonsistent. Damit ist auch $\Sigma \cup \{\neg\alpha\}$ α-tableaukonsistent. Es ist dann $\Sigma \cup \{\neg\alpha\} \in \mathcal{C}_\alpha$, und damit erfüllbar nach dem Modellexistenztheorem.. Es folgt, dass nicht gilt $\Sigma \models_{AL} \alpha$. Insgesamt ist damit gezeigt: $\Sigma \models_{AL} \alpha \ \Rightarrow\ \Sigma \vdash_{AL-T} \alpha$.

■

Fassen wir das Ergebnis von Theorem 1.59 zusammen:
Mit dem Tableau-Kalkül haben wir eine korrekte und vollständige Formalisierung des inhaltlichen Schließens in der Aussagenlogik erzielt.

1.2.2 Der nichtklausale Resolutionskalkül

Wir wollen als nächsten Widerlegungskalkül den Resolutionskalkül einführen. Resolutionskalküle sind von Interesse, da sie einfach zu implementieren und auch sehr effizient sind. Darüber hinaus bilden sie die Grundlage für den Bereich der Logischen Programmierung.
Wir werden hier zunächst den *nichtklausalen* Resolutionskalkül einführen. Dieser heißt nichtklausal, weil die Formeln nicht in Klauselform vorliegen müssen.
Im Zusammenhang mit dem Beweis von Theorem 1.36 (Theorem über konjunktive und disjunktive Normalform) hatten wir die folgenden Reduktionsregeln für Klauselmengen benutzt:

$$\frac{\neg\top}{\bot} \qquad \frac{\neg\bot}{\top} \qquad \frac{\neg\neg\beta}{\beta} \qquad \frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2} \qquad \frac{\boldsymbol{\beta}}{\begin{array}{c}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}}$$

Wir wollen diese Reduktionsregeln im Zusammenhang mit dem nichtklausalen Resolutionskalkül als *Resolutionsexpansionsregeln* bezeichnen. Da im Folgenden häufig verallgemeinerte Konjunktionen und Disjunktionen benötigt werden, sollen zunächst einige einfache Eigenschaften angegeben werden, die leicht unter Verwendung von Folgerung 1.34 gezeigt werden können:

- Für eine α-Formel gilt:

  $$[\alpha_1, \ldots, \alpha_i, \boldsymbol{\alpha}, \alpha_{i+1}, \ldots, \alpha_n] \equiv_{AL} [\alpha_1, \ldots, \alpha_i, \boldsymbol{\alpha}_1, \alpha_{i+1}, \ldots, \alpha_n] \wedge [\alpha_1, \ldots, \alpha_i, \boldsymbol{\alpha}_2, \alpha_{i+1}, \ldots, \alpha_n]\,.$$

  Für eine β-Formel gilt

  $$[\alpha_1, \ldots, \alpha_i, \boldsymbol{\beta}, \alpha_{i+1}, \ldots, \alpha_n] \equiv_{AL} [\alpha_1, \ldots, \alpha_i, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha_{i+1}, \ldots, \alpha_n]$$

- Streicht man in $\alpha := [\alpha_1, \ldots, \alpha_n]$ ($\alpha := \langle\alpha_1, \ldots, \alpha_n\rangle$) außer dem ersten Vorkommen jedes weitere Vorkommen eines Disjunktionsgliedes (Konjunktionsgliedes), so ist die resultierende verallgemeinerte Disjunktion (Konjunktion) α' zu α semantisch äquivalent. Beispielsweise ist also $[A, B, A]$ semantisch äquivalent zu $[A, B]$. Diese so aus $[\alpha_1, \ldots, \alpha_n]$ erhaltene reduzierte verallgemeinerte Disjunktion werden wir mit $[\alpha_1, \ldots, \alpha_n]^{red}$ bezeichnen. Entsprechend ist $\langle\alpha_1, \ldots, \alpha_n\rangle^{red}$ die aus $\langle\alpha_1, \ldots, \alpha_n\rangle$ erhaltene reduzierte verallgemeinerte Konjunktion.

Vereinbarung:
Im Zusammenhang mit Resolutionskalkülen werden wir für reduzierte verallgemeinerte Disjunktionen die Schreibweise $[\alpha_1, \ldots, \alpha_n]$ benutzen, und nur wenn wir den Unterschied zu $[\alpha_1, \ldots, \alpha_n]$ betonen wollen, werden wir $[\alpha_1, \ldots, \alpha_n]^{red}$ verwenden. Weiter führen wir die folgenden *Schreibweisen* ein:

$$[\alpha_1, \ldots, \alpha_n] - [\beta] := \bigvee_{1 \le i \le n,\ \alpha_i \ne \beta} \alpha_i$$

(d.h. wir streichen in der (reduzierten) verallgemeinerten Disjunktion jedes Vorkommen von β).

$$[\alpha_1, \ldots, \alpha_n] \sqcup [\beta_1, \ldots, \beta_m] := [\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m]\,.$$

Definition 1.60 *Seien $\alpha = [\alpha_1, \ldots, \alpha_n]$ und $\beta = [\beta_1, \ldots, \beta_m]$ verallgemeinerte Disjunktionen von aussagenlogischen Formeln.*
Gibt es ein (atomares) $\gamma \in Fm_{AL}$, so dass $\gamma = \alpha_i$ und $\neg\gamma = \beta_j$ oder $\neg\gamma = \alpha_i$ und $\gamma = \beta_j$ für ein i sowie ein j mit $1 \le i \le n$ und $1 \le j \le m$, so heißt

$$([\alpha_1, \ldots, \alpha_n] - [\alpha_i]) \sqcup ([\beta_1, \ldots, \beta_m] - [\beta_j])$$

(atomare) Resolvente *von α und β (notiert als: $R(\alpha, \beta)$).*
Falls $\bot \in \{\alpha_1, \ldots, \alpha_n\}$, so heißt $[\alpha_1, \ldots, \alpha_n] - [\bot]$ triviale Resolvente *von α (notiert als $R(\alpha)$).*

Beispiel 1.61

1. *Resolvente von*
$$[A, A \wedge B, A \rightarrow B]$$
und
$$[B, A \vee B, \neg (A \wedge B), A]$$
ist
$$[A, A \rightarrow B, B, A \vee B] .$$

2. *Triviale Resolvente von*
$$[A, \top, \bot, B \rightarrow A]$$
ist
$$[A, \top, B \rightarrow A] .$$

Um den Resolutionskalkül übersichtlich formulieren zu können, führen wir weiter die folgende **Bezeichnung** ein:
Sei $\beta = [\alpha_1, \ldots, \alpha_i, \alpha, \alpha_{i+1}, \ldots, \alpha_n]$ eine verallgemeinerte Disjunktion. In Abhängigkeit von den Resolutionsexpansionsregeln definieren wir verallgemeinerte Disjunktionen β' bzw. β_1 und β_2 folgendermaßen:

- Fälle $\frac{\neg\top}{\bot}$, $\frac{\neg\bot}{\top}$, $\frac{\neg\neg\gamma}{\gamma}$: kurz notiert als $\frac{\alpha}{\alpha'}$:
$$\beta' := [\alpha_1, \ldots, \alpha_i, \alpha', \alpha_{i+1}, \ldots, \alpha_n]$$

- Fall $\dfrac{\boldsymbol{\beta}}{\begin{array}{c}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}}$: α sei eine β-Formel:
$$\beta' := [\alpha_1, \ldots, \alpha_i, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha_{i+1}, \ldots, \alpha_n]$$

- Fall $\dfrac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2}$: α sei eine α-Formel:
$$\beta_1 := [\alpha_1, \ldots, \alpha_i, \boldsymbol{\alpha}_1, \alpha_{i+1}, \ldots, \alpha_n]$$
und
$$\beta_2 := [\alpha_1, \ldots, \alpha_i, \boldsymbol{\alpha}_2, \alpha_{i+1}, \ldots, \alpha_n] .$$

β' bzw. β_1 und β_2 werden dann als das Resultat der Resolutionsexpansionsregel für β bezeichnet.

Definition 1.62 (Resolutionsexpansion)
Sei $\Sigma = \{\alpha_1, \ldots, \alpha_n\} \subseteq Fm_{AL}$ *endlich.*

1. $[\alpha_1], \ldots, [\alpha_n]$ *ist eine Resolutionsexpansion (die Startexpansion) für* Σ

2. *Ist* $\beta_1, \ldots, \beta_m$ *eine Resolutionsexpansion für* Σ*, und ist* β *bzw. sind* β_1' *und* β_2' *das Resultat einer Resolutionsexpansionsregel für ein* β_i *mit* $1 \leq i \leq m$*, so sind*
$$\beta_1, \ldots, \beta_m, \beta$$
bzw.
$$\beta_1, \ldots, \beta_m, \beta_1', \beta_2'$$
ebenfalls Resolutionsexpansionen für Σ*.*

3. *Ist* $\beta_1, \ldots, \beta_m$ *Resolutionsexpansion für* Σ, *ist* $\beta = R(\beta_i)$ *für ein* β_i *mit* $1 \leq i \leq m$ *oder ist* $\beta = R(\beta_i, \beta_j)$ *für ein* β_i *und ein* β_j *mit* $1 \leq i, j \leq m$, *dann ist*

$$\beta_1, \ldots, \beta_m, \beta$$

Resolutionsexpansion für Σ.

Wir wollen als Nächstes den Beweisbarkeitsbegriff für den Resolutionskalkül definieren.

Definition 1.63

1. *Eine* Resolutionsexpansion $\beta_1, \ldots, \beta_m$ *heißt* geschlossen *genau dann, wenn für ein* i, $1 \leq i \leq m$, $\beta_i = [\,]$.

2. *Eine geschlossene Resolutionsexpansion für* $\{\neg\alpha\}$ *heißt* Resolutionsbeweis *für* α (*notiert als* $\vdash_{AL-R} \alpha$).

Beispiel 1.64 *Resolutionsbeweis für*

$$(A \to B \vee C) \to ((A \to B) \vee (A \to C))$$

(*Wir notieren die Expansion hier vertikal.*)

1	$[\neg((A \to B \vee C) \to ((A \to B) \vee (A \to C)))]$	
2	$[A \to B \vee C]$	*aus 1*
3	$[\neg((A \to B) \vee (A \to C))]$	
4	$[\neg A, B \vee C]$	*aus 2*
5	$[\neg A, B, C]$	*aus 4*
6	$[\neg(A \to B)]$	*aus 3*
7	$[\neg(A \to C)]$	*aus 3*
8	$[A]$	*aus 6*
9	$[\neg B]$	*aus 6*
10	$[A]$	*aus 7*
11	$[\neg C]$	*aus 7*
12	$[B, C]$	$R(5, 8)$
13	$[C]$	$R(9, 12)$
14	$[\,]$	$R(11, 13)$

Korrektheit und Vollständigkeit des nichtklausalen Resolutionskalküls

Wie zuvor für den Tableau-Kalkül werden wir für den Resolutionskalkül die Korrektheit und die Vollständigkeit zeigen. Die Beweismethode entspricht dabei weitestgehend derjenigen, die wir im Falle des Tableau-Kalküls vorgeführt haben, so dass wir hier nur noch die wesentlichen Schritte angeben und die erforderlichen Begriffsbildungen vornehmen und ansonsten auf die Übungsaufgaben verweisen.

Definition 1.65 *Eine* Resolutionsexpansion $\beta_1, \ldots, \beta_m$ *heißt* erfüllbar *genau dann, wenn* $\{\beta_1, \ldots, \beta_m\}$ *erfüllbar ist.*

Wir zeigen zunächst, dass durch die Anwendung einer Resolutionsexpansionsregel oder durch die Bildung einer Resolvente aus einer erfüllbaren Resolutionsexpansion wiederum eine erfüllbare Resolutionsexpansion resultiert. Hieraus kann dann leicht die Korrektheit des Resolutionskalküls gefolgert werden.

Lemma 1.66 *Sei* S *eine erfüllbare Resolutionsexpansion und* S' *resultiere aus* S *durch die Bildung einer Resolvente oder durch die Anwendung einer Resolutionsexpansionsregel. Dann ist* S' *ebenfalls erfüllbar.*

Beweis. Übungsaufgabe. ■

Analog der Vorgehensweise im Tableau-Kalkül kann dann gefolgert werden:

Theorem 1.67 (Korrektheit des nichtklausalen Resolutionskalküls)
Ist die aussagenlogische Formel α im Resolutionskalkül beweisbar, so ist sie eine Tautologie, d.h.

$$\vdash_{AL-R} \alpha \Rightarrow \models_{AL} \alpha.$$

Beweis. Übungsaufgabe. ■

Der Beweis der Vollständigkeit des nichtklausalen Resolutionskalküls verläuft nach dem Schema, das wir im Falle des Tableau-Kalküls schon gesehen haben. Wir definieren hierzu einen geeigneten Konsistenzbegriff und zeigen dann, dass die Kollektion der Formelmengen, die diesem Konsistenzbegriff genügen, eine aussagenlogische Konsistenzeigenschaft bildet. Hieraus kann dann leicht die Vollständigkeit gefolgert werden. Der Nachweis, dass diese Kollektion eine aussagenlogische Konsistenzeigenschaft ist, erfordert aber noch einige Vorbereitungen.

Definition 1.68 *Sei α eine aussagenlogische Formel.*

1. *Sei $\beta = [\alpha_1, \dots, \alpha_n]$ eine verallgemeinerte Disjunktion.*
   Dann heißen
   $$[\alpha_1, \dots, \alpha_n]$$
   und
   $$[\alpha, \alpha_1, \dots, \alpha_n]$$
   α-Erweiterungen von β.

2. *Sei Σ eine Menge von verallgemeinerten Disjunktionen. Geht Σ^* aus Σ dadurch hervor, dass jedes Element in Σ durch eine α-Erweiterung ersetzt wird, so heißt Σ^* α-Erweiterung von Σ.*

3. *Sei $\beta_1, \dots, \beta_m$ eine Resolutionsexpansion. Sei ferner für jedes i mit $1 \leq i \leq m$ β'_i eine α-Erweiterung von β_i. Dann heißt $\beta'_1, \dots, \beta'_m$ α-Erweiterung der Resolutionsexpansion $\beta_1, \dots, \beta_m$.*

Beispiel 1.69 *Sei $\Sigma = \{[\alpha_1, \alpha_2], [\beta_1, \beta_2, \beta_3]\}$. Dann sind*

$$\{[\alpha, \alpha_1, \alpha_2], [\alpha, \beta_1, \beta_2, \beta_3]\}$$

und

$$\{[\alpha_1, \alpha_2], [\alpha, \beta_1, \beta_2, \beta_3]\}$$

α-Erweiterungen von Σ.

Lemma 1.70 *Seien Σ_1, Σ_2 endliche Mengen von verallgemeinerten Disjunktionen. Sei ferner Σ_2 eine α-Erweiterung von Σ_1 für eine Formel $\alpha \in Fm_{AL}$.*
Ist $\beta_1, \dots, \beta_m$ eine Resolutionsexpansion für Σ_1, so gibt es eine α-Erweiterung $\beta'_1, \dots, \beta'_m$ von $\beta_1, \dots, \beta_m$, die Resolutionsexpansion für Σ_2 ist.

Beweis. Sei $n = \#\Sigma_1$.
Wir zeigen durch Induktion nach m die folgende Behauptung:
Für alle $m \geq n$:
ist $\beta_1, \dots, \beta_m$ eine Resolutionsexpansion für Σ_1, so gibt es eine α-Erweiterung $\beta'_1, \dots, \beta'_m$ von $\beta_1, \dots, \beta_m$, die Resolutionsexpansion für Σ_2 ist.

Induktionsanfang: Ohne Beschränkung der Allgemeinheit starten wir den Induktionsanfang mit $m = n$.

In diesem Fall ist die Behauptung erfüllt, da die Startexpansion für Σ_2 eine α-Erweiterung der Startexpansion für Σ_1 ist.

Induktionsvoraussetzung: Ist $\beta_1, \ldots, \beta_k,\ n \leq k \leq m$, eine Resolutionsexpansion für Σ_1, so gibt es eine α-Erweiterung $\beta'_1, \ldots, \beta'_k$ von $\beta_1, \ldots, \beta_k$, die Resolutionsexpansion für Σ_2 ist.

Induktionsschritt: $(m \to m+1)$

Sei $\beta_1, \ldots, \beta_{m+1}$ eine Resolutionsexpansion für Σ_1 mit $m \geq n$. $\beta_1, \ldots, \beta_{m+1}$ ist dann entweder aus $\beta_1, \ldots, \beta_{m-1}$ aufgrund einer Resolutionsexpansionsregel

$$\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2}$$

oder aufgrund einer anderen Resolutionsexpansionsregel bzw. Resolventenbildung aus $\beta_1, \ldots, \beta_m$ erhalten worden.

Wir betrachten exemplarisch 2 der möglichen Fälle.

1. $\beta_1, \ldots, \beta_{m+1}$ ist mit einer Resolutionsexpansionsregel

$$\frac{\boldsymbol{\beta}}{\begin{array}{c}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}}$$

aus $\beta_1, \ldots, \beta_m$ erhalten worden, d.h. es existiert ein β_i, $1 \leq i \leq m$, so dass

$$\beta_i \text{ von der Form } [L_1, \boldsymbol{\beta}, L_2]$$

ist (L_1 sei die Folge der Formeln, die in β_i links der β-Formel stehen, entsprechend sei L_2 die Folge der Formeln, die rechts von $\boldsymbol{\beta}$ stehen. Analoges gilt für die nachfolgend verwendeten L's), und

$$\beta_{m+1} = [L_1, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, L_2]\ .$$

Nach Induktionsvorausetzung existiert eine α-Erweiterung $\beta'_1, \ldots, \beta'_m$ von $\beta_1, \ldots, \beta_m$, die Resolutionsexpansion für Σ_2 ist. Sei $\beta'_i = [L'_1, \boldsymbol{\beta}, L'_2]$ die entsprechende α-Erweiterung von β_i. Dann ist $\beta'_{m+1} = [L'_1, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, L'_2]$ eine α-Erweiterung von β_{m+1} und β'_{m+1} ist das Resultat der Resolutionsexpansionsregel

$$\frac{\boldsymbol{\beta}}{\begin{array}{c}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}}$$

für β'_i.

2. β_{m+1} ist die Resolvente von zwei verallgemeinerten Disjunktionen β_i und β_j mit $1 \leq i, j \leq m$, d.h. $\beta_{m+1} = R(\beta_i, \beta_j)$.

   Dann gibt es eine Formel $\gamma \in Fm_{AL}$, so dass $\beta_i = [L_1, \gamma, L_2]$, $\beta_j = [L_3, \neg\gamma, L_4]$ und $\beta_{m+1} = [L_1, L_2, L_3, L_4]$.

   Wiederum existiert nach Induktionsvorausetzung eine α-Erweiterung $\beta'_1, \ldots, \beta'_m$ von $\beta_1, \ldots, \beta_m$, die Resolutionsexpansion für Σ_2 ist. Sei $\beta'_i = [L'_1, \gamma, L'_2]$ die entsprechende α-Erweiterung von β_i und $\beta'_j = [L'_3, \neg\gamma, L'_4]$ die von β_j. Dann ist $\beta'_{m+1} = [L'_1, L'_2, L'_3, L'_4] = R\left(\beta'_i, \beta'_j\right)$ und β'_{m+1} ist α-Erweiterung von β_{m+1}.

   Damit ist $\beta'_1, \ldots, \beta'_{m+1}$ eine Resolutionsexpansion für Σ_2 und ist ferner eine α-Erweiterung von $\beta_1, \ldots, \beta_{m+1}$.

Damit ist der Induktionsschritt gezeigt. Die Behauptung ist somit bewiesen.

■

Definition 1.71 *Sei $\Sigma \subseteq Fm_{AL}$ eine endliche Formelmenge.*

1. *Σ heißt* resolutionskonsistent *genau dann, wenn keine geschlossene Resolutionsexpansion für Σ existiert.*

2. Σ *heißt* atomar resolutionskonsistent *genau dann, wenn keine geschlossene Resolutionsexpansion, in der Resolventenbildungen nur atomar vorkommen, für* Σ *existiert.*

Analog zum Vollständigkeitsbeweis des Tableau-Kalküls zeigen wir, dass die Kollektion der resolutionskonsistenten Formelmengen eine aussagenlogische Konsistenzeigenschaft bildet. Hiermit kann dann leicht die Vollständigkeit gefolgert werden.

Lemma 1.72 *Sei*

$$\mathcal{C} := \{\Sigma \subseteq Fm_{AL} \mid \Sigma \text{ endlich und (atomar) resolutionskonsistent}\}.$$

Dann ist $\mathcal{C}$ *eine aussagenlogische Konsistenzeigenschaft.*

Beweis. Wir müssen die Bedingungen für eine Konsistenzeigenschaft überprüfen. Wir betrachten nur die interessanteren Fälle, die der $\boldsymbol{\alpha}$ und β-Formeln.

α-Formel: Zu zeigen ist: $\Sigma \in \mathcal{C}$ und $\boldsymbol{\alpha} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$.

Wir zeigen die Kontraposition.

Sei also $\boldsymbol{\alpha} \in \Sigma$ und $\Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \notin \mathcal{C}$, d.h. $\Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ ist nicht resolutionskonsistent.

Sei $\Sigma = \{\alpha_1, \dots, \alpha_n, \boldsymbol{\alpha}\}$. Da $\Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ nicht resolutionskonsistent ist, gibt es für $\Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ o.B.d.A. eine geschlossene Resolutionsexpansion der Gestalt

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\alpha}], [\boldsymbol{\alpha}_1], [\boldsymbol{\alpha}_2], \beta_1, \dots, \beta_m, [\,].$$

(Die β_i's seien verallgemeinerte Disjunktionen). Offensichtlich ist dies aber auch eine Resolutionsexpansion für $\{\alpha_1, \dots, \alpha_n, \boldsymbol{\alpha}\}$ $(= \Sigma)$, da $[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\alpha}], [\boldsymbol{\alpha}_1], [\boldsymbol{\alpha}_2]$ das Resultat der Resolutionsexpansionsregel

$$\frac{\boldsymbol{\alpha}}{\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2}$$

für die Startexpansion von Σ ist. Das bedeutet, dass für Σ eine geschlossene Resolutionsexpansion existiert, Σ also nicht resolutionskonsistent ist. Damit ist dann aber $\Sigma \notin \mathcal{C}$.

β-Formel: Zu zeigen ist: $\Sigma \in \mathcal{C}$ und $\boldsymbol{\beta} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\beta}_1\} \in \mathcal{C}$ oder $\Sigma \cup \{\boldsymbol{\beta}_2\} \in \mathcal{C}$

Wir beweisen wiederum die Kontraposition.

Gelte also $\boldsymbol{\beta} \in \Sigma$ und $\Sigma \cup \{\boldsymbol{\beta}_1\}, \Sigma \cup \{\boldsymbol{\beta}_2\} \notin \mathcal{C}$. Ferner sei $\Sigma = \{\alpha_1, \dots, \alpha_n, \boldsymbol{\beta}\}$ angenommen.

Es gibt dann aufgrund der Voraussetzung geschlossene Resolutionsexpansionen für $\Sigma \cup \{\boldsymbol{\beta}_1\}$:

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_1], \beta_1, \dots, \beta_m, [\,]$$

und für $\Sigma \cup \{\boldsymbol{\beta}_2\}$:

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_2], \beta'_1, \dots, \beta'_k, [\,] \;.$$

$\Sigma \cup \{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]\}$ ist eine $\boldsymbol{\beta}_2$-Erweiterung von $\Sigma \cup \{\boldsymbol{\beta}_1\}$.

Nach Lemma 1.70 gibt es daher eine Resolutionsexpansion für $\Sigma \cup \{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]\}$, die eine $\boldsymbol{\beta}_2$-Erweiterung

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2], \gamma_1, \dots, \gamma_m, \gamma_{m+1}$$

der Resolutionsexpansion

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_1], \beta_1, \dots, \beta_m, [\,]$$

für $\Sigma \cup \{[\boldsymbol{\beta}_1]\}$ ist. Insbesondere ist damit $\gamma_{m+1} = [\,]$ oder $\gamma_{m+1} = [\boldsymbol{\beta}_2]$.

$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]$ ist offensichtlich eine Resolutionsexpansion für Σ.

Damit ist

$$[\alpha_1], \dots, [\alpha_n], [\boldsymbol{\beta}], [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2], \gamma_1, \dots, \gamma_m, \gamma_{m+1}$$

ebenfalls eine Resolutionsexpansion für Σ.

Ist $\gamma_{m+1} = [\,]$, so haben wir eine geschlossene Resolutionsexpansion für Σ. Es gilt dann $\Sigma \notin \mathcal{C}$.
Gilt $\gamma_{m+1} = [\boldsymbol{\beta}_2]$, so bildet

$$\underbrace{[\alpha_1]\,, \ldots, [\alpha_n]\,, [\boldsymbol{\beta}],}_{\text{Resolutionsexpansion}} \qquad [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2]\,, \gamma_1, \ldots, \gamma_m, \quad \underbrace{[\boldsymbol{\beta}_2]\,, \beta'_1, \ldots, \beta'_k, [\,]}_{\text{für } \Sigma \cup \{\boldsymbol{\beta}_2\}}$$

eine geschlossene Resolutionsexpansion für Σ. Also gilt auch in diesem Fall $\Sigma \notin \mathcal{C}$. Dies war für diesen Fall zu zeigen.

Damit ist der Nachweis erbracht, dass $\mathcal{C}$ eine Konsistenzeigenschaft ist. ■

Theorem 1.73 (Vollständigkeit des nichtklausalen Resolutionskalküls)
$\models_{AL} \alpha \;\Rightarrow \vdash_{AL-R} \alpha.$

Beweis. Wir zeigen die Kontraposition.
Angenommen, α hat keinen Resolutionsbeweis. Dann existiert keine geschlossene Resolutionsexpansion für $\{\neg\alpha\}$. $\{\neg\alpha\}$ ist also resolutionskonsistent und nach Lemma 1.72 daher erfüllbar. Folglich ist α keine Tautologie. ■

Bemerkung 1.74 *Schränkt man im Resolutionskalkül die Resolventenbildung so ein, dass statt beliebiger nur atomare Resolventen gebildet werden dürfen, so sind die Beweise von Theorem 1.67, Lemma 1.72 sowie Theorem 1.73 davon nicht betroffen, so dass die folgende stärkere Aussage gilt.*

Theorem 1.75 *α ist eine aussagenlogische Tautologie genau dann, wenn ein Resolutionsbeweis für α existiert, in dem Resolventenbildungen nur atomar vorkommen.*

Der klausale Resolutionskalkül als Spezialfall

In dem bisher betrachteten nichtklausalen Resolutionskalkül waren an die zu beweisende Formel keine Anforderungen hinsichtlich ihrer formalen Gestalt gestellt. Im Rahmen der logischen Programmierung und üblicherweise auch im Zusammenhang mit Theorembeweisern, die auf Resolutionskalkülen basieren, wird von Formeln, die in konjunktiver Normalform (Klauselform) vorliegen, ausgegangen. Für eine Menge von solchen Klauseln kann aber eine Resolutionsexpansion nur durch die Bildung von atomaren Resolventen durchgeführt werden.
Um also eine aussagenlogische Formel β zu beweisen, wird die Negation $\neg\beta$ in eine semantisch äquivalente konjunktive Normalform $\alpha = \langle\alpha_1, \ldots, \alpha_n\rangle$ umgeformt und anschließend eine geschlossene Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$ mittels (atomarer) Resolventenbildungen erzeugt. (Eine solche Resolutionsexpansion besteht dann nur noch aus Klauseln). Existiert eine solche geschlossene Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$, so heißt β im klausalen Resolutionskalkül beweisbar (notiert als $\vdash_{AL-Rk} \beta$).
Der klausale Resolutionskalkül hat somit nur eine einzige Regel: die Resolventenbildung (kurz: Resolutionsregel). Üblich ist im klausalen Resolutionskalkül die Verwendung der Mengenschreibweise für Klauseln, also z.B. $\{A, \neg B, C\}$ statt $[A, \neg B, C]$.
Aus der Korrektheit und Vollständigkeit des nichtklausalen Resolutionskalküls erhalten wir daher sofort

Theorem 1.76 (Korrektheit und Vollständigkeit des klausalen Resolutionskalküls)
$\vdash_{AL-Rk} \alpha \;\Leftrightarrow \models_{AL} \alpha.$

Strenge Korrektheit und Vollständigkeit des Resolutionskalküls

Wie zuvor für den Tableau-Kalkül zeigen wir auch für den Resolutionskalkül, dass wir den Beweisbarkeitsbegriff so zu einem Herleitungsbegriff erweitern können, dass wir in korrekter und vollständiger Weise das semantische Folgern erfassen. Wir verallgemeinern hierzu den Begriff „Resolutionsbeweis” zu „Σ-Resolutionsbeweis” (auch Resolutionsbeweis mit Prämisse(nmenge) Σ)

Definition 1.77 *Sei $\Sigma \subseteq Fm_{AL}$ und $\alpha \in Fm_{AL}$:*

1. Σ-Resolutionsexpansion*:*

   Definition wie die von Resolutionsexpansion, erweitert um den folgenden Fall:

   Ist $\alpha \in \Sigma$ und ist $\beta_1, \ldots, \beta_m$ eine Σ-Resolutionsexpansion, so auch $\beta_1, \ldots, \beta_m, [\alpha]$.

   (Notiert als $\frac{}{\alpha}$).

2. *Eine* Σ-Resolutionsexpansion $\beta_1, \ldots, \beta_m$ *heißt* erfüllbar *genau dann, wenn $\Sigma \cup \{\beta_1, \ldots, \beta_m\}$ erfüllbar ist.*

   Ein Σ-Resolutionsbeweis für α *(auch: Resolutionsherleitung von α aus Σ) ist eine geschlossene Σ-Resolutionsexpansion für $\{\neg\alpha\}$*
   (notiert als $\Sigma \vdash_{AL-R} \alpha$).

3. *Σ heißt* α-resolutionskonsistent *genau dann, wenn nicht gilt $\Sigma \vdash_{AL-R} \alpha$, und resolutionsinkonsistent sonst.*

Theorem 1.78 (Strenge Korrektheit und Vollständigkeit des nichtklausalen Resolutionskalküls)

Sei $\Sigma \subseteq Fm_{AL}$, $\alpha \in Fm_{AL}$:
$\Sigma \vdash_{AL-R} \alpha \;\Leftrightarrow\; \Sigma \models_{AL} \alpha$.

Wir werden diesen Beweis nicht mehr ausführen. Die Vorgehensweise ist analog zu den Beweisen der Korrektheit und Vollständigkeit des Resolutionskalküls (vgl. auch die Beweisskizze zur strengen Korrektheit und Vollständigkeit des Tableau-Kalküls).

1.2.3 Ein Frege-Hilbert-Kalkül

Die bisher betrachteten formalen Beweiskalküle, der Tableaukalkül- und der Resolutionskalkül, sind Widerlegungskalküle. Beide eignen sich, wenn es darum geht, schnell einen Beweis zu finden, und beide werden in verschiedenen Varianten und Verfeinerungen für das automatische Beweisen benutzt. Beide Kalküle besitzen eine Eigenschaft, die sie für das automatische Beweisen geradezu prädestiniert: die Beweissuche in diesen Kalkülen ist einfach. Schwächen dieser Verfahren liegen in ihrer „Erklärungsfähigkeit", d.h. die Vermittlung von Schlussfolgerungen aufgrund von Widerlegungsbeweisen ist eher unüblich und somit gewöhnungsbedürftig.
Historisch gesehen wurde zunächst ein anderer Ansatz verfolgt. Ausgehend von Axiomen werden mit zulässigen Schlussregeln Folgerungen hergeleitet. Der Vorteil in einer solchen Vorgehensweise: Folgerungen können besser vermittelt werden. Kalküle, die in diesem Sinne aufgebaut werden, werden als Frege-Kalküle, manchmal, wenn eine ganz spezielle Axiomatik zugrundegelegt wird, auch als Hilbert-Kalküle bezeichnet. Wir werden im Folgenden die Bezeichnung Frege-Hilbert-Kalkül (kurz: FH-Kalkül) verwenden. Bei einer solchen Vorgehensweise stellt sich dann das Problem, welche Axiome und welche Regeln die Basis des formalen Beweiskalküls bilden sollen. In dem wenig sinnvollen Extremfall könnte man z.B. alle Tautologien als Axiome wählen und käme dann ganz ohne Regeln aus. Wir wollen hier ein Axiomen- und Regelsystem angeben, das α- und β-Formeln mit verwendet (mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ bzw. $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$.)

Axiomenschemata des Frege-Hilbert-Kalküls:

Axiom 1	$\alpha \to (\beta \to \alpha)$	
Axiom 2	$(\alpha \to (\beta \to \gamma)) \to ((\alpha \to \beta) \to (\alpha \to \gamma))$	
Axiom 3	$\bot \to \alpha$	
Axiom 4	$\alpha \to \top$	
Axiom 5	$\neg\neg\alpha \to \alpha$	
Axiom 6	$\alpha \to (\neg\alpha \to \beta)$	
Axiom 7	$(\boldsymbol{\beta}_1 \to \alpha) \to ((\boldsymbol{\beta}_2 \to \alpha) \to (\boldsymbol{\beta} \to \alpha))$	für β-Formeln
Axiom 8	$\boldsymbol{\alpha} \to \boldsymbol{\alpha}_1$	für α-Formeln
Axiom 9	$\boldsymbol{\alpha} \to \boldsymbol{\alpha}_2$	"

Einzige Regel des Frege-Hilbert-Kalküls:

Modus ponens:

$$\frac{\alpha \qquad \alpha \to \beta}{\beta}$$

Wie für den Tableau- und Resolutionskalkül wollen wir als Nächstes einen formalen Beweisbarkeitsbegriff einführen.

Definition 1.79 *Eine Formel $\alpha \in Fm_{AL}$ heißt im Frege-Hilbert-Kalkül beweisbar oder herleitbar(bzw. aus Σ beweisbar oder herleitbar; notiert als $\vdash_{AL-FH} \alpha$ bzw. $\Sigma \vdash_{AL-FH} \alpha$) genau dann, wenn eine endliche Folge $\alpha_1, \ldots, \alpha_n$ von aussagenlogischen Formeln $\alpha_1, \ldots, \alpha_n \in Fm_{AL}$ existiert, so dass $\alpha_n = \alpha$ und für alle i mit $1 \leq i \leq n$ gilt:*

- *α_i ist Axiom (oder aus Σ) oder*
- *α_i ist aus α_j und α_k, $1 \leq j, k < i$, mittels der Modus ponens Regel erhalten worden, d.h. es gibt $\alpha, \beta \in Fm_{AL}$ mit $\alpha_j = \alpha$, $\alpha_k = \alpha \to \beta$ und $\alpha_i = \beta$.*

$\alpha_1, \ldots, \alpha_n$ heißt dann auch Beweis von α (aus Σ).

Ein Beispiel:
Die Folge der Formeln 1. - 5. liefert einen Beweis von $A \to A$:

1.	$(A \to ((A \to A) \to A)) \to ((A \to (A \to A)) \to (A \to A))$	Axiom 2
2.	$A \to ((A \to A) \to A)$	Axiom 1
3.	$(A \to (A \to A)) \to (A \to A)$	Modus ponens $(1, 2)$
4.	$A \to (A \to A)$	Axiom 1
5.	$(A \to A)$	Modus ponens $(3, 4)$

Auch für diesen Beweiskalkül gilt wiederum:

Theorem 1.80 (Strenge Korrektheit und Vollständigkeit des Frege-Hilbert-Kalküls) $\Sigma \models_{AL} \alpha \Leftrightarrow \Sigma \vdash_{AL-FH} \alpha$.

Beweis. Übungsaufgabe. ■

1.2.4 Zur Komplexität von Beweisen

In den vorigen Abschnitten wurden mit dem Tableaukalkül, dem Resolutionskalkül sowie dem Frege-Hilbert-Kalkül drei Beweiskalküle vorgestellt. Darüber hinaus existiert eine Vielzahl von weiteren korrekten und vollständigen Beweiskalkülen, auf die wir hier nicht mehr eingehen, sondern stattdessen den Leser auf die angegebene Literatur verweisen. Einzig der für beweistheoretische Untersuchungen wichtige Sequenzenkalkül wird im Abschnitt über die Prädikatenlogik ganz kurz eingeführt.
Welcher dieser Beweiskalküle ist denn nun geeigneter, um damit Beweise zu führen? Es drängt sich somit natürlich ein Vergleich dieser Kalküle hinsichtlich ihrer „Beweiskomplexität" auf. Wir wollen dieses Thema in diesem Abschnitt kurz streifen und wenigstens ein paar bedeutende Resultate vorstellen.
Wir definieren dazu zunächst den Begriff eines Beweissystems als Abbildung, die einem potentiellen Beweis x eine Tautologie α zuweist.
Sei Σ ein Alphabet.

Definition 1.81 *Ein* Beweissystem *für Taut (die Menge der aussagenlogischen Tautologien) ist eine polyno-mialzeit-berechenbare Abbildung $f : \Sigma^* \to (\Sigma_{AL})^*$ mit Wertebereich Taut.*
Ein Beweissystem f heißt polynomiell beschränkt *genau dann, wenn ein Polynom p existiert, so dass für alle $\alpha \in Taut$ ein $x \in \Sigma^*$ existiert mit $f(x) = \alpha$ und $|x| \leq p(|\alpha|)$. ($|w|$ ist die Länge des Wortes w.)*

Die Intention, die hinter dieser Definition steckt, ist die Folgende: für einen konkreten Beweiskalkül wird x ein Beweis der Formel α sein. Hierbei soll in polynomieller Zeit festgestellt werden können, ob x wirklich ein Beweis darstellt. Da der Wertebereich eines Beweissytems $Taut$ ist, legen wir für den Fall, dass die Zeichenfolge x in dem zugrundeliegenden Beweiskalkül keinen Beweis darstellt, fest, dass $f(x) = \top$ sein soll. Das Beweissystem ist polynomiell beschränkt, wenn es zu jeder aussagenlogischen Tautologie einen Beweis polynomieller Länge in diesem Beweissystem gibt.
Eng verknüpft mit dem Problem der Existenz eines polynomiell beschränkten Beweissystems ist ein bedeutendes komplexitätstheoretisches Problem.

Theorem 1.82 (Cook, Reckhow)
$NP = co - NP$ genau dann, wenn ein polynomiell beschränktes Beweissystem existiert.

Beweis. (Skizze)

„ $\Rightarrow$ " : Es gelte $NP = co - NP$. Dann ist wegen $Taut \in co - NP$ auch $Taut \in NP$. Es gibt also eine nichtdeterministische polynomialzeitbeschränkte Turingmaschine M, die genau $Taut$ akzeptiert. Sei $f : \Sigma^* \to (\Sigma_{AL})^*$ für ein geeignetes Σ eine Abbildung mit der Eigenschaft, dass $f(x) = \alpha$, falls x eine akzeptierende polynomialzeitbeschränkte Berechnung von M für α kodiert, und $\top$ sonst. f ist dann ein polynomiell beschränktes Beweissystem.

„ $\Leftarrow$ " : Sei umgekehrt nun f ein polynomiell beschränktes Beweissystem. Da SAT, die Menge der erfüllbaren aussagenlogischen Formeln, vollständig für NP ist, und $Taut = \{\alpha \mid \neg\alpha \notin SAT\}$, ist $Taut$ vollständig für $co - NP$. Unter Verwendung der zu f existierenden deterministischen polynomiell zeitbeschränkten Turingmaschine kann dann leicht eine nichtdeterministische polynomiell zeitbeschränkte Turingmaschine angegeben werden, die genau $Taut$ akzeptiert, so dass also $Taut \in NP$. Bei Eingabe α rate einen „Beweis" x mit polynomieller Länge in $|\alpha|$ und „verifiziere" anschließend mittels Berechnung von $f(x)$, dass α eine Tautologie ist.

Aus $Taut \in NP$ und $Taut$ vollständig für $co - NP$ folgt aber $NP = co - NP$.

■

Da man vermutet, dass $NP \neq co - NP$, wird man nicht erwarten, ein polynomiell beschränktes Beweissystem zu finden. Kann man aber auch für alle bekannten Beweissysteme eine superpolynomielle oder gar exponentielle untere Schranke zeigen? Dies ist bisher noch nicht für alle aussagenlogischen Beweissysteme geglückt. Bekannt sind aber exponentielle untere Schranken u.a. für den (klausalen) Resolutionskalkül und den Tableaukalkül – wir werden darauf zurückkommen. Hingegen sind z.B. für Frege-Hilbert-Kalküle keine superpolynomiellen unteren Schranken bekannt.

Definition 1.83 *Seien $f_1 : \Sigma_1^* \to (\Sigma_{AL})^*$ und $f_2 : \Sigma_2^* \to (\Sigma_{AL})^*$ Beweissysteme.*
f_2 p-simuliert f_1 (notiert als $f_1 \leq_p f_2$) genau dann, wenn eine polynomialzeit-berechenbare Funktion $g : \Sigma_1^* \to \Sigma_2^*$ existiert mit

$$f_2(g(x)) = f_1(x)$$

für alle $x \in \Sigma_1^*$.

g übersetzt somit Beweise im Beweissytem f_1 in höchstens polynomiell längere Beweise des Beweissystems f_2. Eine unmittelbare Folgerung ist das

Korollar 1.84 *Ist f_1 ein polynomiell beschränktes Beweissystem und p-simuliert das Beweissystem f_2 das Beweissystem f_1, dann ist f_2 ebenfalls polynomiell beschränkt.*

p-Simulierbarkeit ist offensichtlich eine reflexive Relation. Da die Komposition zweier Polynome wiederum ein Polynom ist, ist p-Simulierbarkeit auch eine transitive Relation.

Theorem 1.85 *Für die hier betrachteten Beweissysteme gelten die folgenden Beziehungen hinsichtlich p-Simulierbarkeit:*

1. *Resolutionskalkül* $\leq_p$ *Frege-Hilbert-Kalkül*
2. *Tableaukalkül* $\leq_p$ *Resolutionskalkül*
3. *Wahrheitswerttabellenmethode* $\leq_p$ *Resolutionskalkül*
4. *Wahrheitswerttabellenmethode* $\not\leq_p$ *Tableaukalkül*
5. *Tableaukalkül* $\not\leq_p$ *Wahrheitswerttabellenmethode*

Beweis. Wird hier nicht ausgeführt.
Wir werden lediglich eine Beweisskizze für die vermutlich überraschende Eigenschaft 4. angeben.
Es seien n aussagenlogische Variablen $A_1, \ldots, A_n$ gegeben. Sei $\alpha_i^1 = A_i$ und $\alpha_i^0 = \neg A_i$ für alle i mit $1 \leq i \leq n$.
Betrachte die folgende Menge von Klauseln

$$M_n := \{[\alpha_1^{\varepsilon_1}, \ldots, \alpha_n^{\varepsilon_n}] \mid (\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n\}.$$

Z.B. also im Fall $n = 2$: $\{[A_1, A_2], [\neg A_1, A_2], [A_1, \neg A_2], [\neg A_1, \neg A_2]\}$.
Die Menge M_n ist offensichtlich unerfüllbar, und da in M_n genau n aussagenlogische Variablen vorkommen, kann mit der Wahrheitswerttabellenmethode die Unerfüllbarkeit in Zeit $O\left((2^n)^k\right)$ für ein geeignetes $k \in \mathbb{N}$ festgestellt werden (eine genauere Analyse zeigt, dass $k = 3$ genügt).
Da M_n unerfüllbar ist, existiert somit ein geschlossenes Tableau für M_n. Wir zeigen, dass ein solches geschlossenes Tableau für M_n mindestens $n!$ verschiedene Zweige haben muss. Da $n!$ aber stärker als jedes Polynom in 2^n wächst, kann der Tableaukalkül die Wahrheitswerttabellenmethode nicht p-simulieren.
Für den Nachweis, dass ein geschlossenes Tableau für M_n mindestens $n!$ verschiedene Zweige haben muss, wollen wir die $\boldsymbol{\beta}$-Regel im Falle von Klauseln liberalisieren:
Im Falle einer Klausel $[\beta_1, \ldots, \beta_k]$ werden an das Blatt des aktuell bearbeiteten Zweiges k Söhne markiert mit β_1 rsp. β_2 ... rsp. β_k angehängt. Auf diese Weise werden die Tableaus etwas kleiner, d.h. die Anzahl der Knoten wird reduziert, nicht aber die Anzahl der Zweige.
Betrachte ein minimal geschlossenes Tableau τ für M_n, d.h. ein geschlossenes Tableau, in dem eine Formel entlang eines Zweiges nicht mehrfach bearbeitet wurde. In diesem Tableau hat jeder innere Knoten s, der nicht schon innerer Knoten des Starttableaus ist, n Söhne, von denen der i-te Sohn mit A_i oder mit $\neg A_i$ markiert ist. Sei $\sigma : \{1\ldots, n\} \to \{1\ldots, n\}$ eine beliebige Permutation. Zu dieser Permutation muss für ein $(\varepsilon_1, \ldots, \varepsilon_n) \in \{0,1\}^n$ in τ ein Zweig vorkommen, dessen letzte n Knoten $s_1, \ldots, s_n$ mit $\alpha_{\sigma(1)}^{\varepsilon_1}, \ldots, \alpha_{\sigma(n)}^{\varepsilon_n}$ markiert sind. Für jede solche Permutation liegen diese Pfade in verschiedenen Zweigen, so dass τ also mindestens $n!$ verschiedene Zweige haben muss. Dies wollten wir zeigen. ■

Überraschend ist aber auch, dass der im Rahmen des automatischen Beweisens überwiegend eingesetzte Resolutionskalkül nachweislich ineffizient ist. Dieses Resultat geht auf Haken (1985) zurück. Auf den schwierigen Beweis wollen wir hier nicht eingehen und daher nur das Resultat angeben.

Theorem 1.86 *Es gibt Klauselmengen, für die jede geschlossene Resolutionsexpansion mindestens exponentiell viele Klauseln enthalten muss.*

Korollar 1.87 *Der Resolutionskalkül ist nicht polynomiell beschränkt.*

Kapitel 2

Prädikatenlogik

Mit dem Aufbau eines formalen Kalküls der Aussagenlogik haben wir ein erstes Ziel erreicht: wir haben korrekte und vollständige Beweiskalküle für das logische Folgern erhalten, allerdings unter der Maßgabe, dass sich dieses Folgern nur auf solche Aussagen erstreckte, die über „aussagenlogische” Konnektoren aus „Grundaussagen”, also aus Aussagen, die sich bezüglich dieser Konnektoren nicht weiter zerlegen ließen, aufgebaut waren. Nun gibt es sicherlich viele Anwendungen, in denen nur Aussagen dieser Art benötigt werden. Im Prinzip ist ja eine solche aussagenlogische Sprache immer dann ausreichend, wenn wir es nur mit endlichen Bereichen zu tun haben.
Schwächen der Aussagenlogik offenbaren sich aber dann, wenn wir unendliche Gegenstandsbereiche wie z.B. die natürlichen Zahlen betrachten und das Zutreffen von Eigenschaften für Objekte aus diesem Gegenstandsbereich formulieren wollen. Hier muss eine Aussagenlogik kapitulieren, wenn, was sinnvollerweise gewünscht ist, unendlich lange Verknüpfungen von Aussagen vermieden werden sollen. Aber selbst in einem endlichen Bereich wie z.B. $B = \{c_1, \ldots, c_{10000}\}$, für den das Zutreffen einer bestimmten Eigenschaft E auf alle Elemente dieses Bereichs ausgedrückt werden soll, erscheint uns die zwar denkbare Darstellung als aussagenlogische Verknüpfung von 10000 Aussagen $E(c_1), \ldots, E(c_{10000})$ als wenig wünschenswert und – was vielleicht noch mehr ins Gewicht fällt – entspricht nicht unserem Sprachgebrauch.
Bereits diese kurzen Anmerkungen zeigen aber, worin die Ursachen der Ausdrucksschwäche der Aussagenlogik liegen. In der Aussagenlogik wurde die interne Struktur von Sätzen, die bezüglich der aussagenlogischen Konnektoren nicht weiter zerlegbar waren, nicht berücksichtigt, und – damit verbunden – gab es keine Möglichkeit, über Objektbereiche zu quantifizieren. Wir wollen im Folgenden einen formalen Kalkül aufbauen, der dieses Manko der Aussagenlogik nicht aufweist und für die gängigen Anforderungen im Wissenschaftsbetrieb ausreichend ist.

Beispiel 2.1 (*für nicht unbedingt informatiktypische Anwendungen*)

1. *Im Bereich der Mathematik, aber auch in der Informatik, haben wir es häufig mit Gruppen zu tun, d.h. mit einer algebraischen Struktur* $(A, +, 0)$ *(= wird üblicherweise nicht aufgelistet) mit einer Konstanten* $0 \in A$ *(dem neutralen Element) und einer zweistelligen Verknüpfungsoperation* $+ : A \times A \to A$*, die noch speziellen Axiomen genügt. Für eine solche Gruppe möchten wir dann Aussagen z.B. der Art*

   - *für alle* $a \in A$ *gibt es* $b \in A$*, so dass* $a + b = 0$
   - *für alle* $a \in A$ *gilt:* $(a + 0) + 0 = a$

   formulieren können.

2. *Ein gerichteter Graph* G *ist eine Struktur* $G = (A, R)$ *mit einem Grundbereich* A *(Menge der Knoten) und einer zweistelligen Relation* $R \subseteq A \times A$ *über dem Grundbereich (Menge der gerichteten Kanten).*

   Für einen solchen Graphen ließe sich dann z.B. ausdrücken

   $$\textit{für alle } a, b \in A : \textit{wenn } (a, b) \in R, \textit{ so } (b, a) \in R.$$

Bereits diese beiden einfachen Beispiele zeigen auf, was für Elemente in einer formalen Sprache, die für solche Anwendungen geeignet ist, vorhanden sein sollten. Nämlich Symbole für

- Objekte, über die wir in unbestimmter Weise sprechen können, wie z.B. a, b. Wir sprechen dann von Variablen.
- Relationen

  (In den beiden Beispiel waren dies die Relationen $=$ und R).
- Funktionen

  (Im ersten Beispiel $+$).
- Konstanten

  (Im ersten Beispiel 0).

Darüber hinaus – und dies zeigen die beiden Beispiele ebenfalls – müssen All- und Existenzquantifikationen möglich sein.
Im Folgenden wollen wir eine formale Sprache aufbauen, die unseren Anforderungen in den meisten Fällen genügen wird. Wir werden aus Gründen der einfacheren Darstellung uns dabei zunächst auf den einsortigen Fall beschränken, d.h. wir werden eine formale Sprache aufbauen, mit der wir Strukturen berücksichtigen, die nur einen Grundbereich besitzen. In den meisten Anwendungen der Informatik ist diese Bedingung nicht erfüllt, so dass in dieser Hinsicht unser Ansatz zu kurz greift. Wir werden aber später sehen, dass die Beschränkung auf den einsortigen Fall keine wirkliche Beschränkung darstellt.

2.1 Formalisierung der Prädikatenlogik

Analog zur Aussagenlogik werden wir zunächst eine formale Sprache angeben und anschließend die Semantik für eine solche Sprache definieren.

2.1.1 Syntax der Prädikatenlogik

Wir wollen eine formale Sprache aufbauen, mit der wir über formale Strukturen eines bestimmten Typs sprechen können, z.B. im Falle von Graphen über eine Struktur mit einem Grundbereich und einer zweistelligen Relation. Zu diesem Zweck definieren wir zunächst den Begriff der Signatur.

Definition 2.2 *Eine* Signatur *S ist ein Quadrupel $S = (\mathcal{R}; \mathcal{F}; \mathcal{K}; \mathrm{ar})$, wobei $\mathcal{R}$ eine Menge von Prädikats- oder Relationssymbolen ist, $\mathcal{F}$ eine Menge von Funktionssymbolen, $\mathcal{K}$ eine Menge von Konstantensymbolen und die Funktion $\mathrm{ar} : \mathcal{R} \cup \mathcal{F} \to \mathbb{N}$ die Stelligkeit der Relations- und Funktionssymbole angibt.*

Bemerkung 2.3

1. *Wenn die Stelligkeit der Relations- und Funktionssymbole aus dem Kontext hervorgeht (oder sonst nicht von Bedeutung ist), wird häufig auf die Angabe der Stelligkeitsfunktion ar verzichtet.*
2. *Falls $\mathcal{R}$, $\mathcal{F}$ oder $\mathcal{K}$ endlich ist, wird die betreffende Menge auch als endliche Folge notiert, also z.B. $(R_1, R_2; f_1; 0, 1)$ statt $(\{R_1, R_2\}; \{f_1\}; \{0, 1\})$, wobei hier ar bereits weggelassen wurde. Ist eine dieser Mengen leer, so wird sie i.a. in der Signatur nicht mehr aufgeführt. Auch schreiben wir in einem solchen Fall keine mehrfachen Semikolons, wenn aus dem Kontext hervorgeht, welche Menge leer ist. Z.B. schreiben wir $(R; 0)$ statt $(\{R\}; \emptyset; \{0\})$ und (f) statt $(\emptyset; \{f\}; \emptyset)$.*
3. *Konstantensymbole werden oft auch als nullstellige Funktionssymbole betrachtet. In diesem Fall wird $\mathcal{K}$ in der Regel in der Signatur nicht aufgeführt.*
4. *Ist $\mathcal{K} = \emptyset$, dann heißt S konstantenlos. Ist $\mathcal{K} \cup \mathcal{F} = \emptyset$, so heißt S Relationalstuktur. Und ist $\mathcal{R} \cup \mathcal{F} \cup \mathcal{K}$ endlich, so heißt die Signatur S endlich.*

Beispiel 2.4

1. *In einer formalen Sprache, in der wir Eigenschaften von Gruppen ausdrücken möchten, benötigen wir ein 2-stelliges Symbol* $\cong$ *für die Gleichheitsrelation* $=$, *ein 2-stelliges Symbol für die Verknüpfungsoperation der Gruppe und ein Konstantensymbol für das neutrale Element, so dass z.B.* $(\cong; +; 0)$ *mit* $\mathrm{ar}(\cong) = 2$ *und* $\mathrm{ar}(+) = 2$ *eine geeignete Signatur ist.*

2. *Soll unsere formale Sprache dazu dienen, Eigenschaften von Graphen zu formulieren, dann wird unsere Signatur z.B. so aussehen:* (R) *mit* $\mathrm{ar}(R) = 2$.

Für den Aufbau der prädikatenlogischen Sprache zu einer Signatur S legen wir das Alphabet $\Sigma_{PL}(S)$ zugrunde, das besteht aus:

- Symbolen zur Bildung der (Objekt- oder Individuen-) Variablen: $x, |$
- Zeichen für die logischen Funktoren oder Konnektoren (wie im aussagenlogischen Fall):
  $\neg$ (einstelliger Konnektor)
  $\wedge, \vee, \rightarrow$ (zweistellige Konnektoren)
- den logischen Konstanten $\top$ und $\bot$
- Quantoren $\forall$ und $\exists$
  (unbestimmt angedeutet durch Q)
- den durch die Signatur S festgelegten Relations-, Funktions- und Konstantensymbolen
- Klammersymbolen als technischen Hilfszeichen (,).

Definition 2.5 *Die folgenden Zeichenreihen über dem Alphabet* $\Sigma_{PL}(S)$ *zur Signatur* S *heißen* Variablen*:* x, $x|$, $x||$, $\ldots$.
Var_{PL} *bezeichne die Menge der Variablen.*
Schreibweise: statt $x|\ldots|$ $(n+1$ *Striche*$)$ x_{n+1}.

Als metasprachliche Symbole für Variablen verwenden wir im Folgenden auch die Symbole $x, y, z, v, x_0, y_0, z_0, v_0, x_1, \ldots$.
In Strukturen mit Funktionen können manche Objekte auf verschiedene Weise beschrieben werden, zum einen nämlich direkt, zum anderen als das Resultat der Anwendung einer Funktion auf andere Objekte. So kann z.B. die natürliche Zahl 5 auch als Summe $2 + 3$ dargestellt werden. In dem Aufbau der formalen Sprache berücksichtigen wir dies durch die Einführung von Termen, die zur Bezeichnung von Objekten dienen.

Definition 2.6 *Die Menge* $Tm_{PL}(S)$ *der* Terme der Signatur $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$ *ist folgendermaßen induktiv definiert:*

1. *Jede Variable ist ein Term von* $Tm_{PL}(S)$.

2. *Ist* $c \in \mathcal{K}$, *so ist* c *ein Term von* $Tm_{PL}(S)$.

3. *Ist* $f \in \mathcal{F}$ *ein* n*-stelliges Funktionssymbol und sind* $t_1, \ldots, t_n \in Tm_{PL}(S)$, *so ist* $f(t_1, \ldots, t_n) \in Tm_{PL}(S)$.

4. *Nur Zeichenreihen, die nach* 1. - 3. *erhalten werden, sind Terme von* $Tm_{PL}(S)$.

Terme ohne Variablen werden als variablenfreie Terme oder auch als Grundterme *bezeichnet.*

Beispiel 2.7 *für Terme.*
Sei $S = (\emptyset; f, g; c_1)$ *mit* f *zweistellig und* g *einstellig.*

1. $f(x, c_1)$

2. $f(g(x), g(f(x,y)))$.

Als Nächstes definieren wir die Formeln zur Signatur S.

Definition 2.8 *Sei* $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$.
Atomare Formel der Signatur S *sind*

1. $\top$, $\bot$ *und*
2. *für jedes Relationssymbol* $R \in \mathcal{R}$ *sowie für Terme* $t_1, \ldots, t_n \in Tm_{PL}(S)$, *wobei* n *die Stelligkeit von* R *sei, die* Zeichenreihen $R(t_1, \ldots, t_n)$.

Definition 2.9 (*Menge* $Fm_{PL}(S)$ *der* prädikatenlogischen Formeln der Signatur S)

1. *Jede atomare Formel der Signatur* S *ist in* $Fm_{PL}(S)$.
2. *Falls* $\alpha \in Fm_{PL}(S)$, *dann ist auch* $\neg\alpha \in Fm_{PL}(S)$.
3. *Falls* $\alpha, \beta \in Fm_{PL}(S)$, *und* $\circ \in \{\wedge, \vee, \rightarrow\}$, *dann ist auch* $(\alpha \circ \beta) \in Fm_{PL}(S)$.
4. *Ist* $\alpha \in Fm_{PL}(S)$ *und ist* x *eine Variable, so sind* $(\forall x)\,\alpha$ *und* $(\exists x)\,\alpha$ *Formeln in* $Fm_{PL}(S)$.

Beispiel 2.10 *Die folgenden Zeichenreihen sind Formeln über der Signatur* $S = (P, R; f; 0)$, *wobei* P *und* f *zweistellig und* R *einstellig sein sollen:*

1. $(P(x, f(0,y)) \vee R(f(x,y)))$
2. $((\forall x)\, P(x,y) \rightarrow (\exists z)\,(R(x) \rightarrow (\forall x)\, P(y,x)))$

Hingegen sind die folgenden Zeichenreihen keine Formeln über der Signatur S:

1. $\exists x R(x)$ (*Klammern fehlen*)
2. $(\forall y)\, R(x,y)$ (R *nicht einstellig*)

Wir übernehmen die in der Aussagenlogik eingeführten *Klammerungskonventionen* und ergänzen sie durch die Festlegung, dass ein Quantor (Qx) ebenso schwach wie $\neg$ trennt.

Bezeichnungen:

1. Für eine Formel $(Qx)\,\alpha$ heißt α der Wirkungsbereich des Quantors(Qx).
   (Dieser Wirkungsbereich ist eindeutig bestimmt!)
2. $var(t)$ bezeichne die Menge der im Term t vorkommenden Variablen, und
   $var(\alpha)$ die Menge der in der Formel α vorkommenden Variablen.

Aus der Mathematik kennen wir eine Reihe von Ausdrücken wie z.B. $\sum_{i=1}^{100} i$, in denen formal eine „Variable" i vorkommt, in denen wir aber i nicht durch einen anderen Wert ersetzen können, ohne dass die Bedeutung des Ausdrucks sich ändert. Das i steht somit für eine Ersetzung nicht zur Verfügung, ist also nicht „frei", sondern gebunden. Wir wollen eine analoge Unterscheidung für die Variablen unserer formalen Sprache einführen.

Definition 2.11 *Die Menge* $fvar(\alpha)$ *der* freien Variablen in einer Formel α *ist wie folgt definiert:*

1. *Ist* α *atomar, so ist* $fvar(\alpha) := var(\alpha)$
2. $fvar(\neg\alpha) := fvar(\alpha)$
3. $fvar(\alpha \circ \beta) := fvar(\alpha) \cup fvar(\beta)$,
   wobei $\circ$ *ein zweistelliger Konnektor sein soll mit* $\circ \in \{\wedge, \vee, \rightarrow\}$

4. $fvar((\forall x)\,\alpha) := fvar((\exists x)\,\alpha) := fvar(\alpha) \setminus \{x\}$

Eine Variable x heißt gebunden *in α, falls $x \in var(\alpha)$ und x nicht frei in α.*
x kommt in α vollfrei *vor, falls $x \in fvar(\alpha)$ und in α die Zeichenreihe (Qx) nicht vorkommt.*

Beispiel 2.12

1. *In $(\forall x)\,P(x)$ kommt x gebunden vor.*
2. *In $(\forall x)\,(P(x) \vee R(x,y))$ ist y vollfrei und x gebunden.*
3. *In $(\forall x)\,P(x) \vee R(x,y)$ ist y vollfrei und x frei.*

Definition 2.13 *Eine Formel α heißt* quantorenfrei*, wenn weder $\forall$ noch $\exists$ in α vorkommt.*
α heißt Aussage *(engl. sentence) genau dann, wenn $fvar(\alpha) = \emptyset$.*

In Analogie zur Aussagenlogik steht uns als Beweisprinzip für Terme und Formeln strukturelle Induktion und als Definitionsprinzip strukturelle Rekursion zur Verfügung.

Theorem 2.14 (Strukturelle Induktion für Terme) *E sei eine Eigenschaft für Terme einer Signatur S, für die gilt:*

1. *alle Variablen und alle Konstantensymbole in S haben die Eigenschaft E.*
2. *Haben die Terme $t_1, \ldots, t_n \in Tm_{PL}(S)$ die Eigenschaft E und ist f ein n-stelliges Funktionssymbol der Signatur S, so hat auch $f(t_1, \ldots, t_n)$ die Eigenschaft E.*

Dann gilt:
E trifft auf alle Terme der Signatur S zu.

Theorem 2.15 (Strukturelle Induktion für Formeln) *E sei eine Eigenschaft für Formeln einer Signatur S, für die gilt:*

1. *E trifft auf alle atomaren Formeln zu.*
2. *Trifft E auf α und β zu, ist x eine Variable und $\circ \in \{\wedge, \vee, \rightarrow\}$, dann trifft E auch auf $\neg\alpha$, $(\alpha \circ \beta)$, $(\forall x)\,\alpha$ und $(\exists x)\,\alpha$ zu.*

Dann gilt:
E trifft auf alle Formeln der Signatur S zu.

Auf die einfachen Beweise verzichten wir. Ebenso auf die Angabe der entsprechenden Definitionsprinzipien der strukturellen Rekursion.
In Analogie zum Formelbaum für aussagenlogische Formeln könnten wir im Falle der Prädikatenlogik die Begriffe „Termbaum" und „Formelbaum" definieren. Da klar sein dürfte, was mit diesen Begriffen gemeint ist, verzichten wir hier auf deren Definition, zumal diese Begriffe im weiteren keine Rolle mehr spielen werden. Sie fließen nur in die nachfolgende Bemerkung ein.

Bemerkung 2.16 *Da zu einer Signatur eine eindeutige kontextfreie Grammatik angegeben werden kann, die die Menge der Terme bzw. die Menge der Formeln dieser Signatur erzeugt, gilt:*

1. Zu jedem Term einer Signatur S existiert genau ein Termbaum, und
2. zu jeder Formel von S existiert genau ein Formelbaum.

Das folgende Theorem, das wir nicht beweisen wollen, ist eine leichte Folgerung hieraus.

Theorem 2.17 *Sei S eine Signatur.*

1. *Jeder Term t der Signatur S ist entweder eine Variable, ein Konstantensymbol aus S oder hat eine* eindeutige *Zerlegung in der Form $f(t_1, \ldots, t_n)$. (Insbesondere sind dann f und $t_1, \ldots, t_n$ eindeutig bestimmt.)*

2. *Jede Formel α der Signatur S ist entweder gleich $\top$ oder gleich $\perp$ oder hat eine* eindeutige *Darstellung in einer der der folgenden Formen*

   (a) $R(t_1, \ldots, t_n)$,

   (b) $\neg\beta$,

   (c) $(\beta \circ \gamma)$ *für einen zweistelligen Konnektor* $\circ$,

   (d) $(\forall x)\,\beta$ *oder* $(\exists x)\,\beta$ *für eine Variable* x.

Definition 2.18 *S sei eine Signatur.*

1. *t sei ein Term der Signatur S.*

   Ist der Term t' ein Teilwort von t (d.h. existieren u, v mit $t = ut'v$), so heißt t' Teilterm von t.

2. *α sei eine Formel der Signatur S.*

   α' sei eine Formel, die als Teilwort in α vorkommt. Dann heißt α' Teilformel von α.

2.1.2 Substitutionen

Aus der Mathematik kennen wir viele Fälle, in denen eine Eigenschaft für eine Variable gezeigt wird und anschließend für diese Variable ein Wert eingesetzt (substituiert) wird. Eine entsprechende Eigenschaft soll uns in der formalen Sprache der Logik zur Verfügung stehen. Da wir später – insbesondere im Zusammenhang mit der Logischen Programmierung – den Substitutionsbegriff ganz wesentlich benötigen, wollen wir ihn hier gleich in der nötigen Allgemeinheit einführen.

Definition 2.19 *Sei S eine Signatur.*
Eine Abbildung $\theta : Var_{PL} \to Tm_{PL}(S)$ mit der Eigenschaft, dass für fast alle (d.h. bis auf endlich viele) $x \in Var_{PL}$ gilt $\theta(x) = x$, heißt Substitution *(zur Signatur S).*

Eine Substitution θ mit

$$\theta(x) = \begin{cases} t_i & \text{falls } x = x_i,\ 1 \leq i \leq n \\ x & \text{sonst} \end{cases}$$

notieren wir durch $\theta = \{x_1/t_1, \ldots, x_n/t_n\}$.

Definition 2.20 (Substitution in Termen) *Sei θ eine Substitution. Für Terme $t \in Tm_{PL}(S)$ definieren wir $t\theta$:*

1. *Für ein Konstantensymbol c von S:*

   $c\theta := c$

2. *Für eine Variable x:*

   $x\theta := \theta(x)$

3. $[f(t_1, \ldots, t_n)]\,\theta := f(t_1\theta, \ldots, t_n\theta)$.

Beispiel 2.21 *Sei $\theta = \{x/f(x,y),\ y/g(c),\ z/h(g(x), d)\}$.*
Dann ist $[f(x, g(z))]\,\theta = f(f(x,y), g(h(g(x), d)))$

Definition 2.22 *σ und θ seien Substitutionen (zur gleichen Signatur).*
Die Komposition *von σ und θ (notiert als $\sigma\theta$) ist definiert durch*

$$\sigma\theta\,(x) := (x\sigma)\,\theta$$

Lemma 2.23 *Für alle Terme $t \in Tm_{PL}(S)$ und für alle Substitutionen σ, θ (zur Signatur S) gilt*

$$t\,(\sigma\theta) = (t\sigma)\,\theta.$$

Beweis. Mittels struktureller Induktion für Terme.

Ind.Anf.: Sei x eine Variable. Dann gilt $x\,(\sigma\theta) = (x\sigma)\,\theta$ nach Definition. Im Falle eines Konsantensymbols c aus S gilt $c\,(\sigma\theta) = c = (c\sigma) = (c\sigma)\,\theta$.

Ind.Schr.: Sei f ein n-stelliges Funktionssymbol aus der Signatur S und gelte die Behauptung für $t_1, \ldots, t_n$, d.h. es gilt $t_i\,(\sigma\theta) = (t_i\sigma)\,\theta$, $1 \leq i \leq n$.

Dann gilt

$$\begin{aligned}(f\,(t_1, \ldots, t_n))\,(\sigma\theta) &= f\,(t_1\,(\sigma\theta)\,, \ldots, t_n\,(\sigma\theta))\\ &= f\,((t_1\sigma)\,\theta, \ldots, (t_n\sigma)\,\theta)\\ &= (f\,(t_1\sigma, \ldots, t_n\sigma))\,\theta\\ &= ((f\,(t_1, \ldots, t_n))\,\sigma)\,\theta\end{aligned}$$

■

Ebenso leicht ist zu sehen, dass die Komposition von Substitutionen assoziativ ist, d.h. es gilt das

Lemma 2.24 *Für beliebige Substitutionen σ, θ und ρ gilt $\sigma\,(\theta\rho) = (\sigma\theta)\,\rho$.*

Bemerkung 2.25 *Sind $\sigma = \{x_1/t_1, \ldots, x_n/t_n\}$ und $\theta = \{y_1/u_1, \ldots, y_k/u_k\}$ Substitutionen, so ist*

$$\sigma\theta = \{x_1/\,(t_1\theta)\,, \ldots, x_n/\,(t_n\theta)\,, y_{i_1}/u_{i_1}, \ldots, y_{i_l}/u_{i_l}\} \;\; ,$$

wobei $\{y_{i_1}, \ldots, y_{i_l}\} = \{y_1, \ldots, y_k\} \setminus \{x_1, \ldots, x_n\}$.

Wir wollen als Nächstes den Begriff der Substitution in einer Formel definieren. Hier müssen wir darauf achten, dass eine gebundene Variable durch eine Substitution nicht ersetzt werden darf.

Definition 2.26 *Sei x eine Variable und θ eine Substitution.*
θ_x sei die folgendermaßen definierte Substitution

$$\theta_x\,(y) := \begin{cases} \theta\,(y) & \textit{falls } y \neq x \\ x & \textit{sonst} \end{cases} .$$

Definition 2.27 (Substitution in Formeln) *Sei θ eine Substitution (zur Signatur S). Für Formeln $\alpha \in Fm_{PL}(S)$ ist $\alpha\theta$ definiert durch*

1. $\top\theta := \top \quad \bot\theta := \bot$
2. $[R\,(t_1, \ldots, t_n)]\,\theta := R\,(t_1\theta, \ldots, t_n\theta)$
3. $[\neg\alpha]\,\theta := \neg\,[\alpha\theta]$
4. $(\alpha \circ \beta)\,\theta := (\alpha\theta \circ \beta\theta)$ *für* $\circ \in \{\wedge, \vee, \rightarrow\}$
5. $[(Qx)\,\alpha]\,\theta := (Qx)\,[\alpha\theta_x]$ *für* $Q \in \{\forall, \exists\}$

Beispiel 2.28 *Sei* $\theta = \{x/c,\ y/x\}$. *Dann gilt*

$$\begin{aligned}[(\forall x)\, R(x,y) \to (\exists y)\, R(x,y)]\, \theta &= [(\forall x)\, R(x,y)]\, \theta \to [(\exists y)\, R(x,y)]\, \theta \\ &= (\forall x)\, [R(x,y)]\, \theta_x \to (\exists y)\, [R(x,y)]\, \theta_y \\ &= (\forall x)\, R(x,x) \to (\exists y)\, R(c,y)\, .\end{aligned}$$

Bemerkung 2.29 *Hätten wir in Punkt 4. der Definition der Substitution in einer Formeln nicht* θ_x *sondern* θ *verwendet, so würden wir z.B. im Falle der Gruppentheorie aus der Gleichung*

$$(\exists x)\, (y + x \cong 0)$$

durch die Substitution $\{x/y\}$ *die Gleichung*

$$(\exists x)\, (y + y \cong 0)$$

für eine Variable y *erhalten. Eine wenig wünschenswerte Eigenschaft.*

Bemerkung 2.30 *Für Terme* t *und Substitutionen* σ *und* θ *gilt* $t\,(\sigma\theta) = (t\sigma)\,\theta$.
Für Formeln gilt i.a. nicht $\alpha\,(\sigma\theta) = (\alpha\sigma)\,\theta$. *Für spätere Anwendungen ist jedoch eine solche Eigenschaft für Substitutionen gewünscht. Zu diesem Zweck definieren wir Substitutionen, die diese Eigenschaft haben werden.*

Definition 2.31 *Eine Substitution* θ *(zur Signatur* S*) heißt* frei für $\alpha \in Fm_{PL}(S)$, *falls*

1. α *atomar ist oder*
2. $\alpha = \neg\beta$ *und* θ *frei für* β *ist oder*
3. $\alpha = (\beta \circ \gamma)$ *und* θ *frei für* β *und frei für* γ *ist,* $(\circ \in \{\wedge, \vee, \to\})$ *oder*
4. $\alpha = (Qx)\,\beta$ $(Q \in \{\forall, \exists\})$ *und* θ_x *frei für* β *ist, und außerdem gilt:*
   wenn $y \in fvar(\beta)$ *mit* $y \neq x$, *dann* $x \notin var(y\theta)$.

Beispiel 2.32 *Sei die prädikatenlogische Formel* $\alpha = (\forall x)\, R(x,y)$ *gegeben. Dann ist die Substitution* $\theta = \{x/y,\ y/f(z,c)\}$ *frei für* α, *hingegen ist die Substitution* $\sigma = \{y/f(x,c)\}$ *nicht frei für* α, *da in* $\alpha\sigma = (\forall x)\, R(x, f(x,c))$ *jetzt auch im Term* $f(x,c)$ *die Variable* x *durch den Quantor* $(\forall x)$ *gebunden wird.*

Theorem 2.33 *Sei* $\alpha \in Fm_{PL}(S)$.
σ *sei eine Substitution, die frei für* α *ist und* θ *sei eine Substitution, die frei für* $\alpha\sigma$ *ist. Dann gilt:*

$$(\alpha\sigma)\,\theta = \alpha\,(\sigma\theta)\quad .$$

Beweis. Durch strukturelle Induktion.

Ind.Basis: Für $\top$ und $\bot$ ist die Behauptung offensichtlich.

Sei nun $R(t_1, \ldots, t_n)$ eine atomare Formel der Signatur S. Es gilt dann

$$\begin{aligned}[R(t_1, \ldots, t_n)]\,(\sigma\theta) &= R(t_1\,(\sigma\theta), \ldots, t_n\,(\sigma\theta)) \\ &= R((t_1\sigma)\,\theta, \ldots, (t_n\sigma)\,\theta) \\ &= [R(t_1\sigma, \ldots, t_n\sigma)]\,\theta \\ &= [[R(t_1, \ldots, t_n)]\,\sigma]\,\theta.\end{aligned}$$

Damit ist die Induktionsbasis sichergestellt.

Ind.Schritt: Wir betrachten exemplarisch nur den Fall der Quantifikation $(\forall x)\,\alpha$. Der Fall $(\exists x)\,\alpha$ ist analog. Die verbleibenden Fälle sind weitestgehend trivial.

Sei also σ frei für $(\forall x)\,\alpha$ und θ frei für $[(\forall x)\,\alpha]\,\sigma$.

Es gilt:

ist σ frei für $(\forall x)\,\alpha$, so ist σ_x frei für α, und

ist θ frei für $[(\forall x)\,\alpha]\,\sigma$, so ist θ_x frei für $\alpha\sigma_x$.

Nach Induktionsvoraussetzung gilt

$$[\alpha\sigma_x]\,\theta_x = \alpha\,(\sigma_x\theta_x)\,.$$

Nun gilt $\sigma_x\theta_x\,(y) = (\sigma\theta)_x\,(y)$ für y mit $x \notin var\,(y\sigma)$:

denn für $y = x$ gilt:

$$\begin{aligned}(\sigma_x\theta_x)\,(x) &= x\theta_x\\ &= x\\ &= (\sigma\theta)_x\,(x)\end{aligned}$$

und für $y \neq x$ und $x \notin var\,(y\sigma)$ gilt:

$$\begin{aligned}(\sigma_x\theta_x)\,(y) &= [y\sigma]\,\theta_x \quad \text{da } \sigma_x\,(y) = \sigma\,(y)\\ &= [y\sigma]\,\theta \quad \text{da } x \notin var\,(y\sigma)\\ &= y\,(\sigma\theta) \quad \text{nach Lemma 2.23}\\ &= y\,(\sigma\theta)_x \quad \text{da } (\sigma\theta)\,(y) = (\sigma\theta)_x\,(y)\end{aligned}$$

Damit folgt dann:

$$\begin{aligned}[[(\forall x)\,\alpha]\,\sigma]\,\theta &= [(\forall x)\,(\alpha\sigma_x)]\,\theta\\ &= (\forall x)\,[(\alpha\sigma_x)\,\theta_x]\\ &= (\forall x)\,[\alpha\,(\sigma_x\theta_x)] \quad \text{nach (IV)}\\ &= (\forall x)\,[\alpha\,(\sigma\theta)_x]\,, \text{da } \sigma \text{ frei für } (\forall x)\,\alpha, \text{ ist } x \notin var\,(y\sigma) \text{ für } y \in fvar\,(\alpha)\\ &= [(\forall x)\,\alpha]\,(\sigma\theta)\,.\end{aligned}$$

■

2.1.3 Semantik der Prädikatenlogik

Unsere prädikatenlogischen Formeln dienen zur symbolischen Beschreibung von Aussagen über bzw. in Strukturen. Um solchen Formeln einen Wahrheitswert zuordnen zu können, ist daher eine „Deutung" der in der Formel vorkommenden Komponenten (d.h. im Wesentlichen der Symbole) erforderlich. Diese Deutung geschieht nun in Strukturen, die zu der Signatur passen.

Definition 2.34 *Sei $S = (\mathcal{R};\mathcal{F};\mathcal{K};ar)$ eine Signatur.*
*Eine S-*Struktur *$\mathcal{A}$ ist ein Tupel $\mathcal{A} = \left(A;\mathcal{R}^{\mathcal{A}};\mathcal{F}^{\mathcal{A}};\mathcal{K}^{\mathcal{A}}\right)$ mit*

1. *A nichtleere Menge*
2. *$\mathcal{R}^{\mathcal{A}} = \left(R^{\mathcal{A}}\right)_{R\in\mathcal{R}}$ mit $R^{\mathcal{A}}$ ist $ar\,(R)$-stellige Relation über A*
3. *$\mathcal{F}^{\mathcal{A}} = \left(f^{\mathcal{A}}\right)_{f\in\mathcal{F}}$ mit $f^{\mathcal{A}}$ ist $ar\,(f)$-stellige Funktion von A nach A*
4. *$\mathcal{K}^{\mathcal{A}} = \left(c^{\mathcal{A}}\right)_{c\in\mathcal{K}}$ mit $c^{\mathcal{A}} \in A$.*

A heißt Grundbereich *(auch Träger- oder Individuenbereich) der Struktur $\mathcal{A}$, und wird auch als $|\mathcal{A}|$ notiert.*

Im Falle, dass in der Signatur S keine Funktions- oder Konstantensymbole enthalten sind, werden wir in einer entsprechenden S-Struktur die leere Folge der Funktionen bzw. Konstanten mit zugehörigem Trennsymbol „;" nicht aufführen.

Beispiel 2.35 $Nat := \left(\mathbb{N}; <^{Nat}, =^{Nat}; S^{Nat}, +^{Nat}, \cdot^{Nat}\right)$ *ist eine Struktur zur Signatur* $S = (<, \widetilde{=}; S, +, \cdot)$.
$Real := \left(\mathbb{R}; <^{Real}, =^{Real}; S^{Real}, +^{Real}, \cdot^{Real}\right)$ *ist eine Struktur desselben Typs.*

Definition 2.36 *Eine* $\mathcal{A}$-Belegung (*meist einfach Belegung*) in einer S-Struktur $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ *ist eine Abbildung* $w : Var_{PL} \to A$.
Statt $w(x)$ *schreiben wir auch* x^w

Über Belegungen werden wir den Termen Objekte aus dem Grundbereich der Struktur zuordnen.

Definition 2.37 *Sei* $S = (\mathcal{R}; \mathcal{F}; \mathcal{K}; ar)$ *eine Signatur und* $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ *eine* S*-Struktur.* w *sei eine* $\mathcal{A}$*-Belegung.*
Der Wert $t^{\mathcal{A},w}$ *eines Terms* $t \in Tm_{PL}(S)$ *in der Struktur* $\mathcal{A}$ *und unter der Belegung* w *ist folgendermaßen definiert:*

1. $c^{\mathcal{A},w} := c^{\mathcal{A}}$ *für* $c \in \mathcal{K}$
2. $x^{\mathcal{A},w} := w(x)$ *für* $x \in Var_{PL}$
3. $[f(t_1, \ldots, t_n)]^{\mathcal{A},w} := f^{\mathcal{A}}\left(t_1^{\mathcal{A},w}, \ldots, t_n^{\mathcal{A},w}\right)$

Bemerkung 2.38 *Für Grundterme ist* $t^{\mathcal{A},w}$ *von* w *unabhängig, Wir schreiben in diesem Fall auch* $t^{\mathcal{A}}$.

Relativ zu einer gegebenen Struktur und zu einer Belegung wollen wir einer Formel einen Wahrheitswert zuweisen.

Definition 2.39 *Sei* w *eine* $\mathcal{A}$*-Belegung.*

Eine $\mathcal{A}$*-Belegung* w' *heißt* x-Variante *von* w (*notiert als* $w' \stackrel{(x)}{=} w$), *falls sich* w *und* w' *höchstens in der Belegung von* x *unterscheiden, d.h. es gilt* $w'(y) = w(y)$ *für alle* $y \in Var_{PL}$ *mit* $y \neq x$.

Definition 2.40 *Sei* $S = (\mathcal{R}; \mathcal{F}; \mathcal{K}; ar)$ *eine Signatur und* $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ *eine* S*-Struktur.* w *sei eine* $\mathcal{A}$*-Belegung.*
Jeder Formel $\alpha \in Fm_{PL}(S)$ *wird in* $\mathcal{A}$ *und unter der Belegung* w *ein* Wahrheitswert $\alpha^{\mathcal{A},w}$ *in der folgenden Weise zugeordnet:*

1. *Für atomare Formeln:*
   $[R(t_1, \ldots, t_n)]^{\mathcal{A},w} = T \;:\Leftrightarrow\; \left(t_1^{\mathcal{A},w}, \ldots, t_n^{\mathcal{A},w}\right) \in R^{\mathcal{A}}$
   $\top^{\mathcal{A},w} := T$
   $\bot^{\mathcal{A},w} := F$
2. $[\neg\alpha]^{\mathcal{A},w} = T \;:\Leftrightarrow\; \alpha^{\mathcal{A},w} = F$
3. $(\alpha \wedge \beta)^{\mathcal{A},w} := \begin{cases} T & \textit{falls } \alpha^{\mathcal{A},w} = \beta^{\mathcal{A},w} = T \\ F & \textit{sonst} \end{cases}$

   $(\alpha \vee \beta)^{\mathcal{A},w} := \begin{cases} T & \textit{falls } \alpha^{\mathcal{A},w} = T \textit{ oder } \beta^{\mathcal{A},w} = T \\ F & \textit{sonst} \end{cases}$

   $(\alpha \to \beta)^{\mathcal{A},w} := \begin{cases} T & \textit{falls } \alpha^{\mathcal{A},w} = F \textit{ oder } \beta^{\mathcal{A},w} = T \\ F & \textit{sonst} \end{cases}$

4. $[(\forall x)\,\alpha]^{\mathcal{A},w} = T \;:\Leftrightarrow\; \alpha^{\mathcal{A},w'} = T$ *für alle x-Varianten w' von w.*

5. $[(\exists x)\,\alpha]^{\mathcal{A},w} = T \;:\Leftrightarrow$ *es gibt eine x-Variante w' von w mit $\alpha^{\mathcal{A},w'} = T$.*

Gilt $\alpha^{\mathcal{A},w} = T$, so sagen wir, dass α wahr unter der Belegung w in $\mathcal{A}$ ist. In diesem Fall heißt $(\mathcal{A}, w)$ auch Modell *von α (notiert als $\mathcal{A}, w \models_{PL} \alpha$).*

Enthält α keine freien Variablen, so ist $\alpha^{\mathcal{A},w}$ unabhängig von w (vgl. das prädikatenlogische Koinzidenzlemma 2.49). In diesem Fall notieren wir den Wahrheitswert von α durch $\alpha^{\mathcal{A}}$.

Notation 2.41 *Das Paar $(\mathcal{A}, w)$, bestehend aus einer S-Struktur $\mathcal{A}$ und einer $\mathcal{A}$-Belegung w, wird als* Interpretation *bezeichnet. Ist $I = (\mathcal{A}, w)$ eine Interpretation, so schreiben wir statt $\alpha^{\mathcal{A},w}$ auch α^I. Im Falle $\alpha^I = T$ notieren wir dies auch durch $I \models_{PL} \alpha$.*

Definition 2.42 *Sei $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ eine S-Struktur.*

- *α heißt* allgemeingültig *(oder gültig; notiert als $\mathcal{A} \models_{PL} \alpha$)* in $\mathcal{A}$, *falls für alle $\mathcal{A}$-Belegungen w $\alpha^{\mathcal{A},w} = T$.*
- *α heißt* allgemeingültig *(oder gültig; Notation: $\models_{PL} \alpha$), falls α in allen S-Strukturen $\mathcal{A}$ allgemeingültig ist.*
- *α heißt* erfüllbar in $\mathcal{A}$, *falls eine $\mathcal{A}$-Belegung w mit $\alpha^{\mathcal{A},w} = T$ existiert.*
- *α heißt* erfüllbar, *falls eine S-Struktur $\mathcal{A}$ existiert, so dass α in $\mathcal{A}$ erfüllbar ist.*
- *Eine Menge Σ von Formeln heißt erfüllbar, falls eine S-Struktur $\mathcal{A}$ und eine $\mathcal{A}$-Belegung w existiert, so dass $\alpha^{\mathcal{A},w} = T$ für alle $\alpha \in \Sigma$. In diesem Fall heißt die Interpretation $I = (\mathcal{A}, w)$ auch* Modell *von Σ.*

Beispiel 2.43 *Es sei die Signatur $S = (R; \oplus)$ mit $\mathrm{ar}(R) = \mathrm{ar}(\oplus) = 2$ gegeben. Wir werden im Folgenden für $\oplus$ die Infix-Schreibweise verwenden.*

1. *Es sei die prädikatenlogische Formel $\alpha = (\exists y)\,R(x, y \oplus y)$ gegeben sowie die S-Struktur*

$$\mathcal{A} = \left(\mathbb{N} \setminus \{0\}\,; R^{\mathcal{A}}; \oplus^{\mathcal{A}}\right)$$

mit $R^{\mathcal{A}}$ gleich der Gleichheitsrelation auf $|\mathcal{A}|$ (notiert durch $\cong$) und $\oplus^{\mathcal{A}} = +$.

Sei nun w eine $\mathcal{A}$-Belegung. Es gilt dann:

$$\begin{aligned}
\alpha^{\mathcal{A},w} = T \;&\Leftrightarrow\; [(\exists y)\,R(x, y \oplus y)]^{\mathcal{A},w} = T\\
&\Leftrightarrow\; \textit{es ex. } y\textit{-Variante } w' \textit{ von } w \textit{ mit } [R(x, y \oplus y)]^{\mathcal{A},w'} = T\\
&\Leftrightarrow\; \textit{es ex. } w' \textit{ mit } w' \overset{(y)}{=} w \textit{ und } \left(x^{\mathcal{A},w'}, [y \oplus y]^{\mathcal{A},w'}\right) \in R^{\mathcal{A}}\\
&\Leftrightarrow\; \textit{es ex. } w' \textit{ mit } w' \overset{(y)}{=} w \textit{ und } w'(x) \cong y^{\mathcal{A},w'} \oplus^{\mathcal{A}} y^{\mathcal{A},w'}\\
&\Leftrightarrow\; \textit{es ex. } w' \textit{ mit } w' \overset{(y)}{=} w \textit{ und } w'(x) \cong w'(y) + w'(y)\\
&\Leftrightarrow\; \textit{es ex. } w' \textit{ mit } w' \overset{(y)}{=} w \textit{ und } w(x) \cong w'(y) + w'(y)\,, \textit{ da } w(x) \cong w'(x) \textit{ für } x \neq y\\
&\Leftrightarrow\; \textit{es ex. } a \in \mathbb{N} \setminus \{0\} \textit{ mit } w(x) \cong a + a\\
&\Leftrightarrow\; w(x) \textit{ gerade.}
\end{aligned}$$

2. *Es sei nun die prädikatenlogische Formel $\alpha = (\forall x)\,(\exists y)\,(\exists z)\,R(x \oplus y, z)$ und die S-Struktur $\mathcal{B} = \left(\mathbb{N} \setminus \{0\}\,; R^{\mathcal{B}}; \oplus^{\mathcal{B}}\right)$ mit $R^{\mathcal{B}} = >$ und $\oplus^{\mathcal{B}} = +$ gegeben.*

Es gilt dann

$\alpha^{\mathcal{B},w} = T$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ist* $[(\exists y)(\exists z)\, R(x \oplus y, z)]^{\mathcal{B},w'} = T$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ex.* w'' *mit* $w'' \overset{(y)}{=} w'$ *und* $[(\exists z)\, R(x \oplus y, z)]^{\mathcal{B},w''} = T$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ex.* w'' *mit* $w'' \overset{(y)}{=} w'$ *und ex.* w''' *mit* $w''' \overset{(z)}{=} w''$ *und* $[R(x \oplus y, z)]^{\mathcal{B},w'''} = T$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ex.* w'' *mit* $w'' \overset{(y)}{=} w'$ *und ex.* w''' *mit* $w''' \overset{(z)}{=} w''$ *und* $x^{\mathcal{B},w'''} \oplus^{\mathcal{B}} y^{\mathcal{B},w'''} > z^{\mathcal{B},w'''}$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ex.* w'' *mit* $w'' \overset{(y)}{=} w'$ *und ex.* w''' *mit* $w''' \overset{(z)}{=} w''$ *und* $w'''(x) + w'''(y) > w'''(z)$

$\Leftrightarrow$ *für alle* w' *mit* $w' \overset{(x)}{=} w$ *ex.* w'' *mit* $w'' \overset{(y)}{=} w'$ *und ex.* w''' *mit* $w''' \overset{(z)}{=} w''$ *und* $w'(x) + w''(y) > w'''(z)$

$\Leftrightarrow$ *für alle* $m \in \mathbb{N} \setminus \{0\}$ *existieren* $n, k \in \mathbb{N} \setminus \{0\}$*, so dass* $m + n > k$.

Also erhält $\alpha^{\mathcal{B},w}$ *den Wert* T *für eine beliebige* $\mathcal{B}$*-Belegung* w. α *ist somit allgemeingültig in* $\mathcal{B}$.

3. *Sei* α *wie im vorangegangenen Beispiel festgelegt. Wir betrachten die Struktur* $\mathcal{N} = \left(\mathbb{N}; >^{\mathcal{N}}; \cdot^{\mathcal{N}}\right)$.
   Es gilt dann für eine beliebige $\mathcal{N}$*-Belegung* w:
   $\alpha^{\mathcal{N},w} = T \;\Leftrightarrow$ *für alle* $m \in \mathbb{N}$ *existieren* $n, k \in \mathbb{N}$ *mit* $m \cdot n > k$.
   α *ist somit nicht allgemeingültig in* $\mathcal{N}$ *und damit auch nicht allgemeingültig.*
   $\neg\alpha$ *ist nach dem vorangegangenen Beispiel ebenfalls nicht allgemeingültig.*

2.1.4 Einfache Eigenschaften der Prädikatenlogik

Im folgenden werden wir einige einfache, aber doch sehr grundlegende und immer wieder benutzte Eigenschaften zeigen.

Definition 2.44 *Sei* $I = (\mathcal{A}, w)$ *eine Interpretation zu einer Signatur* S, $a_1, \ldots, a_n$ *seien Elemente von* $|\mathcal{A}|$.
$w\{x_1/a_1, \ldots, x_n/a_n\}$ *ist dann die folgendermaßen definierte Belegung*

$$w\{x_1/a_1, \ldots, x_n/a_n\}(x) := \begin{cases} w(x) & \text{falls } x \notin \{x_1, \ldots, x_n\} \\ a_i & \text{falls } x = x_i \text{ für ein } i \in \{1, \ldots, n\} \end{cases} .$$

$I\{x_1/a_1, \ldots, x_n/a_n\}$ *sei dann die Struktur* $I\{x_1/a_1, \ldots, x_n/a_n\} := (\mathcal{A}, w\{x_1/a_1, \ldots, x_n/a_n\})$.

Mittels dieser „Um-" Belegung bzw. dieser so definierten Struktur läßt sich dann die Semantikdefinition der Quantoren auch auf die folgende Weise charakterisieren.

Lemma 2.45 *Sei* $I = (\mathcal{A}, w)$ *eine Interpretation zur Signatur* S. *Dann gilt:*

1. $[(\forall x)\, \alpha]^I = T \;\Leftrightarrow$ *für alle* $a \in |\mathcal{A}|$ *gilt* $\alpha^{I\{x/a\}} = T$

2. $[(\exists x)\, \alpha]^I = T \;\Leftrightarrow$ *es gibt ein* $a \in |\mathcal{A}|$ *mit* $\alpha^{I\{x/a\}} = T$.

Beweis. Übungsaufgabe. ■

Eine Vielzahl prädikatenlogischer Formeln ist bereits aufgrund ihrer „aussagenlogischen Struktur" allgemeingültig:

Lemma 2.46 (Permanenz der Aussagenlogik in der Prädikatenlogik) *$\beta_1, \ldots, \beta_n \in Fm_{PL}(S)$ seien prädikatenlogische Formeln einer Signatur S, und $\alpha \in Fm_{AL}$ sei eine aussagenlogische Tautologie, in der höchstens die Aussagenvariablen $A_1, \ldots, A_n$ vorkommen.*
α^* sei diejenige prädikatenlogische Formel der Signatur S, die man aus α dadurch erhält, dass jedes Vorkommen von A_i durch β_i, $1 \le i \le n$, ersetzt wird.
Dann gilt: $\models_{PL} \alpha^*$.

Beweis. Skizze:
Für eine Interpretation $I = (\mathcal{A}, w)$ zur Signatur S definieren wir eine aussagenlogische Belegung I^* durch

$$I^*(A) := \begin{cases} \beta_i^I & \text{falls } A = A_i, \text{ für ein } i \in \{1, \ldots, n\} \\ F & \text{sonst.} \end{cases}$$

Über eine strukturelle Induktion über den Aufbau von α zeigt man dann (Übungsaufgabe)

$$\alpha^{I^*} = [\alpha^*]^I \ .$$

Sei nun I eine beliebige Interpretation zur Signatur S und α eine aussagenlogische Tautologie. Dann gilt:

$$[\alpha^*]^I = \alpha^{I^*} = T \ ,$$

da α eine aussagenlogische Tautologie ist. Es gilt somit $\models_{PL} \alpha^*$. ■

Definition 2.47 *Zwei prädikatenlogische Formeln $\alpha, \beta \in Fm_{PL}(S)$ zur Signatur S heißen* semantisch (prädikatenlogisch) äquivalent *(notiert als $\alpha \equiv_{PL} \beta$) genau dann, wenn $\models_{PL} \alpha \leftrightarrow \beta$.*

In dem nächsten Lemma werden die zwischen dem All- und dem Existenzquantor bestehenden Beziehungen wiedergegeben.

Lemma 2.48 (Verallgemeinerungen der de Morgan' schen Gesetze)
Sei $\alpha \in Fm_{PL}(S)$. Es gilt:

1. $\neg(\forall x)\,\alpha \equiv_{PL} (\exists x)\,\neg\alpha$
2. $\neg(\exists x)\,\alpha \equiv_{PL} (\forall x)\,\neg\alpha$
3. $(\forall x)\,\alpha \equiv_{PL} \neg(\exists x)\,\neg\alpha$
4. $(\exists x)\,\alpha \equiv_{PL} \neg(\forall x)\,\neg\alpha$

Beweis. Exemplarisch werden wir die Eigenschaft 2. zeigen. Sei dazu eine beliebige Interpretation $I = (\mathcal{A}, w)$ gegeben. Es gilt

$$\begin{aligned} [\neg(\exists x)\,\alpha]^{\mathcal{A},w} = T \ &\Leftrightarrow\ [(\exists x)\,\alpha]^{\mathcal{A},w} = F \\ &\Leftrightarrow\ \text{es gibt kein } a \in |\mathcal{A}| \text{ mit } \alpha^{\mathcal{A},w\{x/a\}} = T \\ &\Leftrightarrow\ \text{für alle } a \in |\mathcal{A}| \text{ gilt } \alpha^{\mathcal{A},w\{x/a\}} = F \\ &\Leftrightarrow\ \text{für alle } a \in |\mathcal{A}| \text{ gilt } [\neg\alpha]^{\mathcal{A},w\{x/a\}} = T \\ &\Leftrightarrow\ [(\forall x)\,\neg\alpha]^{\mathcal{A},w} = T. \end{aligned}$$

Da dies somit für eine beliebige Interpretation gilt, folgt $\models_{PL} \neg(\exists x)\,\alpha \leftrightarrow (\forall x)\,\neg\alpha$. ■

Das folgende Lemma besagt, dass für den Wahrheitswert einer Formel nur die Belegung der in dieser Formel frei vorkommenden Variablen von Relevanz ist.

Lemma 2.49 (Koinzidenzlemma) *Sei $\mathcal{A}$ eine S-Struktur für eine Signatur S, seien ferner $t \in Tm_{PL}(S)$ und $\alpha \in Fm_{PL}(S)$.*
w_1 und w_2 seien beliebige $\mathcal{A}$-Belegungen, die für alle Variablen von t und für alle freien Variablen von α übereinstimmen.
Dann gilt:

$$t^{\mathcal{A},w_1} = t^{\mathcal{A},w_2}$$
$$\alpha^{\mathcal{A},w_1} = \alpha^{\mathcal{A},w_2} \ .$$

Beweis. (Durch strukturelle Induktion)
Wir wollen für Formeln α 2 Fälle exemplarisch vorführen:

Fall $\alpha = R(t_1, \ldots, t_n)$: Es gilt

$$\begin{aligned} T = \alpha^{\mathcal{A},w_1} & \\ \Leftrightarrow \left(t_1^{\mathcal{A},w_1}, \ldots, t_n^{\mathcal{A},w_1}\right) \in R^{\mathcal{A}} & \\ \Leftrightarrow \left(t_1^{\mathcal{A},w_2}, \ldots, t_n^{\mathcal{A},w_2}\right) \in R^{\mathcal{A}} & \quad \text{nach dem Resultat für Terme} \\ \Leftrightarrow \alpha^{\mathcal{A},w_2} = T \ . & \end{aligned}$$

Fall $\alpha = (\forall x)\,\beta$: Es gilt

$$\begin{aligned} T &= \alpha^{\mathcal{A},w_1} \\ &\Leftrightarrow \text{ für alle } a \in |\mathcal{A}| \text{ gilt } \beta^{\mathcal{A},w_1\{x/a\}} = T \\ &\Leftrightarrow \text{ für alle } a \in |\mathcal{A}| \text{ gilt } \beta^{\mathcal{A},w_2\{x/a\}} = T \\ &\Leftrightarrow [(\forall x)\,\beta]^{\mathcal{A},w_2} = T. \end{aligned}$$

■

Lemma 2.50

1. $(\forall x)(\forall y)\,\alpha \equiv_{PL} (\forall y)(\forall x)\,\alpha$
2. $(\exists x)(\exists y)\,\alpha \equiv_{PL} (\exists y)(\exists x)\,\alpha$
3. $\models_{PL} (\exists x)(\forall y)\,\alpha \to (\forall y)(\exists x)\,\alpha$
4. $(\forall x)\,\alpha \equiv_{PL} \alpha$, *falls* $x \notin fvar(\alpha)$
5. $(\exists x)\,\alpha \equiv_{PL} \alpha$, *falls* $x \notin fvar(\alpha)$.

Beweis. Wir beweisen stellvertetend die Aussagen 3. und 4.

3. Wir betrachten den Fall $y \neq x$. Sei $I = (\mathcal{A}, w)$ eine beliebige Interpretation. Es gelte $[(\exists x)(\forall y)\,\alpha]^I = T$. Es gibt dann ein $a \in |\mathcal{A}|$ mit $[(\forall y)\,\alpha]^{I\{x/a\}} = T$. Damit gibt es dann weiter ein $a \in |\mathcal{A}|$, so dass für alle $b \in |\mathcal{A}|$ gilt $\alpha^{I\{x/a,\,y/b\}} = T$.

Da nun zu einem beliebigen $b \in |\mathcal{A}|$ gilt $\alpha^{I\{y/b\}\{x/a\}} = T$, folgt $[(\exists x)\,\alpha]^{I\{y/b\}} = T$ für ein beliebiges $b \in |\mathcal{A}|$. Damit folgt weiter $[(\forall y)(\exists x)\,\alpha]^I = T$. Dies war zu zeigen.

4. Sei wiederum $I = (\mathcal{A}, w)$ eine beliebige Interpretation.

Es gilt dann $[(\forall x)\,\alpha]^I = T$ genau dann, wenn $\alpha^{(\mathcal{A},w')} = T$ für alle x-Varianten w' von w, also nach dem Koinzidenzlemma genau dann, wenn $\alpha^{(\mathcal{A},w)} = T$.

■

Lemma 2.51

1. $(\forall x)(\alpha \wedge \beta) \equiv_{PL} (\forall x)\alpha \wedge (\forall x)\beta$
2. $(\exists x)(\alpha \vee \beta) \equiv_{PL} (\exists x)\alpha \vee (\exists x)\beta$
3. $(\forall x)(\alpha \vee \beta) \equiv_{PL} (\forall x)\alpha \vee \beta$, *falls* $x \notin fvar(\beta)$
4. $(\exists x)(\alpha \wedge \beta) \equiv_{PL} (\exists x)\alpha \wedge \beta$, *falls* $x \notin fvar(\beta)$.
5. $(\exists x)(\alpha \to \beta) \equiv_{PL} (\forall x)\alpha \to \beta$, *falls* $x \notin fvar(\beta)$.
6. $(\forall x)(\alpha \to \beta) \equiv_{PL} (\exists x)\alpha \to \beta$, *falls* $x \notin fvar(\beta)$.
7. $(\exists x)(\beta \to \alpha) \equiv_{PL} \beta \to (\exists x)\alpha$, *falls* $x \notin fvar(\beta)$.
8. $(\forall x)(\beta \to \alpha) \equiv_{PL} \beta \to (\forall x)\alpha$, *falls* $x \notin fvar(\beta)$.

Das Substitutionslemma

auch als Korrespondenz- oder Überführungslemma bezeichnet, wird häufig bei Beweisen benötigt.

Lemma 2.52 (Substitutionslemma) *Sei S eine Signatur und I eine Interpretation zur Signatur S.*

1. *Für beliebige Terme $t, t_1, \dots, t_n \in Tm_{PL}(S)$ gilt:*
$$[t\{x_1/t_1, \dots, x_n/t_n\}]^I = t^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}}$$
2. *Es sei $\alpha \in Fm_{PL}(S)$. Ferner sei die Substitution $\{x_1/t_1, \dots, x_n/t_n\}$ frei für α.*
Es gilt dann
$$[\alpha\{x_1/t_1, \dots, x_n/t_n\}]^I = \alpha^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}} .$$

Beweis. Beide Teilaussagen werden durch strukturelle Induktion bewiesen. Wir werden jeweils einige Fälle exemplarisch beweisen. Sei $I = (\mathcal{A}, w)$.

1. (a) Fall $t = x_i$ für ein i mit $1 \le i \le n$:
Es ist dann
$$[x_i\{x_1/t_1, \dots, x_n/t_n\}]^I = t_i^I = x_i^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}} .$$
(b) Fall $t = f(s_1, \dots, s_k)$:
Nach Induktionsvoraussetzung gilt dann
$$[s_j\{x_1/t_1, \dots, x_n/t_n\}]^I = s_j^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}}, \ 1 \le j \le k .$$
Damit folgt
$$\begin{aligned}
[[f(s_1, \dots, s_k)]\{x_1/t_1, \dots, x_n/t_n\}]^I &= [f(s_1\{x_1/t_1, \dots, x_n/t_n\}, \dots, s_k\{x_1/t_1, \dots, x_n/t_n\})]^I \\
&= f^{\mathcal{A}}\left([s_1\{x_1/t_1, \dots, x_n/t_n\}]^I, \dots, [s_k\{x_1/t_1, \dots, x_n/t_n\}]^I\right) \\
&= f^{\mathcal{A}}\left(s_1^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}}, \dots, s_k^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}}\right) \\
&= [f(s_1, \dots, s_k)]^{I\left\{x_1/t_1^I, \dots, x_n/t_n^I\right\}} .
\end{aligned}$$

2. (a) $\alpha = R(s_1, \ldots, s_k)$:
Es gilt dann

$$\begin{aligned} T = {} & [\alpha\{x_1/t_1, \ldots, x_n/t_n\}]^I \\ & \Leftrightarrow [R(s_1\{x_1/t_1, \ldots, x_n/t_n\}, \ldots, s_k\{x_1/t_1, \ldots, x_n/t_n\})]^I = T \\ & \Leftrightarrow \left([s_1\{x_1/t_1, \ldots, x_n/t_n\}]^I, \ldots, [s_k\{x_1/t_1, \ldots, x_n/t_n\}]^I\right) \in R^{\mathcal{A}} \\ & \Leftrightarrow \left(s_1^{I\left\{x_1/t_1^I, \ldots, x_n/t_n^I\right\}}, \ldots, s_k^{I\left\{x_1/t_1^I, \ldots, x_n/t_n^I\right\}}\right) \in R^{\mathcal{A}} \quad \text{nach (1)} \\ & \Leftrightarrow [R(s_1, \ldots, s_k)]^{I\left\{x_1/t_1^I, \ldots, x_n/t_n^I\right\}} = T\ . \end{aligned}$$

(b) $\alpha = (\forall x)\,\beta$ und $x = x_i$ für ein $i,\ 1 \leq i \leq n$:
Es gilt (man beachte, dass die Substitution $\{x_1/t_1, \ldots, x_n/t_n\}$ frei für α ist)

$$\begin{aligned} T = {} & [\alpha\{x_1/t_1, \ldots, x_n/t_n\}]^I \\ & \Leftrightarrow [(\forall x_i)\,\beta\{x_1/t_1, \ldots, x_i/x_i, \ldots, x_n/t_n\}]^I = T \\ & \Leftrightarrow \text{für alle } a \in |\mathcal{A}|\ \ [\beta\{x_1/t_1, \ldots, x_i/x_i, \ldots, x_n/t_n\}]^{I\{x_i/a\}} = T \\ & \Leftrightarrow \text{für alle } a \in |\mathcal{A}|\ \ \beta^{I\left\{x_1/t_1^I, \ldots, x_i/a, \ldots, x_n/t_n^I\right\}} = T \\ & \Leftrightarrow [(\forall x)\,\beta]^{I\left\{x_1/t_1^I, \ldots, x_n/t_n^I\right\}} = T\ . \end{aligned}$$

(c) $\alpha = (\forall x)\,\beta$ und $x \neq x_i,\ 1 \leq i \leq n$:
Es gilt

$$\begin{aligned} T = {} & [\alpha\{x_1/t_1, \ldots, x_n/t_n\}]^I \\ & \Leftrightarrow \text{für alle } a \in |\mathcal{A}|\ \ [\beta\{x_1/t_1, \ldots, x_n/t_n\}]^{I\{x/a\}} = T \\ & \Leftrightarrow \text{für alle } a \in |\mathcal{A}|\ \ \beta^{I\left\{x/a, x_1/t_1^{I\{x/a\}}, \ldots, x_n/t_n^{I\{x/a\}}\right\}} = T \quad \text{nach IV} \\ & \Leftrightarrow \text{für alle } a \in |\mathcal{A}|\ \ \beta^{I\left\{x/a, x_1/t_1^I, \ldots, x_n/t_n^I\right\}} = T \quad \text{wegen Koinzidenzlemma} \\ & \Leftrightarrow [(\forall x)\,\beta]^{I\left\{x_1/t_1^I, \ldots, x_n/t_n^I\right\}} = T \end{aligned}$$

■

Als Folgerung aus dem Substitutionslemma ergibt sich das Lemma über gebundene Umbenennung:

Lemma 2.53 (Gebundene Umbenennung) *Sei $\alpha \in Fm_{PL}(S)$ und $x, y \notin fvar(\alpha)$. Ferner seien die beiden Substitutionen $\{z/x\}$ und $\{z/y\}$ frei für α. Dann gilt:*

1. $(\exists x)\,[\alpha\{z/x\}] \equiv_{PL} (\exists y)\,[\alpha\{z/y\}]$
2. $(\forall x)\,[\alpha\{z/x\}] \equiv_{PL} (\forall y)\,[\alpha\{z/y\}]$

Beweis.

1. Sei $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S. Es gilt

$$\begin{aligned} & [(\exists x)\,[\alpha\{z/x\}]]^I = T \\ & \quad \Leftrightarrow \ \text{es ex. } a \in |\mathcal{A}| \text{ mit } [\alpha\{z/x\}]^{I\{x/a\}} = T \\ & \quad \Leftrightarrow \ \text{es ex. } a \in |\mathcal{A}| \text{ mit } \alpha^{I\{z/a\}} = T \quad \text{(nach Subst.Lemma)} \\ & \quad \Leftrightarrow \ \text{es ex. } a \in |\mathcal{A}| \text{ mit } [\alpha\{z/y\}]^{I\{y/a\}} = T \\ & \quad \Leftrightarrow \ [(\exists y)\,[\alpha\{z/y\}]]^I = T. \end{aligned}$$

2. Analog.

■

2.1.5 Invarianzeigenschaften von Strukturhomomorphismen

Bei der Untersuchung algebraischer Strukturen ist Isomorphie von algebraischen Strukturen eine fundamentale Eigenschaft. Isomorphe Strukturen werden in der Algebra identifziert, da sie dieselben Struktureigenschaften aufweisen. Wir wollen hier Isomorphie und allgemeiner Homomorphie von Strukturen betrachten und untersuchen. Der hierbei eingeführte Homomorphiebegriff verallgemeinert den aus der Algebra bekannten Homomorphiebegriff.

Im Folgenden werden wir häufig von der folgenden Notation Gebrauch machen.

Notation 2.54 *Sei $S = (\mathcal{R};\mathcal{F};\mathcal{C})$ eine Signatur. Statt $R \in \mathcal{R}$ schreiben wir auch $R \in S$. Analoges gilt für Funktions- und Konstantensymbole.*
Ist $R^{\mathcal{A}}$ eine n-stellige Relation über dem Grundbereich $|\mathcal{A}|$ und sind $a_1, \ldots, a_n$ Elemente des Grundbereichs $|\mathcal{A}|$, so schreiben wir statt $(a_1, \ldots, a_n) \in R^{\mathcal{A}}$ auch $R^{\mathcal{A}} a_1, \ldots, a_n$.

Definition 2.55 *Sei S eine Signatur, $\mathcal{A}$ und $\mathcal{B}$ seien S-Strukturen.*

1. *Eine Abbildung $\iota : |\mathcal{A}| \to |\mathcal{B}|$ heißt* Homomorphismus *von $\mathcal{A}$ nach $\mathcal{B}$ genau dann, wenn gilt:*

   (a) für $R \in S$ mit $\mathrm{ar}\,(R) = n$ und $a_1, \ldots, a_n \in |\mathcal{A}|$:
   $$R^{\mathcal{A}} a_1 \ldots a_n \Leftrightarrow R^{\mathcal{B}} \iota\,(a_1) \ldots \iota\,(a_n)$$

   (b) für $f \in S$ mit $\mathrm{ar}\,(f) = n$ und $a_1, \ldots, a_n \in |\mathcal{A}|$:
   $$\iota\left(f^{\mathcal{A}}\,(a_1, \ldots, a_n)\right) = f^{\mathcal{B}}\,(\iota\,(a_1), \ldots, \iota\,(a_n))$$

   (c) für $c \in S$:
   $$\iota\left(c^{\mathcal{A}}\right) = c^{\mathcal{B}} \quad .$$

2. *ι heißt* Isomorphismus *von $\mathcal{A}$ auf $\mathcal{B}$ genau dann, wenn die Abbildung $\iota : |\mathcal{A}| \to |\mathcal{B}|$ bijektiv und ein Homomorphismus von $\mathcal{A}$ nach $\mathcal{B}$ ist.*

3. *$\mathcal{A}$ und $\mathcal{B}$ heißen* isomorph *(notiert durch $\mathcal{A} \simeq \mathcal{B}$) genau dann, wenn es einen Isomorphismus von $\mathcal{A}$ auf $\mathcal{B}$ gibt.*

4. *$\mathcal{A}$ heißt* Substruktur *von $\mathcal{B}$ (notiert durch $\mathcal{A} \subset \mathcal{B}$) genau dann, wenn die Identität $\mathrm{id}_{\mathcal{A}} : |\mathcal{A}| \to |\mathcal{B}|$, definiert durch*
   $$\mathrm{id}_{\mathcal{A}}\,(a) := a \qquad \text{für } a \in |\mathcal{A}|\,,$$
   ein Homomorphismus von $\mathcal{A}$ nach $\mathcal{B}$ ist.

Wir beweisen zunächst ein Lemma, mit dem wir dann die zentralen Eigenschaften von isomorphen Strukturen und Substrukturen zeigen können.

Lemma 2.56 *Seien $\mathcal{A}$ und $\mathcal{B}$ S-Strukturen.*
Ist ι ein Isomorphismus (Homomorphismus) von $\mathcal{A}$ auf (bzw. nach) $\mathcal{B}$ und ist w eine $\mathcal{A}$-Belegung, dann gilt:

1. *für Terme $t \in Tm_{PL}\,(S)$:*
   $$\iota\left(t^{\mathcal{A},w}\right) = t^{\mathcal{B},\iota\circ w} \quad ,$$

2. *für alle (quantorenfreien) $\alpha \in Fm_{PL}\,(S)$:*
   $$\mathcal{A}, w \models_{PL} \alpha \ \Leftrightarrow\ \mathcal{B}, \iota \circ w \models_{PL} \alpha$$

(Man beachte, dass die Verkettung $\iota \circ w$ von ι und w wiederum eine Belegung, die „Bildbelegung", ist.)

Beweis. $\mathcal{A}, \mathcal{B}$ seien S-Strukturen, ι ein Isomorphismus (bzw. Homomorphismus) von $\mathcal{A}$ auf (bzw. nach) $\mathcal{B}$, w eine $\mathcal{A}$-Belegung.
Beide Aussagen werden durch strukturelle Induktion bewiesen.
Wir werden jeweils exemplarisch einige Fälle betrachten.

1. (a) $t = x$:
   $\iota\left(x^{\mathcal{A},w}\right) = \iota\left(w\left(x\right)\right) = x^{\mathcal{B},\iota\circ w}$.

   (b) $t = f\left(t_1, \ldots, t_n\right)$:

$$\begin{aligned}\iota\left(t^{\mathcal{A},w}\right) &= \iota\left(f^{\mathcal{A}}\left(t_1^{\mathcal{A},w}, \ldots, t_n^{\mathcal{A},w}\right)\right)\\ &= f^{\mathcal{B}}\left(\iota\left(t_1^{\mathcal{A},w}\right), \ldots, \iota\left(t_n^{\mathcal{A},w}\right)\right)\\ &= f^{\mathcal{B}}\left(t_1^{\mathcal{B},\iota\circ w}, \ldots, t_n^{\mathcal{B},\iota\circ w}\right) \quad \text{(nach IV)}\\ &= \left[f\left(t_1, \ldots, t_n\right)\right]^{\mathcal{B},\iota\circ w} \ .\end{aligned}$$

2. (a) $\alpha = R\left(t_1, \ldots, t_n\right)$:

$$\begin{aligned}\alpha^{\mathcal{A},w} &= T\\ &\Leftrightarrow\ R^{\mathcal{A}}\left(t_1^{\mathcal{A},w}, \ldots, t_n^{\mathcal{A},w}\right)\\ &\Leftrightarrow\ R^{\mathcal{B}}\left(\iota\left(t_1^{\mathcal{A},w}\right), \ldots, \iota\left(t_n^{\mathcal{A},w}\right)\right)\\ &\Leftrightarrow\ R^{\mathcal{B}}\left(t_1^{\mathcal{B},\iota\circ w}, \ldots, t_n^{\mathcal{B},\iota\circ w}\right) \quad \text{(nach Teil (1))}\\ &\Leftrightarrow\ \left[R\left(t_1, \ldots, t_n\right)\right]^{\mathcal{B},\iota\circ w} = T \ .\end{aligned}$$

   (b) Nur für Isomorphismus:
   $\alpha = (\forall x)\,\beta$: (entsprechend $(\exists x)\,\beta$)

$$\begin{aligned}\alpha^{\mathcal{A},w} &= T\\ &\Leftrightarrow\ \text{für alle } a \in |\mathcal{A}| : \beta^{\mathcal{A},w\{x/a\}} = T\\ &\Leftrightarrow\ \text{für alle } a \in |\mathcal{A}| : \beta^{\mathcal{B},\iota\circ(w\{x/a\})} = T \quad \text{(nach IV)}\\ &\Leftrightarrow\ \text{für alle } a \in |\mathcal{A}| : \beta^{\mathcal{B},(\iota\circ w)\{x/\iota(a)\}} = T\\ &\Leftrightarrow\ \text{für alle } b \in |\mathcal{B}| : \beta^{\mathcal{B},(\iota\circ w)\{x/b\}} = T \quad \text{(da } \iota \text{ bijektiv)}\\ &\Leftrightarrow\ \left[(\forall x)\,\beta\right]^{\mathcal{B},\iota\circ w} = T \ .\end{aligned}$$

■

Mit diesem Lemma ergibt sich dann leicht:

Theorem 2.57 (Isomorphiesatz) *Sind $\mathcal{A}$ und $\mathcal{B}$ S-Strukturen mit $\mathcal{A} \simeq \mathcal{B}$, dann gilt für Aussagen $\alpha \in Fm_{PL}(S)$:*

$$\mathcal{A} \models_{PL} \alpha \ \Leftrightarrow\ \mathcal{B} \models_{PL} \alpha \ .$$

Beweis. Sei ι ein Isomorphismus von $\mathcal{A}$ auf $\mathcal{B}$.
Es gilt dann:
$\mathcal{A} \models_{PL} \alpha$
$\Leftrightarrow$ $\mathcal{A}, w \models_{PL} \alpha$ für alle $\mathcal{A}$-Belegungen w
$\Leftrightarrow$ $\mathcal{A}, w \models_{PL} \alpha$ für eine $\mathcal{A}$-Belegung w (nach Koinzidenzlemma)
$\Leftrightarrow$ $\mathcal{B}, \iota\circ w \models_{PL} \alpha$ für eine $\mathcal{A}$-Belegung w (nach Lemma 2.56)
$\Leftrightarrow$ $\mathcal{B}, w' \models_{PL} \alpha$ für eine $\mathcal{B}$-Belegung w' (da ι bijektiv)
$\Leftrightarrow$ $\mathcal{B} \models_{PL} \alpha$ (nach Koinzidenzlemma, da α Aussage). ■

Bemerkung 2.58 *Man könnte vermuten, dass die Umkehrung des Isomorphiesatzes ebenfalls gilt: Sind in zwei S-Strukturen (für eine Signatur S) genau dieselben Formeln allgemeingültig, so sind diese beiden Strukturen bereits isomorph.*
Wir werden später einen einfachen Beweis dafür erhalten, dass diese Vermutung falsch ist. Zwei Strukturen, für die die diese Vermutung nicht gilt, sind beispielsweise $Real := (\mathbb{R}; =; +; 0)$ *und* $Rat := (\mathbb{Q}; =; +; 0)$*, wobei* $=$ *die Gleichheitsrelation und* $+$ *die Addition über dem jeweiligen Grundbereich bezeichnet. Bezeichnet man nämlich für eine S-Struktur* $\mathcal{A}$ *mit* $TH(\mathcal{A})$ *die Menge der in* $\mathcal{A}$ *allgemeingültigen Formeln, d.h.* $TH(\mathcal{A}) := \{\alpha \in Fm_{PL}(S) \mid \mathcal{A} \models_{PL} \alpha\}$*, so gilt* $TH(Real) = TH(Rat)$*. Real und Rat sind aber sicher nicht isomorph.*

Der Isomorphiesatz wird vielfach dazu verwendet, zu zeigen, dass bestimmte Eigenschaften in Strukturen nicht ausdrückbar sind.

Beispiel 2.59

1. *Wir betrachten Strukturen* $\mathcal{A} := \left(\mathbb{R}; >; f^{\mathcal{A}}\right)$ *mit einstelligen Funktionen* $f^{\mathcal{A}}$*. S sei die zugehörige Signatur. Es gibt dann keine Aussage* $\alpha \in Fm_{PL}(S)$*, so dass* α *in genau den Strukturen* $\mathcal{A}$ *mit* $\mathbb{R}$ *als Grundbereich und der* $>$*-Relation allgemeingültig ist, in denen* $f^{\mathcal{A}}$ *positiv ist, d.h. es gibt keine Aussage* α*, so dass gilt:*

   $$\mathcal{A} \models_{PL} \alpha \;\Leftrightarrow\; f^{\mathcal{A}}(b) > 0 \text{ für alle } b \in \mathbb{R}.$$

   Angenommen, es gäbe eine solche Aussage. $\mathcal{A}$ *sei eine solche Struktur mit positivem* $f^{\mathcal{A}}$*. Sei* $k := f^{\mathcal{A}}(0)$*. Betrachte dann die Struktur* $\mathcal{B} := \left(\mathbb{R}; >; f^{\mathcal{B}}\right)$ *mit*

   $$f^{\mathcal{B}}(b) := f^{\mathcal{A}}(b+k) - k$$

   für alle $b \in \mathbb{R}$*.* $f^{\mathcal{B}}$ *ist dann sicherlich nicht positiv, d.h. es gibt* $b \in \mathbb{R}$*, so daß* $f^{\mathcal{B}}(b) > 0$ *nicht gilt. Die Abbildung* $\iota : \mathbb{R} \to \mathbb{R}$ *mit*

   $$\iota(b) := b - k$$

   für alle $b \in \mathbb{R}$ *ist aber ein Isomorphismus von* $\mathcal{A}$ *nach* $\mathcal{B}$*. Es gilt nämlich für beliebige* $a, b \in \mathbb{R}$ *:*

   $$a > b \Leftrightarrow \iota(a) > \iota(b)$$

   und

   $$\iota\left(f^{\mathcal{A}}(b)\right) = f^{\mathcal{A}}(b) - k = f^{\mathcal{A}}(\iota(b) + k) - k = f^{\mathcal{B}}(\iota(b)).$$

   Da ι *ein Isomorphismus ist, ist* α *somit in der Struktur* $\mathcal{B}$ *genau dann allgemeingültig, wenn* α *in der Struktur* $\mathcal{A}$ *allgemeingültig ist. Nach Annahme über* $\mathcal{A}$ *ist daher* α *allgemeingültig in* $\mathcal{B}$*. Somit müsste nach Annahme über* α *die Funktion* $f^{\mathcal{B}}$ *positiv sein. Wir erhalten also einen Widerspruch.*

2. *In der Struktur* $(\mathbb{R}; =; +, \cdot; 0)$ *ist die zweistellige* $<$*-Relation* elementar definierbar*, d.h. es gibt eine Formel* α *der zu* $\mathcal{A}$ *passenden Signatur mit genau 2 (Stelligkeit von* $<$*) freien Variablen, z.B.* $fvar(\alpha) = \{x, y\}$*, und der Eigenschaft*

   $$(\mathbb{R}; =; +, \cdot; 0), w \models_{PL} \alpha \;\Leftrightarrow\; w(x) < w(y)$$

   für alle $\mathcal{A}$*-Belegungen* w*. Hingegen ist die* $<$*-Relation in der Struktur* $(\mathbb{R}; =; +; 0)$ *nicht elementar definierbar. Überlegen Sie sich dazu einen Beweis mit Hilfe des Isomorphiesatzes.*

Wir wollen nun als Nächstes klären, welche Beziehungen zwischen einer Struktur und einer Substruktur (häufig auch Unterstruktur genannt) bestehen.

Definition 2.60 *Eine Formel $\alpha \in Fm_{PL}(S)$ heißt* universal, *falls α eine Aussage von der Gestalt*

$$\alpha = (\forall x_1) \dots (\forall x_n)\, \beta$$

mit quantorenfreiem β ist. Ist α von der Gestalt

$$\alpha = (\exists x_1) \dots (\exists x_n)\, \beta$$

mit quantorenfreiem β, so heißt α existential.

Theorem 2.61 (Substruktursatz) *Sind $\mathcal{A}$ und $\mathcal{B}$ S-Strukturen und ist $\mathcal{A}$ Substruktur von $\mathcal{B}$, dann gilt für universale Aussagen $\alpha \in Fm_{PL}(S)$*

$$\mathcal{B} \models_{PL} \alpha \;\Rightarrow\; \mathcal{A} \models_{PL} \alpha$$

und für existentiale Aussagen $\alpha \in Fm_{PL}(S)$

$$\mathcal{A} \models_{PL} \alpha \;\Rightarrow\; \mathcal{B} \models_{PL} \alpha\ .$$

Beweis. Sei $\alpha = (\forall x_n) \dots (\forall x_1)\, \beta$ und β quantorenfrei. Weiter sei $\alpha_k := (\forall x_k) \dots (\forall x_1)\, \beta$, $0 \leq k \leq n$. Insbesondere ist daher $\alpha_0 = \beta$ und $\alpha_n = \alpha$.
Wir zeigen durch Induktion nach k :
Für alle $k \leq n$: und für alle $\mathcal{A}$-Belegungen w gilt $\mathcal{B}, w \models_{PL} \alpha_k \;\Rightarrow\; \mathcal{A}, w \models_{PL} \alpha_k$.
(Beachten Sie, dass jede $\mathcal{A}$-Belegung auch eine $\mathcal{B}$-Belegung ist, da $\mathcal{A}$ eine Substruktur von $\mathcal{B}$ ist.)

$k = 0$: Die Behauptung gilt aufgrund von Lemma 2.56.

$k \to k+1$: Sei w eine beliebige $\mathcal{A}$-Belegung. Es gilt dann

$$\begin{aligned}
&\mathcal{B}, w \models_{PL} \alpha_{k+1} \\
&\Leftrightarrow \text{für alle } a \in |\mathcal{B}| :\ \mathcal{B}, w\,\{x_{k+1}/a\} \models_{PL} \alpha_k \\
&\Rightarrow \text{für alle } a \in |\mathcal{A}| :\ \mathcal{B}, w\,\{x_{k+1}/a\} \models_{PL} \alpha_k \quad (\text{da } |\mathcal{A}| \subseteq |\mathcal{B}|) \\
&\Rightarrow \text{für alle } a \in |\mathcal{A}| :\ \mathcal{A}, w\,\{x_{k+1}/a\} \models_{PL} \alpha_k \quad (\text{nach IV}) \\
&\Leftrightarrow\ \mathcal{A}, w \models_{PL} (\forall x)\, \alpha_k.
\end{aligned}$$

Damit ist die Induktionsbehauptung bewiesen.
Speziell im Fall $k = n$ erhalten wir für eine beliebige $\mathcal{A}$-Belegung w :

$$\mathcal{B}, w \models_{PL} \alpha \;\Rightarrow\; \mathcal{A}, w \models_{PL} \alpha\ .$$

Damit folgt

$$\mathcal{B} \models_{PL} \alpha \;\Rightarrow\; \mathcal{A} \models_{PL} \alpha\ .$$

Die Behauptung für existentiale Aussagen folgt leicht aus dem Beweis für universale Aussagen. ■

2.1.6 Die prädikatenlogische Folgerungsbeziehung

Entsprechend der Vorgehensweise in der Aussagenlogik wollen wir auch für die Prädikatenlogik einen semantischen Folgerungsbegriff definieren.

Definition 2.62 *Sei S eine Signatur. Ferner seien $\Sigma \subseteq Fm_{PL}(S)$ und $\alpha \in Fm_{PL}(S)$.*
Aus Σ folgt *(semantisch) α (notiert als $\Sigma \models_{PL} \alpha$) genau dann, wenn für jede Interpretation I zur Signatur S gilt:*

ist I Modell von Σ, so erfüllt I α.

(Andere Sprechweise: α ist logische Konsequenz aus Σ)

Analog zur Vorgehensweise in der Aussagenlogik definieren wir auch den Begriff der (prädikatenlogischen) Konsequenzenmenge. Für diesen können wir dann dieselben Eigenschaften zeigen, die wir schon für im Falle der Aussagenlogik erhalten haben.

Definition 2.63 *Für* $\Sigma \subseteq Fm_{PL}(S)$ *ist*

$$Cn(\Sigma) := \{\alpha \in Fm_{PL}(S) \mid \Sigma \models_{PL} \alpha\}$$

die (prädikatenlogische) Konsequenzenmenge von Σ.

Proposition 2.64 Σ, Σ_1 *und* Σ_2 *seien Mengen prädikatenlogischer Formeln zu einer Signatur* S. *Dann gilt:*

1. $\Sigma_1 \subseteq \Sigma_2 \;\Rightarrow\; Cn(\Sigma_1) \subseteq Cn(\Sigma_2)$.
2. $\Sigma \subseteq Cn(\Sigma)$.
3. *Die Menge der allgemeingültigen Formeln ist für jedes* $\Sigma \subseteq Fm_{PL}$ *in* $Cn(\Sigma)$ *enthalten.*
4. $Cn(Cn(\Sigma)) \subseteq Cn(\Sigma)$.

Der Beweis, den wir im Falle der Aussagenlogik geführt haben, kann direkt übertragen werden (siehe Proposition 1.20). Dieselbe Bemerkung gilt auch für das nachfolgende Lemma. Die in diesem Lemma angegebenen Eigenschaften 1. bis 3. haben wir auch im Falle der Aussagenlogik vorgefunden. Und auch hier gilt, dass der im Fall der Aussagenlogik geführte Beweis mit den erforderlichen Anpassungen direkt übernommen werden kann.

Lemma 2.65 *Sei* $\Sigma \subseteq Fm_{PL}(S)$ *und* $\alpha, \beta \in Fm_{PL}(S)$.
Dann gilt:

1. *Aus* $\Sigma \models_{PL} \alpha \to \beta$ *und* $\Sigma \models_{PL} \alpha$ *folgt* $\Sigma \models_{PL} \beta$
   (semantische Modus ponens Regel)
2. *Wenn* $\Sigma \models_{PL} \alpha \to \beta$, *so* $\Sigma \cup \{\alpha\} \models_{PL} \beta$
   (Ableitungstheorem)
3. *Wenn* $\Sigma \cup \{\alpha\} \models_{PL} \beta$, *so* $\Sigma \models_{PL} \alpha \to \beta$.
   (Deduktionstheorem)

In dem folgenden Lemma sind noch einige weitere Eigenschaften angegeben.

Lemma 2.66 *Seien* $\Sigma \subseteq Fm_{PL}(S)$, $\alpha, \beta \in Fm_{PL}(S)$. *Ferner sei* $fvar(\Sigma) := \bigcup_{\beta \in \Sigma} fvar(\beta)$.
Dann gilt:

1. $\Sigma \models_{PL} (\forall x)\,\alpha \;\Rightarrow\; \Sigma \models_{PL} \alpha\{x/t\}$, *falls die Substitution* $\{x/t\}$ *frei für* α *ist.*
2. $\Sigma \models_{PL} \alpha\{x/t\} \;\Rightarrow\; \Sigma \models_{PL} (\exists x)\,\alpha$, *falls die Substitution* $\{x/t\}$ *frei für* α *ist.*
3. $\Sigma \cup \{\beta\} \models_{PL} \alpha \;\Rightarrow\; \Sigma \cup \{(\forall x)\,\beta\} \models_{PL} \alpha$
4. $\Sigma \models_{PL} \alpha \;\Rightarrow\; \Sigma \models_{PL} (\forall x)\,\alpha$ *für* $x \notin fvar(\Sigma)$
5. $\Sigma \cup \{\beta\} \models_{PL} \alpha \;\Rightarrow\; \Sigma \cup \{(\exists x)\,\beta\} \models_{PL} \alpha$ *für* $x \notin fvar(\Sigma \cup \{\alpha\})$

Beweis.

1. Es gelte $\Sigma \models_{PL} (\forall x)\,\alpha$.

   Sei $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S, die Modell von Σ ist. Dann gilt $[(\forall x)\,\alpha]^I = T$. Damit gilt für alle $a \in |\mathcal{A}|$ $\alpha^{I\{x/a\}} = T$. Da $t^I \in |\mathcal{A}|$, ist insbesonde $\alpha^{I\{x/t^I\}} = T$. Da die Substitution $\{x/t\}$ frei für α ist, erhalten wir nach dem Substitutionslemma $[\alpha\,\{x/t\}]^I = T$. Dies war zu zeigen.

2. Es gelte $\Sigma \models_{PL} \alpha\,\{x/t\}$.

   Sei $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S, die Modell von Σ ist. Dann gilt $[\alpha\,\{x/t\}]^I = T$. Da die Substitution $\{x/t\}$ frei für α ist, erhalten wir nach dem Substitutionslemma $\alpha^{I\{x/t^I\}} = T$. Es gibt daher ein $a \in |\mathcal{A}|$ mit $\alpha^{I\{x/a\}} = T$. Es folgt $[(\exists x)\,\alpha]^I = T$. Dies war in diesem Fall zu zeigen.

3. Es gelte $\Sigma \cup \{\beta\} \models_{PL} \alpha$.

   Sei $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S, die Modell von $\Sigma \cup \{(\forall x)\,\beta\}$ ist. Dann ist $[(\forall x)\,\beta]^{\mathcal{A},w} = T$. Damit ist für alle x-Varianten w' von w $\beta^{\mathcal{A},w'} = T$. Insbesondere für w selbst gilt also $\beta^{\mathcal{A},w} = T$. Daher ist also I ein Modell von $\Sigma \cup \{\beta\}$, und damit nach Voraussetzung auch von α.

4. Es gelte $\Sigma \models_{PL} \alpha$ und $x \notin fvar\,(\Sigma)$.

   Sei wiederum $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S, die Modell von Σ ist. Zu zeigen ist $[(\forall x)\,\alpha]^{\mathcal{A},w} = T$. Hierzu genügt es zu zeigen, dass für jede x-Variante w' von w $\alpha^{\mathcal{A},w'} = T$.

   Sei also w' eine beliebige x-Variante von w. Sei β eine beliebige Formel aus Σ. Nach Annahme über I ist $\beta^{\mathcal{A},w} = T$, und da $x \notin fvar\,(\beta)$, $\beta^{\mathcal{A},w'} = T$ nach dem Koinzidenztheorem. Damit ist also $(\mathcal{A}, w')$ ein Modell für Σ. Wegen $\Sigma \models_{PL} \alpha$ gilt dann $\alpha^{\mathcal{A},w'} = T$. Da dies für eine beliebige x-Variante w' von w gilt, folgt $[(\forall x)\,\alpha]^{\mathcal{A},w} = T$.

5. Es gelte $\Sigma \cup \{\beta\} \models_{PL} \alpha$ und $x \notin fvar\,(\Sigma \cup \{\alpha\})$.

   Sei $I = (\mathcal{A}, w)$ eine Interpretation zur Signatur S, die Modell von $\Sigma \cup \{(\exists x)\,\beta\}$ ist. Es gibt dann eine x-Variante w' von w mit $\beta^{\mathcal{A},w'} = T$. Mit dem Koinzidenztheorem folgt wie unter 4., dass $(\mathcal{A}, w')$ ein Modell für Σ und damit auch für $\Sigma \cup \{\beta\}$ ist. Aufgrund der Voraussetzung $\Sigma \cup \{\beta\} \models_{PL} \alpha$ folgt, dass $\alpha^{\mathcal{A},w'} = T$. Da $x \notin fvar\,(\alpha)$ folgt mit dem Koinzidenztheorem weiter $\alpha^{\mathcal{A},w} = T$. Dies war zu zeigen.

■

2.1.7 Signaturerweiterungen

Wir hatten bisher bei der Betrachtung einer Formelmenge immer eine feste Signatur S zugrundegelegt. Wenn wir später für die Prädikatenlogik den Tableaukalkül und den nichtklausalen Resolutionskalkül einführen, wird es erforderlich sein, die vorhandene Signatur S um neue Konstantensymbole zu erweitern. Auch werden wir später noch die Situation vorfinden, dass die Signatur um neue Funktionssymbole erweitert wird. Für diese Fälle wollen wir die später benötigten Eigenschaften herleiten.

Definition 2.67 *Sei $S = (\mathcal{R}; \mathcal{F}; \mathcal{K}; ar)$ eine Signatur.*

1. *Falls $\mathcal{K}' = \mathcal{K} \cup \mathcal{K}_0$ mit $\mathcal{K}_0 \neq \emptyset$ und $\mathcal{K} \cap \mathcal{K}_0 = \emptyset$, so heißt die Signatur $S' = (\mathcal{R}; \mathcal{F}; \mathcal{K}'; ar)$* Konstantenerweiterung der Signatur S *(notiert als $S\,(\mathcal{K}_0)$).*

2. *Falls $\mathcal{F}' = \mathcal{F} \cup \mathcal{F}_0$ mit $\mathcal{F}_0 \neq \emptyset$ und $\mathcal{F} \cap \mathcal{F}_0 = \emptyset$, so heißt die Signatur $S' = (\mathcal{R}; \mathcal{F}'; \mathcal{K}; ar)$* Funktionserweiterung der Signatur S *(notiert als $S\,(\mathcal{F}_0)$).*

In den beiden nächsten Lemmata werden wir auf die Beziehung von Allgemeingültigkeit und Erfüllbarkeit von Formeln in einer Signatur S sowie in einer Signaturerweiterung S' von S eingehen. Strenggenommen sollte daher in diesen Begriffen die Signatur mit berücksichtigt werden und auch in der Notation kenntlich gemacht werden. Wir verzichten im Folgenden darauf, da dies die Notation schwerfälliger machen würde, und meistens auch aus dem Kontext klar ist, welche Signatur zugrundeliegt.
Das erste dieser beiden Lemmata drückt eine ziemlich offensichtliche Eigenschaft aus.

Lemma 2.68 *Sei S' eine Konstantenerweiterung (Funktionserweiterung) der Signatur S.*

1. *Sei $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ eine S-Struktur und w eine $\mathcal{A}$-Belegung. Ist $\mathcal{A}' = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}'^{\mathcal{A}}\right)$ $\left(\mathcal{A}' = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}'^{\mathcal{A}}; \mathcal{K}\right)\right)$ eine Erweiterung der S-Struktur $\mathcal{A}$ zu einer S'-Struktur, dann gilt für Formeln $\alpha \in Fm_{PL}(S)$*
$$\alpha^{\mathcal{A},w} = \alpha^{\mathcal{A}',w}.$$
(Man beachte, dass w auch eine $\mathcal{A}'$-Belegung ist.)

2. *Ist α eine allgemeingültige (erfüllbare) Formel der Signatur S, dann ist α auch eine allgemeingültige (erfüllbare) Formel der Signatur S'.*

Den sehr einfachen Beweis (mittels struktureller Induktion) führen wir nicht aus.

Lemma 2.69 *Sei die Signatur $S' = S(\{c\})$ eine Konstantenerweiterung der Signatur $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$. Ferner sei $\alpha \in Fm_{PL}(S)$ und $\Sigma \subseteq Fm_{PL}(S)$. Es ist dann $\alpha\{x/c\} \in Fm_{PL}(S')$.*
Es gilt:

1. $\models_{PL} (\forall x)\,\alpha \Leftrightarrow\ \models_{PL} \alpha\{x/c\}$

   $\models_{PL} \neg(\exists x)\,\alpha \Leftrightarrow\ \models_{PL} \neg\alpha\{x/c\}$

2. *Hat $\Sigma \cup \{(\exists x)\,\alpha\}$ (bzw. $\Sigma \cup \{\neg(\forall x)\,\alpha\}$) ein Modell zur Signatur S, so hat $\Sigma \cup \{(\exists x)\,\alpha, \alpha\{x/c\}\}$ (bzw. $\Sigma \cup \{\neg(\forall x)\,\alpha, \neg\,\alpha\{x/c\}\}$) ein Modell zur Signatur S'.*

Beweis.

1. **„ $\Rightarrow$ ”:** Wir zeigen die Kontraposition.
   Sei also $\alpha\{x/c\}$ nicht allgemeingültig. Es gibt dann eine S'-Struktur $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; (\mathcal{K} \cup \{c\})^{\mathcal{A}}\right)$ und eine $\mathcal{A}$-Belegung w mit $[\alpha\{x/c\}]^{\mathcal{A},w} = F$. Nach dem Substitutionslemma gilt $\alpha^{\mathcal{A},w\left\{x/c^{\mathcal{A}}\right\}} = F$ mit $c^{\mathcal{A}} \in A$. Sei $\mathcal{B} := \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$. $w\left\{x/c^{\mathcal{A}}\right\}$ ist dann eine $\mathcal{B}$-Belegung. Damit gilt nach Lemma 2.68 $\alpha^{\mathcal{B},w\left\{x/c^{\mathcal{A}}\right\}} = F$, also ist $(\forall x)\,\alpha$ nicht allgemeingültig.

   „ $\Leftarrow$ ”: Wir zeigen wiederum die Kontraposition.
   Sei also $(\forall x)\,\alpha$ nicht allgemeingültig. Es existiert dann eine S-Struktur $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ und ein $a \in A$ mit $\alpha^{\mathcal{A},w\{x/a\}} = F$. Wir betrachten die S'-Struktur $\mathcal{B} := \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}} * \left(c^{\mathcal{A}}\right)\right)$ mit $\mathcal{K}^{\mathcal{A}} * \left(c^{\mathcal{A}}\right) := \left(d^{\mathcal{A}}\right)_{d \in K \cup \{c\}}$ und $c^{\mathcal{A}} := a$, d.h. $\mathcal{B}$ geht aus $\mathcal{A}$ dadurch hervor, dass wir a als neue, dem Konstantensymbol c zugeordnete Konstante aufnehmen. Nach Lemma 2.68 gilt dann $\alpha^{\mathcal{B},w\{x/a\}} = \alpha^{\mathcal{B},w\left\{x/c^{\mathcal{B}}\right\}} = F$, und damit nach dem Substitutionslemma $[\alpha\{x/c\}]^{\mathcal{B},w} = F$. Folglich ist $\alpha\{x/c\}$ nicht allgemeingültig.

2. Ergibt sich sofort aus Lemma 2.68.

■

2.1.8 Herbrand-Modelle

Zu einer Signatur S wollen wir nun spezielle S-Strukturen betrachten, die wir später im Vollständigkeitsbeweis des noch vorzustellenden Tableau-Kalküls (bzw. Resolutionskalküls) benötigen werden und die auch für den Bereich der Logischen Programmierung von großer Bedeutung sind.

Definition 2.70 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält.*

1. *Die Menge der Grundterme in $Tm_{PL}(S)$ heißt* Herbrand-Universum *(zur Signatur S).*

2. *Eine S-Struktur $\mathcal{A} = \left(A; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$ heißt* Herbrand-Struktur *genau dann, wenn gilt*

   (a) A ist das Herbrand-Universum zur Signatur S,

(b) *für alle* $f \in \mathcal{F}$ *und für alle* $t_1, \ldots, t_n \in A$ $(n = ar(f))$ *gilt*

$$f^{\mathcal{A}}(t_1, \ldots, t_n) = f(t_1, \ldots, t_n),$$

(c) $c^{\mathcal{A}} = c$ *für alle* $c \in \mathcal{K}$.

3. *Ist für eine Herbrand-Struktur* $\mathcal{A}$ *und eine* $\mathcal{A}$*-Belegung* w $I = (\mathcal{A}, w)$ *ein Modell einer Formelmenge* Σ, *so heißt* I Herbrand-Modell *von* Σ.

Bei der Angabe einer Herbrand-Struktur zu einer Signatur S sind somit nur noch die den Prädikatensymbolen zugeordneten Prädikate (bzw. Relationen) festzulegen, da die den Funktionssymbolen zugeordneten Funktionen bzw. die den Konstantensymbolen zugeordneten Konstanten bereits eindeutig bestimmt sind.

Beispiel 2.71 *Sei die Signatur* $S = (P; f, g; c, d)$ *mit* $ar(P) = ar(f) = 1$ *und* $ar(g) = 2$ *gegeben. Im Herbrand-Universum* $H(S)$ *zur Signatur* S *sind dann u.a. die folgenden Elemente enthalten:*

$$c,\ d,\ f(c),\ f(d),\ g(c,c),\ g(c,d),\ g(d,c),\ g(d,d),\ f(g(c,c)),\ f(g(c,d)),\ \ldots .$$

$\mathcal{A} = \left(H(S); P^{\mathcal{A}}; f^{\mathcal{A}}, g^{\mathcal{A}}; c, d\right)$ *mit* $P^{\mathcal{A}} = \{c, g(c,d)\}$ *und* $f^{\mathcal{A}}$ *und* $g^{\mathcal{A}}$ *wie in Definition 2.70 festgelegt ist dann eine Herbrand-Struktur zu* S.

Lemma 2.72 *Sei* $\mathcal{A}$ *eine Herbrand-Struktur zur Signatur* S *und* w *eine* $\mathcal{A}$*-Belegung.*

1. *Für Grundterme* $t \in Tm_{PL}(S)$ *gilt* $t^{\mathcal{A}} = t$.
2. *Für beliebige Terme* $t \in Tm_{PL}(S)$ *gilt:*
   ist θ_w *eine Substitution mit* $\theta_w(x) = w(x)$ *für alle* $x \in var(t)$, *so gilt*
   $$t^{\mathcal{A},w} = [t\theta_w]^{\mathcal{A}}.$$
3. *Für beliebige Formeln* $\alpha \in Fm_{PL}(S)$ *gilt:*
   ist θ_w *eine Substitution mit* $\theta_w(x) = w(x)$ *für alle* $x \in fvar(\alpha)$, *so gilt*
   $$\alpha^{\mathcal{A},w} = [\alpha\theta_w]^{\mathcal{A}}.$$

Beweis.

1. Durch strukturelle Induktion.

   Ind.Basis: Für Konstantensymbole $c \in S$ gilt $c^{\mathcal{A}} = c$.

   Ind.Schritt: Sei $f(t_1, \ldots, t_n)$ ein Grundterm zur Signatur S.
   Nach Induktionsvoraussetzung gilt $t_i^{\mathcal{A}} = t_i$, $1 \leq i \leq n$.
   Damit gilt $[f(t_1, \ldots, t_n)]^{\mathcal{A}} = f^{\mathcal{A}}\left(t_1^{\mathcal{A}}, \ldots, t_n^{\mathcal{A}}\right) = f^{\mathcal{A}}(t_1, \ldots, t_n) = f(t_1, \ldots, t_n)$.

2. Ebenfalls durch strukturelle Induktion.

   Ind.Basis: Für Konstantensymbole $c \in S$ gilt $c^{\mathcal{A},w} = c = [c\theta_w]^{\mathcal{A}}$.
   Für eine Variable $x \in Var_{PL}$ gilt $x^{\mathcal{A},w} = w(x)$. Da $\mathcal{A}$ eine Herbrand-Struktur ist, ist $w(x)$ also ein Grundterm. Infolgedessen gilt $[w(x)]^{\mathcal{A}} = w(x)$. Ist nun θ_w eine Substitution mit $\theta_w(x) = w(x)$, so folgt also $x^{\mathcal{A},w} = [x\theta_w]^{\mathcal{A}}$

Ind.Schritt: Sei $t = f(t_1, \ldots, t_n)$ ein Term zur Signatur S und sei θ_w eine Substitution mit $\theta_w(x) = w(x)$ für alle $x \in var(t)$.
Nach Induktionsvoraussetzung gilt $t_i^{\mathcal{A},w} = [t_i\theta_w]^{\mathcal{A}}$, $1 \leq i \leq n$.
Damit gilt

$$\begin{aligned}
t^{\mathcal{A},w} &= [f(t_1, \ldots, t_n)]^{\mathcal{A},w} \\
&= f^{\mathcal{A}}\left(t_1^{\mathcal{A},w}, \ldots, t_n^{\mathcal{A},w}\right) \\
&= f^{\mathcal{A}}\left([t_1\theta_w]^{\mathcal{A}}, \ldots, [t_n\theta_w]^{\mathcal{A}}\right) \quad \text{nach Ind.Vor.} \\
&= f^{\mathcal{A}}(t_1\theta_w, \ldots, t_n\theta_w) \quad \text{da } t_i\theta_w \text{ Grundterme} \\
&= f(t_1\theta_w, \ldots, t_n\theta_w) \\
&= [f(t_1, \ldots, t_n)]\,\theta_w \\
&= [[f(t_1, \ldots, t_n)]\,\theta_w]^{\mathcal{A}}.
\end{aligned}$$

3. Übungsaufgabe.

■

Damit lässt sich dann die folgende Charakterisierung von All- und Existenzquantoren in Herbrand-Strukturen leicht beweisen.

Theorem 2.73 *Sei $\mathcal{A}$ eine Herbrand-Struktur zur Signatur S, w eine $\mathcal{A}$-Belegung und $\alpha \in Fm_{PL}(S)$. Es gilt dann:*

1. $\mathcal{A}, w \models_{PL} (\forall x)\,\alpha \Leftrightarrow$ *für alle* $t \in |\mathcal{A}|$ *gilt* $\mathcal{A}, w \models_{PL} \alpha\{x/t\}$
2. $\mathcal{A}, w \models_{PL} (\exists x)\,\alpha \Leftrightarrow$ *für ein* $t \in |\mathcal{A}|$ *gilt* $\mathcal{A}, w \models_{PL} \alpha\{x/t\}$.

Beweis. Übungsaufgabe. ■

2.1.9 γ- und δ-Formeln

Im Kapitel über die Aussagenlogik hatten wir α- und β-Formeln eingeführt. Wir wollen diese Bezeichnung für prädikatenlogische Formeln, sofern sie von der entsprechenden Gestalt sind, übernehmen und um 2 weitere Kategorien von Formeln ergänzen:

- γ-Formeln,

  das sind „universale" Formeln der Gestalt $(\forall x)\,\alpha$ bzw. $\neg(\exists x)\,\alpha$.

- δ-Formeln,

  das sind „existentiale" Formeln der Gestalt $(\exists x)\,\alpha$ bzw. $\neg(\forall x)\,\alpha$.

Den γ- und δ-Formeln wollen wir sogenannte *Instanzen* zuordnen. Sei dazu α eine Formel einer Signatur S und t ein beliebiger Term derselben Signatur S.

γ-Formel	γ-Instanz $\boldsymbol{\gamma}(t)$
$(\forall x)\,\alpha$	$\alpha\{x/t\}$
$\neg(\exists x)\,\alpha$	$\neg\alpha\{x/t\}$

und

δ-Formel	δ-Instanz $\boldsymbol{\delta}(t)$
$(\exists x)\,\alpha$	$\alpha\{x/t\}$
$\neg(\forall x)\,\alpha$	$\neg\alpha\{x/t\}$

Es ist also für jeden Term t die Formel $\alpha\{x/t\}$ γ-Instanz der γ-Formel $(\forall x)\,\alpha$ und δ-Instanz der δ-Formel $(\exists x)\,\alpha$ usw. .
Im folgenden sei $\boldsymbol{\gamma}$ immer eine γ-Formel und $\boldsymbol{\delta}$ eine δ-Formel. Entsprechend sind dann $\boldsymbol{\gamma}(t)$ und $\boldsymbol{\delta}(t)$ zugehörige Instanzen. Ebenso nehmen wir an, dass $\boldsymbol{\alpha}$ und $\boldsymbol{\beta}$ immer α- bzw. β-Formeln sind mit Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ bzw. $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$.
Eine direkte Folgerung aus Theorem 2.73 bzw. aus Lemma 2.69 ist dann das folgende Lemma.

Lemma 2.74 *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen der Signatur S. Ferner seien $\boldsymbol{\gamma}, \boldsymbol{\delta} \in Fm_{PL}(S)$ Aussagen.*
Es gilt dann

1. *Ist $\Sigma \cup \{\boldsymbol{\gamma}\}$ erfüllbar, so ist $\Sigma \cup \{\boldsymbol{\gamma}, \boldsymbol{\gamma}(t)\}$ erfüllbar für jeden Grundterm $t \in Tm_{PL}(S)$.*

2. *Hat $\Sigma \cup \{\boldsymbol{\delta}\}$ ein Modell zur Signatur S und ist $c \notin S$, so hat $\Sigma \cup \{\boldsymbol{\delta}, \boldsymbol{\delta}(c)\}$ ein Modell zur Konstantenerweiterung $S(\{c\})$.*

Die Einführung der α- und β-Formeln hat uns im Falle der Aussagenlogik ein weiteres Beweisprinzip (wie auch ein Definitionsprinzip) beschert, mit dem Eigenschaften für die Menge der aussagenlogischen Formeln gezeigt werden konnte: nämlich strukturelle Induktion (bzw. Rekursion) über α- und β-Formeln. Analog zum aussagenlogischen Fall steht uns eine strukturelle Induktion (Rekursion) über α-, β-, γ- und δ-Formeln zur Verfügung, mit der wir Eigenschaften für die Menge der prädikatenlogischen Formeln beweisen (definieren) können.

Theorem 2.75 (Strukturelle Induktion über $\alpha-$, $\beta-$, $\gamma-$, $\delta-$Formeln) *E sei eine Eigenschaft von Formeln der Signatur S, für die folgendes gilt:*

1. *Induktionsbasis: E trifft auf jede atomare Formel sowie deren Negation zu.*

2. *Induktionsschritt:*

Trifft E zu auf:	*so auch auf:*
α	$\neg\neg\alpha$
$\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2$	$\boldsymbol{\alpha}$
$\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2$	$\boldsymbol{\beta}$
$\boldsymbol{\gamma}(t)$ *für jeden Term* t	$\boldsymbol{\gamma}$
$\boldsymbol{\delta}(t)$ *für jeden Term* t	$\boldsymbol{\delta}$

Dann gilt, dass E auf alle Formeln der Signatur S zutrifft.

Den Beweis dieses Theorems werden wir nicht ausführen. Er verläuft ganz analog zum Beweis im aussagenlogischen Fall.
Auf die Angabe des zugehörigen Definitionsprinzips der strukturellen Rekursion wird ebenfalls verzichtet.

2.2 Beweiskalküle der Prädikatenlogik

Bei der Betrachtung formaler Beweiskalküle für die Aussagenlogik hatten wir schon das Bedürfnis nach solch formalen Beweiskalkülen motiviert. Unabhängig von semantischen Überlegungen sollte durch rein formales Operieren über Zeichenreihen die Allgemeingültigkeit von Aussagen festgestellt werden können. Die dort vorgestellten Beweiskalküle (Tableau-Kalkül, Resolutionskalkül) hatten zwar diese Eigenschaft, jedoch stand mit der Methode der Wahrheitstabellen – sofern man diese nicht als Beweiskalkül auffasst – auch eine einfache semantische Methode zur Verfügung. Diese Situation ändert sich im Falle der Prädikatenlogik. Der semantische Nachweis der Allgemeingültigkeit erfordert durch die Bezugnahme auf Strukturen mit Grundbereichen beliebig großer Kardinalität und damit auf überabzählbar viele Belegungen eine sehr viel höhere Abstraktion und damit verbunden oft eine viel kompliziertere Argumentation, so dass formale Beweiskalküle nicht nur den Rahmen dafür bilden, dass in ihnen der Beweis für die Allgemeingültigkeit einer Aussage geliefert werden kann, sondern auch, dass ein solcher Beweis in einfacher Weise überprüft werden kann.
Im folgenden werden wir einige der wichtigsten Beweissysteme für die Prädikatenlogik vorstellen. Wir werden diese Beweissysteme gleich so einführen, dass wir mit ihnen den Begriff des semantischen Folgerns erfassen.

2.2.1 Der Tableau-Kalkül

Wir wollen den in der Aussagenlogik eingeführten Tableau-Kalkül so zu einem prädikatenlogischen Beweiskalkül erweitern, dass, falls die prädikatenlogische Aussage $\alpha \in Fm_{PL}(S)$ aus einer Prämissenmenge $\Sigma \subseteq Fm_{PL}(S)$, bestehend aus Aussagen, semantisch folgt, auch ein Tableau-Beweis von α aus der Prämissenmenge Σ existiert. In einem solchen Tableau-Beweis werden dann i.a. auch Formeln aus einer Konstantenerweiterung $S(\mathcal{K}_0)$ vorkommen, wobei $\mathcal{K}_0$ abzählbar unendlich ist.
Die nachfolgende Definition von Σ-Tableaus übernimmt die Definition für den aussagenlogischen Fall (mit den entsprechenden Anpassungen an die prädikatenlogische Sprache) und ergänzt diese um die Tableau-Regeln für $\boldsymbol{\gamma}$- und δ-Formeln. Der Vollständigkeit halber geben wir die Definition ausführlich an.

Definition 2.76 (Σ-Tableau für $\{\alpha_0, \dots, \alpha_n\}$) *Sei $\Sigma \cup \{\alpha_0, \dots, \alpha_n\} \subseteq Fm_{PL}(S)$.*

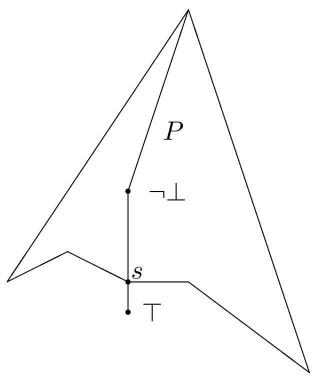
1. *Der mit $\alpha_0, \dots, \alpha_n$ markierte Binärbaum*

$$\begin{array}{c} \alpha_0 \\ | \\ \alpha_1 \\ | \\ \vdots \\ | \\ \alpha_n \end{array}$$

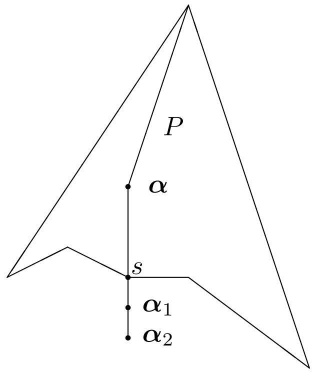
   ist ein Σ-Tableau(das Starttableau) *für $\{\alpha_0, \dots, \alpha_n\}$.*

2. *Ist τ ein Σ-Tableau für $\{\alpha_0, \dots, \alpha_n\}$ und P ein Zweig in τ mit Blatt s, ist ferner ein Knoten in P mit einem Nichtliteral N markiert, dann ist der in Abhängigkeit von N nachfolgend definierte, markierte Binärbaum τ' ebenfalls ein Σ-Tableau für $\{\alpha_0, \dots, \alpha_n\}$:*

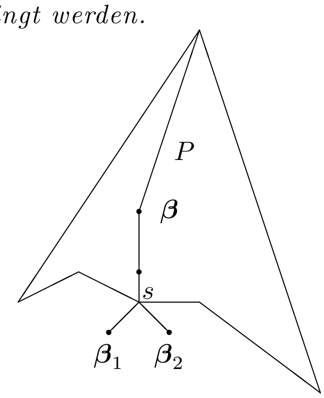
   (a) $N = \neg\bot \quad (\neg\top$ *resp.* $\neg\neg\beta)$:
   τ' resultiert aus τ, indem an das Blatt s ein mit $\top$ ($\bot$ resp. β) markierter Knoten als Sohn angehängt wird.



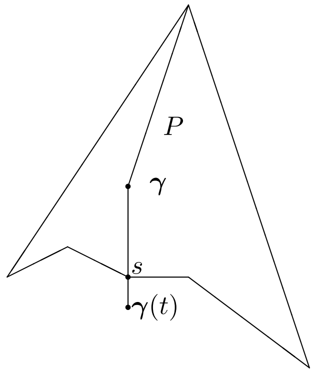
(b) N ist eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$:
τ' resultiert aus τ dadurch, dass das Blatt s von P einen Sohn und einen Enkel erhält, die mit $\boldsymbol{\alpha}_1$ bzw. $\boldsymbol{\alpha}_2$ markiert werden.



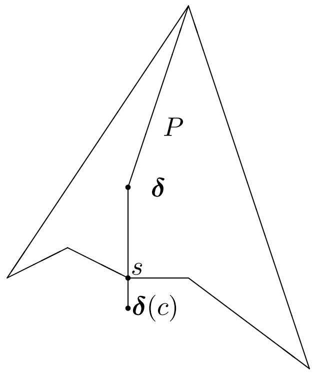
(c) N ist eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$:
τ' wird aus τ dadurch erhalten, dass an das Blatt s von P zwei Söhne, die mit $\boldsymbol{\beta}_1$ bzw. $\boldsymbol{\beta}_2$ markiert werden, angehängt werden.



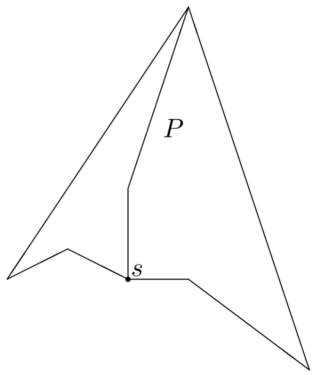
(d) N ist eine γ-Formel mit einer γ-Instanz $\boldsymbol{\gamma}(t)$:
τ' resultiert aus τ, indem an das Blatt s ein mit $\boldsymbol{\gamma}(t)$ markierter Knoten als Sohn angehängt wird.



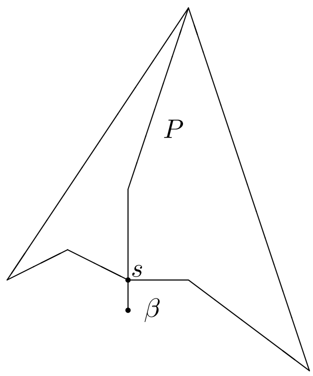
(e) N ist eine δ-Formel mit einer δ-Instanz $\boldsymbol{\delta}(c)$ für ein $c \in \mathcal{K}_0$:
τ' resultiert aus τ, indem an das Blatt s ein mit $\boldsymbol{\delta}(c)$ markierter Knoten als Sohn angehängt wird.



3. *Ist τ ein Σ-Tableau*



und P ein Zweig in τ mit Blatt s, dann ist für eine Aussage $\beta \in \Sigma$ auch der folgende markierte Binärbaum ein Σ-Tableau:



4. *Nur markierte Binärbäume, die nach 1, 2 und 3 erhalten werden, sind Σ-Tableaus für $\{\alpha_0, \ldots, \alpha_n\}$.*

Für das in Punkt 2 bzw. 3 definierte Tableau τ' führen wir die folgende Sprechweise ein:

Wir sagen, das Tableau τ' geht aus dem Tableau τ durch Anwendung der in der folgenden Tabelle angegebenen Tableau-Regel hervor:

$$\begin{cases} \dfrac{\neg\bot}{\top} & \text{falls } N = \neg\bot \\[2ex] \dfrac{\neg\top}{\bot} & \text{falls } N = \neg\top \\[2ex] \dfrac{\neg\neg\beta}{\beta} & \text{falls } N = \neg\neg\beta \\[2ex] \dfrac{\boldsymbol{\alpha}}{\begin{matrix}\boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2\end{matrix}} & \text{falls } N \text{ eine } \alpha - Formel \\[2ex] \dfrac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2} & \text{falls } N \text{ eine } \beta - Formel \\[2ex] \dfrac{\boldsymbol{\gamma}}{\boldsymbol{\gamma}(t)} & \text{falls } N \text{ eine } \gamma - Formel,\ t \text{ Grundterm} \\[2ex] \dfrac{\boldsymbol{\delta}}{\boldsymbol{\delta}(c)} & \text{falls } N \text{ eine } \delta - Formel,\ c \in \mathcal{K}_0 \\[2ex] \dfrac{}{\beta} & \text{falls } \beta \in \Sigma \end{cases}$$

Definition 2.77 *Ein* Zweig P *eines* Σ*-Tableaus heißt* geschlossen *genau dann, wenn*

1. *es Knoten* $s, t \in P$ *sowie eine prädikatenlogische Formel* α *gibt, so dass* s *mit* α *markiert ist und* t *mit* $\neg\alpha$*, oder*

2. *ein mit* $\bot$ *markierter Knoten in* P *existiert.*

Ein Tableau *heißt* geschlossen*, wenn jeder Zweig des Tableaus geschlossen ist.*

Definition 2.78 *Es sei* S *eine Signatur, die wenigstens ein Konstantensymbol enthält.* $S(K_0)$ *sei eine Konstantenerweiterung von* S*, die im Folgenden alle im Tableau neu eingeführten Konstantensymbole enthalten soll.*
Sei ferner $\Sigma \cup \{\alpha\} \subseteq Fm_{PL}(S)$ *eine Menge von* ***Aussagen****.*

1. *Ein* Σ-Tableau-Beweis *für* α *ist ein geschlossenes* Σ*-Tableau für* $\{\neg\alpha\}$ *(notiert als* $\Sigma \vdash_{PL-T} \alpha$*; Sprechweise:* α *ist aus* Σ *herleitbar bzw. ableitbar).*

2. *Ein* Tableau-Beweis *für* α *ist ein* $\emptyset$*-Tableau-Beweis für* α *(notiert als* $\vdash_{PL-T} \alpha$*; „*α *ist beweisbar”).*

Bemerkung 2.79 *Beachten Sie, dass der Tableau-Kalkül nur zum Beweisen von Aussagen bzw. zum Herleiten von Aussagen aus einer Aussagenmenge dient. Damit kann also z.B. keine Tableau-Herleitung von* $P(x)$ *aus* $\{P(x)\}$ *möglich sein, obwohl* $\{P(x)\} \models_{PL} P(x)$ *gilt.*

Beispiel 2.80 *Wir wollen*

$$\vdash_{PL-T} (\forall x)(P(x) \vee Q(x)) \to (\exists x) P(x) \vee (\forall x) Q(x)$$

zeigen.

Ein Tableau-Beweis hierzu:

$$
\begin{array}{ccl}
\neg\left(\left(\forall x\right)\left(P\left(x\right)\vee Q\left(x\right)\right)\rightarrow\left(\exists x\right)P\left(x\right)\vee\left(\forall x\right)Q\left(x\right)\right) & & \textit{1; Starttableau} \\
| & & \\
\left(\forall x\right)\left(P\left(x\right)\vee Q\left(x\right)\right) & & \textit{2; } \boldsymbol{\alpha}_1 \textit{ aus } 1 \\
| & & \\
\neg\left(\left(\exists x\right)P\left(x\right)\vee\left(\forall x\right)Q\left(x\right)\right) & & \textit{3; } \boldsymbol{\alpha}_2 \textit{ aus } 1 \\
| & & \\
\neg\left(\exists x\right)P\left(x\right) & & \textit{4; } \boldsymbol{\alpha}_1 \textit{ aus } 3 \\
| & & \\
\neg\left(\forall x\right)Q\left(x\right) & & \textit{5; } \boldsymbol{\alpha}_2 \textit{ aus } 3 \\
| & & \\
\neg Q\left(c\right) & & \textit{6; } \boldsymbol{\delta}\left(c\right) \textit{ aus } 5\textit{; } c \textit{ neu} \\
| & & \\
\neg P\left(c\right) & & \textit{7; } \boldsymbol{\gamma}\left(c\right) \textit{ aus } 4 \\
| & & \\
P\left(c\right)\vee Q\left(c\right) & & \textit{8; } \boldsymbol{\gamma}\left(c\right) \textit{ aus } 2 \\
\diagup \qquad \diagdown & & \\
P\left(c\right) \qquad\qquad Q\left(c\right) & & \textit{9; } \boldsymbol{\beta}_1 \textit{ bzw. } \boldsymbol{\beta}_2 \textit{ aus } 8
\end{array}
$$

Beide Zweige sind offensichtlich geschlossen, so dass also ein Tableau-Beweis für $\left(\forall x\right)\left(P\left(x\right)\vee Q\left(x\right)\right)\rightarrow\left(\exists x\right)P\left(x\right)\vee\left(\forall x\right)Q\left(x\right)$ *vorliegt.*

Es würde nun naheliegen, die Korrektheit und Vollständigkeit des prädikatenlogischen Tableau-Kalküls zu zeigen. Wir werden zunächst aber erst den nichtklausalen Resolutionskalkül behandeln. Im Anschluss daran werden wir dann die Korrektheit und Vollständigkeit der beiden Beweiskalküle gleichzeitig nachweisen.

2.2.2 Der prädikatenlogische nichtklausale Resolutionskalkül

Wie zuvor den Tableau-Kalkül „erweitern" wir den aussagenlogischen Resolutionskalkül zu einem prädikatenlogischen Resolutionskalkül mit dem Ziel, den semantischen Folgerungsbegriff für Aussagen(mengen) zu erfassen.

Notation 2.81

1. $\beta=\left[\alpha_1,\ldots,\alpha_i,\boldsymbol{\gamma},\alpha_{i+1},\ldots,\alpha_n\right]$ *sei eine verallgemeinerte Disjunktion von Aussagen aus* $Fm_{PL}\left(S\left(\mathcal{K}_0\right)\right)$*, wobei* $\boldsymbol{\gamma}$ *eine* γ*-Formel ist. Sei* $\boldsymbol{\gamma}\left(t\right)$ *eine* γ*-Instanz mit einem Grundterm* t*.*

   $\beta'=\left[\alpha_1,\ldots,\alpha_i,\boldsymbol{\gamma}\left(t\right),\alpha_{i+1},\ldots,\alpha_n\right]$ *wird als Resultat der* γ*-Expansionsregel für* β *bezeichnet.*

2. *Ist* $\beta=\left[\alpha_1,\ldots,\alpha_i,\boldsymbol{\delta},\alpha_{i+1},\ldots,\alpha_n\right]$ *wiederum eine verallgemeinerte Disjunktion von Aussagen aus* $Fm_{PL}\left(S\left(\mathcal{K}_0\right)\right)$*, so wird*

   $$\beta'=\left[\alpha_1,\ldots,\alpha_i,\boldsymbol{\delta}\left(c\right),\alpha_{i+1},\ldots,\alpha_n\right]$$

   für ein $c\in\mathcal{K}_0$ *als Resultat der* δ*-Expansionsregel* $\frac{\boldsymbol{\delta}}{\boldsymbol{\delta}(c)}$ *für* β *bezeichnet.*

Wir wollen als Nächstes den Begriff der Σ-Resolutionsexpansion für die Prädikatenlogik (induktiv) definieren.

Definition 2.82 *Sei* S *eine Signatur, die wenigstens ein Konstantensymbol enthält und* $S\left(\mathcal{K}_0\right)$ *sei eine geeignete Konstantenerweiterung von* S*. Weiter sei* $\Sigma\cup\left\{\alpha_1,\ldots,\alpha_n\right\}\subseteq Fm_{PL}\left(S\right)$ *eine Menge von* ***Aussagen****.*

1. $\left[\alpha_1\right],\ldots,\left[\alpha_n\right]$ *ist* Σ-Resolutionsexpansion für $\left\{\alpha_1,\ldots,\alpha_n\right\}$ *(die* Σ*-Startexpansion)*

2. *Ist* $\beta_1,\ldots,\beta_m$ *eine* Σ*-Resolutionsexpansion für* $\left\{\alpha_1,\ldots,\alpha_n\right\}$*, so gilt:*

   (a) Ist β *das Resultat*

i. *einer Resolutionsexpansionsregel* $\frac{\neg\top}{\bot}$, $\frac{\neg\bot}{\top}$, $\frac{\neg\neg\alpha}{\alpha}$ *oder*

$$\begin{array}{c}\boldsymbol{\beta}\\ \hline \boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\end{array}$$

für ein β_i, $1 \leq i \leq m$, *oder*

ii. $\beta \in \Sigma$ *oder*

iii. $\beta = R(\beta_i)$ *bzw.* $\beta = R(\beta_i, \beta_j)$, $1 \leq i, j \leq m$,

so ist $\beta_1, \ldots, \beta_m, \beta$ *eine* Σ*-Resolutionsexpansion für* $\{\alpha_1, \ldots, \alpha_n\}$.

(b) *Sind* β_1' *und* β_2' *das Resultat einer Resolutionsexpansionsregel*

$$\begin{array}{c}\boldsymbol{\alpha}\\ \hline \boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2\end{array}$$

für ein β_i, $(1 \leq i \leq m)$, *so ist* $\beta_1, \ldots, \beta_m, \beta_1', \beta_2'$ *eine* Σ*-Resolutionsexpansion für* $\{\alpha_1, \ldots, \alpha_n\}$.

(c) *ist* β *das Resultat*

i. *der* γ*-Expansionsregel für ein* β_i, $1 \leq i \leq m$, *oder*

ii. *das Resultat der* δ*-Expansionsregel* $\frac{\boldsymbol{\delta}}{\boldsymbol{\delta}(c)}$ *für ein* β_i, $1 \leq i \leq m$, *wobei das Konstantensymbol* c *in* $\beta_1, \ldots, \beta_m$ *nicht vorkommt,*

so ist $\beta_1, \ldots, \beta_m, \beta$ *ebenfalls eine* Σ*-Resolutionsexpansion für* $\{\alpha_1, \ldots, \alpha_n\}$.

Definition 2.83 *Sei* $\Sigma \cup \{\alpha\} \subseteq Fm_{PL}(S)$ *eine Menge von Aussagen.*

1. *Eine* Σ*-Resolutionsexpansion* $\beta_1, \ldots, \beta_n$ *heißt geschlossen genau dann, wenn für ein* i *mit* $1 \leq i \leq n$ $\beta_i = [\,]$ *gilt.*

2. *Eine geschlossene* Σ*-Resolutionsexpansion für* $\{\neg\alpha\}$ *heißt* Σ-Resolutionsbeweis für α

   (*notiert als* $\Sigma \vdash_{PL-R} \alpha$).

3. *Eine geschlossene* $\emptyset$*-Resolutionsexpansion für* $\{\neg\alpha\}$ *heißt* Resolutionsbeweis für α

   (*notiert als* $\vdash_{PL-R} \alpha$).

Gilt $\vdash_{PL-R} \alpha$ (bzw. $\Sigma \vdash_{PL-R} \alpha$), so sagen wir auch, dass im Resolutionskalkül α (aus Σ) beweisbar oder herleitbar ist

Bemerkung 2.84 *Wie schon für den Tableau-Kalkül gilt auch hier, dass mit dem nichtklausalen Resolutionskalkül nur Aussagen aus einer Aussagenmenge hergeleitet werden können.*

Beispiel 2.85 *Es gilt*

$$\{(\exists x)\, P(x) \wedge (\forall x)\, Q(x)\} \vdash_{PL-R} (\exists x)\,(P(x) \wedge Q(x)).$$

Dies ergibt sich mit der folgenden Resolutionsherleitung:

$[\neg(\exists x)(P(x) \wedge Q(x))]$	1; *Startexpansion*
$[(\exists x) P(x) \wedge (\forall x) Q(x)]$	2; *Prämissenregel*
$[(\exists x) P(x)]$	3; $\boldsymbol{\alpha}_1$ *aus 2*
$[(\forall x) Q(x)]$	4; $\boldsymbol{\alpha}_2$ *aus 2*
$[P(c)]$	5; $\boldsymbol{\delta}(c)$ *aus 3*
$[Q(c)]$	6; $\boldsymbol{\gamma}(c)$ *aus 4*
$[\neg(P(c) \wedge Q(c))]$	7; $\boldsymbol{\gamma}(c)$ *aus 1*
$[\neg P(c), \neg Q(c)]$	8; $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ *aus 7*
$[\neg Q(c)]$	9; $R(5,8)$
$[\,]$	10; $R(6,9)$

Bevor wir mit dem Sequenzenkalkül einen weiteren wichtigen Beweiskalkül vorstellen, werden wir zunächst die Korrektheit und Vollständigkeit des Tableau- und des Resolutionskalküls beweisen.

2.2.3 Korrektheit von Tableau- und nichklausalem Resolutionskalkül

Die Vorgehensweise im Beweis der Korrektheit dieser beiden Kalküle entspricht ganz derjenigen, die wir schon aus dem aussagenlogischen Teil kennen.

Definition 2.86

1. *Ein Zweig P eines Σ-Tableaus heißt* Σ-erfüllbar *genau dann, wenn ein Modell für $\Sigma \cup Fm(P) \subseteq Fm_{PL}(S(K_0))$ existiert.*

   ($Fm(P)$ bezeichne wiederum die Menge der Formeln, mit denen Knoten des Zweiges P markiert sind.)

   Ein Σ-Tableau heißt Σ-erfüllbar *genau dann, wenn im Tableau ein Zweig, der Σ-erfüllbar ist, existiert.*

2. *Eine Σ-Resolutionsexpansion $\beta_1, \ldots, \beta_m$ heißt* Σ-erfüllbar *genau dann, wenn*

   $$\Sigma \cup \{\beta_1, \ldots, \beta_m\} \subseteq Fm_{PL}(S(K_0))$$

   erfüllbar ist.

Lemma 2.87 *Sei $\Sigma \cup \{\alpha\} \subseteq Fm_{PL}(S)$ eine Menge von Aussagen.*

1. *Sei τ ein Σ-erfüllbares Σ-Tableau für $\{\alpha\}$, in dem nur Aussagen vorkommen. τ' sei das Resultat einer einmaligen Anwendung einer Σ-Tableau-Regel auf τ. In τ' kommen dann ebenfalls nur Aussagen vor und τ' ist Σ-erfüllbar.*

2. *Sei S eine Σ-erfüllbare Σ-Resolutionsexpansion, die nur Aussagen enthält, und ist S' das Resultat einer Σ-Resolutionsexpansionsregel oder einer Resolventenbildung für S, so ist S' Σ-erfüllbar und enthält ebenfalls nur Aussagen.*

Beweis.

1. Der Beweis entspricht dem von Lemma 1.45 aus dem Kapitel Aussagenlogik, muss aber noch um die beiden im Falle der Prädikatenlogik zusätzlich möglichen Fälle ergänzt werden.

   Sei also P ein erfüllbarer Zweig des Tableaus τ, und τ' sei durch eine γ-Expansionsregel $\frac{\gamma}{\gamma(t)}$ oder durch eine δ-Expansionsregel $\frac{\boldsymbol{\delta}}{\boldsymbol{\delta}(c)}$ angewandt auf eine Aussage N des Zweiges P hervorgegangen. Der resultierende Zweig sei P'.

   N γ**-Formel:** Nach Voraussetzung ist $Fm(P)$ Σ-erfüllbar. Da in einem Tableau nur endlich viele Konstantensymbole neu eingeführt worden sein konnten, gilt weiter $Fm(P) \subseteq Fm_{PL}(S(\mathcal{K}))$ für ein endliches $\mathcal{K}$. Da $\boldsymbol{\gamma} \in Fm(P)$, ist nach Lemma 2.74 (1) dann auch $Fm(P) \cup \{\boldsymbol{\gamma}(t)\}$ und damit $Fm(P')$ Σ-erfüllbar.

   N δ**-Formel:** Es ist $\boldsymbol{\delta} \in Fm(P)$. Nach Voraussetzung ist $Fm(P)$ wiederum Σ-erfüllbar und es gilt $Fm(P) \subseteq Fm_{PL}(S(\mathcal{K}))$ für ein endliches $\mathcal{K}$. Nach Lemma 2.74 (2) ist dann auch $Fm(P') = Fm(P) \cup \{\boldsymbol{\delta}(c)\}$ für ein bezogen auf $S(\mathcal{K})$ neues Konstantensymbol c ebenfalls Σ-erfüllbar.

   In jedem Fall ist also der Zweig P' Σ-erfüllbar, und somit auch τ'.

2. Der Resolutionskalkül wird ganz analog behandelt.

■

Die strenge Korrektheit von Tableau- und Resolutionskalkül folgt nun wie im Falle der Aussagenlogik.

Theorem 2.88 (Strenge Korrektheit) *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält und sei weiter $\Sigma \cup \{\alpha\} \subseteq Fm_{PL}(S)$ eine Menge von Aussagen.*

1. $\Sigma \vdash_{PL-T} \alpha \quad \Rightarrow \quad \Sigma \models_{PL} \alpha$
2. $\Sigma \vdash_{PL-R} \alpha \quad \Rightarrow \quad \Sigma \models_{PL} \alpha$

Beweis. Einfache Übungsaufgabe. ■

2.2.4 Hintikka's Lemma

Unser nächstes Ziel ist der Beweis der strengen Vollständigkeit von Tableau- und Resolutionskalkül. In Analogie zur Vorgehensweise in der Aussagenlogik definieren wir zunächst Hintikka-Mengen und zeigen dann deren Erfüllbarkeit in Herbrand-Strukturen.

Definition 2.89 *Eine Menge $\Sigma \subseteq Fm_{PL}(S)$ von Aussagen heißt* prädikatenlogische Hintikka-Menge *genau dann, wenn gilt:*

1. *Keine atomare Formel ist zugleich mit ihrer Negation in Σ enthalten.*
2. $\bot \notin \Sigma$, $\neg\top \notin \Sigma$
3. $\neg\neg\beta \in \Sigma \Rightarrow \beta \in \Sigma$
4. $\boldsymbol{\alpha} \in \Sigma \Rightarrow \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Sigma$ *(für α-Formeln)*
5. $\boldsymbol{\beta} \in \Sigma \Rightarrow \boldsymbol{\beta}_1 \in \Sigma$ *oder* $\boldsymbol{\beta}_2 \in \Sigma$ *(für β-Formeln)*
6. $\boldsymbol{\gamma} \in \Sigma \Rightarrow \boldsymbol{\gamma}(t) \in \Sigma$ *für alle Grundterme* $t \in Tm_{PL}(S)$ *(für γ-Formeln)*
7. $\boldsymbol{\delta} \in \Sigma \Rightarrow \boldsymbol{\delta}(t) \in \Sigma$ *für einen Grundterm* $t \in Tm_{PL}(S)$ *(für δ-Formeln)*

Wie im Falle der Aussagenlogik haben im Falle der Prädikatenlogik Hintikka-Mengen die schöne Eigenschaft, erfüllbar zu sein.

Lemma 2.90 (Hintikka) *Sei S eine Signatur mit nichtleerer Konstantensymbolmenge und $\Sigma \subseteq Fm_{PL}(S)$ eine prädikatenlogische Hintikka-Menge. Dann gilt:*
Σ hat ein Herbrand-Modell.

Beweis. Sei $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$. Wir definieren eine Herbrand-Struktur $\mathcal{A}$ zur Signatur S. Da in Herbrand-Strukturen der Grundbereich, die Interpretation der Funktionssymbole und die Interpretation der Konstantensymbole schon festgelegt ist, ist lediglich die Interpretationen der Prädikatensymbole anzugeben.
Sei H das Herbrand-Universum zur Signatur S. Es ist aufgrund der Voraussetzung über S $H \neq \emptyset$.
Für ein n-stelliges Prädikatensymbol $P \in \mathcal{R}$ und für beliebige Elemente (Grundterme) $t_1, \ldots, t_n \in H$ definieren wir

$$(t_1, \ldots, t_n) \in P^{\mathcal{A}} \text{ genau dann, wenn } P(t_1, \ldots, t_n) \in \Sigma.$$

Für die damit dann festgelegte Herbrand-Struktur $\mathcal{A}$ ist dann zu zeigen, dass sie ein Modell von Σ ist. Dies wird über eine strukturelle Induktion über α-, β-, γ-, δ-Formeln erreicht. Sei dazu die in der strukturellen Induktion benutzte Eigenschaft E folgendermaßen festgelegt: E trifft auf eine Formel α zu genau dann, wenn gilt:

$$\alpha \in \Sigma \ \Rightarrow\ \alpha^{\mathcal{A}} = T.$$

Da die Eigenschaft E trivialerweise auf alle Formeln α, die keine Aussagen sind, zutrifft (die „Prämisse" $\alpha \in \Sigma$ ist dann nicht erfüllt, da Σ eine Menge von Aussagen ist), können wir uns im Induktionsbeweis auf die Betrachtung von Aussagen beschränken.

Ind.Basis: Für die atomaren Formeln $\top$, $\bot$, und die Negationen hiervon gilt E offensichtlich.

Sei $P(t_1, \ldots, t_n) \in \Sigma$ eine atomare Aussage. Es gilt

$$\begin{aligned}[P(t_1, \ldots, t_n)]^{\mathcal{A}} = T \ &\Leftrightarrow\ \left(t_1^{\mathcal{A}}, \ldots, t_n^{\mathcal{A}}\right) \in P^{\mathcal{A}} \\ &\Leftrightarrow\ (t_1, \ldots, t_n) \in P^{\mathcal{A}}.\end{aligned}$$

Nach Definition von $P^{\mathcal{A}}$ trifft E somit auf $P(t_1, \ldots, t_n)$ zu.

Gilt $\neg P(t_1, \ldots, t_n) \in \Sigma$, so ist $P(t_1, \ldots, t_n) \notin \Sigma$, da Σ eine Hintikka-Menge ist. Damit gilt dann $(t_1, \ldots, t_n) \notin P^{\mathcal{A}}$, also $[P(t_1, \ldots, t_n)]^{\mathcal{A}} = F$ und damit $[\neg P(t_1, \ldots, t_n)]^{\mathcal{A}} = T$, d.h. E trifft auf $\neg P(t_1, \ldots, t_n)$ zu.

Ind.Schritt: Wir betrachten exemplarisch den Fall einer δ-Formel.

Sei also $\boldsymbol{\delta} \in \Sigma$. Da Σ eine Hintikka-Menge ist, gibt es einen Grundterm $t \in H$ mit $\boldsymbol{\delta}(t) \in \Sigma$. Nach Induktionsvoraussetzung trifft E auf $\boldsymbol{\delta}(t)$ zu, d.h. es gilt $[\boldsymbol{\delta}(t)]^{\mathcal{A}} = T$.

Ist $\boldsymbol{\delta}$ eine Formel der Gestalt $(\exists x)\,\alpha$, so folgt mit Theorem 2.73 (2) sofort $\boldsymbol{\delta}^{\mathcal{A}} = T$.

Ist $\boldsymbol{\delta}$ von der Gestalt $\neg(\forall x)\,\alpha$, also $\boldsymbol{\delta}(t) = \neg\alpha\{x/t\}$, so folgt mit Theorem 2.73 (2) $[(\exists x)\,\neg\alpha]^{\mathcal{A}} = T$ und damit also auch $\boldsymbol{\delta}^{\mathcal{A}} = T$.

■

2.2.5 Das Modellexistenztheorem der Prädikatenlogik

Im Beweis der Vollständigkeit der aussagenlogischen Tableau- bzw. Resolutionskalküle hatten wir in ganz wesentlicher Weise das Modellexistenztheorem der Aussagenlogik benutzt. Dieses besagte, dass jede in einer aussagenlogischen Konsistenzeigenschaft enthaltene Formelmenge erfüllbar ist. Die Beweisidee für das Modellexistenztheorem war dabei die folgende: wir erweitern eine gegebene Formelmenge Σ mit dieser Konsistenzeigenschaft zu einer maximalen Formelmenge Σ', die noch diese Konsistenzeigenschaft besitzt. Aufgrund der Maximalität und aufgrund der AbSchlussbedingungen von Konsistenzeigenschaften musste diese Formelmenge Σ' dann eine Hintikka-Menge und somit erfüllbar sein.
Wir wollen nun für die Prädikatenlogik ebenfalls den Begriff der Konsistenzeigenschaft einführen und anschließend ein Modellexistenztheorem beweisen.
Wir wollen uns – dies sei dann eine generelle Voraussetzung für die nachfolgenden Abschnitte – immer auf eine Signatur S mit höchstens äbzahlbar unendlich vielen Relations-, Funktions- und Konstantensymbolen

beschränken. $S(\mathcal{K}_0)$ sei dann eine Konstantenerweiterung von S um abzählbar unendlich viele neue Konstantensymbole. Diese hier getroffene Voraussetzung bezüglich der Signatur S stellt für die Praxis keine Beschränkung dar. Signaturen mit überabzählbar vielen Relations-, Funktions- oder Konstantensymbolen sind nur für theoretische Betrachtungen von Bedeutung. Die mit dem Modellexistenztheorem noch zu beweisenden Resultate (Vollständigkeitssatz, Kompaktheitssatz und mit geringfügigen Anpassungen auch der Satz von Löwenheim-Skolem) gelten auch für den allgemeineren Fall, sind jedoch etwas schwieriger zu beweisen.

Definition 2.91 *Eine Formelmenge $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$ heißt* $\mathcal{K}_0$-passend *genau dann, wenn es unendliche viele Konstantensymbole $c \in \mathcal{K}_0$ gibt, die in Σ nicht vorkommen.*

Definition 2.92

1. *Es sei $\mathcal{C} \subseteq \mathcal{P}(Fm_{PL}(S(\mathcal{K}_0)))$ eine Menge (Kollektion) von Aussagenmengen.*
   $\mathcal{C}$ heißt prädikatenlogische Konsistenzeigenschaft *genau dann, wenn für alle $\Sigma \in \mathcal{C}$ gilt:*
   (a) *ist α eine atomare Formel und $\alpha \in \Sigma$, so ist $\neg\alpha \notin \Sigma$.*
   (b) $\perp \notin \Sigma$, $\neg\top \notin \Sigma$.
   (c) *ist $\neg\neg\beta \in \Sigma$, so ist $\Sigma \cup \{\beta\} \in \mathcal{C}$.*
   (d) *für eine α-Formel gilt: $\boldsymbol{\alpha} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$.*
   (e) *für eine β-Formel gilt: $\boldsymbol{\beta} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\beta}_1\} \in \mathcal{C}$ oder $\Sigma \cup \{\boldsymbol{\beta}_2\} \in \mathcal{C}$.*
   (f) *für eine γ-Formel gilt: $\boldsymbol{\gamma} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\gamma}(t)\} \in \mathcal{C}$ für jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$.*
   (g) *für eine δ-Formel gilt: $\boldsymbol{\delta} \in \Sigma \;\Rightarrow\; \Sigma \cup \{\boldsymbol{\delta}(c)\} \in \mathcal{C}$ für jedes bzgl. Σ neue Konstantensymbol $c \in \mathcal{K}_0$.*

2. *Die prädikatenlogische Konsistenzeigenschaft $\mathcal{C}$ heißt* teilmengenabgeschlossen *genau dann, wenn aus $\Sigma \in \mathcal{C}$ und $\Sigma_1 \subseteq \Sigma$ folgt $\Sigma_1 \in \mathcal{C}$.*

3. *$\mathcal{C}$ heißt* von endlichem Charakter *genau dann, wenn für $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$ gilt:*
   $\Sigma \in \mathcal{C} \;\Leftrightarrow$ jede endliche Teilmenge von Σ ist in $\mathcal{C}$.

Die Motivation für die Eigenschaft (1) (g) in der Definition 2.92 ist die folgende. Falls in einer maximalen Formelmenge $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$, die diese Konsistenzeigenschaft besitzt, eine δ-Formel aus $Fm_{PL}(S)$ vorkommt, so muss, damit Σ eine Hintikka-Menge sein kann, für einen Grundterm t die δ-Instanz $\boldsymbol{\delta}(t)$ in Σ sein. Dieser Grundterm t soll nun aber genau die „Existenzbehauptung" für $\boldsymbol{\delta}$ sicherstellen und nicht möglicherweise die von weiteren Formeln. Dies erreichen wir dadurch, dass wir ein neues, noch nicht verwendetes Konstantensymbol c einführen, so dass jeder δ-Formel genau ein solches Konstantensymbol entspricht. Nach Lemma 2.74 wissen wir dann, dass mit $\boldsymbol{\delta}$ auch $\{\boldsymbol{\delta}, \boldsymbol{\delta}(c)\}$ für die δ-Instanz $\boldsymbol{\delta}(c)$ ein Modell hat.
Das nachfolgende Lemma hält einige einfache Eigenschaften von prädikatenlogischen Konsistenzeigenschaften fest, die wir im Beweis des Modellexistenztheorems benötigen und die wir auch schon im aussagenlogischen Fall kennengelernt haben.

Lemma 2.93

1. *Jede prädikatenlogische Konsistenzeigenschaft $\mathcal{C}$ kann zu einer prädikatenlogischen Konsistenzeigenschaft, die teilmengenabgeschlossen ist, erweitert werden.*

2. *Jede prädikatenlogische Konsistenzeigenschaft von endlichem Charakter ist teilmengenabgeschlossen.*

3. *Jede prädikatenlogische Konsistenzeigenschaft, die teilmengenabgeschlossen ist, kann zu einer von endlichem Charakter erweitert werden.*

4. *$\mathcal{C}$ sei eine prädikatenlogische Konsistenzeigenschaft von endlichem Charakter.*
   Gilt $\Sigma_0, \Sigma_1, \Sigma_2, \ldots \in \mathcal{C}$ und gilt $\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots$, so ist $\bigcup_{i\in\mathbb{N}} \Sigma_i \in \mathcal{C}$.

Beweis. Verläuft ganz analog zu dem aussagenlogischen Fall. ∎

Theorem 2.94 (Prädikatenlogisches Modellexistenztheorem) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen und $\mathcal{C} \subseteq \mathcal{P}(Fm_{PL}(S(\mathcal{K}_0)))$ eine prädikatenlogische Konsistenzeigenschaft mit $\Sigma \in \mathcal{C}$.*
Dann gilt:
Σ ist erfüllbar in einem Herbrand-Modell zur Signatur $S(\mathcal{K}_0)$.

Beweis. Nach Lemma 2.93 kann angenommen werden, dass $\mathcal{C}$ von endlichem Charakter und teilmengenabgeschlossen ist.
Aufgrund der generellen Voraussetzung über die Signatur S bzw. $S(\mathcal{K}_0)$ ist $Fm_{PL}(S(\mathcal{K}_0))$ abzählbar. Es existiert daher eine Aufzählung $\alpha_0, \alpha_1, \alpha_2, \ldots$ aller Aussagen in $Fm_{PL}(S(\mathcal{K}_0))$. Wir definieren eine Folge von Aussagenmengen $\Sigma_0 \subseteq \Sigma_1 \subseteq \ldots$, die alle in $\mathcal{C}$ liegen:
$\Sigma_0 := \Sigma \in \mathcal{C}$
und für $n \geq 0$:

$$\Sigma_{n+1} := \begin{cases} \Sigma_n & \text{falls} \Sigma_n \cup \{\alpha_n\} \notin \mathcal{C} \\ \Sigma_n \cup \{\alpha_n\} & \text{falls } \Sigma_n \cup \{\alpha_n\} \in \mathcal{C} \text{ und } \alpha_n \text{ keine } \delta - Formel \\ \Sigma_n \cup \{\alpha_n\} \cup \{\boldsymbol{\delta}(c)\} & \text{falls } \Sigma_n \cup \{\alpha_n\} \in \mathcal{C},\ \alpha_n = \boldsymbol{\delta} \text{ für eine } \delta\text{-Formel, } c \text{ neu für } \Sigma_n \cup \{\alpha_n\} \end{cases}$$

(Bemerkung: In der Definition von Σ_{n+1} existiert ein solches „neues" Konstantensymbol c immer, da zuvor höchstens endlich viele Konstantensymbole aus $\mathcal{K}_0$ neu eingeführt worden sind.)
Durch Induktion nach n folgt für alle $n \in \mathbb{N}$ $\Sigma_n \in \mathcal{C}$ und $\Sigma_n \subseteq \Sigma_{n+1}$. Aufgrund von Lemma 2.93 (4) ist daher $H := \bigcup_{n \in \mathbb{N}} \Sigma_n \in \mathcal{C}$. Weiter gilt:

1. H ist bezüglich $\subseteq$ maximales Element in $\mathcal{C}$:

   Der Beweis entspricht dem im aussagenlogischen Fall.

2. H ist eine prädikatenlogische Hintikka-Menge zur Signatur $S(\mathcal{K}_0)$:

   Die Bedingungen (1) - (5) für eine prädikatenlogische Hintikka-Menge werden wie im aussagenlogischen Fall gezeigt.

   Bedingung (6) :

   Für eine γ-Formel sei $\boldsymbol{\gamma} \in H$. Da $\mathcal{C}$ eine prädikatenlogische Konsistenzeigenschaft ist, gilt $H \subseteq H \cup \{\gamma(t)\} \in \mathcal{C}$ für jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$. Aufgrund der Maximalitätseigenschaft von H erhalten wir somit $\boldsymbol{\gamma}(t) \in H$ für jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$. Damit ist die Bedingung für γ-Formeln erfüllt.

   Bedingung (7) :

   Für eine δ-Formel sei $\boldsymbol{\delta} \in H$. Hier können wir nicht wie für die γ-Formel argumentieren. Da $H \in \mathcal{C}$, muss zwar $H \subseteq H \cup \{\delta(c)\} \in \mathcal{C}$ für jedes bezüglich H neue Konstantensymbol $c \in \mathcal{K}_0$ gelten. Möglicherweise kommen aber in H bereits alle Konstantensymbole aus $\mathcal{K}_0$ vor, so dass wir auf diese Weise nicht $\boldsymbol{\delta}(c) \in H$ für ein bezüglich H neues Konstantensymbol c sicherstellen können.

   $\boldsymbol{\delta}$ ist eine Aussage und kommt somit in der Aufzählung $\alpha_0, \alpha_1, \alpha_2, \ldots$ vor. Wegen $\boldsymbol{\delta} \in H = \bigcup_{n \in \mathbb{N}} \Sigma_n$ gibt es damit ein kleinstes $n \in \mathbb{N}$ mit $\alpha_n = \boldsymbol{\delta}$ und $\Sigma_n \cup \{\alpha_n\} \in \mathcal{C}$. Nach Definition von Σ_{n+1} für diesen Fall gilt dann aber auch $\boldsymbol{\delta}(c) \in \Sigma_{n+1} \subseteq H$ für ein Konstantensymbol c, das bezüglich $\Sigma_n \cup \{\alpha_n\}$ neu ist. Damit ist auch in diesem Fall die Hintikka-Eigenschaft nachgewiesen.

Es folgt somit, dass H eine prädikatenlogische Hintikka-Menge zur Signatur $S(\mathcal{K}_0)$ ist. Nach dem Lemma von Hintikka ist dann H und damit auch Σ in einem Herbrand-Modell zur Signatur $S(\mathcal{K}_0)$ erfüllbar. ∎

2.2.6 Strenge Vollständigkeit von Tableau- und Resolutions-Kalkül

Mit dem Modellexistenztheorem haben wir nun das zentrale Resultat, um die strenge Vollständigkeit von Tableau- und Resolutionskalkül zeigen zu können. Die Vorgehensweise entspricht dabei derjenigen, die wir im Falle der Aussagenlogik kennengelernt haben.
Wir benötigen noch ein Lemma, das besagt, dass in einem Σ-Tableau neu eingeführte Konstantensymbole durch andere, nicht vorkommende Konstantensymbole ersetzt werden können.

Lemma 2.95 *Ist τ ein Σ-Tableau, in dem das Konstantensymbol $c \in \mathcal{K}_0$ vorkommt, und geht τ' aus τ dadurch hervor, dass (jedes Vorkommen von) c durch ein im Σ-Tableau τ nicht vorkommendes Konstantensymbol $d \in \mathcal{K}_0$ ersetzt wird, so ist τ' ebenfalls ein Σ-Tableau.*
Ferner gilt:
τ ist geschlossen genau dann, wenn τ' geschlossen ist.

Beweis. Einfache Übungsaufgabe. ■

Definition 2.96 *Sei $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$ eine Menge von Aussagen, Σ $\mathcal{K}_0$-passend, und $\alpha \in Fm_{PL}(S(\mathcal{K}_0))$ eine Aussage.*
Σ heißt α-tableaukonsistent genau dann, wenn kein geschlossenes Σ-Tableau für $\{\neg\alpha\}$ existiert.

Lemma 2.97 *Sei $\alpha \in Fm_{PL}(S)$ eine Aussage und*

$$\mathcal{C}_\alpha := \{\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0)) \mid \Sigma \text{ Menge von Aussagen, } \Sigma\ \alpha\text{-tableaukonsistent und } \mathcal{K}_0\text{-passend}\}.$$

Dann gilt:
$\mathcal{C}_\alpha$ ist eine prädikatenlogische Konsistenzeigenschaft.

Beweis. Wir müssen für $\mathcal{C}_\alpha$ die Eigenschaften einer prädikatenlogischen Konsistenzeigenschaft nachweisen. Die meisten Fälle werden wie im aussagenlogischen Fall bewiesen. Wir werden exemplarisch einige Fälle behandeln.

1. Sei $\Sigma \in \mathcal{C}_\alpha$ und $\beta \in \Sigma$ für eine atomare Aussage β. Damit kann $\neg\beta$ kein Element von Σ sein, da sonst trivialerweise ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$ existiert und Σ dann nicht α-tableaukonsistent wäre.

2. Der „kompliziertteste" Fall ist der der β-Formeln.

   Sei $\Sigma \in \mathcal{C}_\alpha$ und $\boldsymbol{\beta} \in \Sigma$ eine β-Formel. Angenommen weder $\Sigma \cup \{\boldsymbol{\beta}_1\}$ noch $\Sigma \cup \{\boldsymbol{\beta}_2\}$ sind α-tableaukonsistent. Wir zeigen, dass dann auch Σ nicht α-tableaukonsistent ist im Widerspruch zu $\Sigma \in \mathcal{C}_\alpha$.

   Da $\Sigma \cup \{\boldsymbol{\beta}_i\}$, $i = 1, 2$, nach Annahme α-tableau-inkonsistent ist, existieren geschlossene $\Sigma \cup \{\boldsymbol{\beta}_i\}$-Tableaus, $i = 1, 2$, für $\{\neg\alpha\}$:

$$\begin{array}{c} \neg\alpha \\ | \\ \boldsymbol{\beta}_i \\ | \\ T_i \end{array}$$

   Aufgrund von Lemma 2.95 können wir annehmen, dass es in $\mathcal{K}_0$ kein Konstantensymbol gibt, das sowohl in T_1 als auch in T_2 vorkommt. Ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$ erhalten wir dann folgendermaßen:

$$\begin{array}{ccc} & \neg\alpha & \\ & | & \\ & \boldsymbol{\beta} & \\ / & & \backslash \\ \boldsymbol{\beta}_1 & & \boldsymbol{\beta}_2 \\ | & & | \\ T_1 & & T_2 \end{array}$$

Aufgrund der Voraussetzungen über T_1 und T_2 sind die Bedingungen für ein korrekt gebildetes Tableau erfüllt. Außerdem ist dieses Tableau geschlossen, so dass also ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$ vorliegt. Es folgt, dass Σ α-tableau-inkonsistent ist im Widerspruch zur Voraussetzung $\Sigma \in \mathcal{C}_\alpha$.

3. Sei $\Sigma \in \mathcal{C}_\alpha$ und $\boldsymbol{\delta} \in \Sigma$.

   Angenommen, $\Sigma \cup \{\boldsymbol{\delta}(c)\}$ für ein bezüglich Σ neues Konstantensymbol $c \in \mathcal{K}_0$ sei α-tableau-inkonsistent. Es existiert dann ein geschlossenes $\Sigma \cup \{\boldsymbol{\delta}(c)\}$-Tableau T für $\{\neg\alpha\}$. Betrachte

$$\begin{array}{c} \neg\alpha \\ | \\ \boldsymbol{\delta} \\ | \\ \boldsymbol{\delta}(c) \\ | \\ T' \end{array}$$

   wobei T' aus T dadurch hervorgeht, dass in T jedes $\boldsymbol{\delta}(c)$, das aus einer Prämissenregel $\frac{}{\boldsymbol{\delta}(c)}$ resultiert, durch eine neue $\boldsymbol{\delta}$-Instanz $\boldsymbol{\delta}(d)$ ersetzt wird, wobei $d \in \mathcal{K}_0$ ein Konstantensymbol ist, das bis dahin noch nicht verwendet wurde.

   Es ist leicht zu verifizieren, dass hierdurch ein korrekt gebildetes Σ-Tableau erhalten wird, das genau dann geschlossen ist, wenn T geschlossen ist. Da T nach Annahme geschlossen ist, existiert somit ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$ im Widerspruch zu $\Sigma \in \mathcal{C}_\alpha$.

■

Nun ist es leicht, die strenge Vollständigkeit des Tableau-Kalküls zu zeigen.

Theorem 2.98 (Strenge Vollständigkeit des Tableau-Kalküls) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen und $\alpha \in Fm_{PL}(S)$ eine Aussage.*
Dann gilt:
$\Sigma \models_{PL} \alpha \;\Rightarrow\; \Sigma \vdash_{PL-T} \alpha$.

Beweis. Wir zeigen die Kontraposition:

$$\Sigma \nvdash_{PL-T} \alpha \;\Rightarrow\; \Sigma \not\models_{PL} \alpha.$$

Es gelte $\Sigma \nvdash_{PL-T} \alpha$.
Dann ist Σ α-tableaukonsistent. Damit ist dann auch $\Sigma \cup \{\neg\alpha\}$ tableaukonsistent für α. Nach Lemma 2.97 ist die dort angegebene Menge $\mathcal{C}_\alpha$ eine prädikatenlogische Konsistenzeigenschaft. Ferner ist $\Sigma \cup \{\neg\alpha\} \in \mathcal{C}_\alpha$. Nach dem Modellexistenztheorem ist $\Sigma \cup \{\neg\alpha\}$ dann erfüllbar, also gilt $\Sigma \not\models_{PL} \alpha$. ■

Die strenge Vollständigkeit des Resolutionskalküls wird mit entsprechenden Anpassungen wie im aussagenlogischen Fall bzw. im Fall des prädikatenlogischen Tableau-Kalküls gezeigt. Hier soll ohne Beweis nur noch das Ergebnis festgehalten werden. Der Vollständigkeit halber geben wir die noch dazu erforderlichen Begriffsbildungen und Lemmata an.

Lemma 2.99 *Sei $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$ eine Menge von Aussagen. Ferner seien Σ_1, Σ_2 endliche Mengen von verallgemeinerten Disjunktionen, bestehend aus Aussagen aus $Fm_{PL}(S(\mathcal{K}_0))$, wobei Σ_2 eine α-Erweiterung von Σ_1 sei.*
Ist $\beta_1, \ldots, \beta_m$ eine Σ-Resolutionsexpansion für Σ_1 und enthält α kein Konstantensymbol aus $\mathcal{K}_0$, das in $\beta_1, \ldots, \beta_m$ vorkommt, dann gibt es eine α-Erweiterung $\beta_1', \ldots, \beta_m'$ von $\beta_1, \ldots, \beta_m$, die eine Σ-Resolutionsexpansion für Σ_2 ist.

Definition 2.100 *Sei $\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0))$ eine Menge von Aussagen, Σ $\mathcal{K}_0$-passend, und*
$\alpha \in Fm_{PL}(S(\mathcal{K}_0))$ eine Aussage.
Σ heißt α-resolutionskonsistent genau dann, wenn keine geschlossene Σ-Resolutionsexpansion für $\{\neg\alpha\}$ existiert.

Lemma 2.101 *Sei $\alpha \in Fm_{PL}(S)$ eine Aussage und*

$$\mathcal{C}_\alpha := \{\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0)) \mid \Sigma \text{ Menge von Aussagen, } \alpha\text{-resolutionskonsistent und } \mathcal{K}_0\text{-passend}\}$$

Dann gilt:
$\mathcal{C}_\alpha$ ist eine prädikatenlogische Konsistenzeigenschaft.

Theorem 2.102 (Strenge Vollständigkeit des Resolutionskalküls) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen und $\alpha \in Fm_{PL}(S)$ eine Aussage.*
Dann gilt:
$\Sigma \models_{PL} \alpha \ \Rightarrow\ \Sigma \vdash_{PL-R} \alpha$.

2.2.7 Der Sequenzenkalkül

Wir wollen nun als einen weiteren Beweiskalkül den Sequenzenkalkül einführen. Waren die bisher betrachteten prädikatenlogischen Beweiskalküle Widerlegungskalküle insofern, als ein Beweis einer Formel in der Widerlegung der Negation dieser Formel bestand, so ist der Sequenzenkalkül ein positiver Beweiskalkül in dem Sinne, dass ausgehend von Axiomen und Prämissen die zu beweisende Formel unter Verwendung der zulässigen Schlussregeln hergeleitet wird. Über diesen Aspekt hinaus bietet die Betrachtung des Sequenzenkalküls den Vorteil, in einfacher Weise einen Beweiskalkül für die intuitionistische Logik zu erhalten, einen Zugang zu den sogenannten substrukturellen Logiken, insbesondere der Linearen Logik, zu liefern und auch die Bedeutung der Schnittregel diskutieren und demonstrieren zu können.

Definition 2.103 *Eine* Sequenz *ist ein geordnetes Paar (Γ, Δ) von zwei endlichen (möglicherweise auch leeren) Folgen Γ und Δ von prädikatenlogischen Formeln. Eine solche Sequenz notieren wir als $\Gamma \rhd \Delta$ mit $\rhd$ als* Sequenzsymbol. *Γ wird auch als das* Antecedens *und Δ als das* Succedens *der Sequenz bezeichnet.*

Informell ist eine solche Sequenz so zu verstehen: wenn jede Formel im „Antecedens" Γ wahr ist, dann ist wenigstens eine Formel im „Succedens" Δ wahr.

Definition 2.104 *Eine* Schlussregel *ist ein Ausdruck der Form*

$$\frac{S_1}{S} \quad \textit{oder} \quad \frac{S_1 \quad S_2}{S},$$

wobei S_1, S_2 und S Sequenzen sind. S_1 bzw. S_1 und S_2 heißen Obersequenzen *und S heißt* Untersequenz *einer solchen Schlussregel.*

Im Sequenzenkalkül haben wir die folgenden Axiome und Schlussregeln:

1. Axiome:
$$(Ax) \quad \alpha \rhd \alpha \qquad (L\bot) \quad \bot \rhd \qquad (R\top) \quad \rhd \top$$

2. Abschwächungsregeln:
$$(wL) \quad \frac{\Gamma \rhd \Delta}{\alpha, \Gamma \rhd \Delta} \qquad (wR) \quad \frac{\Gamma \rhd \Delta}{\Gamma \rhd \Delta, \alpha}$$

3. Vertauschungsregeln:
$$(eL) \quad \frac{\Gamma_1, \alpha, \beta, \Gamma_2 \rhd \Delta}{\Gamma_1, \beta, \alpha, \Gamma_2 \rhd \Delta} \qquad (eR) \quad \frac{\Gamma \rhd \Delta_1, \alpha, \beta, \Delta_2}{\Gamma \rhd \Delta_1, \beta, \alpha, \Delta_2}$$

4. Kontraktionsregeln:
$$(cL) \quad \frac{\Gamma_1, \alpha, \alpha, \Gamma_2 \rhd \Delta}{\Gamma_1, \alpha, \Gamma_2 \rhd \Delta} \qquad (cR) \quad \frac{\Gamma \rhd \Delta_1, \alpha, \alpha, \Delta_2}{\Gamma \rhd \Delta_1, \alpha, \Delta_2}$$

5. $\neg$-Einführungsregeln:
$$(\neg L)\quad \frac{\Gamma \rhd \Delta, \alpha}{\neg\alpha, \Gamma \rhd \Delta} \qquad\qquad (\neg R)\quad \frac{\alpha, \Gamma \rhd \Delta}{\Gamma \rhd \Delta, \neg\alpha}$$

6. $\wedge$-Einführungsregeln:
$$(\wedge L)\quad \frac{\Gamma, \alpha, \beta \rhd \Delta}{\Gamma, \alpha \wedge \beta \rhd \Delta} \qquad\qquad (\wedge R)\quad \frac{\Gamma \rhd \Delta, \alpha \qquad \Gamma \rhd \Delta, \beta}{\Gamma \rhd \Delta, \alpha \wedge \beta}$$

7. $\vee$-Einführungsregeln:
$$(\vee L)\quad \frac{\Gamma, \alpha \rhd \Delta \qquad \Gamma, \beta \rhd \Delta}{\Gamma, \alpha \vee \beta \rhd \Delta} \qquad\qquad (\vee R)\quad \frac{\Gamma \rhd \Delta, \alpha, \beta}{\Gamma \rhd \Delta, \alpha \vee \beta}$$

8. $\rightarrow$-Einführungsregeln:
$$(\rightarrow L)\quad \frac{\Gamma \rhd \Delta, \alpha \qquad \Gamma, \beta \rhd \Delta}{\Gamma, \alpha \rightarrow \beta \rhd \Delta} \qquad\qquad (\rightarrow R)\quad \frac{\Gamma, \alpha \rhd \Delta, \beta}{\Gamma \rhd \Delta, \alpha \rightarrow \beta}$$

9. Schnittregel:
$$(cut)\quad \frac{\Gamma \rhd \Delta, \alpha \qquad \alpha, \Gamma \rhd \Delta}{\Gamma \rhd \Delta}$$

10. $\forall$-Einführungsregeln:
$$(\forall L)\quad \frac{\alpha\{x/t\}, \Gamma \rhd \Delta}{(\forall x)\,\alpha, \Gamma \rhd \Delta} \qquad\qquad (\forall R)\quad \frac{\Gamma \rhd \Delta, \alpha\{x/y\}}{\Gamma \rhd \Delta, (\forall x)\,\alpha}$$
wobei für die Regel $(\forall R)$ gilt: die Variable y kommt in den Formeln von Γ und Δ nicht frei vor; ferner gilt $x = y$ oder $y \notin fvar(\alpha)$.

11. $\exists$-Einführungsregeln:
$$(\exists L)\quad \frac{\alpha\{x/y\}, \Gamma \rhd \Delta}{(\exists x)\,\alpha, \Gamma \rhd \Delta} \qquad\qquad (\exists R)\quad \frac{\Gamma \rhd \Delta, \alpha\{x/t\}}{\Gamma \rhd \Delta, (\exists x)\,\alpha}$$
wobei für die Regel $(\exists L)$ gilt: die Variable y kommt in den Formeln von Γ und Δ nicht frei vor; ferner gilt $x = y$ oder $y \notin fvar(\alpha)$.

Axiome werden wir auch als *Ausgangssequenzen* bezeichnen.

Als nächstes wollen wir für den Sequenzenkalkül den Beweisbegriff definieren. Wir werden dies hier mehr informell tun.

Definition 2.105

1. *Ein* Beweis im Sequenzenkalkül LK *ist ein mit Sequenzen markierter Binärbaum, der den folgenden Bedingungen genügt:*
   (a) *die Blätter des Baumes sind mit Ausgangssequenzen markiert.*
   (b) *jeder Knoten des Baumes, der kein Blatt ist, ist mit der Untersequenz einer Schlussregel markiert, und der Sohn ist mit der Obersequenz (bzw. die Söhne sind mit den Obersequenzen) markiert.*

   Die Sequenz, mit der die Wurzel markiert ist, heißt Endsequenz.

   Eine Sequenz $\Gamma \rhd \Delta$ heißt beweisbar in *LK genau dann, wenn es einen Beweis in LK mit Endsequenz* $\Gamma \rhd \Delta$ *gibt.*

   Ein Beweis heißt schnittfrei, *wenn die Schnittregel im Beweis nicht vorkommt.*

   Eine Formel α heißt im Sequenzenkalkül beweisbar *genau dann,wenn die Sequenz* $\rhd \alpha$ *beweisbar ist (notiert als* $\vdash_{PL-LK} \alpha$).

2. *Eine* Herleitung einer Sequenz $\Gamma \rhd \Delta$ aus einer Menge Σ von Sequenzen *ist dann ein Beweis mit Endsequenz* $\Gamma \rhd \Delta$*, in dem zusätzlich auch jede Sequenz aus* Σ *als Ausgangssequenz zugelassen ist.*

   Eine Herleitung von α aus einer Menge Σ von Formeln *ist dann eine Herleitung der Sequenz* $\rhd \alpha$ *aus der Menge* $\{\rhd \beta \mid \beta \in \Sigma\}$ *(notiert als* $\Sigma \vdash_{PL-LK} \alpha$*).*

Beweise werden wir üblicherweise in der Form aufschreiben, dass die Wurzel des Baumes unten und die Blätter oben liegen.
Ein bemerkenswertes Resultat über den Sequenzenkalkül ist das folgende Schnitteliminationstheorem, auch als Gentzen's Hauptsatz bezeichnet:

Theorem 2.106 (Schnitteliminationstheorem)
Ist eine Sequenz in LK beweisbar, dann ist sie auch schnittfrei in LK beweisbar.

Auf den technisch aufwendigen Beweis können wir hier nicht eingehen. Eine eingehendere Analyse des Schnitteliminationstheorems zeigt aber, dass ein schnittfreier Beweis nur mit sehr viel mehr Aufwand erhalten werden kann. Orevkov (in: Lower bounds for increasing complexity of derivation after cut elimination, Journal of Soviet Mathematics 20, 1982, S. 2337-2350) gibt ein Beispiel einer Sequenz $\rhd B_k$ an, deren Beweisbaum linear in k viele Knoten enthält, wohingegen ein schnittfreier Beweis nicht in der Größenordnung

$$2^{2^{\cdot^{\cdot^{\cdot^{2^c}}}}} \qquad k \text{ Exponenten } 2$$

gefunden werden kann. Dieses Beispiel macht klar, dass den Formeln, die in einer Schnittregel eliminiert werden, eine ganz wesentliche Rolle im Beweis zukommt. Sie haben gewissermaßen die Funktion von Lemmata, mit deren Hilfe Beweise sehr viel übersichtlicher und kürzer geführt werden können. Im schnittfreien Kalkül ist die Beweissuche allerdings sehr viel einfacher. Dies wird u.a. durch das nachfolgende Theorem begründet. Wir verallgemeinern hierzu nur für diesen Abschnitt den Begriff der Teilformel, so wie wir ihn in Definition 2.18 eingeführt haben. Im Falle einer Formel $(\forall x)\,\alpha$ bzw. $(\exists x)\,\alpha$ der Signatur S ist für jeden Term $t \in Tm_{PL}(S)$ auch jede Teilformel von $\alpha\,\{x/t\}$ eine Teilformel von $(\forall x)\,\alpha$ bzw. $(\exists x)\,\alpha$. Insbesondere ist also $\alpha\,\{x/t\}$ eine Teilformel von $(\forall x)\,\alpha$ und $(\exists x)\,\alpha$.

Theorem 2.107 (Teilformeleigenschaft)
In einem schnittfreien Beweis sind alle auftretenden Formeln Teilformeln einer Formel der Endsequenz.

Zwischen schnittfreien Beweisen im Sequenzenkalkül und Tableau-Beweisen besteht ein sehr enger Zusammenhang. Wir wollen hier nicht weiter darauf eingehen, da wir, um diesen Zusammenhang beschreiben zu können, eine Variante des Tableau-Kalküls benötigen würden, die mit sogenannten markierten Tableaus arbeitet.
Für den Sequenzenkalkül gilt nun ebenfalls, dass er ein streng korrekter und vollständiger Beweiskalkül ist. Aus Zeitgründen werden wir einen Beweis, der z.B. der Vorgehensweise der entsprechenden Beweise für die beiden anderen Kalküle folgen könnte, nicht ausführen, sondern das Ergebnis nur noch in dem folgenden Theorem explizit formulieren.

Theorem 2.108 (Strenge Korrektheit und Vollständigkeit) *Sei* $\Sigma \subseteq Fm_{PL}(S)$ *eine Menge von Aussagen und* $\alpha \in Fm_{PL}(S)$ *eine Aussage. Dann gilt:*
$\Sigma \models_{PL} \alpha \;\Leftrightarrow\; \Sigma \vdash_{PL-LK} \alpha$.

2.2.8 Anwendungen des Modellexistenztheorems

Wie im Falle der Aussagenlogik ergeben sich mit dem Modellexistenztheorem einige wichtige Resultate. Hier ist zunächst der Kompaktheitssatz zu erwähnen, mit dem, wie schon im aussagenlogischen Fall gezeigt, Eigenschaften, die für alle endlichen Bereiche gelten, auf unendliche Bereiche übertragen werden können. Es gibt aber auch andere schöne Anwendungen, von denen wir eine hier behandeln werden.

Theorem 2.109 (Kompaktheitssatz der Prädikatenlogik) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen.*
Σ ist genau dann erfüllbar, wenn jede endliche Teilmenge von Σ erfüllbar ist.

Beweis.

„ $\Rightarrow$ “ : Offensichtlich.

„ $\Leftarrow$ “ : Sei

$\mathcal{C} := \{\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0)) \mid \Sigma\ \mathcal{K}_0$-passende Aussagenmenge, jede endliche Teilmenge von Σ erfüllbar$\}$.
$\mathcal{C}$ ist eine prädikatenlogische Konsistenzeigenschaft. Wir zeigen dies exemplarisch für den Fall der γ-Formel (vgl. Definition 2.92 (f)):

$\Sigma \in \mathcal{C},\ \boldsymbol{\gamma} \in \Sigma\ \Rightarrow\ \Sigma \cup \{\boldsymbol{\gamma}(t)\} \in \mathcal{C}$ für jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$.

Annahme: Es gibt einen Grundterm t mit $\Sigma \cup \{\boldsymbol{\gamma}(t)\} \notin \mathcal{C}$.

Dann existiert eine endliche Teilmenge Σ' von $\Sigma \cup \{\boldsymbol{\gamma}(t)\}$, die nicht erfüllbar ist. Da nach Voraussetzung $\Sigma \in \mathcal{C}$, ist somit $(\Sigma' \setminus \{\boldsymbol{\gamma}(t)\}) \cup \{\boldsymbol{\gamma}\}$ als endliche Teilmenge von Σ erfüllbar in einer Struktur $\mathcal{A}$. Damit gilt nach Theorem 2.73 für jeden Grundterm t $[\boldsymbol{\gamma}(t)]^{\mathcal{A}} = T$, also ist damit auch $\Sigma' \subseteq \Sigma' \cup \{\boldsymbol{\gamma}\}$ in $\mathcal{A}$ erfüllbar. Wir erhalten also einen Widerspruch. Es gilt daher $\Sigma \cup \{\boldsymbol{\gamma}(t)\} \in \mathcal{C}$ für alle Grundterme t.

Wir wissen somit, dass $\mathcal{C}$ eine prädikatenlogische Konsistenzeigenschaft ist. Σ ist aber offensichtlich in $\mathcal{C}$ und damit nach dem Modellexistenztheorem erfüllbar.

■

Beispiel 2.110 *Sei $S = (R)$ eine Signatur mit $ar(R) = 2$. S-Strukturen, also Strukturen der Gestalt $\mathcal{A} = (A; R^{\mathcal{A}})$ mit einem nichtleeren Grundbereich A und einer 2-stelligen Relation $R^{\mathcal{A}} \subseteq A \times A$, sind dann gerichtete Graphen. Die transitive Hülle $(R^{\mathcal{A}})^+$ von $R^{\mathcal{A}}$ ist als die Menge aller Paare $(a,b) \in A \times A$ definiert, für die ein endlicher $R^{\mathcal{A}}$-Pfad von a nach b existiert, d.h.*

$$\left(R^{\mathcal{A}}\right)^+ := \left\{(a,b) \mid ex.\ n > 0,\ ex.\ a_0, \ldots, a_n \in A\ mit\ a_0 = a,\ a_n = b,\ (a_i, a_{i+1}) \in R^{\mathcal{A}},\ i < n\right\}$$

Kann der Begriff der transitiven Hülle in einer prädikatenlogischen Sprache zur Signatur S definiert werden, d.h. gibt es eine Formel $\alpha \in Fm_{PL}(S)$ mit genau 2 freien Variablen x, y, so dass für jede S-Struktur $\mathcal{A} = (A; R^{\mathcal{A}})$ und für jede $\mathcal{A}$-Belegung w gilt

$$\left(R^{\mathcal{A}}\right)^+ := \left\{(w(x), w(y)) \mid \alpha^{\mathcal{A},w} = T\right\}\ ?$$

In diesem Fall ist die transitive Hülle elementar definierbar *durch α.*
Wir zeigen unter Verwendung des Kompaktheitssatzes, dass die transitive Hülle nicht elementar definierbar ist.
Hierzu definieren wir induktiv die folgenden Formeln:
$\beta_1 := R(x,y)$
und für $n > 0$:
$\beta_{n+1} := (\exists z_{n+1})\,(R(x, z_{n+1}) \wedge \beta_n\,\{x/z_{n+1}\})$.
Durch Induktion nach $i \in \mathbb{N} \setminus \{0\}$ folgt sofort, dass für jedes i β_i eine Formel zur Signatur S mit genau den freien Variablen x und y ist.
Ebenfalls durch Induktion nach $i \in \mathbb{N} \setminus \{0\}$ folgt, dass $\mathcal{A}, w \models_{PL} \beta_i$ genau dann, wenn ein $R^{\mathcal{A}}$-Pfad der Länge i von $w(x)$ nach $w(y)$ existiert.

Angenommen die transitive Hülle wäre elementar definierbar durch α. Wir betrachten die Formelmenge $\Sigma := \{\alpha\} \cup \{\neg\beta_n \mid n \geq 1\}$. Sei Σ' eine beliebig gewählte endliche Teilmenge von Σ. Es existiert dann ein größtes $n \in \mathbb{N} \setminus \{0\}$, so dass $\neg\beta_n \in \Sigma'$ oder aber $\Sigma' \subseteq \{\alpha\}$. Falls $\Sigma' \subseteq \{\alpha\}$, dann sei $n := 0$. Es gilt dann in jedem Fall $\neg\beta_{n+1} \notin \Sigma'$.
Sei die Struktur $\mathcal{A}$ der folgende unendliche, gerichtete Graph, den wir hier bildlich darstellen:

$$0 \to 1 \to 2 \to 3 \to \ldots$$

Sei w eine $\mathcal{A}$-Belegung mit $w(x) = 0$ und $w(y) = n+1$. Nach Annahme über α gilt $\mathcal{A}, w \models_{PL} \alpha$. Offensichtlich gilt aber auch für jedes $\neg\beta_i \in \Sigma'$ $\mathcal{A}, w \models_{PL} \neg\beta_i$, da kein Pfad der Länge i von $w(x) = 0$ nach $w(y) = n+1$ führt. Folglich ist Σ' erfüllbar. Da dies für jede endliche Teilmenge Σ' von Σ gilt, ist Σ aufgrund des Kompaktheitssatzes erfüllbar durch eine Interpretation $(\mathcal{A}, w)$.
Da $\alpha \in \Sigma$, gilt somit $\mathcal{A}, w \models_{PL} \alpha$. Es gibt daher einen $R^{\mathcal{A}}$-Pfad (einer Länge $n > 0$ für ein geeignetes $n \in \mathbb{N} \setminus \{0\}$) von $w(x)$ nach $w(y)$.
Da für jedes $i \in \mathbb{N} \setminus \{0\}$ $\neg\beta_i \in \Sigma$, gilt also $\mathcal{A}, w \models_{PL} \neg\beta_i$ für alle $i > 0$, d.h. für jedes $i > 0$ gibt es keinen Pfad der Länge i von $w(x)$ nach $w(y)$. Wir erhalten somit einen Widerspruch. Daher ist die Annahme, dass die transitive Hülle elementar definierbar ist, falsch.

Mit Hilfe des Kompaktheitssatzes erhalten wir wie im aussagenlogischen Fall die Endlichkeitseigenschaft der Folgerungsbeziehung.

Theorem 2.111 (Endlichkeitseigenschaft der Folgerungsbeziehung) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen, $\alpha \in Fm_{PL}(S)$ eine Aussage. Es gilt dann:*
$\Sigma \models_{PL} \alpha$ genau dann, wenn eine endliche Teilmenge $\Sigma' \subseteq \Sigma$ mit $\Sigma' \models_{PL} \alpha$ existiert.

Ein weiteres Resultat, das sich sofort aus dem Modellexistenztheorem ergibt, und das auf den ersten Blick einige paradox anmutende Konsequenzen mit sich bringt, ist das

Theorem 2.112 (Löwenheim und Skolem) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen. Ist Σ erfüllbar, so ist Σ erfüllbar in einer Struktur mit abzählbarem Grundbereich.*

Beweis. Sei $\mathcal{C} := \{\Sigma \subseteq Fm_{PL}(S(\mathcal{K}_0)) \mid \Sigma\ \mathcal{K}_0\text{-passende, erfüllbare Aussagenmenge}\}$.
$\mathcal{C}$ ist eine prädikatenlogische Konsistenzeigenschaft (leicht zu verifizieren). Ist nun Σ eine erfüllbare Aussagenmenge zur Signatur S, so gilt $\Sigma \in \mathcal{C}$. Nach dem Modellexistenztheorem hat Σ dann ein Herbrand-Modell zur Signatur $S(\mathcal{K}_0)$. Ein solches Herbrand-Modell hat aber einen abzählbar unendlichen Grundbereich. ■

Bemerkung 2.113 *Für diese Form des Theorems von Löwenheim und Skolem ist ganz wesentlich, dass das Gleichheitssymbol nicht als logisches Symbol behandelt wird, d.h. es kann in unterschiedlichen Strukturen verschieden interpretiert werden.*
Im Vorgriff sei schon darauf hingewiesen, dass im Falle der Prädikatenlogik mit Gleichheit der Satz von Löwenheim-Skolem nur in einer abgeschwächten Fassung gültig ist.

Bemerkung 2.114 *In diesem Abschnitt haben wir die generelle Voraussetzung getroffen, dass die zugrundeliegenden Signaturen höchstens abzählbar unendlich viele Relations-, Funktions- und Konstantensymbole enthalten. Im Beweis des Modellexistenztheorems haben wir ganz wesentlich von dieser Tatsache Gebrauch gemacht. Unter Verwendung von stärkeren Beweisprinzipien kann man sich von dieser Voraussetzung befreien und die angegebenen Sätze (strenge Vollständigkeit, Kompaktheitssatz, Endlichkeitseigenschaft sowie einen leicht modifizierten Satz von Löwenheim-Skolem) auch im allgemeinen Fall beweisen.*
Den Satz von Löwenheim-Skolem kann man dann im allgemeinen Fall in verschiedenen Varianten angeben:
Sei $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$ eine Signatur. $\Sigma \subseteq Fm_{PL}(S)$ sei eine erfüllbare Menge von Aussagen.

„Absteigende Variante“: *Σ ist in einer Struktur $\mathcal{A}$ erfüllbar, für deren Grundbereich $|\mathcal{A}|$ gilt:*

$|\mathcal{A}|$ hat höchstens die Mächtigkeit von $Fm_{PL}(S)$.

„Aufsteigende Variante": *Für jede Menge M gilt:*

> *Σ ist in einer Struktur $\mathcal{A}$ erfüllbar, deren Grundbereich $|\mathcal{A}|$ mindestens so viele Elemente wie M enthält. (Es gibt also Modelle beliebig großer Mächtigkeit)*

Beispiel 2.115 (*zum Satz von Löwenheim und Skolem*)
Im folgenden werden wir jeweils Mengen Σ von Aussagen einer Signatur S betrachten. Wir wollen in diesen Fällen eine S-Struktur $\mathcal{A}$ als Modell von Σ bezeichnen, wenn für eine (und nach dem Koinzidenztheorem damit für alle) $\mathcal{A}$-Belegungen w $(\mathcal{A}, w)$ ein Modell für Σ ist.

1. *Sei $\mathcal{R} := (\mathbb{R}; <, =; +, \cdot; 0, 1)$. S sei eine dazu passende Signatur.*

   $Th(\mathcal{R}) := \{\alpha \in Fm_{PL}(S) \mid \alpha \text{ Aussage}, \mathcal{R} \models_{PL} \alpha\}$ wird dann als die Theorie der reellen Zahlen bezeichnet.

   Nach dem Satz von Löwenheim und Skolem besitzt $Th(\mathcal{R})$ ein abzählbares Modell, d.h. es existiert eine S-Struktur $\mathcal{A}$ mit abzählbarem Grundbereich und $\mathcal{A} \models_{PL} \alpha$ für alle $\alpha \in Th(\mathcal{R})$.

2. *Die Mengenlehre wird üblicherweise axiomatisch eingeführt. Ein ganz bedeutende Axiomatisierung geht dabei auf Zermelo und Fraenkel zurück. Diese Axiomatisierung – hier kurz mit $\mathcal{ZF}$ bezeichnet – ist in der Sprache zur Signatur $(\in, =)$ formuliert, wobei beide Relationssymbole 2-stellig sind. Es wird gemeinhin angenommen, dass dieses Axiomensystem widerspruchsfrei und somit erfüllbar ist. Nach dem Satz von Löwenheim und Skolem besitzt $\mathcal{ZF}$ dann auch ein abzählbares Modell, so dass alle in diesem Modell betrachtbaren Mengen höchstens abzählbar sind. Aus dem Axiomensystem der Zermelo-Fraenkelschen Mengenlehre folgt aber semantisch, dass es eine überabzählbar unendliche Menge gibt (z.B. die Potenzmenge jeder abzählbar unendlichen Menge), deren Elemente im Grundbereich des Modells enthalten sein müssen, so dass im Grundbereich des Modells überabzählbar viele Elemente vorhanden sind. Wie kann $\mathcal{ZF}$ dann aber ein abzählbares Modell besitzen* (Skolem'sches Paradox)*?*

   Antwort:

   Abzählbarkeit innerhalb und außerhalb eines Modells bedeuten nicht dasselbe. Um die Abzählbarkeit zeigen zu können, benötigt man eine injektive Abbildung in die Menge der natürlichen Zahlen. Ein abzählbares Modell für $\mathcal{ZF}$ ist jedoch so schwach, d.h. so „ausgedünnt", dass gerade nur die Axiome erfüllt werden, während die für den Nachweis der Abzählbarkeit erforderlichen Mengen und Funktionen in diesem Modell nicht alle vorhanden sind. Für das Modell ist somit die betreffende Menge überabzählbar. Mit Mitteln außerhalb des Modells lässt sich aber die Abzählbarkeit zeigen.

3. *Sei $\mathcal{N} := (\mathbb{N}; <, =; +, \cdot; 0, 1)$ und S eine passende Signatur.*

   $Th(\mathcal{N}) := \{\alpha \in Fm_{PL}(S) \mid \alpha \text{ Aussage}, \mathcal{N} \models_{PL} \alpha\}$ sei die Theorie der natürlichen Zahlen. $Th(\mathcal{N})$ besitzt dann nach der aufsteigenden Variante des Satzes von Löwenheim und Skolem ein Modell von der Mächtigkeit der reellen Zahlen, das somit zu $\mathcal{N}$ nicht isomorph ist.

2.3 Die Unentscheidbarkeit der Prädikatenlogik

Für aussagenlogische Formeln hatten wir z.B. mit der Methode der Wahrheitswerttabellen ein einfaches Verfahren zur Hand, mit dem entschieden werden konnte, ob eine Formel allgemeingültig (bzw. erfüllbar) ist oder nicht. Wir werden im Folgenden zeigen, dass bereits bei sehr einfachen Anforderungen an die zugrundeliegende Signatur eine entsprechende Eigenschaft für die Prädikatenlogik nicht gilt.
S sei im Folgenden wiederum ein höchstens abzählbar unendliche Signatur.
Die folgende Problemstellung

- Gegeben ist eine Aussage $\alpha \in Fm_{PL}(S)$
- Frage: Ist α allgemeingültig?

bezeichnet man als das Entscheidungsproblem der Prädikatenlogik.
Formal ist damit das Entscheidungsproblem für die Prädikatenlogik definiert als ein Entscheidungsproblem für die Menge

$$\mathrm{VALID}_S := \{\alpha \in Fm_{PL}(S) \mid \alpha \text{ Aussage und } \models_{PL} \alpha\}$$

Unter Verwendung der Unentscheidbarkeit des Postschen Korrespondenzproblems werden wir die Unentscheidbarkeit von VALID_S, d.h. somit der Prädikatenlogik, zeigen.

Definition 2.116 *Es sei das Alphabet* $ALPH = \{0, 1, (,)\}$ *zugrundegelegt.*
Ein Postsches Korrespondenzproblem über $\{0, 1\}$ *ist das folgende Entscheidungsproblem:*

- *Gegeben ist ein endliches Tupel* $K = ((x_1, y_1), \dots, (x_n, y_n))$ *„korrespondierender" Wortpaare* (x_i, y_i), $1 \leq i \leq n$, *über dem Alphabet* $\{0, 1\}$. *(*K *wird als Korrespondenzsystemüber* $\{0, 1\}$ *bezeichnet.)*
- *Frage:*

  Gibt es eine Folge $i_1, i_2, \dots, i_k \in \{1, \dots, n\}$, $k \geq 1$, *mit* $x_{i_1} x_{i_2} \dots x_{i_k} = y_{i_1} y_{i_2} \dots y_{i_k}$*?*

Falls es eine solche Folge $i_1, i_2, \dots, i_k$ *gibt, so heißt sie* Lösung des Postschen Korrespondenzproblems *(für das Korrespondenzsystem)* K. K *heißt in diesem Fall auch lösbar.*
Das Postsche Korrespondenzproblem über $\{0, 1\}$ *ist somit das Entscheidungsproblem für die Menge*

$$PKP := \{K \mid K \textit{ ist lösbares Postsches Korrespondenzsystem über } \{0,1\}\}.$$

Es gilt (siehe z.B. die Vorlesung GTI):

Theorem 2.117 *Das Postsche Korrespondenzproblem* PKP *ist unentscheidbar.*

Das folgende Theorem von Church besagt, dass es kein algorithmisches Verfahren geben kann, das, angesetzt auf eine beliebige Formel der Prädikatenlogik, entscheidet, ob diese Formel allgemeingültig ist. Bevor wir zu seiner Formulierung und zu seinem Beweis übergehen, sollen noch einmal einige grundlegende Begriffe aus der Berechenbarkeitstheorie, die aus der Vorlesung GTI bekannt sein sollten und die wir im Folgenden benötigen, in Erinnerung gerufen werden.
Seien A, A' Alphabete. Weiter ist mit Turingmaschine im Folgenden immer eine deterministische Turingmaschine mit Eingabe-Alphabet A und Ausgabe-Alphabet A' gemeint. (In GTI wurde nur ein Alphabet zugrundegelegt. Die Verallgemeinerung auf eine Turingmaschine mit eventuell unterschiedlichem Ein- und Ausgabealphabet ergibt sich in kanonischer Weise. Für unsere Darstellung ist es vorteilhaft, ein solches Turing-Maschinenmodell zu benutzen.)

Definition 2.118

1. *Für eine Turingmaschine* M *ist* $L(M) := \{x \in A^* \mid M \textit{ akzeptiert } x\}$ *die von* M akzeptierte Sprache*.*
2. $L \subseteq A^*$ *heißt* rekursiv aufzählbar *genau dann, wenn eine Turingmaschine* M *mit* $L = L(M)$ *existiert.*
3. *Eine Sprache* $L \subseteq A^*$ *heißt* rekursiv *(oder* entscheidbar*) genau dann, wenn eine Turingmaschine* M *existiert, die für jede Eingabe hält und für* $x \in L$ *das leere Wort* ε *und für* $x \notin L$ *ein Wort* $\neq \varepsilon$ *ausgibt.*

   Es gilt:

   $L \subseteq A^*$ *ist rekursiv genau dann, wenn* L *und* $A^* \setminus L$ *rekursiv aufzählbar sind.*
4. *Die von einer Turingmaschine* M berechnete Funktion $f_M : A^* \dashrightarrow (A')^*$ *besteht aus der Menge aller Paare* $(x, y) \in A^* \times (A')^*$, *so dass die Turingmaschine* M *bei Eingabe* x *mit der Ausgabe* y *auf dem Ausgabeband hält.*

   ($\dashrightarrow$ *bedeutet, dass die Funktion auch partiell sein kann, d.h. eventuell nicht überall definiert ist. Die Schreibweise* $f : A^* \to (A')^*$ *soll ausdrücken, dass* f *total, d.h. für alle* $x \in A^*$ *definiert ist.)*

5. *$f : A^* \dashrightarrow (A')^*$ heißt* berechenbar (*oder* rekursiv) *genau dann, wenn eine Turingmaschine mit Eingabe-Alphabet A, Ausgabe-Alphabet A' und $f = f_M$ existiert.*

6. *Sei $M \subseteq A^*$ und $N \subseteq (A')^*$.*

   $M \leq N$ (in Worten: M ist (many-one) reduzierbar *auf N) genau dann, wenn eine berechenbare Funktion $f : A^* \to (A')^*$ existiert, so dass für alle $x \in A^*$ gilt:*

   $$x \in M \;\Leftrightarrow\; f(x) \in N.$$

   Es gilt dann der:

   Reduktionssatz:

   Sei $M \subseteq A^*$ und $N \subseteq (A')^*$ und gelte $M \leq N$.

   (a) M nicht entscheidbar $\Rightarrow$ N nicht entscheidbar.

   (b) N entscheidbar $\Rightarrow$ M entscheidbar.

Mit diesen Vorbereitungen können wir nun das Theorem von Church beweisen.

Theorem 2.119 (*Church*)
Sei $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$ eine Signatur, welche mindestens ein Konstantensymbol c, mindestens zwei einstellige Funktionssymbole f_0 und f_1 und mindestens ein zweistelliges Relationssymbol P enthält.
Dann gilt:
Die Menge $VALID_S$ der allgemeingültigen Aussagen zur Signatur S ist unentscheidbar.

Beweis. Formeln in $Fm_{PL}(S)$ sind Zeichenreihen über einem endlichen Alphabet A'. Damit wird es möglich, eine (many-one) Reduktion des Postschen Korrespondenzproblems PKP auf VALID_S anzugeben, d.h. wir zeigen $PKP \leq \mathrm{VALID}_S$.
Wir definieren dazu folgendermaßen eine rekursive Abbildung $F : ALPH^* \to (A')^*$ mit der Eigenschaft

$$K \in PKP \;\Leftrightarrow\; F(K) \in \mathrm{VALID}_S.$$

1. K ist kein Postsches Korrespondenzsystem: Wir setzen $F(K) := \perp$.

2. $K = ((x_1, y_1), \ldots, (x_n, y_n))$ ist ein Postsches Korrespondenzsystem:
   Wir führen zunächst eine Bezeichnung ein.
   Es sei $g_{d_1 \ldots d_k}(x) := f_{d_1}(f_{d_2}(\ldots(f_{d_k}(x))\ldots))$ mit $d_i \in \{0,1\}$, $1 \leq i \leq k$.
   (Z.B. $g_{0010}(x) = f_0(f_0(f_1(f_0(x))))$.)
   Zur Motivation der nachfolgenden Definition von $F(K)$ geben wir die intendierte Interpretation der Symbole c, f_0, f_1 und P in einer S-Struktur $\mathcal{A}$ mit Grundbereich $\{0,1\}^*$ und $\mathcal{A}$-Belegung w an:
   $c^{\mathcal{A}}$: das leere Wort in $|\mathcal{A}| = \{0,1\}^*$
   $f_0^{\mathcal{A}}(a) = 0a$ (d.h. $[f_0(x)]^{\mathcal{A},w} = 0w(x)$; „Linksmultiplikation" mit 0)
   $f_1^{\mathcal{A}}(a) = 1a$ (d.h. $[f_1(x)]^{\mathcal{A},w} = 1w(x)$; „Linksmultiplikation" mit 1)
   $(a,b) \in P^{\mathcal{A}} \;\Leftrightarrow\; a = x_{i_1} \ldots x_{i_k},\ b = y_{i_1} \ldots y_{i_k}$ für ein $k \geq 1$ und $i_1, \ldots, i_k \in \{1, \ldots, n\}$
   Es sei dann $F(K) := \alpha_1 \wedge \alpha_2 \to \beta$ mit
   $\alpha_1 := P(g_{x_1}(c), g_{y_1}(c)) \wedge \ldots \wedge P(g_{x_n}(c), g_{y_n}(c))$ „Anfangsbedingung"
   $\alpha_2 := (\forall x)(\forall y)(P(x,y) \to P(g_{x_1}(x), g_{y_1}(y)) \wedge \ldots \wedge P(g_{x_n}(x), g_{y_n}(y)))$ „Induktionsbedingung"
   $\beta := (\exists z)\, P(z,z)$ „Existenz einer Lösung"

Die so festgelegte Funktion F ist offensichtlich rekursiv. Es muss noch die Äquivalenz

$$K \in PKP \;\Leftrightarrow\; F(K) \in \mathrm{VALID}_S$$

gezeigt werden.

„ $\Leftarrow$ " : Sei $F(K) \in \mathrm{VALID}_S$.

Sei $\mathcal{A}$ eine S-Struktur mit

$|\mathcal{A}| = \{0,1\}^*$

$c^{\mathcal{A}} = \varepsilon$ (das leere Wort in $\{0,1\}^*$)

$f_0^{\mathcal{A}}(a) = 0a$

$f_1^{\mathcal{A}}(a) = 1a$

$P^{\mathcal{A}} = \{(a,b) \mid \text{ex. } m > 0,\ i_1, \ldots, i_m \in \{1, \ldots, n\} \text{ mit } a = x_{i_1} \ldots x_{i_n}, b = y_{i_1} \ldots y_{i_n}\}$.

Es gilt offensichtlich

$$\mathcal{A} \models_{PL} \alpha_1$$
$$\mathcal{A} \models_{PL} \alpha_{2.}$$

Da nach Annahme $F(K)$ allgemeingültig ist, folgt somit $\mathcal{A} \models_{PL} \beta$, d.h. es gibt ein $a \in |\mathcal{A}|$ mit $(a,a) \in P^{\mathcal{A}}$. Aufgrund der Wahl von $P^{\mathcal{A}}$ bedeutet dies aber, dass eine Lösung für das Postsche Korrespondenzsystem K existiert, also $K \in PKP$.

„ $\Rightarrow$ " : Sei $K \in PKP$.

Wir zeigen, dass $F(K)$ allgemeingültig ist.

Sei $\mathcal{A}$ eine beliebige S-Struktur. Da α_1, α_2 und β Aussagen sind, genügt es, zu zeigen:

$$(\mathcal{A} \models_{PL} \alpha_1 \text{ und } \mathcal{A} \models_{PL} \alpha_2) \ \Rightarrow\ \mathcal{A} \models_{PL} \beta.$$

Wir definieren induktiv eine Abbildung $h : \{0,1\}^* \to |\mathcal{A}|$ durch

$$\begin{aligned} h(\varepsilon) &:= c^{\mathcal{A}} \\ h(0w) &:= f_0^{\mathcal{A}}(h(w)) \\ h(1w) &:= f_1^{\mathcal{A}}(h(w)). \end{aligned}$$

Es gilt dann für alle $w \in \{0,1\}^*$

$$h(w) = [g_w(c)]^{\mathcal{A}}. \tag{1}$$

(Dies folgt leicht mittels Induktion über die Länge von w.)

Wegen $\mathcal{A} \models_{PL} \alpha_1$ gilt

$$\left([g_{x_i}(c)]^{\mathcal{A}}, [g_{y_i}(c)]^{\mathcal{A}}\right) \in P^{\mathcal{A}}, \qquad 1 \leq i \leq n.$$

Wegen $\mathcal{A} \models_{PL} \alpha_2$ gilt für beliebige Grundterme s und t

$$\left(s^{\mathcal{A}}, t^{\mathcal{A}}\right) \in P^{\mathcal{A}} \ \Rightarrow\ \left([g_{x_i}(s)]^{\mathcal{A}}, [g_{y_i}(t)]^{\mathcal{A}}\right) \in P^{\mathcal{A}}, \qquad 1 \leq i \leq n.$$

Durch Induktion nach $m \in \mathbb{N}$ folgt:

für alle $m > 0$ und für alle $i_1, \ldots, i_m \in \{1, \ldots, n\}$ gilt

$$\left(\left[g_{x_{i_m}}\left(g_{x_{i_{m-1}}}\left(\ldots\left(g_{x_{i_1}}(c)\right)\ldots\right)\right)\right]^{\mathcal{A}}, \left[g_{y_{i_m}}\left(g_{y_{i_{m-1}}}\left(\ldots\left(g_{y_{i_1}}(c)\right)\ldots\right)\right)\right]^{\mathcal{A}}\right) \in P^{\mathcal{A}}. \tag{2}$$

Da $K \in PKP$, existieren $m \geq 1$ und $i_m, \ldots, i_1 \in \{1, \ldots, n\}$ mit $x_{i_m} \ldots x_{i_1} = y_{i_m} \ldots y_{i_1}$. Hieraus folgt dann $h(x_{i_m} \ldots x_{i_1}) = h(y_{i_m} \ldots y_{i_1})$. Aufgrund von (1) folgt

$$\left[g_{x_{i_m} \ldots x_{i_1}}(c)\right]^{\mathcal{A}} = \left[g_{y_{i_m} \ldots y_{i_1}}(c)\right]^{\mathcal{A}}$$

und damit

$$\left[g_{x_{i_m}}\left(\ldots\left(g_{x_{i_1}}(c)\right)\ldots\right)\right]^{\mathcal{A}} = \left[g_{y_{i_m}}\left(\ldots\left(g_{y_{i_1}}(c)\right)\ldots\right)\right]^{\mathcal{A}}.$$

Wegen (2) gilt somit $\mathcal{A} \models_{PL} \beta$. Dies war zu zeigen.

■

Bemerkung 2.120
Für die Menge $VALID_S$ der allgemeingültigen Aussagen lässt sich zeigen, dass sie rekursiv aufzählbar ist.

Beweis. (Idee)
Ein Weg, diese Behauptung zu zeigen, besteht in der Angabe einer Turingmaschine, die genau für alle Aussagen in VALID_S hält. Eine solche Turingmaschine könnte in der folgenden Weise arbeiten:
Bei Eingabe einer Aussage α werden auf einem Arbeitsband systematisch und in einer normierten Weise Tableaus für $\{\neg\alpha\}$ erzeugt. Existiert für $\{\neg\alpha\}$ ein geschlossenes Tableau, so wird in diesem systematischen Verfahren insbesondere ebenfalls ein (im allgemeinen aber anderes) geschlossenes Tableau erzeugt. Zunächst wird das Tableau mit einem Knoten generiert, dann alle Tableaus mit 2 Knoten, in denen Grundterme mit einer Zeichenlänge ≤ 2 neu vorkommen, dann alle Tableaus mit 3 Knoten, in denen Grundterme mit einer Zeichenlänge ≤ 3 neu vorkommen usw., bis möglicherweise ein geschlossenes Tableau erhalten wird.
Falls $\alpha \in \mathrm{VALID}_S$, wird in diesem Verfahren ein geschlossenes Tableau für $\{\neg\alpha\}$ gefunden und das Verfahren stoppt. Falls $\alpha \notin \mathrm{VALID}_S$, läuft dieses Verfahren endlos. ■

In der hier angegebenen Fassung des Churchschen Theorems war vorausgesetzt, dass die Signatur neben einem zweistelligen Prädikatensymbol wenigstens noch zwei einstellige Funktionssymbole sowie ein Konstantensymbol enthält. Diese Voraussetzung kann noch abgeschwächt werden. Die Unentscheidbarkeit kann bereits dann gezeigt werden, wenn in der Signatur nur ein zweistelliges Prädikatensymbol vorhanden ist. Sind hingegen nur einstellige Prädikatensymbole und höchstens einstellige Funktionssymbole vorhanden, ist die Prädikatenlogik dieser Signatur entscheidbar. (Einen ganz ausgezeichneten und praktisch vollständigen Überblick über entscheidbare und unentscheidbare Formelklassen der Prädikatenlogik findet man in dem Buch von Börger, Grädel und Gurevich: The classical decision problem, Springer Verlag 1997.)

2.4 Weitere Eigenschaften der Prädikatenlogik

In diesem Abschnitt werden wir neben dem Ersetzungstheorem die Skolemisierung und den Satz von Herbrand behandeln, die abgesehen von ihrer eigenständigen Bedeutung eine wichtige Rolle im Rahmen der logischen Programmierung und des automatischen Beweisens spielen.

2.4.1 Das prädikatenlogische Ersetzungstheorem

Wir hatten bereits im Falle der Aussagenlogik ein Ersetzungstheorem bewiesen. Wir wollen hier noch eine allgemeinere Fassung für die Prädikatenlogik angeben.

Notation 2.121 *Seien $\alpha, \alpha_1, \alpha_2 \in Fm_{PL}(S)$ und α_1 eine atomare Formel.*
$\alpha\left(\dfrac{\alpha_1}{\alpha_2}\right)$ *sei diejenige Formel, die aus α dadurch resultiert, dass jedes Vorkommen von α_1 in α durch α_2 ersetzt wird.*

Die Menge der Teilformeln von α sei mit $TF(\alpha)$ bezeichnet. Wir definieren als Nächstes das positive und negative Vorkommen von Teilformeln in einer Formel α.

Definition 2.122 (Positives und negatives Vorkommen einer Formel) *Wir definieren induktiv zu einer Formel α simultan Formelmengen $TF^+(\alpha) \subseteq TF(\alpha)$ und $TF^-(\alpha) \subseteq TF(\alpha)$, die diejenigen Teilformeln, die positiv bzw. negativ vorkommen, enthalten.*

1. *Falls α eine atomare Formel:*
   $TF^+(\alpha) := \{\alpha\} \qquad TF^-(\alpha) := \emptyset$
2. $TF^+(\neg\alpha) := TF^-(\alpha) \cup \{\neg\alpha\} \qquad TF^-(\neg\alpha) := TF^+(\alpha)$

3. $TF^+(\alpha \circ \beta) := TF^+(\alpha) \cup TF^+(\beta) \cup \{\alpha \circ \beta\}$ *und*
   $TF^-(\alpha \circ \beta) := TF^-(\alpha) \cup TF^-(\beta)$ *für* $\circ \in \{\wedge, \vee\}$
4. $TF^+(\alpha \to \beta) := TF^-(\alpha) \cup TF^+(\beta) \cup \{\alpha \to \beta\}$ *und*
   $TF^-(\alpha \to \beta) := TF^+(\alpha) \cup TF^-(\beta)$
5. $TF^+((Qx)\,\alpha) := TF^+(\alpha) \cup \{(Qx)\,\alpha\}$ *und*
   $TF^-((Qx)\,\alpha) := TF^-(\alpha)$ *für* $Q \in \{\forall, \exists\}$

Eine Teilformel β *von* α *kommt* positiv in α (negativ in α) *vor genau dann, wenn* $\beta \in TF^+(\alpha)$ $(\beta \in TF^-(\alpha))$.
Ist jedes Vorkommen von β *in* α *positiv (negativ), so sagen wir, dass* β nur positiv (nur negativ) *in* α *vorkommt.*

Eine unmittelbare Folgerung aus dieser Definition ist das nachfolgende Korollar, mit dessen Hilfe der Beweis des Ersetzungstheorems etwas einfacher wird.

Korollar 2.123 *Sei* β *eine atomare Formel.*

1. *Sei* α *eine atomare Formel.*
   β *kommt entweder nur positiv oder überhaupt nicht in* α *vor,*
   und β *kommt entweder nur negativ oder überhaupt nicht in* $\neg\alpha$ *vor.*
2. β *kommt in* α *genau dann nur positiv (negativ) vor, wenn* β *auch in* $\neg\neg\alpha$ *nur positiv (negativ) vorkommt.*
3. β *kommt in den Komponenten* $\boldsymbol{\alpha}_1$ *und* $\boldsymbol{\alpha}_2$ *einer* α*-Formel genau dann nur positiv (negativ) vor, wenn* β *auch in der Formel* $\boldsymbol{\alpha}$ *selbst nur positiv (negativ) vorkommt.*
4. β *kommt in den Komponenten* $\boldsymbol{\beta}_1$ *und* $\boldsymbol{\beta}_2$ *einer* β*-Formel genau dann nur positiv (negativ) vor, wenn* β *auch in der Formel* $\boldsymbol{\beta}$ *selbst nur positiv (negativ) vorkommt.*
5. *Ist eine* γ*-Formel der Gestalt* $(\forall x)\,\gamma'$ *oder* $\neg(\exists x)\,\gamma'$ *gegeben, dann kommt* β *genau dann nur positiv (negativ) in der* $\boldsymbol{\gamma}$*-Instanz* $\boldsymbol{\gamma}(x)$ *vor, wenn* β *auch in* $\boldsymbol{\gamma}$ *selbst nur positiv (negativ) vorkommt.*
6. *Ist eine* δ*-Formel der Gestalt* $(\exists x)\,\delta'$ *oder* $\neg(\forall x)\,\delta'$ *gegeben, dann kommt* β *genau dann nur positiv (negativ) in der* δ*-Instanz* $\boldsymbol{\delta}(x)$ *vor, wenn* β *auch in* $\boldsymbol{\delta}$ *selbst nur positiv (negativ) vorkommt.*

Beweis. Einfache Übungsaufgabe. ■

Theorem 2.124 (Ersetzungstheorem) *Seien* $\alpha, \alpha_1, \alpha_2, \beta \in Fm_{PL}(S)$ *und* β *eine atomare Formel.* $\mathcal{A}$ *sei eine* S*-Struktur.*

1. *Gilt* $\mathcal{A} \models_{PL} \alpha_1 \to \alpha_2$ *und kommt* β *nur positiv (bzw. nur negativ) in* α *vor, dann gilt*
   $\mathcal{A} \models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \to \alpha\left(\frac{\beta}{\alpha_2}\right)$ (*bzw.* $\mathcal{A} \models_{PL} \alpha\left(\frac{\beta}{\alpha_2}\right) \to \alpha\left(\frac{\beta}{\alpha_1}\right)$).
2. $\mathcal{A} \models_{PL} \alpha_1 \leftrightarrow \alpha_2 \;\Rightarrow\; \mathcal{A} \models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \leftrightarrow \alpha\left(\frac{\beta}{\alpha_2}\right)$

Beweis.

1. Wir beweisen die Aussage für α mittels struktureller Induktion über α-, ... , δ-Formeln, wobei wir nur den Fall des positiven Vorkommens behandeln.

   Ind.Basis: α atomare Formel oder $\alpha = \neg\alpha'$ und α' atomare Formel:
   β kann in diesem Fall in α gar nicht oder aber nur positiv vorkommen.

(a) β komme in α vor, aber nur positiv.
Dies ist nur dann möglich, wenn α atomare Formel und $\alpha = \beta$. In diesem Fall ist $\alpha\left(\frac{\beta}{\alpha_1}\right) = \alpha_1$ und $\alpha\left(\frac{\beta}{\alpha_2}\right) = \alpha_2$. Wegen $\mathcal{A} \models_{PL} \alpha_1 \to \alpha_2$ gilt also $\mathcal{A} \models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \to \alpha\left(\frac{\beta}{\alpha_2}\right)$.

(b) β komme nicht in α vor.
In diesem Fall ist $\alpha\left(\frac{\beta}{\alpha_i}\right) = \alpha,\ i = 1,2$. Wegen $\mathcal{A} \models_{PL} \alpha \to \alpha$ gilt $\mathcal{A} \models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \to \alpha\left(\frac{\beta}{\alpha_2}\right)$.

Ind.Schritt: Wir betrachten exemplarisch 2 Fälle.

(a) Es liege eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$ vor.
Da nach Voraussetzung β in $\boldsymbol{\alpha}$ nur positiv vorkommt, ist somit nach Korollar 2.123 auch jedes Vorkommen von β in $\boldsymbol{\alpha}_1$ und in $\boldsymbol{\alpha}_2$ nur positiv. Nach Induktionsvoraussetzung gilt daher $\mathcal{A} \models_{PL} \boldsymbol{\alpha}_i\left(\frac{\beta}{\alpha_1}\right) \to \boldsymbol{\alpha}_i\left(\frac{\beta}{\alpha_2}\right),\ i = 1,2$.
Es gilt $\models_{PL} \boldsymbol{\alpha}\left(\frac{\beta}{\alpha_i}\right) \leftrightarrow \boldsymbol{\alpha}_1\left(\frac{\beta}{\alpha_i}\right) \wedge \boldsymbol{\alpha}_2\left(\frac{\beta}{\alpha_i}\right),\ i = 1,2$.
Damit folgt $\mathcal{A} \models_{PL} \boldsymbol{\alpha}\left(\frac{\beta}{\alpha_1}\right) \to \boldsymbol{\alpha}_i\left(\frac{\beta}{\alpha_2}\right),\ i = 1,2$.
Hieraus folgt weiter $\mathcal{A} \models_{PL} \boldsymbol{\alpha}\left(\frac{\beta}{\alpha_1}\right) \to \boldsymbol{\alpha}_1\left(\frac{\beta}{\alpha_2}\right) \wedge \boldsymbol{\alpha}_2\left(\frac{\beta}{\alpha_2}\right)$, und damit $\mathcal{A} \models_{PL} \boldsymbol{\alpha}\left(\frac{\beta}{\alpha_1}\right) \to \boldsymbol{\alpha}\left(\frac{\beta}{\alpha_2}\right)$.

(b) Es liege eine γ-Formel der Form $(\forall x)\,\gamma'$ oder $\neg\,(\exists x)\,\gamma'$ vor mit γ-Instanz $\boldsymbol{\gamma}(x)$.
Nach Induktionsvoraussetzung gilt für eine beliebige $\mathcal{A}$-Belegung w'

$$\mathcal{A}, w' \models_{PL} [\boldsymbol{\gamma}(x)]\left(\frac{\beta}{\alpha_1}\right) \to [\boldsymbol{\gamma}(x)]\left(\frac{\beta}{\alpha_2}\right).$$

Es gilt nun für eine beliebige $\mathcal{A}$-Belegung w

$$\begin{aligned}
\left[\boldsymbol{\gamma}\left(\frac{\beta}{\alpha_1}\right)\right]^{\mathcal{A},w} = T \quad &\Leftrightarrow \quad \left[[\boldsymbol{\gamma}(x)]\left(\frac{\beta}{\alpha_1}\right)\right]^{\mathcal{A},w\{x/a\}} = T \qquad \text{für alle } a \in |\mathcal{A}| \\
&\Rightarrow \quad \left[[\boldsymbol{\gamma}(x)]\left(\frac{\beta}{\alpha_2}\right)\right]^{\mathcal{A},w\{x/a\}} = T \qquad \text{für alle } a \in |\mathcal{A}|, \text{ nach (IV)} \\
&\Leftrightarrow \quad \left[\boldsymbol{\gamma}\left(\frac{\beta}{\alpha_2}\right)\right]^{\mathcal{A},w} = T.
\end{aligned}$$

Damit folgt $\mathcal{A} \models_{PL} \boldsymbol{\gamma}\left(\frac{\beta}{\alpha_1}\right) \to \boldsymbol{\gamma}\left(\frac{\beta}{\alpha_2}\right)$.

2. Die Behauptung folgt mittels struktureller Induktion analog zum Teil 1).

■

Korollar 2.125 *Seien $\alpha, \alpha_1, \alpha_2, \beta \in Fm_{PL}(S)$ und β eine atomare Formel. Dann gilt:*

1. $\models_{PL} \alpha_1 \leftrightarrow \alpha_2 \ \Rightarrow \models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \leftrightarrow \alpha\left(\frac{\beta}{\alpha_2}\right)$.
2. *Gilt $\models_{PL} \alpha_1 \to \alpha_2$ und kommt β nur positiv (bzw. nur negativ) in α vor, dann gilt*
$\models_{PL} \alpha\left(\frac{\beta}{\alpha_1}\right) \to \alpha\left(\frac{\beta}{\alpha_2}\right)$ *(bzw.* $\models_{PL} \alpha\left(\frac{\beta}{\alpha_2}\right) \to \alpha\left(\frac{\beta}{\alpha_1}\right)$ *).*

2.4.2 Die Skolem-Form einer Formel

Wir wollen als Nächstes zeigen, dass für Fragen der Erfüllbarkeit in Formeln implizit vorhandene Existenzquantoren (das sind $\forall$- oder $\exists$-Quantoren, die sich in ihrer semantischen Funktion wie $\exists$-Quantoren verhalten) eliminiert werden können. Dies hat u.a. zur Folge, dass in Tableau- bzw. Resolutionsbeweisen keine $\boldsymbol{\delta}$-Regeln mehr erforderlich sind, wenn in den vorkommenden Formeln zuvor alle solche implizit vorhandenen Existenzquantoren eliminiert worden sind.
Für das Folgende sei angenommen, dass die Signatur S für jedes $n \in \mathbb{N}$ unendlich viele Funktionssymbole der Stelligkeit n enthält. Es sei daran erinnert, dass Konstantensymbole als nullstellige Funktionssymbole aufgefasst werden können. Wir machen hiervon im weiteren Gebrauch.

Lemma 2.126
Sei $\alpha \in Fm_{PL}(S)$ mit $fvar(\alpha) \subseteq \{x, y_1, \ldots, y_n\}$. f sei ein Funktionssymbol aus S, das in α nicht vorkommt. $\mathcal{A}$ sei eine S-Struktur.
Es gibt dann S-Strukturen $\mathcal{A}_1$ und $\mathcal{A}_2$, die sich von $\mathcal{A}$ höchstens in der Interpretation von f unterscheiden mit

1. $\mathcal{A}_1 \models_{PL} (\exists x)\,\alpha \to \alpha\,\{x/f(y_1, \ldots, y_n)\}$
2. $\mathcal{A}_2 \models_{PL} \alpha\,\{x/f(y_1, \ldots, y_n)\} \to (\forall x)\,\alpha$.

Beweis.

1. Wir müssen eine Struktur $\mathcal{A}_1$ angeben, die sich von $\mathcal{A}$ höchstens in der Interpretation von f unterscheidet und für die $\mathcal{A}_1 \models_{PL} (\exists x)\,\alpha \to \alpha\,\{x/f(y_1, \ldots, y_n)\}$ gilt. Hierzu ist somit nur noch $f^{\mathcal{A}_1}$ festzulegen und die Allgemeingültigkeit der angegebenen Implikation in $\mathcal{A}_1$ zu zeigen.

   Seien $a_1, \ldots, a_n$ beliebige Elemente aus $|\mathcal{A}|$ $(= |\mathcal{A}_1|)$. Es ist $f^{\mathcal{A}_1}(a_1, \ldots, a_n)$ zu definieren. Sei w eine beliebige

   $\mathcal{A}$-Belegung (also auch $\mathcal{A}_1$-Belegung) mit $w(y_i) = a_i$, $1 \le i \le n$.

   (a) Falls $[(\exists x)\,\alpha]^{\mathcal{A},w} = F$, wähle zu den Elementen a_1, …, a_n ein Element $a \in |\mathcal{A}| = |\mathcal{A}_1|$ aus und setze
   $$f^{\mathcal{A}_1}(a_1, \ldots, a_n) := a.$$
   Da f in α nicht vorkommt, gilt in diesem Fall $[(\exists x)\,\alpha]^{\mathcal{A}_1,w} = F$ und damit $\mathcal{A}_1, w \models_{PL} (\exists x)\,\alpha \to \alpha\,\{x/f(y_1, \ldots, y_n)\}$.

   (b) Falls $[(\exists x)\,\alpha]^{\mathcal{A},w} = T$, so existiert ein $a \in |\mathcal{A}| = |\mathcal{A}_1|$ mit $\alpha^{\mathcal{A},w\{x/a\}} = T$. Wir definieren
   $$f^{\mathcal{A}_1}(a_1, \ldots, a_n) := a.$$
   Es gilt dann
   $$\begin{aligned} T &= \alpha^{\mathcal{A},w\{x/a\}} \\ &= \alpha^{\mathcal{A}_1,w\{x/a\}} && \text{da } f \text{ in } \alpha \text{ nicht vorkommt} \\ &= \alpha^{\mathcal{A}_1,w\left\{x/[f(y_1,\ldots,y_n)]^{\mathcal{A}_1,w}\right\}} \\ &= [\alpha\,\{x/f(y_1, \ldots, y_n)\}]^{\mathcal{A}_1,w} && \text{nach Substitutionslemma.} \end{aligned}$$
   Damit gilt dann $\mathcal{A}_1, w \models_{PL} (\exists x)\,\alpha \to \alpha\,\{x/f(y_1, \ldots, y_n)\}$.
   Da w beliebig gewählt war, erhalten wir insgesamt $\mathcal{A}_1 \models_{PL} (\exists x)\,\alpha \to \alpha\,\{x/f(y_1, \ldots, y_n)\}$.

2. Folgt leicht aus 1).

■

Notation 2.127 *Seien $\alpha, \alpha_1, \beta \in Fm_{PL}(S)$ und α_1 eine atomare Formel.*
Statt $\alpha\begin{pmatrix}\alpha_1\\ \beta\end{pmatrix}$ schreiben wir im folgenden auch einfach $\alpha(\beta)$, und statt α schreiben wir dann auch $\alpha(\alpha_1)$, wenn wir Ersetzungen für α_1 betrachten wollen.

Theorem 2.128 (Skolem) *Seien $\alpha(\alpha_1), \beta \in Fm_{PL}(S)$, α_1 eine atomare Formel,*
$fvar(\beta) \subseteq \{x, y_1, \ldots, y_n\}$ und $\alpha((\exists x)\,\beta)$, $\alpha((\forall x)\,\beta)$ sowie $\alpha(\beta\,\{x/f(y_1, \ldots, y_n)\})$ Aussagen. f sei ein Funktionssymbol, das in $\alpha((\exists x)\,\beta)$ bzw. in $\alpha((\forall x)\,\beta)$ nicht vorkommt.

1. *Kommt* α_1 *nur positiv in* α *vor, dann gilt*:

$$\alpha\left((\exists x)\,\beta\right) \text{ ist erfüllbar} \Leftrightarrow \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right) \text{ ist erfüllbar}$$

und

$$\alpha\left((\forall x)\,\beta\right) \text{ ist allgmeingültig} \Leftrightarrow \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right) \text{ ist allgemeingültig.}$$

2. *Kommt* α_1 *nur negativ in* α *vor, dann gilt*:

$$\alpha\left((\forall x)\,\beta\right) \text{ ist erfüllbar} \Leftrightarrow \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right) \text{ ist erfüllbar}$$

und

$$\alpha\left((\exists x)\,\beta\right) \text{ ist allgmeingültig} \Leftrightarrow \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right) \text{ ist allgemeingültig.}$$

Beweis.
Wir zeigen die Behauptung im Falle der Erfüllbarkeit, wobei wir uns auf (1) beschränken. Die Behauptung für den Fall der Allgemeingültigkeit folgt dann sofort.
α_1 komme nur positiv in α vor.

„ $\Leftarrow$ " : Da $\alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right)$ nach Voraussetzung erfüllbar und eine Aussage ist, gibt es eine S-Struktur $\mathcal{A}$ mit

$$\mathcal{A} \models_{PL} \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right). \tag{2.1}$$

Weiter gilt $\models_{PL} \beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\} \rightarrow (\exists x)\,\beta$.

Nach Korollar 2.125 (2) gilt somit

$$\models_{PL} \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right) \rightarrow \alpha\left((\exists x)\,\beta\right).$$

Wegen (2.1) folgt $\mathcal{A} \models_{PL} \alpha\left((\exists x)\,\beta\right)$, d.h. $\alpha\left((\exists x)\,\beta\right)$ ist erfüllbar.

„ $\Rightarrow$ " : Es gelte $\mathcal{A} \models_{PL} \alpha\left((\exists x)\,\beta\right)$ (Man beachte: $\alpha\left((\exists x)\,\beta\right)$ ist Aussage!)

Nach Lemma 2.126 existiert eine Struktur $\mathcal{A}_1$, die sich von $\mathcal{A}$ höchstens in der Interpretation von f unterscheidet, mit $\mathcal{A}_1 \models_{PL} (\exists x)\,\beta \rightarrow \beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}$.

Nach Korollar 2.125 (2) gilt daher $\mathcal{A}_1 \models_{PL} \alpha\left((\exists x)\,\beta\right) \rightarrow \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right)$.

Da f in $\alpha\left((\exists x)\,\beta\right)$ nicht vorkommt, gilt auch $\mathcal{A}_1 \models_{PL} \alpha\left((\exists x)\,\beta\right)$, damit dann also

$$\mathcal{A}_1 \models_{PL} \alpha\left(\beta\left\{x\diagup f\left(y_1,\ldots,y_n\right)\right\}\right).$$

Dies war zu zeigen.

■

Nach dem Theorem von Skolem können wir für Fragen der Erfüllbarkeit (und Allgemeingültigkeit) Quantoren unter Einführung eines neuen Funktionssymbols, wie im obigen Theorem angegeben, eliminieren. Wir bezeichnen einen solchen Schritt als (dualen) Skolemisierungsschritt.

Definition 2.129

1. *Ist* $(\exists x)\,\beta$ *(bzw.* $(\forall x)\,\beta$*) eine Teilformel von* α*, die in* α *positiv (bzw. negativ) vorkommt, so wird für dieses Vorkommen der Quantor* $(\exists x)$ *(bzw.* $(\forall x)$*) als* existentialer Quantor *in* α *bezeichnet. Ist der Quantor nicht existential, so heißt er* universal *in* α.

2. *Ist* β *eine Formel, die keine existentialen Quantoren mehr enthält, und ist* β *aus* α *unter Durchführung von ein oder mehreren Skolemisierungsschritten erhalten worden, so heißt* β Skolemisierte (Skolem-Form *oder auch* Erfüllbarkeitsfunktionalform) *von* α *(notiert als:* $\beta \in F^s(\alpha)$*).*

3. *Ist β eine Formel, die keine universalen Quantoren mehr enthält, und ist β aus α unter Durchführung von ein oder mehreren dualen Skolemisierungsschritten erhalten worden, so heißt β* duale Skolem-Form (*oder auch* Gültigkeitsfunktionalform) *von α (notiert als: $\beta \in F^v(\alpha)$).*

Wir werden im Folgenden die Bezeichnungen Erfüllbarkeits- und Gültigkeitsfunktionalform anstelle von Skolem- bzw. dualer Skolem-Form verwenden.

Bemerkung 2.130 *Eine Erfüllbarkeitsfunktionalform einer Formel enthält nur universale Quantoren und eine Gültigkeitsfunktionalform nur existentiale Quantoren.*
Offensichtlich ist die Negation einer Erfüllbarkeitsfunktionalform eine Gültigkeitsfunktionalform und die Negation einer Gültigkeitsfunktionalform eine Erfüllbarkeitsfunktionalform.

Beispiel 2.131 *Sei $\beta =$*

$$(\forall x)(\exists y)((\forall z)(\exists u)P(x,z,u) \rightarrow (\forall z)(\exists u)Q(x,y,z,u)).$$

Dann ist

1. $\alpha_1 =$
$$(\forall x)((\exists u)P(x,g(x),u) \rightarrow (\forall z)Q(x,f(x),z,h(x,z)))$$
eine Erfüllbarkeitsfunktionalform von β, und

2. $\alpha_2 =$
$$(\exists y)((\forall z)P(c,z,f(z)) \rightarrow (\exists u)Q(c,y,g(y),u))$$
eine Gültigkeitsfunktionalform von β.

Eine direkte Verallgemeinerung von 2.128 ist dann das folgende Theorem.

Theorem 2.132

1. *Sei α' eine Gültigkeitsfunktionalform von α.*

   Dann gilt:

   α ist allgemeingültig genau dann, wenn α' allgemeingültig ist.

2. *Sei α' eine Erfüllbarkeitsfunktionalform von α.*

   Dann gilt:

   α ist erfüllbar genau dann, wenn α' erfüllbar ist.

2.4.3 Pränexe Normalform

Wir wollen als Nächstes für prädikatenlogische Formeln eine Normalform einführen, die für viele Zwecke einfacher zu handhaben ist, und die insbesondere eine wichtige Rolle für den prädikatenlogischen (klausalen) Resolutionskalkül spielt.

Definition 2.133 *Eine Formel $\alpha \in Fm_{PL}(S)$ heißt in* pränexer Normalform *genau dann, wenn α von der Gestalt $\alpha = (Q_1x_1)\dots(Q_nx_n)\beta$ ist, wobei $Q_i \in \{\forall, \exists\}$ für $1 \leq i \leq n$ und β quantorenfrei ist. $(Q_1x_1)\dots(Q_nx_n)$ ist das* Präfix *von α und β der* Kern *von α.*

Es gilt, dass jede prädikatenlogische Formel semantisch äquivalent in eine pränexe Normalform umgeformt werden kann.

Theorem 2.134 *Zu jeder Formel $\alpha \in Fm_{PL}(S)$ gibt es eine Formel $\beta \in Fm_{PL}(S)$ in pränexer Normalform mit $\alpha \equiv_{PL} \beta$.*

Beweis. Übungsaufgabe. ■

Eine direkte Folgerung aus diesem Theorem ist das

Korollar 2.135 *Zu jeder Formel $\alpha \in Fm_{PL}(S)$ gibt es eine semantisch äquivalente Formel $\beta \in Fm_{PL}(S)$ in pränexer Normalform mit einem Kern in konjunktiver (bzw. disjunktiver) Normalform.*

2.4.4 Der Satz von Herbrand

Im folgenden Abschnitt wollen wir den Satz von Herbrand behandeln, der von J. Herbrand 1930 als sogenanntes Fundamentaltheorem bewiesen wurde. Dieses Theorem stellt eine Beziehung zwischen der Allgemeingültigkeit einer prädikatenlogischen Formel α und der „aussagenlogischen Allgemeingültigkeit” einer Herbrand-Expansion dieser Formel α her, so dass das Beweisen in der Prädikatenlogik zurückgeführt wird auf eine abzählbar unendliche Menge von aussagenlogischen Tautologie-Problemen. Insbesondere die ersten automatischen Theorembeweiser machten von dieser Tatsache Gebrauch.

Definition 2.136 *Eine nichtleere, endliche Teilmenge des Herbrand-Universums zur Signatur S heißt* Herbrand-Bereich *(zur Signatur S).*

Für das Folgende wollen wir zur besseren Lesbarkeit die Schreibweise $\alpha_1 \vee \ldots \vee \alpha_n$ als Abkürzung für $\bigvee_{i=1}^n \alpha_i$ und entsprechend $\alpha_1 \wedge \ldots \wedge \alpha_n$ als Abkürzung für $\bigwedge_{i=1}^n \alpha_i$ einführen.

Definition 2.137 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält.*
$D = \{t_1, \ldots, t_n\}$ sei ein Herbrand-Bereich zur Signatur S.
Die Herbrand-Expansion *einer Aussage $\alpha \in Fm_{PL}(S)$ über D ist wie folgt definiert:*

1. *Ist α eine atomare Formel oder die Negation einer atomaren Formel, dann ist $E(\alpha, D) := \alpha$.*
2. *$E(\neg\neg\alpha, D) := E(\alpha, D)$.*
3. *$E(\boldsymbol{\alpha}, D) := E(\boldsymbol{\alpha}_1, D) \wedge E(\boldsymbol{\alpha}_2, D)$ für eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$.*
4. *$E(\boldsymbol{\beta}, D) := E(\boldsymbol{\beta}_1, D) \vee E(\boldsymbol{\beta}_2, D)$ für eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$.*
5. *$E(\boldsymbol{\gamma}, D) := E(\boldsymbol{\gamma}(t_1), D) \wedge \ldots \wedge E(\boldsymbol{\gamma}(t_n), D)$ für eine γ-Formel.*
6. *$E(\boldsymbol{\delta}, D) := E(\boldsymbol{\delta}(t_1), D) \vee \ldots \vee E(\boldsymbol{\delta}(t_n), D)$ für eine δ-Formel.*

Im folgenden Lemma halten wir einige Eigenschaften fest, die wir im Beweis des Satzes von Herbrand benötigen.

Lemma 2.138 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält. Sei ferner $\alpha \in Fm_{PL}(S)$ eine Aussage und D ein Herbrand-Bereich zur Signatur S.*
Es gilt dann:

1. *$E(\neg\alpha, D) \equiv_{PL} \neg E(\alpha, D)$.*
2. *Sei D' ein Herbrand-Bereich zur Signatur S mit $D \subseteq D'$.*
   (a) *Ist α eine Gültigkeitsfunktionalform, dann gilt*
   $$\models_{PL} E(\alpha, D) \to E(\alpha, D').$$
   (b) *Ist α eine Erfüllbarkeitsfunktionalform, dann gilt*
   $$\models_{PL} E(\alpha, D') \to E(\alpha, D).$$

3. Sei α eine Gültigkeitsfunktionalform. Dann gilt

$$\models_{PL} E(\alpha, D) \to \alpha.$$

Beweis. Alle 3 Teilaussagen werden durch strukturelle Induktion bewiesen. Wir zeigen stellvertretend (3.).

Ind.Basis: Ist α eine atomare Aussage oder die Negation einer atomaren Aussage, so folgt aufgrund der Definition der Expansion $\models_{PL} E(\alpha, D) \to \alpha$.

Ind.Schritt: Wir betrachten exemplarisch 2 Fälle. (Man beachte, dass α keine γ-Formel sein kann. Da α insbesondere eine Gültigkeitsfunktionalform ist, also keine universalen Quantoren enthält, kann der Fall, dass α eine γ-Formel ist, nicht eintreten).

1. Sei $\boldsymbol{\alpha}$ eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}$ Aussage.
   Nach Induktionsvoraussetzung gilt $\models_{PL} E(\boldsymbol{\alpha}_i, D) \to \boldsymbol{\alpha}_i$, $i = 1, 2$.
   Es gilt $\models_{PL} E(\boldsymbol{\alpha}, D) \to E(\boldsymbol{\alpha}_1, D) \wedge E(\boldsymbol{\alpha}_2, D)$, also $\models_{PL} E(\boldsymbol{\alpha}, D) \to \boldsymbol{\alpha}_1 \wedge \boldsymbol{\alpha}_2$ und damit auch $\models_{PL} E(\boldsymbol{\alpha}, D) \to \boldsymbol{\alpha}$.
2. Sei eine δ-Formel in Gültigkeitsfunktionalform gegeben. Sei weiter $D = \{t_1, \ldots, t_n\}$ angenommen. Nach Induktionsvoraussetzung gilt für jedes $t_i \in D$, $1 \leq i \leq n$,

   $$\models_{PL} E(\boldsymbol{\delta}(t_i), D) \to \boldsymbol{\delta}(t_i) \; .$$

   Wegen

   $$\models_{PL} E(\boldsymbol{\delta}, D) \to E(\boldsymbol{\delta}(t_1), D) \vee \ldots \vee E(\boldsymbol{\delta}(t_n), D)$$

   folgt somit

   $$\models_{PL} E(\boldsymbol{\delta}, D) \to \boldsymbol{\delta}(t_1) \vee \ldots \vee \boldsymbol{\delta}(t_n) \, .$$

   Weiter gilt $\models_{PL} \boldsymbol{\delta}(t_i) \to \boldsymbol{\delta}$ für $1 \leq i \leq n$. Damit folgt also weiter

   $$\models_{PL} E(\boldsymbol{\delta}, D) \to \boldsymbol{\delta} \vee \ldots \vee \boldsymbol{\delta}$$

   und folglich auch

   $$\models_{PL} E(\boldsymbol{\delta}, D) \to \boldsymbol{\delta}.$$

Damit ist der Induktionsschritt abgeschlossen und die Aussage somit bewiesen.

■

Nach dieser Vorbereitung nun der Satz von Herbrand:

Theorem 2.139 (Satz von Herbrand) *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält. Sei ferner $\alpha \in Fm_{PL}(S)$ eine Aussage und $\beta \in F^v(\alpha)$. Dann gilt:*
α ist allgemeingültig genau dann, wenn für einen Herbrand-Bereich D $E(\beta, D)$ allgemeingültig ist.

Beweis. Zunächst gilt nach Theorem 2.132, dass α genau dann allgemeingültig ist, wenn β allgemeingültig ist. Es genügt daher zu zeigen, dass gilt
$\models_{PL} \beta \Leftrightarrow$ es existiert ein Herbrand-Bereich D mit $\models_{PL} E(\beta, D)$.

„ $\Leftarrow$ " : Ist für einen Herbrand-Bereich D $E(\beta, D)$ allgemeingültig, so folgt wegen $\models_{PL} E(\beta, D) \to \beta$ (siehe Lemma 2.138 (3)) die Allgemeingültigkeit von β.

„ $\Rightarrow$ " : Sei nun β allgemeingültig.

Es existiert dann ein geschlossenes Tableau für $\{\neg\beta\}$. Da β eine Gültigkeitsfunktionalform ist, sind alle Quantoren in $\neg\beta$ universal. In einem Tableau-Beweis für β können somit keine „δ-Regeln”, das sind Tableau-Regeln $\frac{\delta}{\delta(t)}$, vorkommen.

Sei τ ein geschlossenes Tableau für $\{\neg\beta\}$.

Ein Tableau heiße *parameterfrei*, wenn in ihm keine Konstantensymbole aus $\mathcal{K}_0$ vorkommen. (Man beachte: $S(\mathcal{K}_0)$ ist die zugrundeliegende Konstantenerweiterung von S.)

Aufgrund von γ-Regeln (Tableau-Regeln der Art $\frac{\gamma}{\gamma(t)}$) können in T durchaus Konstantensymbole aus $\mathcal{K}_0$ enthalten sein. Ersetzt man in τ alle Vorkommen solcher Konstantensymbole durch ein Konstantensymbol c aus S, so resultiert ein Tableau τ', das ebenfall geschlossen ist (dies folgt über einen einfachen Induktionsbeweis über die Anzahl der Knoten des Tableaus).

Für die weitere Argumentation benötigen wir die folgende

Hilfsbehauptung:

Sei $\Sigma \subseteq Fm_{PL}(S)$ eine endliche Menge von Aussagen, die keine existentialen Quantoren enthalten und somit also Erfüllbarkeitsfunktionalformen darstellen. Sei weiter τ ein geschlossenes, parameterfreies Tableau für Σ. $D \subseteq Tm_{PL}(S)$ sei ein Herbrandbereich. Ist die Menge der Grundterme, die bei der Bildung von γ-Instanzen in T verwendet wurden, in D enthalten, dann gilt:

$$\models_{PL} \neg E\left(\bigwedge \Sigma, D\right).$$

Hierbei ist für $\Sigma = \{\alpha_1, \ldots, \alpha_n\}$ $\bigwedge \Sigma$ eine Abkürzung für $\bigwedge_{1\leq i\leq n} \alpha_i$.

Beweis der Hilfsbehauptung:

Wir führen den Beweis durch Induktion über die Anzahl $B(\tau, \Sigma)$ der Knoten von τ, die im Tableau τ unterhalb des Starttableaus für Σ liegen.

$B(\tau, \Sigma) = 0$:

In diesem Fall ist das Starttableau schon geschlossen, d.h es gilt $\beta, \neg\beta \in \Sigma$ für ein β. Damit gilt

$$\neg E\left(\bigwedge \Sigma, D\right) \equiv_{PL} \neg E\left(\bigwedge \Sigma \wedge \beta \wedge \neg\beta, D\right),$$

also

$$\neg E\left(\bigwedge \Sigma, D\right) \equiv_{PL} \neg E\left(\bigwedge \Sigma, D\right) \vee \neg E(\beta, D) \vee \neg E(\neg\beta, D).$$

Es folgt weiter

$$\neg E\left(\bigwedge \Sigma, D\right) \equiv_{PL} \neg E\left(\bigwedge \Sigma, D\right) \vee \neg E(\beta, D) \vee \neg\neg E(\beta, D).$$

Da in der letzten semantischen Äquivalenz die rechte Seite allgemeingültig ist, gilt somit

$$\models_{PL} \neg E\left(\bigwedge \Sigma, D\right).$$

Induktionsvoraussetzung:

Die Behauptung gelte für alle Tableaus τ' und Aussagenmengen Σ', die die Voraussetzungen der Hilfsbehauptung erfüllen und für die $B(\tau', \Sigma') < B(\tau, \Sigma)$ gilt.

Induktionsschritt:

Wir zeigen, dass die Behauptung dann auch für τ und Σ zutrifft.

Sei nun τ ein geschlossenes Tableau für $\Sigma = \{\alpha_1, \ldots, \alpha_n\}$.

Wir unterscheiden nach der ersten Regel, die im Aufbau von τ nach dem Starttableau angewendet wurde. Exemplarisch betrachten wir 2 Fälle.

Fall $\dfrac{\boldsymbol{\beta}}{\boldsymbol{\beta}_1 \mid \boldsymbol{\beta}_2}$:

Σ enthält eine β-Formel. Das Tableau τ hat somit die Gestalt

$$
\begin{array}{ccc}
 & \alpha_1 & \\
 & \vdots & \\
 & \boldsymbol{\beta} & \\
 & \vdots & \\
 & \alpha_n & \\
\diagup & & \diagdown \\
\boldsymbol{\beta}_1 & & \boldsymbol{\beta}_2 \\
| & & | \\
T'_1 & & T'_2 \\
* & & *
\end{array}
$$

($*$ deute an, dass alle Zweige geschlossen sind.)

Sei τ_i, $i = 1, 2$, das Tableau

$$
\begin{array}{c}
\alpha_1 \\
\vdots \\
\boldsymbol{\beta} \\
\vdots \\
\alpha_2 \\
| \\
\boldsymbol{\beta}_i \\
| \\
T'_i \\
*
\end{array}
$$

$\boldsymbol{\tau}_i$ ist somit ein geschlossenes Tableau für $\Sigma \cup \{\boldsymbol{\beta}_i\}$, $i = 1, 2$.

Wir nehmen an, dass die Menge der Grundterme, die bei der Bildung von γ-Instanzen in τ und damit auch in τ_1 und τ_2 verwendet wurden, in D enthalten ist (andernfalls ist nichts zu zeigen).

Es ist $B(\tau, \Sigma) = B(\tau_1, \Sigma \cup \{\boldsymbol{\beta}_1\}) + B(\tau_2, \Sigma \cup \{\boldsymbol{\beta}_2\}) + 2$, also ist $B(\tau_i, \Sigma \cup \{\boldsymbol{\beta}_i\}) < B(\tau, \Sigma)$, $i = 1, 2$.

Nach Induktionsvoraussetzung gilt daher $\models_{PL} \neg E(\bigwedge \Sigma \wedge \boldsymbol{\beta}_i, D)$, $i = 1, 2$. Damit gilt dann auch $\models_{PL} \neg E(\bigwedge \Sigma \wedge \boldsymbol{\beta}_1, D) \wedge \neg E(\bigwedge \Sigma \wedge \boldsymbol{\beta}_2, D)$.

Es gilt weiter

$$
\begin{aligned}
&\neg E\left(\bigwedge \Sigma \wedge \boldsymbol{\beta}_1, D\right) \wedge \neg E\left(\bigwedge \Sigma \wedge \boldsymbol{\beta}_2, D\right) \\
&\equiv_{PL} \neg \left(E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\beta}_1, D) \vee E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\beta}_2, D)\right) \\
&\equiv_{PL} \neg \left(E\left(\bigwedge \Sigma, D\right) \wedge (E(\boldsymbol{\beta}_1, D) \vee E(\boldsymbol{\beta}_2, D))\right) \\
&\equiv_{PL} \neg \left(E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\beta}, D)\right) \\
&\equiv_{PL} \neg E\left(\bigwedge \Sigma \wedge \boldsymbol{\beta}, D\right) \\
&\equiv_{PL} \neg E\left(\bigwedge \Sigma, D\right).
\end{aligned}
$$

Es folgt

$$
\models_{PL} \neg E\left(\bigwedge \Sigma, D\right).
$$

Fall $\dfrac{\boldsymbol{\gamma}}{\boldsymbol{\gamma}(t)}$:

τ hat dann die Gestalt

$$\begin{array}{c} \alpha_1 \\ \vdots \\ \boldsymbol{\gamma} \\ \vdots \\ \alpha_2 \\ | \\ \boldsymbol{\gamma}(t) \\ | \\ T' \\ * \end{array}$$

τ ist somit auch ein geschlossenes Tableau für $\Sigma \cup \{\boldsymbol{\gamma}(t)\}$ und es gilt $B(\tau, \Sigma \cup \{\boldsymbol{\gamma}(t)\}) < B(\tau, \Sigma)$.

Nach Induktionsvoraussetzung gilt $\models_{PL} \neg E(\bigwedge \Sigma \wedge \boldsymbol{\gamma}(t), D)$.

Weiter gilt, da $\boldsymbol{\gamma} \in \Sigma$:

$$\models_{PL} E\left(\bigwedge \Sigma, D\right) \leftrightarrow E\left(\bigwedge \Sigma \wedge \boldsymbol{\gamma}, D\right),$$

also

$$\models_{PL} E\left(\bigwedge \Sigma, D\right) \leftrightarrow E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\gamma}, D)$$

Wegen

$$\models_{PL} E(\boldsymbol{\gamma}, D) \to E(\boldsymbol{\gamma}(t), D)$$

erhalten wir

$$\models_{PL} E\left(\bigwedge \Sigma, D\right) \to E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\gamma}(t), D)$$

und damit

$$\models_{PL} \neg\left(E\left(\bigwedge \Sigma, D\right) \wedge E(\boldsymbol{\gamma}(t), D)\right) \to \neg E\left(\bigwedge \Sigma, D\right)$$

Ist also $\neg(E(\bigwedge \Sigma, D) \wedge E(\boldsymbol{\gamma}(t), D))$ allgemeingültig, so auch $\neg E(\bigwedge \Sigma, D)$.

Dies schließt den Beweis des Induktionsschrittes und damit auch den Beweis der Hilfsbehauptung ab.

Der Beweis der Richtung „$\Rightarrow$" unseres Theorems kann nun leicht abgeschlossen werden.
$\{\neg\beta\}$ zusammen mit dem Tableau τ' und einem Herbrand-Bereich D, der alle in γ-Instanzen in τ' eingeführten Grundterme enthält, erfüllt die Voraussetzungen der Hilfsbehauptung. Daher ist $\neg E(\neg\beta, D)$ allgemeingültig. Es folgt

$$\models_{PL} E(\beta, D).$$

Dies war zu zeigen. ■

Korollar 2.140 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält. Sei ferner $\alpha \in Fm_{PL}(S)$ eine Aussage und β eine Erfüllbarkeitsfunktionalform von α Dann gilt:*
α ist unerfüllbar genau dann, wenn für einen Herbrand-Bereich D $E(\beta, D)$ unerfüllbar ist.

Bemerkung 2.141 *Zur Anwendung des Satzes von Herbrand im Fall endlich axiomatisierbarer Theorien kann die Deduktionseigenschaft*

$$\{\alpha_1, \ldots, \alpha_n\} \models_{PL} \beta \Leftrightarrow \models_{PL} \alpha_1 \wedge \ldots \wedge \alpha_n \to \beta$$

herangezogen werden ($\alpha_1, \ldots, \alpha_n$ seien die Axiome, als Aussagen formuliert).
Sind $\alpha_1, \ldots, \alpha_n$ universale Aussagen, also Erfüllbarkeitsfunktionalformen, und ist β' eine Gültigkeitsfunktionalform von β, so gilt für einen Herbrand-Bereich D

$$\{\alpha_1, \ldots, \alpha_n\} \models_{PL} \beta \Leftrightarrow \{\alpha_1, \ldots, \alpha_n\} \models_{PL} E(\beta', D).$$

2.5 Prädikatenlogik mit Identität

Die meisten Relationen sind nur sinnvoll in einem bestimmten Kontext. Gleichheit jedoch ist eine universelle Kategorie. Unabhängig davon, welchen Bereich wir betrachten, der Begriff der Gleichheit ist in einem solchen Bereich immer anwendbar. Will man diesen Aspekt im Aufbau einer formalen Sprache berücksichtigen, so darf das Gleichheitssymbol in Strukturen aber auch nicht mehr willkürlich interpretiert werden, sondern soll immer für die Identitätsrelation stehen. Das Gleichheitsysmbol wird gewissermaßen zu einem *logischen* Symbol. Wir reflektieren diesen Sachverhalt auf der semantischen Ebene dadurch, dass wir für die Prädikatenlogik mit Identität nur solche Strukturen zulassen, in denen das Gleichheitssymbol als Identität interpretiert ist.
In diesem Abschnitt setzen wir voraus, dass alle Signaturen das Gleichheitssymbol $\approx$ als zweistelliges Relationssymbol enthalten, so dass wir aus diesem Grund $\approx$ in der Signatur nicht mehr aufführen, sondern als zusätzliches logisches Symbol unserer prädikatenlogischen Sprache behandeln. Im Aufbau der Formeln benutzen wir für eine Formel $\approx (s, t)$ die infix-Schreibweise, d.h wir schreiben $s \approx t$.
Bei der Definition der Semantik müssen wir dann beachten, dass diesem Gleichheitssymbol immer die Identitätsrelation über dem jeweiligen Grundbereich zugeordnet ist. Die Identitätsrelation werden wir daher in der Struktur nicht mehr aufführen.

Definition 2.142 *Sei S eine Signatur.*
Eine S-Struktur $\mathcal{A}$ heißt normal *genau dann, wenn $\approx^{\mathcal{A}}$ die Identitätsrelation auf $|\mathcal{A}|$ ist* *(d.h. $\approx^{\mathcal{A}} = \{(a, a) \mid a \in |\mathcal{A}|\}$).*

Notation 2.143 *Um anzudeuten, dass $\mathcal{A}$ eine normale Struktur ist, schreiben wir anstelle von $\mathcal{A}$ auch $\mathcal{A}_{\approx}$.*

Die Semantik der Prädikatenlogik mit Gleichheitist wie im Falle der Prädikatenlogik definiert – allerdings mit der Beschränkung auf normale Strukturen. Ist $\mathcal{A}$ eine normale Struktur, so schreiben wir statt $\alpha^{\mathcal{A},w} = T$ auch $\mathcal{A}, w \models_{PLI} \alpha$. Die Begriffe „Allgemeingültigkeit einer Formel α" (notiert durch $\models_{PLI} \alpha$), Erfüllbarkeit einer Formel α wie auch der Begriff der semantischen Folgerung von α aus der Prämissenmenge Σ (notiert als $\Sigma \models_{PLI} \alpha$) ergeben sich aufgrund der Beschränkung auf normale Strukturen direkt aus der Prädikatenlogik.

Bemerkung 2.144 *Aufgrund der Definition der Semantik für die Prädikatenlogik mit Identität gilt natürlich $\Sigma \models_{PL} \alpha \;\Rightarrow\; \Sigma \models_{PLI} \alpha$.*
Die Umkehrung gilt aber nicht. Ist zum Beispiel $\Sigma = \{c \approx d, P(c)\}$, dann gilt $\Sigma \models_{PLI} P(d)$, aber nicht $\Sigma \models_{PL} P(d)$.

Durch die feste Interpretation des Gleichheitssymbols als Identitätsrelation wird die Ausdrucksfähigkeit der Prädikatenlogik mit Identität gegenüber der Prädikatenlogik erhöht. In der Prädikatenlogik mit Identität sind jetzt Anzahlaussagen formulierbar. Konnten wir in der Prädikatenlogik noch ausdrücken, dass im Grundbereich mindestens k, $k \in \mathbb{N} \setminus \{0\}$, verschiedene Elemente vorhanden sind, so können wir jetzt sogar formulieren, dass der Grundbereich genau k Elemente besitzt.

Beispiel 2.145 *Die Aussage*

$$\alpha := (\exists x)\,(\exists y)\,(\exists z)\,(\neg x \approx y \wedge \neg x \approx z \wedge \neg y \approx z)$$

ist nur in solchen normalen Strukturen erfüllt, deren Grundbereich mindestens 3 Elemente enthält. Die Aussage

$$\beta := (\forall x)\,(\forall y)\,(\forall z)\,(\forall u)\,(x \approx y \vee x \approx z \vee x \approx u \vee y \approx z \vee y \approx u \vee z \approx u)$$

ist nur in solchen normalen Strukturen erfüllt, deren Grundbereich höchstens 3 verschiedene Elemente besitzt. Für eine Aussage γ mit $\{\alpha, \beta\} \models_{PLI} \gamma$ gilt daher, dass sie genau in allen 3-elementigen Strukturen allgemeingültig ist.

Eine Folge, hiervon ist, dass der Satz von Löwenheim und Skolem in der Form von Theorem 2.112 für die Prädikatenlogik mit Identität nicht gilt.
Bekanntermaßen ist die Identitätsrelation eine Äquivalenzrelation. Darüberhinaus ist sie aber auch eine Kongruenzrelation. Dies drückt sich in den Eigenschaften aus, die in dem nachfolgenden Lemma formuliert sind.

Lemma 2.146 *Sei S eine Signatur, f ein n-stelliges Funktionssymbol und P ein n-stelliges Relationssymbol der Signatur S.*
Die folgenden Formeln sind dann (in normalen Strukturen) allgemeingültig:

1. $(\forall x)\, x \approx x \qquad (Ref)$
2. $Ers\,(f) := (\forall x_1) \ldots (\forall x_n)\,(\forall y_1) \ldots (\forall y_n)\,(x_1 \approx y_1 \wedge \ldots \wedge x_n \approx y_n \to f\,(x_1, \ldots, x_n) \approx f\,(y_1, \ldots, y_n))$
   „Ersetzungsaxiom für f"
3. $Ers\,(P) := (\forall x_1) \ldots (\forall x_n)\,(\forall y_1) \ldots (\forall y_n)\,(x_1 \approx y_1 \wedge \ldots \wedge x_n \approx y_n \to (P\,(x_1, \ldots, x_n) \to P\,(y_1, \ldots, y_n)))$
   „Ersetzungsaxiom für P"

Weiter gilt:
$\{Ref,\, Ers\,(\approx)\} \models_{PL} (\forall x)\,(\forall y)\,(x \approx y \to y \approx x) \qquad (Sym)$
$\{Ref,\, Ers\,(\approx)\} \models_{PL} (\forall x)\,(\forall y)\,(\forall z)\,(x \approx y \wedge y \approx z \to x \approx z) \qquad (Trans)$

Beweis. Wir begründen

$$\{Ref,\, Ers\,(\approx)\} \models_{PL} (\forall x)\,(\forall y)\,(x \approx y \to y \approx x)\,. \tag{Sym}$$

Sei $I = (\mathcal{A}, w)$ ein prädikatenlogisches Modell für $\{Ref,\, Ers\,(\approx)\}$. Es gilt dann

$$[(\forall x_1)\,(\forall x_2)\,(\forall y_1)\,(\forall y_2)\,(x_1 \approx y_1 \wedge x_2 \approx y_2 \to (x_1 \approx x_2 \to y_1 \approx y_2))]^I = T.$$

Daraus folgt für beliebige a, b aus dem Grundbereich $|\mathcal{A}|$:

$$[(x_1 \approx y_1 \wedge x_2 \approx y_2 \to (x_1 \approx x_2 \to y_1 \approx y_2))]^{I\{x_1/a, x_2/a, y_1/b, y_2/a\}} = T,$$

also wenn $a \approx^{\mathcal{A}} b$ und $a \approx^{\mathcal{A}} a$, dann gilt mit $a \approx^{\mathcal{A}} a$ auch $b \approx^{\mathcal{A}} a$.
Da I auch ein Modell von Ref ist, gilt für alle $a \in |\mathcal{A}|$ $a \approx^{\mathcal{A}} a$. Wir erhalten, dass für alle $a, b \in |\mathcal{A}|$ gilt: wenn $a \approx^{\mathcal{A}} b$ dann auch $b \approx^{\mathcal{A}} a$, also $[x \approx y \to y \approx x]^{I\{x/a, y/b\}} = T$. Es folgt, dass I die Formel $(\forall x)\,(\forall y)\,(x \approx y \to y \approx x)$ erfüllt. ■

Sei S eine Signatur. Mit $\mathrm{fun}\,(S) := \{Ers\,(f) \mid f$ Funktionssymbol in $S\}$ bezeichnen wir die Menge der Ersetzungsaxiome für Funktionssymbole in S und mit $\mathrm{rel}\,(S) := \{Ers\,(P) \mid P$ Relationssymbol in $S\}$ die Menge der Ersetzungsaxiome für Relationssymbole in S.
Als Menge der *Gleichheitsaxiome* wird dann die Formelmenge $Eq\,(S) := \{Ref\} \cup \mathrm{fun}\,(S) \cup \mathrm{rel}\,(S)$ bezeichnet.

Lemma 2.147 *Sei $\alpha \in Fm_{PL}\,(S)$ mit $fvar\,(\alpha) \subseteq \{x_1, \ldots, x_n\}$. Sei ferner $\beta := \alpha\,\{x_1/y_1, \ldots, x_n/y_n\}$, wobei $y_1, \ldots, y_n$ neue, in α nicht vorkommende Variablen sind. Dann gilt*

$$Eq\,(S) \models_{PL} (\forall x_1) \ldots (\forall x_n)\,(\forall y_1) \ldots (\forall y_n)\,(x_1 \approx y_1 \wedge \ldots \wedge x_n \approx y_n \to (\alpha \to \beta))\,.$$

Beweis. Durch strukturelle Induktion. ■

Bemerkung 2.148
Das „Formelschema" $(\forall x_1) \ldots (\forall x_n)\,(\forall y_1) \ldots (\forall y_n)\,(x_1 \approx y_1 \wedge \ldots \wedge x_n \approx y_n \to (\alpha \to \beta))$, wobei β wie oben aus α gebildet ist, wird als Leibniz-Prinzip *bezeichnet. Es ist zu $Eq\,(S)$ gleichwertig und wird auch im Tableau- bzw. Resolutionskalkül der Prädikatenlogik mit Identität zugrundegelegt.*

Über diese Gleichheitsaxiome kann man den folgenden Zusammenhang zwischen der Prädikatenlogik und der Prädikatenlogik mit Identität herstellen.

Theorem 2.149 *Sei* $\alpha \in Fm_{PL}(S)$ *und* $\Sigma \subseteq Fm_{PL}(S)$. *Es gilt dann*

$$\Sigma \models_{PLI} \alpha \;\Leftrightarrow\; \Sigma \cup Eq(S) \models_{PL} \alpha$$

Den Beweis dieses Theorems werden wir nicht führen. Üblicherweise erhält man dieses Resultat aus dem Vollständigkeitssatz für die Prädikatenlogik mit Identität, dessen Beweis wir später skizzieren werden.

Die zentralen Resultate der Prädikatenlogik – Lemma von Hintikka, Modellexistenztheorem, Kompaktheitssatz, Satz von Löwenheim und Skolem und der Vollständigkeitssatz – gelten mit geringfügigen Änderungen auch für die Prädikatenlogik mit Identität. In den meisten Fällen kann der entsprechende Beweis aus der Prädikatenlogik direkt übertragen werden, so dass wir nur für die Fälle, in denen es größere Abweichungen im Beweis gibt, eine Beweisskizze liefern werden.

2.5.1 Hintikka's Lemma für die Prädikatenlogik mit Identität

Herbrand-Strukturen spielten eine wesentliche Rolle im Beweis der Vollständigkeit der Prädikatenlogik (ohne Identität): Hintikka's Lemma garantierte dort die Existenz eines Herbrand-Modells. Enthält nun aber die Signatur beispielsweise wenigstens ein Funktionssymbol, so ist das Herbrand-Universum unendlich. Eine Formel wie

$$(\forall x)\,(\forall y)\, x \approx y,$$

die nur in normalen Strukturen mit einelementigem Grundbereich gilt, kann daher im Falle der Prädikatenlogik mit Identität in einer solchen Herbrand-Struktur nicht erfüllbar sein, so dass wir das Lemma von Hintikka diesen Gegebenheiten anpassen werden. Hierbei wird die Rolle der Herbrand-Strukturen von den kanonischen Strukturen übernommen werden:

Definition 2.150 *Eine* S*-Struktur* $\mathcal{A}$ *heißt* kanonisch *genau dann, wenn für jedes* $a \in |\mathcal{A}|$ *ein Grundterm* $t \in Tm_{PL}(S)$ *mit* $t^{\mathcal{A}} = a$ *existiert.*

Bemerkung 2.151 *Jede Herbrand-Struktur ist kanonisch, aber nicht jede kanonische Struktur ist eine Herbrand-Struktur.*
Beispiel:
Sei die Signatur $Sig = (S; 0)$ *gegeben. Die (normale)* Sig*-Struktur* $\mathcal{N} := \left(\mathbb{N}; S^{\mathcal{N}}; 0\right)$ *(wir schreiben hier* 0 *statt* $0^{\mathcal{N}}$*) mit* $S^{\mathcal{N}}(n) := n+1$ *für alle* $n \in \mathbb{N}$ *ist dann kanonisch.*

Definition 2.152 *Eine Menge* $\Sigma \subseteq Fm_{PL}(S)$ *von Aussagen heißt* prädikatenlogische Hintikka-Menge mit Gleichheit *genau dann, wenn gilt:*

1. *Keine atomare Formel ist zugleich mit ihrer Negation in* Σ *enthalten.*
2. $\bot \notin \Sigma$, $\neg\top \notin \Sigma$
3. $\neg\neg\beta \in \Sigma \;\Rightarrow\; \beta \in \Sigma$
4. $\boldsymbol{\alpha} \in \Sigma \;\Rightarrow\; \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2 \in \Sigma$ *(für* α*-Formeln mit Komponenten* $\boldsymbol{\alpha}_1$ *und* $\boldsymbol{\alpha}_2$*)*
5. $\boldsymbol{\beta} \in \Sigma \;\Rightarrow\; \boldsymbol{\beta}_1 \in \Sigma$ *oder* $\boldsymbol{\beta}_2 \in \Sigma$ *(für* β*-Formeln mit Komponenten* $\boldsymbol{\beta}_1$ *und* $\boldsymbol{\beta}_2$*)*
6. $\boldsymbol{\gamma} \in \Sigma \;\Rightarrow\; \boldsymbol{\gamma}(t) \in \Sigma$ *für alle Grundterme* $t \in Tm_{PL}(S)$ *und* γ*-Formeln*
7. $\boldsymbol{\delta} \in \Sigma \;\Rightarrow\; \boldsymbol{\delta}(t) \in \Sigma$ *für eine* δ*-Formel und einen Grundterm* $t \in Tm_{PL}(S)$
   Neu:
8. $t \approx t \in \Sigma$ *für jeden Grundterm* $t \in Tm_{PL}(S)$
9. $s_1 \approx t_1, \ldots, s_n \approx t_n \in \Sigma \;\Rightarrow\; f(s_1, \ldots, s_n) \approx f(t_1, \ldots, t_n) \in \Sigma$ *für jedes* n*-stellige Funktionssymbol in* S

10. $s_1 \approx t_1, \ldots, s_n \approx t_n, R(s_1, \ldots, s_n) \in \Sigma \Rightarrow R(t_1, \ldots, t_n) \in \Sigma$ *für jedes n-stellige Relationssymbol in S*

Lemma 2.153 (Hintikka's Lemma für die Prädikatenlogik mit Identität) *Sei S eine Signatur mit nichtleerer Konstantensymbolmenge und $\Sigma \subseteq Fm_{PL}(S)$ eine prädikatenlogische Hintikka-Menge mit Gleichheit.*
Dann gilt:
Σ hat ein normales Modell, das kanonisch ist.

Beweis. Skizze.
Wir geben nur die Schritte an, in denen der Beweis vom Beweis des prädikatenlogischen Lemmas von Hintikka abweicht.
Wir betrachten zunächst Σ als eine prädikatenlogische Formelmenge und bestimmen nach dem prädikatenlogischen Lemma von Hintikka für Σ ein Herbrand-Modell $\mathcal{A} = \left(A; \approx^{\mathcal{A}}, (R_i^{\mathcal{A}})_{i \in I}; (f_j^{\mathcal{A}})_{j \in J}; (c_k^{\mathcal{A}})_{k \in K}\right)$ zur Signatur S (d.h. $\mathcal{A}$ ist im allgemeinen keine normale Struktur; es können in $\mathcal{A}$ Grundterme s und t mit $s \neq t$ und $(s,t) \in \approx^{\mathcal{A}}$ vorhanden sein).
Nun besagen nach Lemma 2.146 die Bedingungen (8) - (10) in der Definition einer prädikatenlogischen Hintikka-Menge mit Gleichheit insbesondere, dass $\approx^{\mathcal{A}}$ eine Äquivalenzrelation auf A ist.
Wir bilden daher die Faktorstruktur

$$\mathcal{A}' := \left(A'; \approx^{\mathcal{A}'}, \left(R_i^{\mathcal{A}'}\right)_{i \in I}; \left(f_j^{\mathcal{A}'}\right)_{j \in J}; \left(c_k^{\mathcal{A}'}\right)_{k \in K}\right)$$

mit
$A' := \{[t] \mid t \in A\}$ (d.h. $A' = A/\approx^{\mathcal{A}}$; $[t]$ ist die Äquivalenzklasse von t unter $\approx^{\mathcal{A}}$)
$f_j^{\mathcal{A}'}([t_1], \ldots, [t_n]) := \left[f_j^{\mathcal{A}}(t_1, \ldots, t_n)\right]$
$([t_1], \ldots, [t_n]) \in R_i^{\mathcal{A}'} :\Leftrightarrow (t_1, \ldots, t_n) \in R_i^{\mathcal{A}}$
$c_k^{\mathcal{A}'} := \left[c_k^{\mathcal{A}}\right]$
$[t_1] \approx^{\mathcal{A}'} [t_2] :\Leftrightarrow t_1 \approx^{\mathcal{A}} t_2$ (d.h. $\approx^{\mathcal{A}'}$ ist die Identität auf A')
Es ist leicht zu sehen, dass diese Definitionen von der jeweiligen Wahl des Repräsentatenten unabhängig sind.
Es folgt:

- $\mathcal{A}'$ ist normal, da $\approx^{\mathcal{A}'}$ die Identität auf A' ist
- $\left[t^{\mathcal{A}}\right] = t^{\mathcal{A}'}$ für jeden Grundterm t, also ist $\mathcal{A}'$ kanonisch

Ist nun w eine $\mathcal{A}$-Belegung und definiert man eine $\mathcal{A}'$-Belegung w' durch

$$w'(x) := [w(x)],$$

so gilt für beliebige Terme t in $Tm_{PL}(S)$

$$t^{\mathcal{A}',w'} = \left[t^{\mathcal{A},w}\right]$$

und für beliebige Formeln α in $Fm_{PL}(S)$

$$\alpha^{\mathcal{A}',w'} = \alpha^{\mathcal{A},w}.$$

Da $\mathcal{A}$ ein Herbrand-Modell von Σ ist, folgt somit, dass $\mathcal{A}'$ ein normales, kanonisches Modell von Σ. ■

Bemerkung 2.154 *Man beachte, dass der Grundbereich von $\mathcal{A}'$ aufgrund der Faktorisierung von A nach $\approx^{\mathcal{A}}$ höchstens abzählbar unendlich viele Elemente haben kann.*

2.5.2 Das Modellexistenztheorem für die Prädikatenlogik mit Identität

Mit Hilfe des eben angegebenen Lemmas von Hintikka können wir für die Prädikatenlogik mit Identität ebenfalls ein Modellexistenztheorem zeigen. Wir legen dabei dieselben Voraussetzungen hinsichtlich der Signatur S wie im Fall der Prädikatenlogik zugrunde, d.h. es sind höchstens abzählbar unendlich viele Relationssymbole, Funktionssymbole und Konstantensymbole vorhanden und $\mathcal{K}_0$ ist eine abzählbar unendliche Menge von Konstantensymbolen, die in der Signatur S nicht vorkommen.

Definition 2.155 *Es sei $\mathcal{C} \subseteq \mathcal{P}(Fm_{PL}(S(\mathcal{K}_0)))$ eine Menge (Kollektion) von Aussagenmengen.*
$\mathcal{C}$ heißt prädikatenlogische Konsistenzeigenschaft mit Gleichheit*genau dann, wenn für alle $\Sigma \in \mathcal{C}$ gilt:*

1. *α atomare Formel und $\alpha \in \Sigma \Rightarrow \neg\alpha \notin \Sigma$.*
2. *$\bot \notin \Sigma$, $\neg\top \notin \Sigma$.*
3. *$\neg\neg\beta \in \Sigma \Rightarrow \Sigma \cup \{\beta\} \in \mathcal{C}$.*
4. *$\boldsymbol{\alpha} \in \Sigma \Rightarrow \Sigma \cup \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\} \in \mathcal{C}$ (für eine α-Formel mit Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$).*
5. *$\boldsymbol{\beta} \in \Sigma \Rightarrow \Sigma \cup \{\boldsymbol{\beta}_1\} \in \mathcal{C}$ oder $\Sigma \cup \{\boldsymbol{\beta}_2\} \in \mathcal{C}$ (für eine β-Formel mit Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$).*
6. *$\boldsymbol{\gamma} \in \Sigma \Rightarrow \Sigma \cup \{\boldsymbol{\gamma}(t)\} \in \mathcal{C}$ für eine γ-Formel und jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$.*
7. *$\boldsymbol{\delta} \in \Sigma \Rightarrow \Sigma \cup \{\boldsymbol{\delta}(c)\} \in \mathcal{C}$ für eine δ-Formel und für jedes bzgl. Σ neue Konstantensymbol $c \in \mathcal{K}_0$.*
   Neu:
8. *$\Sigma \cup \{t \approx t\} \in \mathcal{C}$ für jeden Grundterm $t \in Tm_{PL}(S(\mathcal{K}_0))$.*
9. *$s_1 \approx t_1, \ldots, s_n \approx t_n \in \Sigma \Rightarrow \Sigma \cup \{f(s_1, \ldots, s_n) \approx f(t_1, \ldots, t_n)\} \in \mathcal{C}$ für jedes n-stellige Funktionssymbol in S.*
10. *$s_1 \approx t_1, \ldots, s_n \approx t_n, R(s_1, \ldots, s_n) \in \Sigma \Rightarrow \Sigma \cup \{R(t_1, \ldots, t_n)\} \in \mathcal{C}$ für jedes n-stellige Relationssymbol in S.*

Wie im Falle der Prädikatenlogik kann gezeigt werden, dass jede solche prädikatenlogische Konsistenzeigenschaft mit Gleichheit zu einer erweitert werden kann, die teilmengenabgeschlossen und von endlichem Charakter ist. Ferner gilt, dass für eine aufsteigende, abzählbar unendliche Folge (bzgl. Mengeninklusion) von Elementen (d.h. Mengen) in einer solchen prädikatenlogischen Konsistenzeigenschaft mit Gleichheit die Vereinigung dieser Mengen ebenfalls ein Element der Konsistenzeigenschaft ist. Damit ergibt sich wie im Falle der Prädikatenlogik mit praktisch unverändertem Beweis (nur für den Nachweis der Hintikka-Menge sind zusätzliche Fälle zu betrachten) das folgende Modellexistenztheorem.

Theorem 2.156 (Prädikatenlogisches Modellexistenztheorem mit Gleichheit) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen und $\mathcal{C} \subseteq \mathcal{P}(Fm_{PL}(S(\mathcal{K}_0)))$ eine prädikatenlogische Konsistenzeigenschaft mit Gleichheit und mit $\Sigma \in \mathcal{C}$.*
Dann gilt:
Σ ist erfüllbar in einem normalen kanonischen Modell zur Signatur $S(\mathcal{K}_0)$.

Als Anwendungen dieses Modellexistenztheorems erhalten wir wiederum den Kompaktheitssatz und den Satz von Löwenheim und Skolem.

Theorem 2.157 (Kompaktheitssatz der Prädikatenlogik mit Identität) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen.*
Σ ist genau dann erfüllbar, wenn jede endliche Teilmenge von Σ erfüllbar ist.

Theorem 2.158 (Löwenheim und Skolem) *Sei $\Sigma \subseteq Fm_{PL}(S)$ eine Menge von Aussagen.*
Ist Σ erfüllbar, so ist Σ erfüllbar in einer Struktur mit höchstens abzählbarem *Grundbereich.*

2.5.3 Tableau- und Resolutionskalkül für die Prädikatenlogik mit Gleichheit

Wir geben hier nur noch die gegenüber der Prädikatenlogik neu hinzukommenden Tableau-Regeln bzw. Resolutionsexpansionsregeln an. Die neuen Regeln resultieren aus der Reflexivitätseigenschaft von = und dem Leibniz-Prinzip.

Tableau-Kalkül

Die zusätzlichen Regeln sind:

1. Tableau-Reflexivitätsregel: (informal)

   An einen Zweig eines Tableaus kann ein Knoten angehängt werden, der mit einer Aussage $t \approx t$ markiert ist. (notiert als $\frac{}{t \approx t}$). t ist dabei ein Grundterm aus $S(\mathcal{K}_0)$.

2. Tableau-Ersetzungsregel: (informal)

- Wenn in einem Zweig eines Tableaus ein Knoten mit $s \approx t$ und ein anderer Knoten mit $\alpha\{u/s\}$ markiert sind,

  dann kann an das Ende des Zweiges ein Knoten angehängt werden, der mit $\alpha\{u/t\}$ markiert ist. Wir notieren diese Regel in der Form

$$\frac{\begin{array}{c} s \approx t \\ \alpha\{u/s\} \end{array}}{\alpha\{u/t\}}$$

Resolutionskalkül

In diesem Fall lauten die ergänzenden Regeln:

1. Resolutions-Reflexivitätsregel: (informal)

   Eine Resolutionsexpansion kann um die Klausel $[t \approx t]$ erweitert werden, wobei t ein Grundterm ist.

2. Resolutions-Ersetzungsregel: (informal)

   Eine Resolutionsexpansion, in der die Klauseln $[\alpha_1, \ldots, \alpha_n, s \approx t]$ und $[\alpha'_1, \ldots, \alpha'_m, \beta\{x/s\}]$ vorhanden sind, kann um die Klausel $[\alpha_1, \ldots, \alpha_n, \alpha'_1, \ldots, \alpha'_m, \beta\{x/t\}]$ erweitert werden. Wir notieren diese Resolutionsexpansionsregel in der Form

$$\frac{\begin{array}{c} [\alpha_1, \ldots, \alpha_n, s \approx t] \\ [\alpha'_1, \ldots, \alpha'_m, \beta\{x/s\}] \end{array}}{[\alpha_1, \ldots, \alpha_n, \alpha'_1, \ldots, \alpha'_m, \beta\{x/t\}]}$$

Es gilt dann

Theorem 2.159 (Korrektheit und Vollständigkeit) *Die um die angegebenen Regeln erweiterten Tableau- und Resolutionskalküle sind korrekt und vollständig für die Prädikatenlogik mit Identität.*

Beweis. Übungsaufgabe.
Der Nachweis der Korrektheit und Vollständigkeit entspricht ganz der Vorgehensweise im rein prädikatenlogischen Fall. ∎

2.6 Mehrsortige Prädikatenlogik

Im Aufbau der Prädikatenlogik hatten wir bisher immer nur einen Grundbereich, d.h. eine Sorte von Objekten, zugrundegelegt. Vielen Anwendungen der Informatik liegen jedoch Strukturen zugrunde, in denen Objekte verschiedenen Typs vorhanden sind. Im folgenden sollen zwei Möglichkeiten angedeutet werden, mit denen solche mehrsortigen Strukturen im prädikatenlogischen Formalismus behandelt werden können.

2.6.1 Sortenreduktion

Bei dieser Methode wird ein Grundbereich, der die Objekte der verschiedenen Typen enthält, zugrundegelegt. Um dann über Objekte eines bestimmten Typs sprechen zu können, sind daher zusätzliche Sortenprädikate erforderlich, die genau auf die Objekte dieses Typs zutreffen, so dass Objekte dieses Typs mit diesen Prädikaten ausgewählt werden können.

Beispiel 2.160 *Bei der formalen Behandlung minimaler Spannbäume für gerichtete Graphen wird eine Struktur benötigt für kantenmarkierte gerichtete Graphen, deren Kanten (a,b) mit Kantengewichten $g(a,b)$ in Form von reellen Zahlen markiert sind.*
Eine geeignete Struktur für gerichtete Graphen ist z.B. $\mathcal{A} = \left(A; R^{\mathcal{A}}\right)$ mit einer zweistelligen Relation $R^{\mathcal{A}}$ über A. Für die reellen Zahlen ist beispielsweise die Struktur $\mathfrak{R} = (\mathbb{R}; =, <; +, \cdot; 0, 1)$ passend.
Der Grundbereich B für kantenmarkierte gerichtete Graphen besteht dann aus der Menge der Knoten A und der Menge der reellen Zahlen $\mathbb{R}$, d.h. $B = A \cup \mathbb{R}$. Um nun gezielt über Knoten bzw. reelle Zahlen sprechen zu können, benötigen wir daher Sortenprädikate z.B. $V^{\mathcal{B}}$ (für die Knoten) und $Real^{\mathcal{B}}$ (für die reellen Zahlen), so dass man für kantenmarkierte gerichtete Graphen z.B. die Struktur

$$\mathcal{B} = \left(B; V^{\mathcal{B}}, Real^{\mathcal{B}}, R^{\mathcal{B}}, =^{\mathcal{B}}, <^{\mathcal{B}}; g^{\mathcal{B}}, +^{\mathcal{B}}, \cdot^{\mathcal{B}}; 0^{\mathcal{B}}, 1^{\mathcal{B}}\right)$$

verwenden kann, wobei auf $\mathbb{R}$ die Relationen, Funktionen und Konstanten $=^{\mathcal{B}}, <^{\mathcal{B}}; +^{\mathcal{B}}, \cdot^{\mathcal{B}}; 0^{\mathcal{B}}, 1^{\mathcal{B}}$ mit den entsprechenden Relationen, Funktionen und Konstanten der Struktur $\mathfrak{R}$ übereinstimmen. Eine passende Signatur ist dann $S = (\approx, <, V, Real, R; g, +, \cdot; 0, 1)$.
Möchte man dann z.B. ausdrücken, dass die Kantenrelation nur zwischen Knoten besteht, so kann dies in der folgenden Form geschehen:

$$(\forall x)(\forall y)(R(x,y) \to [x] \wedge [y]),$$

d.h. unter Hinzufügung dieser Formel erreichen wir, dass $R(x,y)$ immer eine Aussage über Knoten x und y macht. Im allgemeinen werden jedoch Sortenprädikate verwendet, um gezielt über Objekte einer Sorte zu sprechen wie z.B. in den folgenden Aussagen:
„Der Graph hat genau einen Knoten":

$$(\forall x)(\forall y)([x] \wedge [y] \to x \approx y)$$

„Zu je zwei reellen Zahlen x und y gibt es eine reelle Zahl z, die verschieden von x und y ist":

$$(\forall x)(\forall y)(Real(x) \wedge Real(y) \to (\exists z)(Real(z) \wedge \neg x \approx z \wedge \neg y \approx z)).$$

All- und Existenzaussagen werden also in der Regel immer über die Relativierung auf die Sorte ausgedrückt:
„Für alle Knoten x gilt α" durch

$$(\forall x)([x] \to \alpha)$$

und „es gibt eine reelle Zahl z, so dass β" durch

$$(\exists z)(Real(z) \wedge \beta).$$

2.6.2 Mehrsortige Sprachen und Strukturen

Bei diesem Ansatz wird von vornherein eine mehrsortige prädikatenlogische Sprache gewählt.

Beispiel 2.161 *Einen Vektorraum $\mathcal{V}$ über einem Körper $\mathcal{K}$ können wir durch die folgende zweisortige Struktur repräsentieren:*

$$\mathcal{V} = \left(K, V; =^K, =^V; +^K, \cdot^K, \circ^V, *^{K,V}; 0^K, 1^K, e^V\right),$$

wobei
$\mathcal{K} = \left(K; =^K; +^K, \cdot^K; 0^K, 1^K\right)$ ein Körper,
$\left(V; =^V; \circ^V; e^V\right)$ die additive Gruppe der Vektoren und
$*^{K,V} : K \times V \to V$ die skalare Multiplikation von Vektoren ist.
Eine Sprache zur Behandlung einer solchen zweisortigen Struktur enthält dann

- *Individuenvariablen für jede Sorte*
- *Prädikatssymbole und Funktionssymbole vom passenden „Typ", z.B.*

  $+$ *vom Typ* $k \times k \to k$

  $*$ *vom Typ* $k \times v \to v$,

  falls k *der Typ von* K *und* v *der Typ von* $\mathcal{V}$ *ist.*
- *Quantoren für jede Sorte*
- *Der Aufbau von Termen und Formeln muss typkompatibel erfolgen.*

  Einige Formeln sollen dies verdeutlichen. Wir verwenden x_i, $i \in \mathbb{N}$, *als Variable vom Typ* k *(also für Elemente des Körpers) und* y_j, $j \in \mathbb{N}$, *als Variable vom Typ* v*:*

  $(\forall x_0)\,(\forall x_1)\,(\forall x_2)\,(x_0 \cdot (x_1 + x_2) \approx x_0 \cdot x_1 + x_0 \cdot x_2)$

  $(\forall x_0)\,(\forall x_1)\,(\forall y_0)\,((x_0 \cdot x_1) * y_0 \approx x_0 * (x_1 * y_0))$

Der formale Aufbau einer mehrsortigen Prädikatenlogik in der hier angedeuteten Weise bereitet keine prinzipiellen Schwierigkeiten. Lediglich der technische Aufwand erhöht sich. Wir wollen hier wenigstens noch kurz skizzieren, wie der Signatur- und der Strukturbegriff im mehrsortigen Fall aussehen könnten.

Definition 2.162 *Im mehrsortigen Fall ist eine Signatur* S *ein Tupel* $S = (So; \mathcal{R}; \mathcal{F}; \mathcal{K}; \mathrm{ar})$ *mit*

1. *So ist nichtleere Menge (die Menge der Sortensymbole)*

   (o.B.d.A. sei angenommen, dass die Symbole $\times$ *und* $\to$ *nicht in So vorkommen.)*

   Wir führen für die nachfolgend angegebenen Worte über dem Alphabet $So \cup \{\times, \to\}$ *mit* $k_1, k_2, \ldots, k_n, k \in So, (n \in \mathbb{N})$ *Bezeichnungen ein:*

   (a) $k_1 \times k_2 \times \ldots \times k_n$ *heiße* Relationstyp *und*

   (b) $k_1 \times k_2 \times \ldots \times k_n \to k$ Funktionstyp.

2. $\mathcal{R}, \mathcal{F}, \mathcal{K}$ *seien wie im einsortigen Fall die Mengen der Relations-, Funktions- und Konstantensymbole, wobei diese Mengen jeweils paarweise disjunkt sind.*

3. ar *ist jetzt eine Abbildung mit der Eigenschaft*

   (a) *für* $P \in \mathcal{R}$ *ist* $\mathrm{ar}\,(P)$ *ein Relationstyp*

   (b) *für* $f \in \mathcal{F}$ *ist* $\mathrm{ar}\,(f)$ *ein Funktionstyp, und*

   (c) *für* $c \in \mathcal{K}$ *ist* $\mathrm{ar}\,(c) \in So$.

Der Aufbau der Terme einer Sorte und der atomaren Formeln erfolgt dann wie zuvor im unsortierten Fall, aber typkompatibel.

Definition 2.163

1. *Jede Variable der Sorte* $k \in So$ *ist ein Term der Sorte* k.

   Ist f *ein Funktionssymbol mit* $\mathrm{ar}\,(f) = k_1 \times k_2 \times \ldots \times k_n \to k$ *und ist* t_i *ein Term der Sorte* k_i *,* $1 \leq i \leq n$, *so ist* $f\,(t_1, \ldots, t_n)$ *ein Term der Sorte* k.

2. *Ist* P *ein Prädikatensymbol mit* $\mathrm{ar}\,(P) = k_1 \times k_2 \times \ldots \times k_n$ *und ist* t_i *ein Term der Sorte* k_i *,* $1 \leq i \leq n$, *dann ist* $P\,(t_1, \ldots, t_n)$ *atomare Formel.*

   Der Aufbau der weiteren Formeln erfolgt entsprechend dem einsortigen Fall.

Definition 2.164 *Eine S-Struktur$\mathcal{A}$ ist ein Tupel $\mathcal{A} = \left(\Delta; \mathcal{R}^{\mathcal{A}}; \mathcal{F}^{\mathcal{A}}; \mathcal{K}^{\mathcal{A}}\right)$, wobei*

1. *Δ eine Abbildung ist, die jedem $k \in So$ eine nichtleere Menge Δ_k zuordnet (die Menge der Objekte der Sorte k),*
2. *für jedes Relationssymbol $R \in \mathcal{R}$ mit $\operatorname{ar}(R) = k_1 \times \ldots \times k_n$ $R^{\mathcal{A}}$ eine Relation mit $R^{\mathcal{A}} \subseteq \Delta_{k_1} \times \ldots \times \Delta_{k_n}$ ist,*
3. *für jedes Funktionssymbol $f \in \mathcal{F}$ mit $\operatorname{ar}(f) = k_1 \times \ldots \times k_n \rightarrow k$ $f^{\mathcal{A}}$ eine Abbildung von $\Delta_{k_1} \times \ldots \times \Delta_{k_n}$ nach Δ_k ist, und*
4. *für ein Konstantensymbol $c \in \mathcal{K}$ mit $\operatorname{ar}(c) = k$ $c^{\mathcal{A}} \in \Delta_k$ ist.*

Eine $\mathcal{A}$-Belegung w ist schließlich eine Abbildung, die jeder Variablen der Sorte k ein Objekt aus Δ_k zuordnet.

Die Semantikdefinition egibt sich dann kanonisch. Auch all die verschiedenen Resultate zur Prädikatenlogik lassen sich ohne Schwierigkeiten auf die mehrsortige Prädikatenlogik übertragen.
(Für eine ausführliche Darstellung einer mehrsortigen Prädikatenlogik siehe das Vorlesungsskript „Logische Systeme der Informatik" von Herrn Thiele oder das Buch „Logic for Computer Science" von J. Gallier, erschienen bei Harper und Row, New York 1986)

Kapitel 3

Resolutionskalkül und Logische Programmierung

Im Aufbau der Prädikatenlogik hatten wir als die zentralen Beweiskalküle den Tableaukalkül und den (nichtklausalen) Resolutionskalkül eingeführt. In der Weise, wie dies geschehen ist, sind diese Beweiskalküle jedoch nur für das manuelle Beweisen brauchbar. Wir wollen als Nächstes den Resolutionskalkül in einer Form präsentieren, in der er mit entsprechenden Verfeinerungen für das automatische Beweisen geeignet ist und in der er die Grundlage für das Logische Programmieren bildet. Insbesondere der zweite Punkt, das Logische Programmieren, d.h. die Verwendung der Prädikatenlogik als Programmiersprache, spielt eine ganz wichtige Rolle im Bereich der Künstlichen Intelligenz und auch für einige Datenbanktheorien.

Wir legen im Folgenden wieder die Prädikatenlogik ohne Identität zugrunde.

3.1 Die prädikatenlogische Resolution

Wir formulieren zunächst als ein spezielles Verfahren die Grundresolution, um so später die Vollständigkeit der prädikatenlogischen Resolution einfach beweisen zu können.

3.1.1 Grundresolution

Der folgende Abschnitt fasst bereits bekannte Resultate zum Resolutionskalkül und zum Satz von Herbrand zusammen zu dem Verfahren der Grundresolution. Ausgangspunkt für unsere Überlegungen ist dabei das Korollar 2.140 zum Satz von Herbrand. (Mit entsprechenden Anpassungen und Ergänzugen hätten wir für diese Betrachtungen auch direkt die Korrektheit und Vollständigkeit des Resolutionskalküls heranziehen können.)
Nach diesem Korollar ist eine Aussage α genau dann unerfüllbar, wenn es zu einer Erfüllbarkeitsfunktionalform $\beta \in F^s(\alpha)$ einen Herbrand-Bereich D gibt, so dass $E(\beta, D)$ unerfüllbar ist.
Wir übernehmen die Begriffe Literal, Klausel und konjunktive Normalform aus dem aussagenlogischen Teil (vgl. die Definitionen 1.35 und 1.36):

Definition 3.1 (*Literal*, *Klausel*, *konjunktive Normalform*) *Sei S eine Signatur und sei $\alpha \in Fm_{PL}(S)$.*

1. *Ist α eine atomare Formel oder ist $\alpha = \neg\beta$ für eine atomare Formel $\beta \neq \top$ und $\beta \neq \bot$, so ist α ein* Literal.

2. *Eine* Klausel *ist eine verallgemeinerte Disjunktion $[\alpha_1, \ldots, \alpha_n]$ von Literalen $\alpha_1, \ldots, \alpha_n$.*

3. *Eine* konjunktive Normalform *ist eine verallgemeinerte Konjunktion $\langle\beta_1, \ldots, \beta_k\rangle$ von Klauseln $\beta_1, \ldots, \beta_k$.*

Definition 3.2 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält.*

1. *Für eine quantorenfreie Formel $\alpha \in Fm_{PL}(S)$ mit $fvar(\alpha) \subseteq \{x_1, \ldots, x_n\}$ und Grundterme $t_1, \ldots, t_n$ (aus dem Herbrand-Universum) heißt $\alpha\{x_1/t_1, \ldots, x_n/t_n\}$* Herbrand-Instanz *von α.*

   Ist α speziell eine Klausel, so wird eine Herbrand-Instanz von α auch als Grundklausel *bezeichnet.*

2. *α sei eine Aussage in Erfüllbarkeitsfunktionalform. β sei diejenige Formel, die aus α dadurch resultiert, dass alle Quantoren (die nur noch universal sind) gestrichen werden. Jede Herbrand-Instanz von β ist dann auch eine Herbrand-Instanz von α.*

Sei α eine Aussage und $\alpha' \in F^s(\alpha)$ eine Erfüllbarkeitsfunktionalform von α. Mit dem Theorem von Skolem (Theorem 2.128) folgt, dass α genau dann erfüllbar ist, wenn α' erfüllbar ist. Zu α' existiert dann nach Theorem 2.134 eine semantisch äquivalente pränexe Normalform der Gestalt $(\forall x_1)\ldots(\forall x_n)\,\beta$ mit einem Kern β in konjunktiver Normalform. Für Fragen der Erfüllbarkeit bzw. Unerfüllbarkeit einer Aussage können wir uns somit auf Erfüllbarkeitsfunktionalformen der Gestalt $(\forall x_1)\ldots(\forall x_n)\,\beta$ mit einem Kern β in konjunktiver Normalform beschränken.

Theorem 3.3 (Grundresolution) *Eine Aussage $(\forall x_1)\ldots(\forall x_n)\,\alpha$ in Erfüllbarkeitsfunktionalform und mit einem quantorenfreien Kern $\alpha = \langle \alpha_1, \ldots, \alpha_k \rangle$ in konjunktiver Normalform ist unerfüllbar $\Leftrightarrow$ es existiert eine Folge $K_1, \ldots, K_j$ von Grundklauseln mit*

- *$K_j = [\,]$ und*
- *für $i = 1, \ldots, j$ gilt:*
  entweder ist K_i Herbrand-Instanz einer Klausel α_l für ein $l \in \{1, \ldots, k\}$ oder
  $K_i = R(K_r, K_s)$ für $1 \leq r < s < i$ oder $K_i = R(K_s)$ für $1 \leq s < i$.

Beweis. Es gilt:
$(\forall x_1)\ldots(\forall x_n)\,\alpha$ ist unerfüllbar

$\Leftrightarrow$ es existiert ein Herbrand-Bereich D, so dass $E((\forall x_1)\ldots(\forall x_n)\,\alpha, D)$ unerfüllbar ist (nach Satz von Herbrand).

$\Leftrightarrow$ es gibt für ein geeignetes $m \geq 1$ Herbrand-Instanzen $\beta_1, \ldots, \beta_m$ von α, so dass $\{\beta_1, \ldots, \beta_m\}$ unerfüllbar ist. Sei $\beta_i = \langle \alpha_{1,i}, \ldots, \alpha_{k,i} \rangle$, wobei die Klausel $\alpha_{j,i}$ eine Herbrand-Instanz der Klausel α_j ist.

$\Leftrightarrow$ $\{\alpha_{1,1}, \ldots, \alpha_{k,1}, \ldots, \alpha_{1,m}, \ldots, \alpha_{k,m}\}$ ist unerfüllbar.

$\Leftrightarrow$ aus der Startexpansion für $\{\alpha_{1,1}, \ldots, \alpha_{k,1}, \ldots, \alpha_{1,m}, \ldots, \alpha_{k,m}\}$ ist unter alleiniger Verwendung der Resolventenbildung die leere Klausel $[\,]$ herleitbar.

$\Leftrightarrow$ es existiert eine Folge $K_1, \ldots, K_j$ von Klauseln mit

- $K_j = [\,]$ und
- für $i = 1, \ldots, j$ gilt:
  entweder ist K_i Herbrand-Instanz einer Klausel α_l für ein $l \in \{1, \ldots, k\}$ oder $K_i = R(K_r, K_s)$ für $1 \leq r < s < i$ oder $K_i = R(K_s)$ für $1 \leq s < i$.

∎

Beispiel 3.4

1. *Gegeben sei die folgende Aussage* $\alpha := (\forall x)\,(P\,(x) \wedge \neg P\,(f\,(x)))$*, die eine Erfüllbarkeitsfunktionalform von* $(\forall x)\,(\exists y)\,(P\,(x) \wedge \neg P\,(y))$ *darstellt.* α *ist offensichtlich unerfüllbar. Sei ferner der Herbrand-Bereich* $D = \{c, f\,(c)\}$ *gewählt. Die Herbrand-Expansion von* α *über dem Herbrand-Bereich* D *ist dann*
$$P\,(c) \wedge \neg P\,(f\,(c)) \wedge P\,(f\,(c)) \wedge \neg P\,(f\,(f\,(c)))\,,$$
oder in der Schreibweise mittels verallgemeinerter Konjunktionen und Disjunktionen
$$\langle [P\,(c)]\,, [\neg P\,(f\,(c))]\,, [P\,(f\,(c))]\,, [\neg P\,(f\,(f\,(c)))]\rangle\,.$$
Es ist ersichtlich, dass nach mehreren Resolutionsexpansionen u.a. die beiden Klauseln $[\neg P\,(f\,(c))]$ *und* $[P\,(f\,(c))]$ *vorkommen, so dass nach einem weiteren Schritt über eine Resolventenbildung die leere Klausel* $[\,]$ *erhalten wird.*
Nach dem Verfahren der Grundresolution können wir so vorgehen, dass wir zu den im Kern
$$\langle [P\,(x)]\,, [\neg P\,(f\,(x))]\rangle$$
von α *vorkommenden Klauseln durch geeignete Substitutionen Grundklauseln bilden und dann Resolventenbildungen vornehmen:*
$$\begin{tikzcd}
{[P\,(x)]} \arrow[d, "{\{x/f\,(c)\}}"'] & & {[\neg P\,(f\,(x))]} \arrow[d, "{\{x/c\}}"] \\
{[P\,(f\,(c))]} \arrow[dr] & & {[\neg P\,(f\,(c))]} \arrow[dl] \\
& {[\,]} &
\end{tikzcd}$$

2. *Sei* $\alpha = (\forall x)\,(\forall y)\,((\neg P\,(x) \vee \neg P\,(f\,(c)) \vee Q\,(y)) \wedge (P\,(y)) \wedge (\neg P\,(g\,(d,x)) \vee \neg Q\,(d)))\,.$
Der Kern von α *ist in konjunktiver Normalform und besteht aus den folgenden Klauseln*
$$[\neg P\,(x)\,, \neg P\,(f\,(c))\,, Q\,(y)]\,,\ [P\,(y)]\,,\ [\neg P\,(g\,(d,x))\,, \neg Q\,(d)]\,.$$
Mit dem Verfahren der Grundresolution erhalten wir
$$\begin{tikzcd}
{[\neg P\,(x)\,, \neg P\,(f\,(c))\,, Q\,(y)]} \arrow[d, "{\{x/f\,(c)\,, y/d\}}"'] & & & {[P\,(y)]} \arrow[dl, "{\{y/f\,(c)\}}"'] \arrow[dr, "{\{y/g\,(d,c)\}}"] & & {[\neg P\,(g\,(d,x))\,, \neg Q\,(d)]} \arrow[d, "{\{x/c\}}"] \\
{[\neg P\,(f\,(c))\,, Q\,(d)]} \arrow[dr] & & {[P\,(f\,(c))]} \arrow[dl] & & {[\neg P\,(g\,(d,c))]} \arrow[dr] & {[\neg P\,(g\,(d,c))\,, \neg Q\,(d)]} \arrow[d] \\
& {[Q\,(d)]} \arrow[drr] & & & & {[\neg Q\,(d)]} \arrow[dll] \\
& & & {[\,]} & &
\end{tikzcd}$$

Das Theorem über die Grundresolution besagt somit, dass wir die Unerfüllbarkeit einer Aussage
$$(\forall x_1) \ldots (\forall x_n)\,\alpha$$

in Erfüllbarkeitsfunktionalform und mit einem quantorenfreien Kern alleine durch die Bildung von Grundklauseln und unter Verwendung der Resolventenbildung zeigen können.
Ein wesentlicher Nachteil dieser als „Grundresolution" bezeichneten Methode ist, dass ein systematisches, mit erträglichem Aufwand arbeitendes Verfahren zur Ezeugung von Herbrand-Instanzen, die im Falle einer unerfüllbaren Aussage schließlich zu einer Resolutionsherleitung von $[\,]$ führen, nur schwer realisierbar zu sein scheint. Die Substitutionen zur Erzeugung von Herbrand-Instanzen müssen ja gewissermaßen „vorausschauend" gemacht werden. Auch ist der Aufwand noch unnötig groß im Vergleich zu dem folgenden Ansatz.
Dieser Ansatz, in dem Substitutionen zur Bildung von Resolventen zurückhaltender durchgeführt werden, geht auf J. A. Robinson zurück und führt uns zu einem neuen, klausalen Resolutionskalkül, der schließlich auch die Basis der logischen Programmierung bildet.

3.1.2 Unifikation

Im Verfahren der Grundresolution konnte eine (nichttriviale) Resolventenbildung dann vorgenommen werden, wenn in einer Grundklausel eine atomare Formel vorhanden war, die in einer anderen Grundklausel negiert vorkam. Grundklauseln hatten wir dabei durch Substitutionen von Grundtermen erhalten. Mit der Methode der Unifikation wird nun versucht, diese „Anpassung" von atomaren Formeln durch möglichst allgemein gehaltene Substitutionen, also nicht durch Substitution von Grundtermen, so zu erreichen, dass für die zugehörigen resultierenden Klauseln eine Resolventenbildung ermöglicht wird.

Notation 3.5 *Für eine Menge M von Literalen und für eine Substitution σ sei $M\sigma := \{l\sigma \mid l \in M\}$.*

Definition 3.6 *Sei M eine endliche Menge von Literalen (einer Signatur S).*
Eine Substitution σ heißt

1. Unifikator *für M genau dann, wenn $\|M\sigma\| = 1$, d.h. wenn $M\sigma$ einelementig ist.*
2. allgemeinster Unifikator *für M genau dann, wenn*
   (a) σ Unifikator für M ist, und
   (b) wenn gilt: ist θ ein weiterer Unifikator für M ist, so existiert eine Substitution σ' mit $\theta = \sigma\sigma'$.
3. Variablenumbenennung *genau dann, wenn $\sigma(V) \subseteq V$ und σ injektiv ist.*

Bemerkung 3.7 *Der allgemeinste Unifikator einer endlichen Menge von Literalen ist bis auf Variablenumbenennung eindeutig bestimmt.*

Beweis. Übungsaufgabe. ∎

Beispiel 3.8

1. *Für die Literalmenge $\{P(g(x,c),y), P(g(f(d),z),f(z))\}$ ist $\{x/f(d), y/f(c), z/c\}$ ein Unifikator, da*
$$\begin{aligned}[P(g(x,c),y)]\{x/f(d), y/f(c), z/c\} &= P(g(f(d),c),f(c)) \\ &= [P(g(f(d),z),f(z))]\{x/f(d), y/f(c), z/c\}.\end{aligned}$$
   In diesem Fall ist der Unifikator $\{x/f(d), y/f(c), z/c\}$ bereits auch der allgemeinste Unifikator.
2. *Die Literalmenge $\{P(x), P(f(x))\}$ besitzt offensichtlich keinen Unifikator.*

Theorem 3.9 (Unifikationstheorem) *M sei eine endliche Menge von Literalen (einer Signatur S). Besitzt M einen Unifikator, so besitzt M auch einen allgemeinsten Unifikator.*

Beweis. Konstruktiv unter Verwendung des folgenden Unifikationsalgorithmus, der zu der Menge M einen allgemeinsten Unifikator bestimmt, sofern einer existiert.
Unifikationsalgorithmus:

EINGABE**:** Eine nichtleere, endliche und geordnete Menge M von Literalen.

$\sigma := \{\,\}$ (d.h. σ ist die identische Substitution, die jede Variable auf sich selbst abbildet)

WHILE $\|M\sigma\| > 1$ DO

BEGIN

Suche im ersten Paar (l_1, l_2) von Literalen aus $M\sigma$ die erste Stelle, wo sich l_1 und l_2 unterscheiden.

IF keines der beiden an dieser ersten Stelle stehenden Symbole ist eine Variable

THEN stop „nicht unifizierbar"

ELSE BEGIN

Sei an dieser ersten Unterscheidungsstelle x die Variable in dem einen Literal und t der Term im anderen Literal.

IF x kommt in t vor

THEN stop „nicht unifizierbar"

ELSE $\sigma' := \{x/t\}$; $\quad \sigma := \sigma\sigma'$

END

END;

Gib σ als allgemeinsten Unifikator aus.

Wir zeigen die Korrektheit des Unifikationsalgorithmus.

Zunächst gilt, dass der Unifikationsalgorithmus terminiert, da nach jedem Schleifendurchlauf die Anzahl der in $M\sigma$ vorkommenden verschiedenen Variablen um eins vermindert wurde. Es kann daher höchstens so viele Schleifendurchläufe geben, wie zu Beginn in M verschiedene Variablen vorhanden sind.
Es ist ferner offensichtlich, dass M nicht unifizierbar ist, wenn der Algorithmus mit der Meldung „nicht unifizierbar" stoppt.
Gibt der Algorithmus eine Substitution σ aus, so ist wegen $\|M\sigma\| = 1$ klar, dass σ ein Unifikator ist. Es muss also nur noch gezeigt werden, dass σ auch allgemeinster Unifikator ist.
Sei n die Anzahl der Schleifendurchläufe. Sei $\sigma_0 = \{\,\}$ und sei σ_i, $0 < i \leq n$, die im i-ten Schleifendurchlauf erhaltene Substitution σ', so dass $\sigma = \sigma_0\sigma_1 \ldots \sigma_n$ ist. Sei θ ein beliebiger Unifikator für M. Wir zeigen $\theta = \sigma\theta$. Hieraus folgt, dass σ ein allgemeinster Unifikator ist.
Wir zeigen hierzu durch Induktion nach i, dass für alle $i \leq n$ $\theta = \sigma_0 \ldots \sigma_i\theta$ gilt. Speziell für $i = n$ bedeutet dies, dass $\theta = \sigma\theta$ gilt.

$i = 0:$ Da σ_0 die identische Substitution ist, gilt $\sigma_0\theta = \theta$.

$i \to i+1:$ Sei $\sigma_{i+1} = \{x/t\}$, d.h. im $(i+1)$-ten Schleifendurchlauf ist $\sigma' = \{x/t\}$. Es gibt damit in $M(\sigma_0 \ldots \sigma_i)$ zwei Literale $l_1(\sigma_0 \ldots \sigma_i)$ und $l_2(\sigma_0 \ldots \sigma_i)$, so dass an der ersten Stelle, an der sie sich unterscheiden, x bzw. t vorkommt.

Für $y \neq x$ gilt $y\theta = y(\sigma_{i+1}\theta)$.

Im Falle $y = x$ gilt ebenfalls $y\theta = y(\sigma_{i+1}\theta)$:

Nach Voraussetzung ist θ ein Unifikator von M und nach Induktionsvoraussetzung gilt $\theta = \sigma_0 \ldots \sigma_i\theta$. Daher ist $\|M(\sigma_0 \ldots \sigma_i\theta)\| = \|M\theta\| = 1$. Folglich ist $l_1((\sigma_0 \ldots \sigma_i)\theta) = l_2((\sigma_0 \ldots \sigma_i)\theta)$. Insbesondere muss dann aber $x\theta = t\theta$ gelten. Es ist aber $t\theta = (x\sigma_{i+1})\theta = x(\sigma_{i+1}\theta)$. Es gilt also $x\theta = x(\sigma_{i+1}\theta)$.

Damit folgt dann $\theta = \sigma_{i+1}\theta$ und somit schließlich $\theta = \sigma_0 \ldots \sigma_i\theta = \sigma_0 \ldots \sigma_i\sigma_{i+1}\theta$. Dies schließt den Beweis der Induktionsbehauptung ab.

■

Beispiel 3.10 *Zum Unifikationsalgorithmus.*
Gegeben sei die folgende Menge $M = \{P(f(z, g(c, y)), h(z)), P(f(f(u, v), w), h(f(c, d)))\}$ *von Literalen.*

1. unterscheidende Stelle:

`P(f(z,g(c,y)),h(z))`

`P(f(f(u,v),w),h(f(c,d)))`

↑

führt zu $\sigma' = \{z/f(u, v)\}$ *und* $\sigma = \{z/f(u, v)\}$.

Nach Durchführung der Substitution σ' *erhalten wir die*

nächste unterscheidende Stelle:

`P(f(f(u,v),g(c,y)),h(f(u,v)))`

`P(f(f(u,v),w),h(f(c,d)))`

↑

Folglich ist $\sigma' = \{w/g(c, y)\}$ *und* $\sigma = \{z/f(u, v), w/g(c, y)\}$.

Anwenden der Substitution σ' *liefert die*

nächste unterscheidende Stelle:

`P(f(f(u,v),g(c,y)),h(f(u,v)))`

`P(f(f(u,v),g(c,y)),h(f(c,d)))`

↑

Damit ist $\sigma' = \{u/c\}$ *und* $\sigma = \{z/f(c, v), w/g(c, y), u/c\}$.

Wenden wir wiederum die Substitution σ' *an, so erhalten wir die*

nächste unterscheidende Stelle:

`P(f(f(c,v),g(c,y)),h(f(c,v)))`

`P(f(f(c,v),g(c,y)),h(f(c,d)))`

↑

Es ist dann $\sigma' = \{v/d\}$ *und* $\sigma = \{z/f(c, d), w/g(c, y), u/c, v/d\}$.

Nach Durchführung der Substitution σ' *erhalten wir schließlich*

$$M\sigma = \{P(f(f(c, d), g(c, y)), h(f(c, d)))\}.$$

Da $M\sigma$ *einelementig ist, haben wir mit* $\sigma = \{z/f(c, d), w/g(c, y), u/c, v/d\}$ *den allgemeinsten Unifikator gefunden.*

3.1.3 Prädikatenlogischer klausaler Resolutionskalkül

Einen klausalen Resolutionskalkül hätten wir in Analogie zur Vorgehensweise in der Aussagenlogik einführen können. Wir wählen hier jedoch einen allgemeineren und insbesondere effizienteren Ansatz. Unter Verwendung der Unifikation definieren wir eine allgemeinere Form der Resolventenbildung, die dann die Grundlage für den nachfolgenden prädikatenlogischen klausalen Resolutionskalkül bilden wird.
Wir gehen im Folgenden jeweils von einer festen Signatur S aus.

Notation 3.11 *Ein Literal l wird als* echtes Literal *bezeichnet, falls $l \neq \top$ und $l \neq \bot$.*
Für ein echtes Literal l bezeichne

$$\bar{l} := \begin{cases} \neg\beta & \text{falls } l = \beta \text{ und } \beta \text{ atomare Formel} \\ \beta & \text{falls } l = \neg\beta. \end{cases}$$

$\bar{l}$ wird auch als duales Literal *von l bezeichnet.*

Definition 3.12 *Seien α_1, α_2 und β Klauseln.*
β heißt prädikatenlogische Resolvente *von α_1 und α_2 (notiert als $\beta = \mathrm{Res}(\alpha_1, \alpha_2)$) genau dann, wenn gilt:*

1. *es gibt Variablenumbenennungen σ_1 und σ_2, so dass $\alpha_1\sigma_1$ und $\alpha_2\sigma_2$ keine gemeinsamen Variablen enthalten.*
2. *es existieren echte Literale $l_1, \ldots, l_m$ in $\alpha_1\sigma_1$, $m \geq 1$, und $l'_1, \ldots, l'_n$ in $\alpha_2\sigma_2$, $n \geq 1$ sowie eine Substitution θ, so dass $M = \left\{\overline{l_1}, \ldots, \overline{l_m}, l'_1, \ldots, l'_n\right\}$ mit θ als allgemeinstem Unifikator unifizierbar ist zu $M\theta = \{l\}$.*
3. $\beta = ((\alpha_1\sigma_1 - [l_1, \ldots, l_m]) \sqcup (\alpha_2\sigma_2 - [l'_1, \ldots, l'_n]))\,\theta$

   (oder in anderer, äquivalenter Darstellung: $\beta = \left(\alpha_1\sigma_1\theta - \left[\bar{l}\right]\right) \sqcup (\alpha_2\sigma_2\theta - [l])$)

β heißt triviale prädikatenlogische Resolvente *von α_1 genau dann, wenn $\bot$ als Literal in α_1 vorkommt und*

$$\beta = \alpha_1 - [\bot].$$

(notiert als $\beta = \mathrm{Res}(\alpha_1)$)

Sprechweise*: β_1 ist aus α_1 und α_2 (bzw. aus α_1) mittels Resolution (oder unter Anwendung der Resolutionsregel) erhalten worden.*

Beispiel 3.13 *Gegeben sind die beiden Klauseln $\alpha = [\neg Q(z), P(f(x)), P(z)]$ und*
$\beta = [R(g(x), c), \neg P(x)]$. Wir wollen eine Resolvente zu diesen beiden Klauseln bestimmen.
Da in beiden Klauseln die Variable x vorkommt, führen wir für die beiden Klauseln eine Variablenumbenennung durch mit $\sigma_\alpha = \{\,\}$ und $\sigma_\beta = \{x/u\}$.
Es sei $\alpha' = \alpha\sigma_\alpha$ und $\beta' = \beta\sigma_\beta$. Somit ist $\alpha' = [\neg Q(z), P(f(x)), P(z)]$ und $\beta' = [R(g(u), c), \neg P(u)]$.
Die Menge $= \{P(f(x)), P(z), P(u)\}$ ist unifizierbar mit $\theta = \{u/f(x), z/f(x)\}$ als allgemeinstem Unifikator.
Damit ist $\alpha'\theta = [\neg Q(f(x)), P(f(x))]$ und $\beta'\theta = [R(g(f(x)), c), \neg P(f(x))]$. Resolvente von α und β ist dann

$$\mathrm{Res}(\alpha, \beta) = [\neg Q(f(x)), R(g(f(x)), c)].$$

Lemma 3.14 *Seien $\alpha_1, \alpha_2, \beta$ Klauseln zur Signatur S mit $\beta = \mathrm{Res}(\alpha_1, \alpha_2)$ (bzw. $\beta = \mathrm{Res}(\alpha_1)$).*
Es gilt dann für eine beliebige S-Struktur $\mathcal{A}$:
α_1, α_2 sind allgemeingültig in $\mathcal{A}$ $\Leftrightarrow$ $\alpha_1, \alpha_2, \beta$ sind allgemeingültig in $\mathcal{A}$.

Beweis. Die Beweisrichtung „$\Leftarrow$" ist trivialerweise erfüllt. Wir zeigen daher nur die Richtung „$\Rightarrow$":
Es gelte $\mathcal{A} \models_{PL} \alpha_i$ für $i = 1, 2$. Wir müssen $\mathcal{A} \models_{PL} \beta$ zeigen.
Da $\beta = \mathrm{Res}\,(\alpha_1, \alpha_2)$, existieren Variablenumbenennungen σ_1, σ_2 und ein allgemeinster Unifikator θ sowie ein Literal l mit $\beta = \left(\alpha_1\sigma_1\theta - \left[\bar{l}\right]\right) \sqcup (\alpha_2\sigma_2\theta - [l])$.
Da $\mathcal{A} \models_{PL} \alpha_i$, folgt sofort mit dem Substitutionslemma $\mathcal{A} \models_{PL} \alpha_i\sigma_i\theta$ für $i = 1, 2$.
Sei w eine beliebige $\mathcal{A}$-Belegung. Es gilt dann $l^{\mathcal{A},w} = T$ oder $\bar{l}^{\mathcal{A},w} = T$.
Wir betrachten den Fall $l^{\mathcal{A},w} = T$. Der andere Fall ist analog zu behandeln.
Es ist somit $\bar{l}^{\mathcal{A},w} = F$. Wegen $\mathcal{A} \models_{PL} \alpha_1\sigma_1\theta$ existiert in der Klausel $\alpha_1\sigma_1\theta - \left[\bar{l}\right]$ ein Literal l_1 mit $l_1^{\mathcal{A},w} = T$.
Es folgt $\mathcal{A}, w \models_{PL} \left(\alpha_1\sigma_1\theta - \left[\bar{l}\right]\right) \sqcup (\alpha_2\sigma_2\theta - [l])$ und somit $\mathcal{A}, w \models_{PL} \beta$. Da w beliebig gewählt war, folgt $\mathcal{A} \models_{PL} \beta$. ■

In Analogie zum Begriff der Resolutionsexpansion im nichtklausalen Resolutionskalkül bzw. im klausalen Resolutionskalkül der Aussagenlogik führen wir den Begriff der prädikatenlogischen Resolutionsexpansion ein. Da nur Klauseln vorhanden sind, ist lediglich die Resolventenbildung durch die prädikatenlogische Resolventenbildung zu ersetzen.

Definition 3.15 *Sei Σ eine Menge von Klauseln und $\{\alpha_1, \ldots, \alpha_n\}$ eine endliche Menge von Klauseln. Der Begriff der* prädikatenlogischen Σ-Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$ *wird induktiv definiert:*

1. *$\alpha_1, \ldots, \alpha_n$ ist eine prädikatenlogische Σ-Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$ (die Startexpansion)*
2. *Ist*
$$\beta_1, \ldots, \beta_m$$
eine prädikatenlogische Σ-Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$ und ist $\beta \in \Sigma$ oder $\beta = \mathrm{Res}\,(\beta_i, \beta_j)$ oder $\beta = \mathrm{Res}\,(\beta_i)$, $1 \leq i \leq j \leq m$, so ist
$$\beta_1, \ldots, \beta_m, \beta$$
eine prädikatenlogische Σ-Resolutionsexpansion für $\{\alpha_1, \ldots, \alpha_n\}$.

Eine prädikatenlogische Σ-Resolutionsexpansion $\beta_1, \ldots, \beta_m$ für $\emptyset$ mit $m \geq 1$ und $\beta_m = \beta$ heißt prädikatenlogische Resolutionsherleitung von β aus Σ
(notiert als $\Sigma \vdash_R \beta$).

Als Vorbereitung für den Nachweis der Widerlegungsvollständigkeit zeigen wir, dass „aussagenlogische Resolventenbildungen" in gewisser Weise zu prädikatenlogischen Resolventenbildungen „geliftet" werden können mit dem Resultat, dass die Widerlegungsvollständigkeit auf die Widerlegungsvollständigkeit der Grundresolution (siehe Grundresolutionssatz) zurückgeführt werden kann.

Theorem 3.16 (Lifting-Lemma) *α_1, α_2 seien Klauseln, β_1 und β_2 seien Herbrand-Instanzen von α_1 bzw. α_2 und es sei $\beta = R\,(\beta_1, \beta_2)$ eine nichttriviale Resolvente von β_1 und β_2.*
Dann existiert eine nichttriviale prädikatenlogische Resolvente $\alpha = \mathrm{Res}\,(\alpha_1, \alpha_2)$, so dass β eine Herbrand-Instanz von α ist.

Graphisch veranschaulicht:

$$\begin{tikzcd}
\alpha_1 \arrow[d] & & \alpha_2 \arrow[d] & & \alpha_1 \arrow[dr] & & \alpha_2 \arrow[dl] \\
\beta_1 \arrow[dr] & & \beta_2 \arrow[dl] & \Longrightarrow & & \alpha \arrow[d] & \\
& \beta & & & & \beta &
\end{tikzcd}$$

Beweis. Es seien θ_1 und θ_2 Variablenumbenennungen, so dass $var\,(\alpha_1\theta_1) \cap var\,(\alpha_2\theta_2) = \emptyset$.
β_1 und β_2 sind Herbrand-Instanzen von α_1 bzw. α_2. Daher sind β_1 und β_2 auch Herbrand-Instanzen von $\alpha_1\theta_1$ bzw. $\alpha_2\theta_2$.
Seien σ_1, σ_2 Substitutionen mit $\beta_i = \alpha_i\theta_i\sigma_i$, $i = 1, 2$. O.B.d.A. gilt für $x \notin var\,(\alpha_i\theta_i)$ $\sigma_i\,(x) = x$, $i = 1, 2$.
Wir setzen $\sigma := \sigma_1\sigma_2$.

Es gilt dann für $i = 1, 2$ $\beta_i = \alpha_i\theta_i\sigma$, da in β_1 und β_2 keine freien Variablen vorkommen und die Variablenmengen in $\alpha_1\theta_1$und $\alpha_2\theta_2$ disjunkt sind. Nach Voraussetzung existiert ein Literal l in β_1, so dass $\overline{l}$ in β_2 vorkommt und $\beta = (\beta_1 - [l]) \sqcup \left(\beta_2 - \left[\,\overline{l}\,\right]\right)$. Es seien $l_{11}, \ldots, l_{1m}$, $m \geq 1$, die Literale in $\alpha_1\theta_1$ mit $l_{11}\sigma = \ldots = l_{1m}\sigma = l$ und $l_{21}, \ldots, l_{2n}$, $n \geq 1$, die Literale in $\alpha_2\theta_2$ mit $l_{21}\sigma = \ldots = l_{2n}\sigma = \overline{l}$. (Es gilt also $\overline{l_{2i}}\sigma = l$, $1 \leq i \leq n$.)
Da σ ein Unifikator für die Literalmenge $M = \left\{l_{11}, \ldots, l_{1m}, \overline{l_{21}}, \ldots, \overline{l_{2n}}\right\}$ ist, existiert nach dem Unifikationstheorem ein allgemeinster Unifikator θ für M. Sei $M\theta = \{l_\theta\}$. Nach Definition ist daher $\alpha = (\alpha_1\theta_1\theta - [l_\theta]) \sqcup \left(\alpha_2\theta_2\theta - \left[\,\overline{l_\theta}\,\right]\right)$ eine prädikatenlogische Resolvente von α_1 und α_2.
Da θ ein allgemeinster Unifikator für M und σ ein Unifikator für M ist, existiert eine Substitution σ' mit $\theta\sigma' = \sigma$. Damit folgt

$$
\begin{aligned}
\beta &= (\beta_1 - [l]) \sqcup \left(\beta_2 - \left[\,\overline{l}\,\right]\right) \\
&= (\alpha_1\theta_1\sigma - [l]) \sqcup \left(\alpha_2\theta_2\sigma - \left[\,\overline{l}\,\right]\right) \\
&= ((\alpha_1\theta_1 - [l_{11}, \ldots, l_{1m}]) \sqcup (\alpha_2\theta_2 - [l_{21}, \ldots, l_{2n}]))\,\sigma \\
&= ((\alpha_1\theta_1 - [l_{11}, \ldots, l_{1m}]) \sqcup (\alpha_2\theta_2 - [l_{21}, \ldots, l_{2n}]))\,\theta\sigma' \\
&= \left((\alpha_1\theta_1\theta - [l_\theta]) \sqcup \left(\alpha_2\theta_2\theta - \left[\,\overline{l_\theta}\,\right]\right)\right)\sigma' \\
&= \alpha\sigma',
\end{aligned}
$$

d.h. β ist eine Herbrand-Instanz von α. ■

Theorem 3.17 (Resolutionstheorem der Prädikatenlogik) *Sei die Aussage*

$$\alpha = (\forall x_1) \ldots (\forall x_n)\, \beta$$

eine Erfüllbarkeitsfunktionalform mit einem Kern $\beta = \langle\alpha_1, \ldots, \alpha_k\rangle$ *in konjunktiver Normalform. Es gilt dann:*

$$\alpha \textit{ ist unerfüllbar genau dann, wenn } \{\alpha_1, \ldots, \alpha_k\} \vdash_R [\,].$$

Beweis.

„$\Leftarrow$": Es gelte $\{\alpha_1, \ldots, \alpha_k\} \vdash_R [\,]$.

Sei $\beta_1, \ldots, \beta_m$ mit $m \geq 1$ und $\beta_m = [\,]$ eine *prädikatenlogische Resolutionsherleitung von* $[\,]$ *aus* $\{\alpha_1, \ldots, \alpha_k\}$.

Angenommen, α ist erfüllbar. Da α eine Aussage ist, gibt es also eine Struktur $\mathcal{A}$ mit $\mathcal{A} \models_{PL} \alpha$. Damit sind dann auch β und folglich $\alpha_1, \ldots, \alpha_k$ allgemeingültig in $\mathcal{A}$.

Da $\beta_1 \in \{\alpha_1, \ldots, \alpha_k\}$, ist also β_1 allgemeingültig in $\mathcal{A}$. Durch Induktion und unter Verwendung von Lemma 3.14 folgt, dass alle β_i, $1 \leq i \leq m$, in $\mathcal{A}$ allgemeingültig sind, insbesondere auch die leere Klausel $[\,]$. Wir erhalten also einen Widerspruch. α ist folglich nicht erfüllbar.

„$\Rightarrow$": Sei α unerfüllbar.

Nach dem Theorem über die Grundresolution (Theorem 3.3) existiert eine Folge von Grundklauseln $\beta'_1, \ldots, \beta'_m$ mit $m \geq 1$ und $\beta'_m = [\,]$, so dass für $i = 1, 2, \ldots, m$ gilt: β'_i ist entweder Herbrand-Instanz einer Klausel aus $\{\alpha_1, \ldots, \alpha_k\}$ oder $\beta'_i = R\left(\beta'_j\right)$ oder $\beta'_i = R\left(\beta'_j, \beta'_l\right)$ mit $1 \leq j \leq l < i$. Für jedes $i = 1, \ldots, m$ definieren wir nun induktiv ein β_i so, dass $\beta_1, \ldots, \beta_i$ eine prädikatenlogische Resolutionsherleitung von β_i ist.

Ind.Anf.: β'_1 kann nur Herbrand-Instanz einer Klausel γ aus $\{\alpha_1, \ldots, \alpha_k\}$ sein. Setze $\beta_1 := \gamma$.

Ind.Vor.: Für alle $j < i$ sei β_j so definiert, dass β'_j eine Herbrand-Instanz von β_j und $\beta_1, \ldots, \beta_j$ eine prädikatenlogische Resolutionsherleitung von β_j ist.

Ind.Schritt: Wir unterscheiden nach den möglichen Fällen:

Falls β'_i eine Herbrand-Instanz einer Klausel $\gamma \in \{\alpha_1, \ldots, \alpha_k\}$ ist, so setze $\beta_i := \gamma$.

Falls $\beta'_i = R\left(\beta'_j\right)$ oder $\beta'_i = R\left(\beta'_j, \beta'_l\right)$ mit $1 \leq j \leq l < i$, so sind nach Induktionsvoraussetzung β_j bzw. β_j und β_l bereits definiert. Nach dem Lifting-Lemma (Lemma 3.16) gibt es dann eine prädikatenlogische Resolvente $\mathrm{Res}(\beta_j)$ bzw. $\mathrm{Res}(\beta_j, \beta_l)$, so dass β_i' Herbrand-Instanz dieser Resolvente ist. Wir setzen $\beta_i := \mathrm{Res}(\beta_j)$ bzw. $\beta_i := \mathrm{Res}(\beta_j, \beta_l)$. Damit ist dann β_i' Herbrand-Instanz von β_i und $\beta_1, \ldots, \beta_i$ eine prädikatenlogische Resolutionsherleitung von β_i.

Da insbesondere also $\beta_1, \ldots, \beta_m$ eine prädikatenlogische Resolutionsherleitung von β_m ist und β_m gleich der leeren Klausel sein muss, folgt somit $\{\alpha_1, \ldots, \alpha_k\} \vdash_R [\,]$. ∎

Beispiel 3.18 *Sei $\alpha = (\forall x)(\forall y)\,\beta$ eine Erfüllbarkeitsfunktionalform mit dem Kern*

$$\beta = \langle [\neg P(x), Q(x), R(x, f(x))], [\neg P(x), Q(x), S(f(x))], [T(c)], [P(c)], [\neg R(c, x), T(x)], [\neg T(x), \neg Q(x)], [\neg T(y), \neg S(y)] \rangle .$$

β ist offensichtlich in konjunktiver Normalform.
Wir zeigen, dass aus der Menge der in β vorkommenden Klauseln die leere Klausel herleitbar ist, so dass α also unerfüllbar ist.

1	$[T(c)]$	*in β*
2	$[\neg T(x), \neg Q(x)]$	*in β*
3	$[\neg Q(c)]$	*Resolvente aus 1,2*
4	$[\neg P(x), Q(x), S(f(x))]$	*in β*
5	$[\neg P(c), S(f(c))]$	*Resolvente aus 3,4*
6	$[P(c)]$	*in β*
7	$[S(f(c))]$	*Resolvente aus 5,6*
8	$[\neg P(x), Q(x), R(x, f(x))]$	*in β*
9	$[Q(c), R(c, f(c))]$	*Resolvente aus 6,8*
10	$[R(c, f(c))]$	*Resolvente aus 3,9*
11	$[\neg R(c, x), T(x)]$	*in β*
12	$[T(f(c))]$	*Resolvente aus 10,11*
13	$[\neg T(y), \neg S(y)]$	*in β*
14	$[\neg S(f(c))]$	*Resolvente aus 12,13*
15	$[\,]$	*Resolvente aus 7,14.*

3.2 Verfeinerung der Resolution

Die Verwendung der Unifikation im prädikatenlogischen klausalen Resolutionskalkül bringt für die Beweissuche gegenüber der Grundresolution eine erhebliche Reduzierung des Suchraumes mit sich. Nach wie vor sind aber die kombinatorischen Möglichkeiten für die Bildung von Resolutionsherleitungen immens groß, da es i.a. viele Paare von Klauseln gibt, zu denen eine Resolvente gebildet werden kann. Insbesondere können Resolventen länger als die Ausgangsklauseln werden, was das Potential für Resolventenbildungen erhöht. Man war daher schon frühzeitig bestrebt, den Resolutionskalkül so zu verfeinern, dass die Menge der Klauseln (und damit der Suchraum für Klauseln), die für eine Resolventenbildung in Frage kommen, eingeschränkt wird. Im folgenden sollen einige bedeutende Resolutionsrestriktionen vorgestellt werden.

Zur Vereinfachung der Darstellung und der Beweise sei angenommen, dass $\top$ und $\bot$ in Klauseln nicht vorkommen.

3.2.1 Lineare Resolution

Die Idee ist hier, die Resolutionsherleitung so einzuschränken, dass jeweils die zuletzt erhaltene Klausel im nächsten Schritt für eine Resolventenbildung verwendet wird.

Definition 3.19 *Sei Σ eine Menge von Klauseln und α eine Klausel mit $\alpha \notin \Sigma$.*

1. *Eine* lineare (prädikatenlogische) Resolutionsherleitung *von α aus Σ (notiert als $\Sigma \vdash_{LR} \alpha$) ist eine Folge*
$$(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n)$$
von Paaren von Klauseln α_i, β_i, $0 \leq i \leq n$, so dass
   (a) $\alpha_0 \in \Sigma$ *(Startklausel).*
   (b) $\beta_i \in \Sigma$ *oder* $\beta_i = \alpha_j$ *für ein j mit* $0 \leq j < i \leq n$.
   (c) $\alpha_{i+1} = \mathrm{Res}(\alpha_i, \beta_i), \quad 0 \leq i < n.$
   (d) $\alpha = \mathrm{Res}(\alpha_n, \beta_n)$

   α_i *heißt dann* Verbindungsklausel *(engl.: center clauses) und* β_i Seitenklausel.
   $L(\Sigma) := \{\alpha \mid \Sigma \vdash_{LR} \alpha\}.$

2. *Eine* lineare Resolutionswiderlegung *von Σ ist eine lineare Resolutionsherleitung von $[\,]$ aus Σ.*

Beispiel 3.20 *Sei* $\Sigma = \{[P(x), Q(x)], [P(x), \neg Q(x)], [\neg P(x), Q(x)], [\neg P(x), \neg Q(x)]\}.$
Eine lineare Resolutionsherleitung von $[\,]$ aus Σ ist dann z.B.

$$\begin{array}{ccc}
[P(x), Q(x)] & & [P(x), \neg Q(x)] \\
| & / & \\
[P(x)] & & [\neg P(x), Q(x)] \\
| & / & \\
[Q(x)] & & [\neg P(x), \neg Q(x)] \\
| & / & \\
[\neg P(x)] & & [P(x)] \\
| & / & \\
[\,] & &
\end{array}$$

Wir wollen als Nächstes für die lineare Resolution die Korrektheit und Vollständigkeit ebenfalls in der folgenden Form zeigen:
Ist die Aussage
$$\alpha = (\forall x_1) \ldots (\forall x_n)\, \beta$$
eine Erfüllbarkeitsfunktionalform mit einem Kern $\beta = \langle \alpha_1, \ldots, \alpha_k \rangle$ in konjunktiver Normalform, dann gilt:
$$\alpha \text{ ist unerfüllbar genau dann, wenn } \{\alpha_1, \ldots, \alpha_k\} \vdash_{LR} [\,].$$
Die Korrektheitsaussage „$\{\alpha_1, \ldots, \alpha_k\} \vdash_{LR} [\,] \Rightarrow \alpha$ ist unerfüllbar" ist offensichtlich, da die lineare Resolution eine Einschränkung der vollen Resolution darstellt. Wir müssen also nur noch die Vollständigkeit zeigen. Hierzu sind noch einige Vorbereitungen erforderlich.

Definition 3.21 *Sei Σ eine Klauselmenge und $\Delta \subseteq \Sigma$.*
Δ heißt Stützmenge *für Σ genau dann, wenn $\Sigma \setminus \Delta$ erfüllbar ist.*
Eine lineare Resolutionsherleitung $(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n)$ von α aus Σ besitzt die Stützmenge *Δ genau dann, wenn $\alpha_0 \in \Delta$.*

Die Intuition, die mit dem Begriff der Stützmenge verbunden ist, ist die, dass die Stützmenge den „eigentlichen Grund" für die Unerfüllbarkeit einer Formelmenge darstellt. Beispielsweise möchte man zeigen, dass aus einem Axiomensystem Σ eine bestimmte Aussage α folgt, so dass also $\Sigma \cup \{\neg\alpha\}$ unerfüllbar ist. Nimmt man also das Axiomensystem als gegeben an, so wird man den „Grund" für die Unerfüllbarkeit in der Formel $\neg\alpha$ lokalisieren wollen.

Definition 3.22 *Sei Σ eine Klauselmenge.*
Σ heißt minimal unerfüllbar *genau dann, wenn Σ unerfüllbar und jede echte Teilmenge von Σ erfüllbar ist.*

Lemma 3.23 *Sei Σ eine unerfüllbare Klauselmenge und Δ eine Stützmenge für Σ.*
Es gibt dann eine (bzgl. $\subseteq$) minimale unerfüllbare Teilmenge $\Sigma' \subseteq \Sigma$ und $\Delta \cap \Sigma'$ ist eine Stützmenge für Σ'.

Beweis. Sei Σ unerfüllbar. Σ ist somit ungleich der leeren Menge. Damit existiert eine bzgl. Mengeninklusion minimale Teilmenge Σ' von Σ'', die unerfüllbar ist.
Sei nun Δ eine Stützmenge für Σ. Wäre $\Delta \cap \Sigma' = \emptyset$, so würde $\Sigma' \subseteq \Sigma \setminus \Delta$ gelten. Somit wäre $\Sigma \setminus \Delta$ unerfüllbar, da Σ' unerfüllbar ist. Wir erhalten dann einen Widerspruch, da Δ eine Stützmenge für Σ und $\Sigma \setminus \Delta$ daher erfüllbar ist.
Also gilt $\Delta \cap \Sigma' \neq \emptyset$. Dann aber ist $\Sigma' \setminus (\Delta \cap \Sigma') \subsetneq \Sigma'$, somit erfüllbar aufgrund der Minimalitätseigenschaft von Σ'. Es folgt, dass $\Delta \cap \Sigma'$ Stützmenge für Σ' ist. ■

Notation 3.24 *Sei Σ eine Menge von Klauseln und l ein (echtes) Literal mit $var(\Sigma) = var(l) = \emptyset$.*
$\Sigma^l := \left\{\alpha - \left[\bar{l}\right] \mid \alpha \in \Sigma \text{ und } l \text{ kommt in } \alpha \text{ nicht vor}\right\}$.

Bemerkung 3.25

1. *Informell gesprochen ist Σ^l diejenige Klauselmenge, die aus Σ dadurch entsteht, dass der Wahrheitswert von l mit T fixiert wird, wobei dann zusätzliche Vereinfachungen vorgenommen werden. Klauseln, die aufgrund dieser Festlegung den Wahrheitswert T erhalten, werden gestrichen. In Klauseln, die $\bar{l}$ enthalten, wird $\bar{l}$ entfernt, da die resultierende Klausel semantisch äquivalent ist.*

2. *Ist Σ eine unerfüllbare Klauselmenge mit $var(\Sigma) = \emptyset$, so sind für ein variablenfreies Literal l Σ^l und $\Sigma^{\bar{l}}$ ebenfalls unerfüllbar.*

Beispiel 3.26 *Sei $\Sigma = \{[P(c), \neg R(d), R(c)], [\neg P(c), \neg R(c)], [P(d), R(c)]\}$.*
Dann ist $\Sigma^{P(c)} = \{[\neg R(c)], [P(d), R(c)]\}$ und $\Sigma^{\overline{P(c)}} = \{[\neg R(d), R(c)], [P(d), R(c)]\}$.

Theorem 3.27 (Vollständigkeit der linearen Resolution)
Sei die Aussage $(\forall x_1) \ldots (\forall x_n)\, \alpha$ eine Erfüllbarkeitsfunktionalform mit einem Kern $\alpha = \langle \alpha_1, \ldots, \alpha_m \rangle$ in konjunktiver Normalform.
Es gilt dann:
Ist $(\forall x_1) \ldots (\forall x_n)\, \alpha$ unerfüllbar, so gilt $\{\alpha_1, \ldots, \alpha_m\} \vdash_{LR} [\,]$.
Ferner:
Ist Δ eine Stützmenge für die unerfüllbare Klauselmenge $\{\alpha_1, \ldots, \alpha_m\}$, so gibt es eine lineare Resolutionswiderlegung von $\{\alpha_1, \ldots, \alpha_m\}$ mit Stützmenge Δ.

Beweis. Aus dem Beweis des Grundresolutionssatzes (Theorem 3.3) geht hervor, dass $(\forall x_1)\dots(\forall x_n)\,\alpha$ genau dann unerfüllbar ist, wenn es eine endliche Menge Σ von Herbrand-Instanzen von $\alpha_1,\dots,\alpha_m$ gibt, die unerfüllbar ist.
Insbesondere gibt es dann auch immer eine minimal unerfüllbare Menge Σ von solchen Herbrand-Instanzen. Sei $\beta \in \Sigma$ beliebig gewählt. Wir zeigen, dass es eine lineare Resolutionswiderlegung von Σ mit Stützmenge $\{\beta\}$ (d.h. mit Startklausel β) gibt.
Wir beweisen dies durch Induktion nach der Anzahl k der in Σ vorkommenden Literale (jedes Vorkommen wird also gezählt). Da wir $\top$ und $\bot$ ausgeschlossen haben, muss eine unerfüllbare Klauselmenge Σ mindestens zwei verschiedene Literale enthalten, so dass wir den Induktionsbeweis mit $k = 2$ starten.

Induktionsanfang $k = 2:$ Es ist dann $[l], \left[\,\bar{l}\,\right] \in \Sigma$ und $\beta = [l]$ für ein Literal l. Die Behauptung gilt offensichtlich.

Induktionsschritt $k \to k+1:$ Σ enthalte $k+1$ Literale, $k \geq 2$.

Fall 1: $\beta = [l]$ für ein Literal l.

Nach Bemerkung 3.25 2. ist Σ^l unerfüllbar. Σ^l enthält höchstens k Literale. Sei Σ' eine minimale unerfüllbare Teilmenge von Σ^l. Es gilt dann:

In Σ' existiert eine Klausel β', so dass $\beta' \sqcup \left[\,\bar{l}\,\right] \in \Sigma$. Würde dies nämlich nicht gelten, so wäre $\Sigma' \subseteq \Sigma \setminus \{\beta\}$. Da Σ' unerfüllbar ist, wäre Σ also nicht minimal unerfüllbar, im Widerspruch zur Voraussetzung, dass Σ minimal unerfüllbar ist.

Nach Induktionsvoraussetzung gibt es für Σ' eine lineare Resolutionswiderlegung $(\alpha_1, \beta_1), \dots, (\alpha_n, \beta_n)$ mit Stützmenge $\{\beta'\}$, d.h. $\alpha_1 = \beta'$. Es sei $\alpha_{n+1} := \mathrm{Res}\,(\alpha_n, \beta_n)$, also $\alpha_{n+1} = [\,]$.

Wir konstruieren aus dieser Resolutionswiderlegung für Σ' eine für Σ mit Stützmenge $\{\beta\}$ in $n+1$ Schritten.

Schritt $0:$

Setze $\gamma_0 := \beta\ \ (= [l])$, $\delta_0 := \beta' \sqcup \left[\,\bar{l}\,\right]$. Es gilt dann $\alpha_1 = \mathrm{Res}\,(\gamma_0, \delta_0) = \beta'$.

Es seien für $1 \leq i \leq n$ die Schritte $0, \dots, i-1$ bereits ausgeführt.

Schritt $i:$

Wir ersetzen in Abhängigkeit von den möglichen Fällen (α_i, β_i) durch (γ_i, δ_i) bzw. durch $(\gamma_i, \delta_i), (\gamma_i', \delta_i')$, so dass $\alpha_{i+1} = \mathrm{Res}\,(\gamma_i, \delta_i)$ bzw. $\alpha_{i+1} = \mathrm{Res}\,(\gamma_i', \delta_i')$.

$\gamma_i := \alpha_i$.

1. Falls $\beta_i = \alpha_j\ \ (= \gamma_j)$ für ein $j < i$, setze $\delta_i := \alpha_j$.
   Es gilt dann $\alpha_{i+1} = \mathrm{Res}\,(\alpha_i, \beta_i) = \mathrm{Res}\,(\gamma_i, \delta_i)$.
2. Falls $\beta_i \in \Sigma' \subseteq \Sigma^l$ unterscheiden wir 2 weitere Fälle:
   $\beta_i \in \Sigma:$ Setze $\delta_i := \beta_i$. Es gilt dann $\alpha_{i+1} = \mathrm{Res}\,(\alpha_i, \beta_i) = \mathrm{Res}\,(\gamma_i, \delta_i)$.

   $\beta_i \notin \Sigma:$
   In diesem Fall gilt $\beta_i \sqcup \left[\,\bar{l}\,\right] \in \Sigma$.
   Wir setzen
   $\delta_i := \beta_i \sqcup \left[\,\bar{l}\,\right]$,
   $\gamma_i' := \mathrm{Res}\,(\gamma_i, \delta_i) \quad \left(= \mathrm{Res}\left(\alpha_i, \beta_i \sqcup \left[\,\bar{l}\,\right]\right) = \alpha_{i+1} \sqcup \left[\,\bar{l}\,\right]\right)$
   $\delta_i' := \beta \quad (= [l])$.
   Es gilt dann für diesen Fall $\alpha_{i+1} = \mathrm{Res}\left(\alpha_{i+1} \sqcup \left[\,\bar{l}\,\right], [l]\right) = \mathrm{Res}\,(\gamma_i', \delta_i')$.

   Es ist leicht zu verifizieren, dass die so definierte Folge eine lineare Resolutionsherleitung aus Σ mit Stützmenge $\{\beta\}$ darstellt. Da speziell für $i = n$ $\alpha_{n+1} = \mathrm{Res}\,(\gamma_n, \delta_n)$ bzw. $\alpha_{n+1} = \mathrm{Res}\,(\gamma_n', \delta_n')$ gilt, und da $\alpha_{n+1} = [\,]$, folgt, dass diese lineare Resolutionsherleitung bereits eine Resolutionswiderlegung ist.

Fall 2: $\beta = [l, \ldots]$, d.h. β enthält mehr als ein Literal.

Setze $\beta' := \beta - [l]$. Es gilt, dass die Klauselmenge $\Sigma^{\bar{l}}$ unerfüllbar ist und dass $\beta' \in \Sigma^{\bar{l}}$.

$\Sigma^{\bar{l}} \setminus \{\beta'\}$ ist erfüllbar.

Begründung:

Da Σ minimal unerfüllbar ist, existiert eine Struktur $\mathcal{A}$ mit

$\mathcal{A} \models_{PL} \alpha$ für alle $\alpha \in \Sigma \setminus \{\beta\}$ und

$\beta^{\mathcal{A}} = F$.

Hieraus folgt $l^{\mathcal{A}} = F$, da $\beta = [l, \ldots]$.

Damit folgt weiter $\mathcal{A} \models_{PL} \alpha$ für alle $\alpha \in \Sigma^{\bar{l}} \setminus \{\beta'\}$, d.h. $\Sigma^{\bar{l}} \setminus \{\beta'\}$ ist erfüllbar.

Sei nun Σ' eine minimal unerfüllbare Teilmenge von $\Sigma^{\bar{l}}$.

Es ist dann $\beta' \in \Sigma'$, da $\Sigma^{\bar{l}} \setminus \{\beta'\}$ erfüllbar ist.

Die Anzahl der Literale in Σ' ist kleiner als die in Σ, also $\leq k$. Nach Induktionsvoraussetzung existiert damit eine lineare Resolutionswiderlegung $(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n)$ von Σ' mit Stützmenge $\{\beta'\}$.

Setzt man für $0 \leq i \leq n$

$\alpha_i' := \alpha_i \sqcup [l]$

$$\beta_i' := \begin{cases} \beta_i & \text{falls } \beta_i \in \Sigma \\ \beta_i \sqcup [l] & \text{sonst (d.h. } \beta_i \in \Sigma^{\bar{l}}, \beta_i \notin \Sigma), \end{cases}$$

dann folgt

$\alpha_0' = \beta$

$(\alpha_0', \beta_0'), \ldots, (\alpha_n', \beta_n')$ ist lineare Resolutionsherleitung aus Σ und mit Stützmenge $\{\beta\}$ und

$\alpha_{n+1}' := \mathrm{Res}(\alpha_n', \beta_n')$ ist gleich $[l]$.

Nun ist $(\Sigma \setminus \{\beta\} \cup \{[l]\})$ unerfüllbar. Denn wäre diese Klauselmenge erfüllbar durch eine Interpretation I, so würde diese Interpretation auch Σ erfüllen, da $\beta = [l, \ldots]$. Σ ist aber unerfüllbar.

Die Anzahl der Literale in $(\Sigma \setminus \{\beta\} \cup \{[l]\})$ ist kleiner als die von Σ, also $\leq k$, da β mindestens zwei Literale enthält. Nach Induktionsvoraussetzung existiert dann eine lineare Resolutionswiderlegung $(\alpha_0'', \beta_0''), \ldots, (\alpha_m'', \beta_m'')$ von $(\Sigma \setminus \{\beta\} \cup \{[l]\})$ mit Stützmenge $\{[l]\}$.

Damit ist dann aber $(\alpha_0', \beta_0'), \ldots, (\alpha_n', \beta_n'), (\alpha_0'', \beta_0''), \ldots, (\alpha_m'', \beta_m'')$ eine lineare Resolutionswiderlegung von Σ mit Stützmenge $\{\beta\}$.

Exakt wie im Beweis des prädikatenlogischen Resolutionstheorems kann schließlich mit Hilfe des Lifting-Lemmas zu dieser linearen Resolutionswiderlegung von Σ eine lineare (prädikatenlogische) Resolutionswiderlegung von $\{\alpha_1, \ldots, \alpha_m\}$ konstruiert werden. Damit ist dann die Behauptung gezeigt. ■

3.2.2 Input- und SLD-Resolution

Neben der linearen Resolution gibt es eine Vielzahl weiterer Resolutionsrestriktionen, die im Rahmen des automatischen Beweisens entwickelt worden sind und die ebenfalls zu vollständigen Beweisverfahren für die Prädikatenlogik führen. (Eine ausgezeichnete Darstellung solcher Verfahren findet sich in dem Buch von A. Leitsch: The Resolution Calculus, Springer Verlag 1997.)
Wir wollen als Nächstes zwei Resolutionsrestriktionen behandeln, die von zentraler Bedeutung für das Logische Programmieren sind: die Input-Resolution und die SLD-Resolution. Beide Resolutionsrestriktionen führen dazu, dass sie nicht mehr vollständig für die volle Prädikatenlogik sind, d.h. es gilt nicht, dass aus einer unerfüllbaren Klauselmenge die leere Klausel mittels dieser Resolutionsrestriktionen herleitbar ist. Für die wichtige Klasse der Hornformeln sind sie jedoch vollständig.

Lineare Input-Resolution

Definition 3.28

1. *Eine* (prädikatenlogische) lineare Input-Resolutionsherleitung *von α aus Σ (notiert als $\Sigma \vdash_{LIR} \alpha$) ist eine lineare Resolutionsherleitung von α aus Σ, in der jede Seitenklausel aus Σ ist.*

2. *Eine* lineare Input-Resolutionswiderlegung von Σ *ist eine lineare Input-Resolutionsherleitung von $[\,]$ aus Σ.*

Bemerkung 3.29 *Lineare Input-Resolution ist nicht vollständig. Zum Beispiel besitzt die unerfüllbare Klauselmenge*

$$\Sigma = \{[P(c), R(c)], [\neg P(c), R(c)], [P(c), \neg R(c)], [\neg P(c), \neg R(c)]\}$$

keine lineare Input-Resolutionswiderlegung.

Definition 3.30

1. *Ein Literal heißt*

   (a) positiv*, falls es eine atomare Formel ist, und*

   (b) negativ*, falls es eine Negation einer atomaren Formel ist.*

2. *Eine Klausel $\alpha = [l_1, \ldots, l_m]$ heißt*

   (a) definite Klausel*, falls genau ein Literal l_i, $1 \le i \le m$, positiv ist.*

   (b) negative Klausel*, falls $l_1, \ldots, l_m$ negativ sind.*

   (c) Horn-Klausel*, falls sie definit oder negativ ist.*

3. *Eine Formel α in pränexer Normalform*

$$\alpha = (Qx_1) \ldots (Qx_n)\, \beta,$$

   $\beta = \langle \alpha_1, \ldots, \alpha_m \rangle$ in konjunktiver Normalform, heißt Horn-Formel*, wenn jede Klausel α_i, $1 \le i \le m$, eine Horn-Klausel ist.*

Beispiel 3.31 $[\neg P(x, f(y), c), \neg Q(x)]$, $[\neg P(x, x, y), Q(y)]$ *und* $[Q(c)]$ *sind Horn-Klauseln, hierbei ist die erste Klausel $[\neg P(x, f(y), c), \neg Q(x)]$ eine negative Klausel. Die beiden anderen Klauseln sind definit.*
$(\forall x)(\exists y) \langle [\neg P(x, f(y), c), \neg Q(x)], [\neg P(x, x, y), Q(y)], [Q(c)] \rangle$ *ist daher eine Horn-Formel.*

SLD-Resolution

Diese Resolutionsrestriktion ist nur für Horn-Klauseln definiert. In ihr wird die Wahl des Literals, hinsichtlich dessen die nächste Resolvente gebildet werden soll, mittels einer „Selektionsfunktion" festgelegt. Da wir hier mehr an theoretischen Betrachtungen interessiert sind, ignorieren wir diesen Aspekt und legen stattdessen eine nichtdeterministische Auswahl zugrunde.

Definition 3.32 *Sei Σ eine Menge von Hornklauseln.*

1. *Eine lineare Input-Resolutionsherleitung von α aus Σ mit negativer Startklausel heißt* SLD-Resolutionsherleitung *von α aus Σ (notiert als $\Sigma \vdash_{SLDR} \alpha$).*

2. *Eine* SLD-Resolutionswiderlegung von Σ *ist eine SLD-Resolutionsherleitung von $[\,]$ aus Σ.*

Bemerkung 3.33

1. *Die Abkürzung „SLD" steht für „Linear resolution with Selection function for Definite clauses".*

2. *Als Spezialfall der linearen Input-Resolution ist die SLD-Resolution als Widerlegungsverfahren für beliebige prädikatenlogische Klauselmengen ebenfalls nicht vollständig.*

3. *Die hier gegebene Definition der SLD-Resolution ist allgemeiner als die insbesondere Implementierungen zugrunde liegenden Fassungen. Dort werden geordnete Klauseln (Listen, in denen Literale mehrfach auftreten können) verwendet, und für die Resolventenbildung wird jeweils nur ein Literal aus den beiden Klauseln herangezogen, auch wenn dies mehrfach auftritt bzw. unifiziert werden könnte. Da dies für die Vollständigkeitsbetrachtungen keinerlei Auswirkungen hat, gehen wir nicht darauf ein.*

4. *Als Spezialfälle der linearen Resolution sind die lineare Input-Resolution und die SLD-Resolution natürlich ebenfalls korrekt.*

Beispiel 3.34 *Wir geben eine SLD-Resolutionswiderlegung der Klauselmenge*

$$\Sigma = \{[\neg P(c), \neg P(b), P(a)], [P(b)], [P(c), P(e)], [P(c), \neg P(d)], [P(d)], [\neg P(a)]\}$$

an. a, b, c, d *und* e *sind hierbei Konstantensymbole.*

$$\begin{array}{ccl}
[\neg P(a)] & & [\neg P(c), \neg P(b), P(a)] \\
| & / & \\
[\neg P(c), \neg P(b)] & & [P(c), \neg P(d)] \\
| & / & \\
[\neg P(b), \neg P(d)] & & [P(d)] \\
| & / & \\
[\neg P(b)] & & [P(b)] \\
| & / & \\
[\,] & &
\end{array}$$

Theorem 3.35 (Vollständigkeit der linearen Input- und der SLD-Resolution) *Sei* $\alpha = (\forall x_1)\ldots(\forall x_n)\,\beta$ *eine Hornformel (in Erfüllbarkeitsfunktionalform) mit dem Kern* $\beta = \langle \alpha_1, \ldots, \alpha_m \rangle$. *Ist* α *unerfüllbar, so gilt*

1. $\{\alpha_1, \ldots, \alpha_m\} \vdash_{LIR} [\,]$
2. $\{\alpha_1, \ldots, \alpha_m\} \vdash_{SLDR} [\,]$.

Beweis. Da eine SLD-Resolutionswiderlegung insbesondere eine lineare Input-Resolutionswiderlegung ist, genügt es, $\{\alpha_1, \ldots, \alpha_m\} \vdash_{SLDR} [\,]$ zu zeigen.
Wie im Beweis der Vollständigkeit der linearen Resolution schon gezeigt, genügt es, den folgenden Spezialfall zu betrachten, um dann mit dem Lifting-Lemma das Resultat auf den allgemeinen Fall zu übertragen:
$\Sigma = \{\alpha_1, \ldots, \alpha_m\}$ sei eine minimal unerfüllbare Menge von Hornklauseln, die keine Variablen enthalten.
Wir werden für diesen Spezialfall zeigen, dass aus Σ mittels SLD-Resolution die leere Klausel herleitbar ist.
Für Σ gilt, dass Σ wenigstens eine negative Klausel enthalten muss. Andernfalls wäre eine Herbrand-Struktur, in der jede atomare Formel den Wahrheitswert T erhält, ein Modell von Σ, da jedes positive Literal einer in

Σ vorkommenden Klausel in dieser Herbrand-Struktur den Wahrheitswert T erhält. Sei somit α eine solche negative Klausel in Σ. Aufgrund der Vollständigkeit der linearen Resolution existiert dann eine lineare Resolutionswiderlegung von Σ mit der Stützmenge $\{\alpha\}$. In dieser Resolutionsherleitung ist dann jede Resolvente, die verschieden von der leeren Klausel [] ist, eine negative Klausel (einfacher Induktionsbeweis). Da zu zwei negativen Klauseln keine Resolvente gebildet werden kann, muss die Seitenklausel aus Σ sein, d.h. es liegt eine lineare Input-Resolution vor, in der die Startklausel negativ ist, also eine SLD-Resolutionswiderlegung von Σ. ■

Notation 3.36

1. *Eine SLD-Resolutionsherleitung*
$$(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n)$$
von α aus Σ notieren wir auch in der Weise:
$$\alpha_0 \overset{\beta_0,\sigma_0}{\longrightarrow} \alpha_1 \overset{\beta_1,\sigma_1}{\longrightarrow} \alpha_2 \ldots \overset{\beta_{n-1},\sigma_{n-1}}{\longrightarrow} \alpha_n \overset{\beta_n,\sigma_n}{\longrightarrow} \alpha glqq$$
wobei die σ_i, $0 \leq i \leq n$, die in der Resolventenbildung von α_i und β_i benutzten Substitutionen sind.

2. *Ist $(\alpha_0, \beta_0), \ldots, (\alpha_n, \beta_n)$ eine SLD-Resolutionswiderlegung von Σ, so heißt die Verkettung*
$$\sigma_0\sigma_1 \ldots \sigma_n$$
dieser Substitutionen Antwortsubstitution dieser Resolutionswiderlegung.

3.2.3 Logik-Programme

Die Bestimmung von Antwortsubstitutionen kann für die Lösung von Problemen, die in Form von Klauselmengen gegeben sind, genutzt werden. Diese Eigenschaft führt uns zu den Logik-Programmen, die üblicherweise in Form von Horn-Klauseln formuliert sind. Die Beschränkung auf Horn-Klauseln hat hierbei die folgenden Gründe.

1. Fast alle interessanten mathematischen Theorien lassen sich mittels Horn-Formeln axiomatisieren, so dass die Beschränkung auf Horn-Klauseln in den meisten Fällen keine Einschränkung darstellt.

2. Verfahren zur Erzeugung von Antworten bei Vorliegen von Klauseln, die keine Horn-Klauseln sind, sind schwierig.

3. Für Horn-Klauseln existieren effizientere Algorithmen. Z.B. ist das aussagenlogische Erfüllbarkeitsproblem für Horn-Klauseln in quadratischer Zeit lösbar. Die prädikatenlogische Erfüllbarkeit bzw. Allgemeingültigkeit von Horn-Formeln ist allerdings ebenfalls unentscheidbar, wie man sofort dem Beweis des Satzes von Church entnehmen kann: die dort auftretenden Formeln können leicht semantisch äquivalent in Horn-Formeln überführt werden.

Definition 3.37 *Ein* Logik-Programm *ist eine endliche Menge von definiten Horn-Klauseln.*
Wir unterscheiden hierin zwei Arten von Programm-Klauseln:

1. Fakten.

   Das sind Klauseln, die nur ein Literal, das folglich positiv ist, enthalten.

   (*In der logischen Programmiersprache* PROLOG *wird ein solches Fakt $[P(x, y)]$ in der Form $p(X, Y)$ notiert, d.h. für Variablen werden große Anfangsbuchstaben und für Prädikats- und Funktionssymbole kleine Anfangsbuchstaben verwendet.*).

2. Prozedur-Klauseln.

Das sind die Klauseln, die wenigstens zwei Literale enthalten (genau ein positives und mindestens ein negatives).

(Notation einer Prozedur-Klausel $[\alpha, \neg\alpha_1, \ldots, \neg\alpha_n]$ *mit atomaren Formeln* $\alpha, \alpha_1, \ldots, \alpha_n$ *in* PROLOG*:*

$$\alpha :- \alpha_1, \ldots, \alpha_n.$$

α *heißt dann der (Prozedur-) Kopf und* $\alpha_1, \ldots, \alpha_n$ *der (Prozedur-) Rumpf.)*

Ein Logik-Programm wird aufgerufen durch eine Zielklausel*. Hierbei ist eine Zielklausel als eine negative Klausel definiert.*
Notation einer Zielklausel $[\neg\alpha_1, \ldots, \neg\alpha_n]$ *mit atomaren Formeln* $\alpha_1, \ldots, \alpha_n$ *in* PROLOG*:*

$$?- \ \alpha_1, \ldots, \alpha_n.$$

Mit dem Begriff der Prozedur-Klausel

$$\alpha :- \alpha_1, \ldots, \alpha_n.$$

verbinden wir anschaulich gesprochen die folgende prozedurale Interpretation:
Um α erfolgreich auszuführen, muss zunächst α_1, dann α_2, usw. und schließlich α_n erfolgreich ausgeführt werden.

Notation 3.38 *Für eine Klausel* α *mit* $var(\alpha) = \{x_1, \ldots, x_m\}$ *sei* $\forall\alpha := (\forall x_1) \ldots (\forall x_m)\, \alpha$*.*
Für ein Logik-Programm Σ *sei* $\forall\Sigma := \{\forall\alpha \mid \alpha \in \Sigma\}$*.*

Theorem 3.39 *Sei* Σ *ein Logik-Programm,* $\alpha = [\neg\alpha_1, \ldots, \neg\alpha_n]$ *eine Zielklausel für* Σ *und* θ *eine Antwortsubstitution einer SLD-Resolutionswiderlegung von*

$$\Sigma \cup \{\alpha\}\,.$$

Dann gilt:

$$\forall\Sigma \models_{PL} (\alpha_1 \wedge \ldots \wedge \alpha_n)\, \theta\,.$$

Beweis. Wir zeigen durch Induktion über die Länge k der *SLD*-Resolutionswiderlegung von $\Sigma \cup \{\alpha\}$, dass gilt:
Für alle $k \geq 1$, für alle Zielklauseln α der Gestalt $\alpha = [\neg\alpha_1, \ldots, \neg\alpha_n]$ für ein $n \geq 1$: besitzt $\Sigma \cup \{\alpha\}$ eine *SLD*-Resolutionswiderlegung der Länge k, so gilt $\forall\Sigma \models_{PL} (\alpha_1 \wedge \ldots \wedge \alpha_n)\, \theta$.

$k = 1$: Es gilt dann für die Zielklausel $\alpha = [\neg\alpha_1, \ldots, \neg\alpha_n]$ und für ein $\beta \in \Sigma$

$$\alpha \xrightarrow{\beta,\sigma} [\,]\ .$$

Dies ist nur dann der Fall, wenn $n = 1$, also $\alpha = [\neg\alpha_1]$ ist und β ein Fakt ist mit $\beta\sigma = \alpha_1\sigma$.

Es gilt $\Sigma \models_{PL} \beta$. Damit gilt $\forall\Sigma \models_{PL} \beta\sigma$, also $\forall\Sigma \models_{PL} \alpha_1\sigma$.

$k \to k+1$: Es sei $\alpha \xrightarrow{\beta,\sigma} \underbrace{\gamma_1 \xrightarrow{\beta_1,\sigma_1} \gamma_2 \xrightarrow{\beta_2,\sigma_2} \ldots \to \gamma_k \xrightarrow{\beta_k,\sigma_k} [\,]}_{\text{Länge } k}$ eine Resolutionswiderlegung von $\Sigma \cup \{\alpha\}$ mit $\alpha = [\neg\alpha_1, \ldots, \neg\alpha_n]$ als Zielklausel und

$\beta = [\alpha', \neg\delta_1, \ldots, \neg\delta_m]$.

Es gelte für die Substitution σ : $\quad \alpha'\sigma = \alpha_i\sigma$ für ein i mit $1 \leq i \leq n$.

Damit ist

$$\gamma_1 = ([\neg\delta_1, \ldots, \neg\delta_m, \neg\alpha_1, \ldots, \neg\alpha_n] - [\neg\alpha_i])\, \sigma.$$

Für $\Sigma \cup \{\gamma_1\}$ existiert eine Resolutionswiderlegung der Länge k. Nach Induktionsvoraussetzung für (Σ und) γ_1 gilt mit $\theta' = \sigma_1 \ldots \sigma_k$

$$\forall\Sigma \models_{PL} \left(\bigwedge_{1\le j\le m} \delta_j\sigma \wedge \bigwedge_{\substack{1\le j\le n\\ j\ne i}} \alpha_j\sigma \right) \theta' \tag{3.1}$$

Die Antwortsubstitution für die Resolutionswiderlegung von $\Sigma \cup \{\alpha\}$ ist $\theta := \sigma\theta'$.

Wegen $\beta \in \Sigma$ gilt

$$\forall\Sigma \models_{PL} [\alpha'\theta, \neg\delta_1\theta, \ldots, \neg\delta_m\theta] \tag{3.2}$$

Wegen (3.1) gilt $\forall\Sigma \models_{PL} \delta_j\theta, \quad 1 \le j \le m$ und $\forall\Sigma \models_{PL} \alpha_j\theta, \quad 1 \le j \le n$ und $j \ne i$.

Mit (3.2) folgt somit $\forall\Sigma \models_{PL} \alpha'\theta$. Es ist aber $\alpha'\theta = \alpha_i\theta$. Damit folgt schließlich insgesamt

$$\forall\Sigma \models_{PL} \bigwedge_{1\le j\le n} \alpha_j\theta.$$

Dies war zu zeigen.

■

Bemerkung 3.40 *Die prozedurale Interpretation der Prozedur-Klauseln zusammen mit der Bestimmung der Antwortsubstitution bildet die Grundlage für die Verwendung von Logik-Programmen als logische Programmiersprachen. Das folgende, allgemein gehaltene Beispiel soll hierbei die Rolle der Antwortsubstitution verdeutlichen.*

Sei $(\forall x_1)\ldots(\forall x_m)(\alpha_1 \wedge \ldots \wedge \alpha_n)$ *eine universale Hornformel mit definiten Klauseln* $\alpha_1, \ldots, \alpha_n$ *(d.h.* $\{\alpha_1, \ldots, \alpha_n\}$ *ist ein Logik-Programm). Es soll festgestellt werden ob für die atomare Formel* β *mit den freien Variablen* $y_1, \ldots, y_k$

$$(\forall x_1)\ldots(\forall x_m)(\alpha_1 \wedge \ldots \wedge \alpha_n) \models_{PL} (\exists y_1)\ldots(\exists y_k)\,\beta$$

gilt.
Wir nutzen den folgenden Zusammenhang aus:
$(\forall x_1)\ldots(\forall x_m)(\alpha_1 \wedge \ldots \wedge \alpha_n) \models_{PL} (\exists y_1)\ldots(\exists y_k)\,\beta$
$\Leftrightarrow$ $(\forall x_1)\ldots(\forall x_m)(\alpha_1 \wedge \ldots \wedge \alpha_n) \wedge \neg(\exists y_1)\ldots(\exists y_k)\,\beta$ *unerfüllbar*
$\Leftrightarrow$ $\{\alpha_1, \ldots, \alpha_n, \neg\beta\} \vdash_{SLDR} []$ *; hierbei sei* θ *die Antwortsubstitution.*
$\Leftrightarrow$ $(\forall x_1)\ldots(\forall x_m)(\alpha_1 \wedge \ldots \wedge \alpha_n) \models_{PL} \beta\{y_1/\theta(y_1), \ldots, y_k/\theta(y_k)\}$.
Die Antwortsubstitution θ *liefert uns somit Elemente* $\theta(y_1), \ldots, \theta(y_k)$*, für die die Formel* β *behauptet wird.*

Beispiel 3.41 *Wir haben das folgende Logik-Programm*

$$\begin{array}{ll} 1 & [Q(x,z), \neg P(x,y), \neg Q(y,z)] \\ 2 & [Q(u,u)] \\ 3 & [P(b,c)] \end{array}$$

und die Zielklausel (Anfrage)

$$4 \quad [\neg Q(x,c)]$$

gegeben. Wir wollen somit also feststellen, ob

$$(\forall x)(\forall y)(\forall z)(\forall u)((Q(x,z) \vee \neg P(x,y) \vee \neg Q(y,z)) \wedge Q(u,u) \wedge P(b,c)) \models_{PL} (\exists x)\,Q(x,c)$$

gilt. Wir erhalten die folgende Resolutionsherleitung (Achtung: wir benutzen eine nichtdeterminierte Selektion)

4	$[\neg Q(x,c)]$	
	$\downarrow$	$\sigma_1 = \{z_1/c,\ x/x_1\}$
5	$\left[\neg P(x_1,y_1), \underline{\neg Q(y_1,c)}\right]$	*Resolvente aus 1,4; das nächste selektierte Literal unterstrichen*
	$\downarrow$	$\sigma_2 = \{y_1/c,\ u/c\}$
6	$[\neg P(x_1,c)]$	*Resolvente aus 2,5*
	$\downarrow$	$\sigma_3 = \{x_1/b\}$
7	$[\,]$	*Resolvente aus 3,6*

Die Antwortsubstitution $\sigma_1\sigma_2\sigma_3$ liefert ein gesuchtes Element, nämlich $x(\sigma_1\sigma_2\sigma_3) = b$.
(Die Indizes für die Variablen in den Substitutionen resultieren aus den Umbenennungen, die bei den Resolventenbildungen vorzunehmen sind.)

Notation 3.42 *G bezeichne im Folgenden eine Zielklausel für ein Logik-Programm.*

Definition 3.43 *Sei Σ ein Logik-Programm, G eine Zielklausel.*

1. *Eine* Selektionsfunktion *f ist eine Funktion, die jeder Zielklausel G' ein Literal in dieser Zielklausel zuordnet.*
2. *Ein SLD-*Baum für *Σ, G und mit Selektionsfunktion f (notiert als $SLD(\mathbf{\Sigma}, G, f)$) ist ein Baum, dessen Knoten mit Zielklauseln markiert sind, so dass gilt:*
   (a) *Die Wurzel ist mit G markiert.*
   (b) *Ist s ein Knoten des Baumes mit Markierung $G' = [\neg\alpha_1, \ldots, \neg\alpha_k]$ und ist $f([\neg\alpha_1, \ldots, \neg\alpha_k]) = \neg\alpha_i$, wobei $1 \le i \le k$, dann hat s für jede Programmklausel $[\beta, \neg\beta_1, \ldots, \neg\beta_n] \in \Sigma$, für die zu $\{\alpha_i, \beta\}$ ein allgemeinster Unifikator θ existiert, einen Sohn, der mit*
   $$[\neg\alpha_1, \ldots, \neg\alpha_{i-1}, \neg\beta_1, \ldots, \neg\beta_n, \neg\alpha_{i+1}, \ldots, \neg\alpha_k]\,\theta$$
   markiert ist.
   (c) *Weitere Knoten existieren nicht.*

Anmerkungen zur Programmiersprache PROLOG

Wenngleich logische Programmiersprachen nicht Gegenstand der Vorlesung sind, soll doch wenigstens auf einige Besonderheiten der gegenwärtig wohl wichtigsten logischen Programmiersprache, nämlich PROLOG, hingewiesen werden.

1. In PROLOG ist die Selektionsfunktion in der Weise implementiert, dass sie immer das erste Literal der Zielklausel auswählt. Dies stellt, wie man sich leicht überlegen kann, keine Beschränkung des Verfahrens dar. Auf die Möglichkeit, in einer Resolventenbildung mehrere Literale einer Klausel zu unifizieren, wird verzichtet.
2. PROLOG's „Standard-Auswertungsstrategie" sieht so aus, dass immer die erste Klausel im Programm gewählt wird, für die eine Resolventenbildung möglich ist, d.h. es wird im SLD-Baum aus Effizienzgründen ein Tiefendurchlauf vorgenommen mit der Konsequenz, dass u.U. eine vorhandene Lösung nicht gefunden wird.

3. In fast allen Implementierungen von PROLOG wird im Unifikationsalgorithmus auf den Test verzichtet, ob die Variable, die mit einem Term unifiziert werden soll, in diesem bereits vorkommt („*occur check*"). Die Folge ist: Das Programm kann u.U. eine „Lösung" liefern, obwohl gar keine existiert, so dass in einem solchen Fall das PROLOG-Programm nicht korrekt ist. Es liegt am Programmierer, auf die Problematik des occur check zu achten und damit die Korrektheit des Programms zu sichern.

4. PROLOG bietet die Möglichkeit, mittels eines „*cut*" Teile des *SLD*-Baumes beim Backtracking von einem Durchlauf auszugrenzen.

   Positiv gesehen kann dies zu einer Effizienzsteigerung führen bzw. überhaupt erst das Finden einer Lösung ermöglichen, stellt andererseits aber auch eine weitere Quelle für Unvollständigkeit dar, da vorhandene Lösungen eventuell nicht gefunden werden.

Deklarative Semantik von Logik-Programmen

Das Thema „Semantik von Programmiersprachen" und damit auch die Semantik von logischen Programmiersprachen ist nicht Gegenstand dieser Vorlesung. Aufgrund der engen Verflechtung mit dem Gebiet der Logik und hier insbesondere mit den Herbrand-Modellen wollen wir dieses Thema aber wenigstens flüchtig streifen.

Notation 3.44 *Sei S eine Signatur, die wenigstens ein Konstantensymbol enthält.*
Die Herbrand-Basis B_S zur Signatur S ist dann die Menge aller atomaren variablenfreien Formeln aus $Fm_{PL}(S)$.

Beispiel 3.45 *Sei $S = (P, Q; f; c)$ mit $ar(P) = ar(f) = 1$ und $ar(Q) = 1$. Dann ist $\left\{f^i(c) \mid i \in \mathbb{N}\right\}$ das Herbrand-Universum und*

$$\left\{P\left(f^i(c)\right), Q\left(f^i(c)\right) \mid i \in \mathbb{N}\right\}$$

die Herbrand-Basis.
Hierbei ist

$$\begin{aligned} f^0(c) &:= c \\ f^1(c) &:= f(c) \ \ \textit{und} \\ f^{n+1}(c) &:= f(f^n(c)) \ \ \textit{für } n > 0. \end{aligned}$$

Bemerkung 3.46 *In Herbrand-Strukturen ist der Grundbereich wie auch die Interpretation der Funktions- und Konstantensymbole bereits durch die Signatur bestimmt. Verschiedene Herbrand-Strukturen können sich nur durch die Interpretation der Prädikatensymbole unterscheiden. Wir können nun Herbrand-Strukturen mit der Menge der Formeln aus der Herbrand-Basis identifizieren, die sie erfüllen.*
Es sei für eine Herbrand-Struktur $\mathcal{A}$ zur Signatur S $M_{\mathcal{A}} := \{\alpha \in B_S \mid A \models_{PL} \alpha\}$. $M_{\mathcal{A}}$ ist somit eine Teilmenge von B_S. Im folgenden werden wir häufig von dieser Identifizierung von einer Herbrand-Struktur A mit der Teilmenge $M_{\mathcal{A}}$ der Herbrand-Basis Gebrauch machen.

Beispiel 3.47 *Es sei die Signatur aus Beispiel 3.45 zugrundegelegt. Ferner sei die folgende Herbrand-Struktur $\mathcal{A} = \left\{H(S); P^{\mathcal{A}}, Q^{\mathcal{A}}; f; c\right\}$ mit $P^{\mathcal{A}} = \{c, f(f(c))\}$ und $Q^{\mathcal{A}} = \{f(c)\}$ gegeben. Dann können wir $\mathcal{A}$ mit der Menge $\{P(c), P(f(f(c))), Q(f(c))\}$ identifizieren.*

Sei Σ nun ein Logik-Programm der Signatur S. Σ besteht also aus einer Menge von definiten Klauseln. Unter der eben genannten Identifizierung ist dann B_S ein Herbrand-Modell von $\forall\Sigma$. Da jedes andere Herbrand-Modell (via Identifizierung) eine Teilmenge von B_S ist, ist B_S somit das größte Herbrand-Modell von $\forall\Sigma$.

Theorem 3.48 *Sei Σ ein Logik-Programm der Signatur S. $\forall\Sigma$ besitzt ein kleinstes Herbrand-Modell.*

Beweis. Sei $M_i,\ i \in I$ für eine geeignete Indexmenge I, die Menge der Herbrand-Modelle von $\forall\Sigma$. Da B_S ein Herbrand-Modell von $\forall\Sigma$ ist, ist $I \neq \emptyset$. Wir setzen $M_\Sigma := \bigcap_{i\in I} M_i$. M_Σ ist somit eine Herbrand-Struktur (via Identifizierung). Sei α eine beliebige definite Klausel aus Σ. Wir zeigen, dass M_Σ $\forall\alpha$ erfüllt, so dass also M_Σ ein Modell von $\forall\Sigma$ ist.
Wir betrachten den Fall $\forall\alpha = (\forall x_1)\ldots(\forall x_k)\,(l \vee \neg l_1 \vee \ldots \vee \neg l_m)$ mit positiven Literalen $l, l_1, \ldots, l_m$ und $m \geq 1$. Der Fall, dass α ein positives Literal ist, ist einfacher und wird hier nicht bewiesen.
Sei w eine beliebige Belegung. Es gilt dann
$M_\Sigma, w \models_{PL} \forall\alpha$
$\Leftrightarrow$ $M_\Sigma, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l \vee \neg l_1 \vee \ldots \vee \neg l_m$ für alle Terme $t_1, \ldots, t_k \in H(S)$ ($H(S)$ ist das Herbrand-Universum zur Signatur S)
$\Leftrightarrow$ ($M_\Sigma, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l_1 \wedge \ldots \wedge l_m \quad\Rightarrow\quad M_\Sigma, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l$ für alle Terme $t_1, \ldots, t_k \in H(S)$.)
Es gelte nun für beliebig gewählte Terme $t_1, \ldots, t_k \in H(S)$

$$M_\Sigma, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l_1 \wedge \ldots \wedge l_m$$

Aufgrund des Substitutionstheorems gilt

$$M_\Sigma, w \models_{PL} l_1\,\{x_1/t_1, \ldots, x_k/t_k\} \wedge \ldots \wedge l_m\,\{x_1/t_1, \ldots, x_k/t_k\}$$

Damit gilt dann für jedes $i \in I$

$$M_i, w \models_{PL} l_1\,\{x_1/t_1, \ldots, x_k/t_k\} \wedge \ldots \wedge l_m\,\{x_1/t_1, \ldots, x_k/t_k\},$$

und folglich auch

$$M_i, w \models_{PL} l\,\{x_1/t_1, \ldots, x_k/t_k\}$$

nach Voraussetzung über die Herbrand-Modelle M_i.
Wir erhalten somit für alle $i \in I$

$$M_i, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l$$

und damit nach Definition von M_Σ

$$M_\Sigma, w\,\{x_1/t_1, \ldots, x_k/t_k\} \models_{PL} l.$$

Dies war aber zu zeigen. ■

Dieses nach dem eben bewiesenen Theorem 3.48 vorhandene und eindeutig bestimmte kleinste Herbrand-Modell M_Σ wird als die *deklarative Semantik* des Logik-Programms Σ bezeichnet. Für diese erhalten wir noch eine weitere Charakterisierung.

Theorem 3.49 *Sei Σ ein Logik-Programm der Signatur S. Dann gilt:*

$$M_\Sigma = \{\alpha \in B_S \mid \forall\Sigma \models_{PL} \alpha\}.$$

Beweis. Es gilt
$\forall\Sigma \models_{PL} \alpha$
$\Leftrightarrow$ $\forall\Sigma \cup \{\neg\alpha\}$ ist unerfüllbar
$\Leftrightarrow$ $\forall\Sigma \cup \{\neg\alpha\}$ besitzt kein Herbrand-Modell
$\Leftrightarrow$ $[\neg\alpha]^{\mathcal{A}} = F$ für jedes Herbrand-Modell von $\forall\Sigma$
$\Leftrightarrow$ $\alpha^{\mathcal{A}} = T$ für jedes Herbrand-Modell von $\forall\Sigma$
$\Leftrightarrow$ $\alpha \in M_\Sigma$. ■

Für die deklarative Semantik kann noch eine elegante Fixpunktcharakterisierung erhalten werden. Dies soll hier aber nicht mehr behandelt werden. Siehe hierzu z.B. J.W. Lloyd, Foundations of Logic Programming, Springer Verlag 1993 (2. Auflage).

Anmerkungen zu „negation by (finite) failure”

Logik-Programme der hier betrachteten Art haben die Eigenschaft, dass aus ihnen keine negativen Literale gefolgert werden können. Dies ergibt sich z.B. aus der Tatsache, dass für ein Logik-Programm Σ die Herbrand-Basis immer ein Herbrand-Modell von $\forall\Sigma$ darstellt. In manchen Kontexten – man denke z.B. an Datenbanken – möchte man jedoch aus der Tatsache, dass ein positives Fakt nicht gefolgert werden kann, schließen, dass dieses Fakt nicht zutrifft. Bezogen auf unsere Logik-Programme heißt dies, dass man aus der Tatsache, dass ein positives Literal l aus einem Logik-Programm nicht folgt, folgert, dass $\neg l$ gilt.

Beispiel 3.50 *Betrachten wir das folgende Logik-Programm* Σ :

$$[P(c)]$$
$$[P(d)].$$

Angenommen, wir wollen ermitteln, dass $P(e)$ *nicht zutrifft. Weder* $P(e)$ *noch* $\neg P(e)$ *ist eine semantische Folgerung aus* $\{P(c), P(d)\}$. *Unterstellt man, dass dieses Logik-Programm eine vollständige Beschreibung aller positiven Sachverhalte und der daraus ableitbaren positiven Sachverhalte beinhaltet, dann wird man aus der Tatsache, dass* $P(e)$ *nicht gefolgert werden kann, schließen, dass* $\neg P(e)$ *zutrifft.*

Das Postulat (bzw. die Schlussregel), dass eine Aussage α nicht zutrifft, wenn α aus dem Logik-Programm Σ nicht gefolgert werden kann, wird als closed world assumption (CWA) bezeichnet. Diese zugrundeliegende Schlussregel ist ein Beispiel für eine nicht-monotone Schlussregel. Das Hinzufügen von zusätzlichen Prämissen kann bewirken, dass Aussagen, die zuvor gefolgert werden konnten, nach dem Hinzufügen dieser Prämissen nicht mehr gefolgert werden können. Würden wir zu dem im Beispiel 3.50 angegebenen Logik-Programm Σ das Fakt $[P(e)]$ hinzufügen, so könnten wir mit der CWA-Regel nicht mehr auf $\neg P(e)$ schließen, da ja $P(e)$ jetzt eine semantische Folgerung ist.
Die Feststellung, dass für ein Logik-Programm Σ ein positives Literal l nicht aus Σ semantisch folgt, ist äquivalent dazu, daß $\Sigma \cup \{\neg l\}$ erfüllbar ist, also äquivalent dazu, daß keine SLD-Resolutionswiderlegung von $\Sigma \cup \{\neg l\}$ existiert. Man spricht in diesem Zusammenhang daher auch von *negation by failure*, d.h. $\neg l$ wird aufgrund der Tatsache gefolgert, daß keine SLD-Resolutionswiderlegung von $\Sigma \cup \{\neg l\}$ existiert. Man würde nun gerne Logik-Programme um dieses „negation by failure” Prinzip erweitern. Hier ist man aber mit dem Problem konfrontiert, dass es kein algorithmisches Entscheidungsverfahren gibt, mit dem man ermitteln kann, ob ein positives Literal l aus einem Logik-Programm Σ semantisch folgt. Man verwendet daher ein schwächeres Prinzip: *negation by finite failure*. Nach diesem Prinzip darf $\neg l$ dann gefolgert werden, wenn für eine Selektionsfunktion f ein endlicher SLD-Baum $SLD(\Sigma, [\neg l], f)$ existiert, der nicht erfolgreich ist (d.h. kein Blatt enthält die leere Klausel). Dieses Prinzip des *negation by finite failure* ist in den gängigen PROLOG-Implementierungen realisiert.

Kapitel 4

Modale Logiken

Die klassische Prädikatenlogik (erster Stufe), die wir im zweiten Kapitel und im Zusammenhang mit der Logischen Programmierung auch noch im dritten Kapitel behandelt haben, ist für die meisten Anwendungen ausreichend. Anhand von Beispielen ergab sich aber auch, dass bestimmte Eigenschaften nicht mit Mitteln der Prädikatenlogik (erster Stufe) ausdrückbar sind. Wir wollen uns nun speziellen Anwendungen zuwenden, für die die Ausdrucksmittel der Prädikatenlogik in den meisten Fällen zwar genügen würden, bei denen aber eine Beschränkung auf die Prädikatenlogik oft zu einer umständlichen, zum Teil auch Probleme verdeckenden Behandlung führt.

Betrachten wir die Aussagen $A =$ „im Jahre 2050 ist 1 Euro mehr wert als 1 Dollar" und $B =$ „im Jahre 2050 ist 1 Euro mehr wert als 1 DM". α bzw. β seien die formalen Repräsentationen dieser Aussagen in einer geeigneten prädikatenlogischen Sprache. Es ist offensichtlich, dass α und β keine allgemeingültige, wohl aber erfüllbare Aussagen darstellen, da die in diesen Aussagen enthaltenen Konstantensymbole in unterschiedlichen Strukturen unterschiedliche Interpretationen erfahren können. Wenn wir nun ausdrücken wollen, dass die Aussage A möglicherweise zutrifft und die Aussage B (aufgrund des festen Wechselkurses) aber notwendigerweise gilt, so können wir dies folglich nicht gleichsetzen mit der Erfüllbarkeit von α bzw. Allgemeingültigkeit von β im prädikatenlogischen Sinn. Für die Beurteilung, ob die in unserer Alltagssprache formulierten Aussagen A bzw. B möglicherweise bzw. notwendigerweise zutreffen, kommen für uns nur solche Situationen (Welten, Strukturen) in Betracht, in denen die beteiligten Größen immer dieselbe Bedeutung wie in unserer aktuellen Situation haben, d.h. Konstantensymbole werden immer gleich interpretiert. Abhängig von einer aktuellen Welt sind somit immer nur bestimmte andere Welten vorstellbar oder – wie wir auch sagen wollen – möglich. „Notwendigerweise gilt A" werden wir daher in einer Welt p (beispielsweise der aktuellen) dann als wahr ansehen, wenn in allen in p als möglich vorstellbaren Welten q die Aussage A gilt. So können wir uns also für das Jahr 2050 eine Welt vorstellen, in denen 1 Euro mehr wert ist als 1 Dollar. Und ebenso ist z.B. denkbar, dass im Jahr 2050 1 Dollar mindestens soviel wert ist wie 1 Euro. Beide Aussagen „möglicherweise gilt A" und „möglicherweise gilt nicht A" werden wir daher als wahr ansehen.

Um nun solche *modalen Bestimmungen* von Aussagen wie „A gilt möglicherweise" und „A gilt notwendigerweise" behandeln zu können, erweitern wir die Prädikatenlogik zur sogenannten Modallogik.

Historisch gesehen hat die Modallogik einen rein philosophischen Ursprung. Es war hier das Ziel, insbesondere die Logik der Begriffe „notwendigerweise", „möglicherweise", „wissen", „glauben" u.a. zu erhellen. Durch die Entwicklung der Informatik hat der Bereich der Modallogik ganz wesentliche Impulse erfahren. Nicht nur, dass mit Programmverifikation und -spezifikation neue Anwendungsbereiche erschlossen wurden. Auch eher traditionelle Themenbereiche wie die epistemische Logik fanden Eingang in die Gebiete der Künstlichen Intelligenz und der verteilten Systeme.

4.1 Formalisierung der Modallogik

Wir werden im Folgenden eine formale Sprache für die Modallogik angeben.
Wir erweitern hierzu die Ausdrucksmittel der Prädikatenlogik. Wir übernehmen daher den Begriff der Signatur, setzen zur Vereinfachung der Darstellung für das Folgende zunächst aber voraus, dass in einer Signatur keine Funktionssymbole enthalten sind.

Das *modallogische Alphabet* $\Sigma_{ML}(S)$ zur Signatur S besteht aus dem prädikatenlogischen Alphabet $\Sigma_{PL}(S)$ erweitert um die beiden zusätzlichen einstelligen *Modaloperatoren* $\Box, \Diamond$ (ausgesprochen: box und diamond). Variablen und Terme sind wie im Falle der Prädikatenlogik definiert. Die Menge der Variablen bezeichnen wir mit Var.
Die modallogischen Formeln werden wie im Falle der Prädikatenlogik definiert, ergänzt aber um eine weitere Regel, die die Operatoren $\Box$ und $\Diamond$ betrifft. Wir geben hier die induktive Definition der besseren Lesbarkeit wegen vollständig an.

Definition 4.1 (Menge $Fm_{ML}(S)$ der *modallogischen Formeln der Signatur* S)

1. *Eine* atomare modallogische Formel der Signatur S *ist eine Zeichenreihe der Gestalt* $P(t_1,\ldots,t_n)$, *wobei* P *ein* n*-stelliges Prädikatensymbol der Signatur* S *ist und* $t_1,\ldots,t_n \in Tm_{ML}(S)$ *oder aber* $\top$ *bzw.* $\bot$.
2.
   (a) *Jede atomare modallogische Formel der Signatur* S *ist in* $Fm_{ML}(S)$.
   (b) *Falls* $\alpha \in Fm_{ML}(S)$, *dann ist auch* $\neg\alpha \in Fm_{ML}(S)$.
   (c) *Falls* $\alpha, \beta \in Fm_{ML}(S)$ *und* $\circ \in \{\wedge, \vee, \rightarrow\}$, *dann ist auch* $(\alpha \circ \beta) \in Fm_{ML}(S)$.
   (d) *Ist* $\alpha \in Fm_{ML}(S)$ *und ist* x *eine Variable, so sind* $(\forall x)\,\alpha$ *und* $(\exists x)\,\alpha$ *Formeln in* $Fm_{ML}(S)$.
   (e) *Ist* $\alpha \in Fm_{ML}(S)$, *so sind auch* $\Box\alpha, \Diamond\alpha \in Fm_{ML}(S)$.

Beispiel 4.2 *Die folgenden Zeichenreihen sind modallogische Formeln der Signatur* $(P,Q;c)$ *mit* $ar(P) = 1$ *und* $ar(Q) = 2$:

1. $(\forall x)\,P(x) \vee \neg\Box\,(\exists x)\,Q(x,c)$.
2. $\Diamond\,(\forall x)\,(\Box\Box P(x) \rightarrow (\exists y)\,\Diamond\,(\neg\Box Q(x,y) \wedge P(y)))$.

Die Definition von Begriffen wie „freie und gebundene Variable”, „Teilformel”, „Aussage”, „Substitution” usw. kann wörtlich übernommen werden, so dass wir sie hier nicht mehr angeben. Ebenfalls übernehmen wir die für die Prädikatenlogik eingeführten Klammerungskonventionen, ergänzt um eine Festlegung für die beiden Modaloperatoren $\Box$ und $\Diamond$. Für diese legen wir fest, dass sie die gleiche Priorität wie $\neg$ haben.
Wie bereits in der Aussagen- und der Prädikatenlogik steht uns auch für die Modallogik eine strukturelle Induktion als Beweisprinzip und eine strukturelle Rekursion als Definitionsprinzip zur Verfügung. Wir geben hier ohne Beweis das betreffende Beweisprinzip der strukturellen Induktion an.

Theorem 4.3 (Strukturelle Induktion für modallogische Formeln) *E sei eine Eigenschaft für modallogische Formeln einer Signatur S, für die gilt:*

1. *E trifft auf alle atomaren modallogischen Formeln zu.*
2. *Trifft E auf α und β zu, ist x eine Variable und* $\circ \in \{\wedge, \vee, \rightarrow\}$, *dann trifft E auch auf* $\neg\alpha$, $\Box\alpha$, $\Diamond\alpha$, $(\alpha \circ \beta)$, $(\forall x)\,\alpha$ *und* $(\exists x)\,\alpha$ *zu.*

Dann gilt:
E trifft auf alle modallogischen Formeln der Signatur S zu.

4.1.1 Semantik der Modallogik

Die Interpretation von modallogischen Formeln basiert auf der von Kripke eingeführten Relationalsemantik (häufig als *Kripke-Semantik* oder auch *mögliche Welten Semantik* bezeichnet). Mit ihr wird die Vorstellung, dass in einer gegebenen Welt nur bestimmte andere Welten als möglich anzusehen sind, formalisiert. Auch wollen wir diese Weltensemantik so modellieren, dass wir nicht in allen Welten „Namen" für bestimmte Objekte haben müssen. Zu diesem Zweck führen wir die folgende Notation ein.

Notation 4.4 *Sei* $S = (\mathcal{R}; \mathcal{K})$ *eine Signatur und* $\mathcal{K}' \subseteq \mathcal{K}$.
Für die Signatur $(\mathcal{R}; \mathcal{K}')$ *schreiben wir dann* $S_{\mathcal{K}'}$.

Definition 4.5 *Sei* $S = (\mathcal{R}; \mathcal{K})$ *eine Signatur.*

1. *Ein* Kripke-Rahmen *ist ein Tupel* (W, R), *wobei*
   (a) W *eine nichtleere Menge (von (möglichen) Welten, auch als Zustände bezeichnet) und*
   (b) $R \subseteq W \times W$ *die* Erreichbarkeitsrelation *ist.*
   (Statt $(p, q) \in R$ *schreiben wir im allgemeinen* pRq*.)*
2. *Eine* $(S-)$ Kripke-Struktur *ist ein Tripel* $(W, R, \mathfrak{I})$, *wobei*
   (a) (W, R) *ein Kripke-Rahmen und*
   (b) $\mathfrak{I}$ *eine Abbildung ist, die jeder Welt* $p \in W$ *eine prädikatenlogische* $S_{\mathcal{K}_p}$*-Struktur* $\mathfrak{I}(p)$ *zuordnet mit* $\mathcal{K}_p \subseteq \mathcal{K}$.
   $\mathfrak{I}$ *wird als* Welteninterpretation *bezeichnet. (Statt* $\mathfrak{I}(p)$ *schreiben wir auch* $\mathfrak{I}_p$*. Ist* $c \in \mathcal{K}_p$*, so sagen wir, dass das Konstantensymbol* c *in der Welt* p *ein Objekt bezeichnet.)*

Beispiel 4.6 *Sei* $S = (P; c)$.
$(\{1,2,3\}, \{(1,2), (1,3), (2,3), (3,2)\})$ *ist ein Kripke-Rahmen und*
$(\{1,2,3\}, \{(1,2), (1,3), (2,3), (3,2)\}, \mathfrak{I})$ *mit*
$\mathfrak{I}_1 := (\{a_0, a_1\}; \{a_0\}; a_0)$
$\mathfrak{I}_2 := (\{a_0, a_1, a_2\}; \{a_0, a_2\}; a_0)$
$\mathfrak{I}_1 := (\{a_0, a_1, a_2\}; \{a_0, a_1\}; a_0)$
ist eine Kripke-Struktur.

Aus diesem Beispiel ist schon ersichtlich, dass über die Grundbereiche in den verschiedenen möglichen Welten (Zuständen) zunächst gar nichts festgelegt ist, sie insbesondere also verschieden sein können. In vielen wichtigen Anwendungen hat man jedoch Beschränkungen bezüglich der Objekte, die in den verschiedenen Welten existieren. Zwei Fälle sind dabei von größerer Bedeutung: einmal die Annahme von überall gleichen Grundbereichen, auch als *constant domain Semantik* bezeichnet, zum anderen die Annahme, dass mit dem Übergang von einer Welt in eine andere der Grundbereich nicht abnimmt Wir wollen uns hier auf den letzteren Fall beschränken, wir treffen also die folgende Voraussetzung für S- Kripke-Strukturen mit $S = (\mathcal{R}; \mathcal{K})$ (man beachte, dass wir im Falle der Modallogik zunächst keine Funktionssymbole zugelassen haben) :

- Ist $M = (W, R, \mathfrak{I})$ eine Kripke-Struktur, so gilt für alle $p, q \in W$, wenn pRq, so $|\mathfrak{I}_p| \subseteq |\mathfrak{I}_q|$ und $\mathcal{K}_p \subseteq \mathcal{K}_q$.

Darüberhinaus verlangen wir, dass Konstantensymbole $c \in \mathcal{K}$, sofern sie in Welten ein Objekt bezeichnen, dann auch immer dasselbe Objekt bezeichnen, d.h es gilt

$$c \in \mathcal{K}_p \cap \mathcal{K}_q \Rightarrow c^{\mathfrak{I}_p} = c^{\mathfrak{I}_q}.$$

Sie sind gewissermaßen Konstanten in einem globalen Sinne (englisch: rigid designators). Wir werden später noch auf Konstanten im lokalen Sinne eingehen.

- Die nach der eben getroffenen Vereinbarung einem Konstantensymbol c in einer Kripke-Struktur $M = (W, R, \mathfrak{I})$ eindeutig zugeordnete Konstante werde mit $c^{\mathfrak{I}}$ notiert.

Definition 4.7 *Sei* $M = (W, R, \mathfrak{J})$ *eine Kripke-Struktur.*
Eine Abbildung $w : Var \to \bigcup_{p \in W} |\mathfrak{J}_p|$ *heißt* M-Belegung.

Notation 4.8 *Sei* $M = (W, R, \mathfrak{J})$ *eine Kripke-Struktur. Statt* $p \in W$ *schreiben wir auch* $p \in M$.

Wir ordnen nun jeder modallogischen Formel $\alpha \in Fm_{ML}(S)$ in der Welt p der Kripke-Struktur M und unter der M-Belegung w einen Wahrheistwert $\alpha^{M,p,w}$ zu, wobei wir den Fall $\alpha^{M,p,w} = T$ durch

$$M, p, w \Vdash_{ML} \alpha$$

notieren. (Sprechweise: M, p, w erzwingt α.)

Definition 4.9 *Sei* $S = (\mathcal{R}; \mathcal{K})$, $M = (W, R, \mathfrak{J})$ *eine* S*-Kripke-Struktur,* $p \in M$ *und* w *eine* M*-Belegung.*

1. *Der* Wert $t^{M,p,w}$ eines Terms t *(d.h. einer Variablen oder eines Konstantensymbols) in der Kripke-Struktur* M *und unter der Belegung* w *ist wie folgt definiert:*

   (a) *Für* $x \in V$: $\quad x^{M,p,w} := w(x)$.

   (b) *Für ein Konstantensymbol* $c \in \mathcal{K}$: $\quad c^{M,p,w} := \begin{cases} c^{\mathfrak{J}} & \text{falls } c \in \mathcal{K}_p \\ \text{nicht definiert} & \text{sonst.} \end{cases}$

2.

   (a) *Für eine atomare Formel* $Q(t_1, \ldots, t_n)$:
   $M, p, w \Vdash_{ML} Q(t_1, \ldots, t_n)$ *genau dann, wenn* $t_1^{M,p,w}, \ldots, t_n^{M,p,w}$ *definiert sind und* $\left(t_1^{M,p,w}, \ldots, t_n^{M,p,w}\right) \in Q^{\mathfrak{J}_p}$ *gilt.*
   $\top^{M,p,w} := True$
   $\bot^{M,p,w} := False$

   (b) $M, p, w \Vdash_{ML} \neg\alpha$ *genau dann, wenn nicht* $M, p, w \Vdash_{ML} \alpha$

   (c) $M, p, w \Vdash_{ML} \alpha \wedge \beta$ *genau dann, wenn* $M, p, w \Vdash_{ML} \alpha$ *und* $M, p, w \Vdash_{ML} \beta$

   (d) $M, p, w \Vdash_{ML} \alpha \vee \beta$ *genau dann, wenn* $M, p, w \Vdash_{ML} \alpha$ *oder* $M, p, w \Vdash_{ML} \beta$

   (e) $M, p, w \Vdash_{ML} \alpha \to \beta$ *genau dann, wenn* ($M, p, w \Vdash_{ML} \alpha$ *impliziert* $M, p, w \Vdash_{ML} \beta$)

   (f) $M, p, w \Vdash_{ML} (\forall x)\,\alpha$ *genau dann, wenn für alle* x*-Varianten* w' *mit* $w'(x) \in |\mathfrak{J}_p|$ *gilt* $M, p, w' \Vdash_{ML} \alpha$

   (g) $M, p, w \Vdash_{ML} (\exists x)\,\alpha$ *genau dann, wenn es eine* x*-Variante* w' *mit* $w'(x) \in |\mathfrak{J}_p|$ *und* $M, p, w' \Vdash_{ML} \alpha$ *gibt.*

   (h) $M, p, w \Vdash_{ML} \Box\alpha$ *genau dann, wenn für alle* $q \in W$ *gilt: wenn* pRq, *so* $M, q, w \Vdash_{ML} \alpha$

   (i) $M, p, w \Vdash_{ML} \Diamond\alpha$ *genau dann, wenn es ein* $q \in W$ *mit* pRq *und* $M, q, w \Vdash_{ML} \alpha$ *gibt.*

3. α *ist* allgemeingültig (*oder gültig*) in der Welt p der Kripke-Struktur M (*in Zeichen:* $M, p \Vdash_{ML} \alpha$) *genau dann, wenn für jede* M*-Belegung* w *mit* $w(Var) \subseteq |\mathcal{I}_p|$ *gilt* $M, p, w \Vdash_{ML} \alpha$.

4. α *ist* allgemeingültig (*oder gültig*) in der Kripke-Struktur M (*in Zeichen:* $M \Vdash_{ML} \alpha$) *genau dann, wenn für jede Welt* $p \in M$ *gilt* $M, p \Vdash_{ML} \alpha$.

5. α *ist* allgemeingültig (*oder gültig*) (*in Zeichen:* $\Vdash_{ML} \alpha$) *genau dann, wenn für jede Kripke-Struktur* M *gilt* $M \Vdash_{ML} \alpha$.

Mit der Definition der Semantik von $\Box\alpha$ wird unsere Vorstellung präzisiert, dass in einer Welt p die Aussage „notwendigerweise gilt α" genau dann als wahr anzusehen ist, wenn in allen in der Welt p möglichen alternativen Welten (das sind diejenigen Welten q mit pRq) α gilt.

Notation 4.10 *Im Folgenden wollen wir bei graphischen Veranschaulichungen von $S-$Kripke-Strukturen $M = (W, R, \mathfrak{I})$ von der folgenden Konvention Gebrauch machen:*

Eine Welt $q \in W$ wird graphisch dagestellt durch ein Rechteck, in dem für jedes $n-$stellige Prädikatensymbol P der Signatur S und jedes $(a_1, \ldots, a_n) \in P^{\mathfrak{I}_q}$ $P(a_1, \ldots, a_n)$ aufgelistet wird. D.h. statt der früher schon eingeführten Schreibweise $P^{\mathfrak{I}_q}(a_1, \ldots, a_n)$ für das Zutreffen von $P^{\mathfrak{I}_q}$ auf das $n-$Tupel $(a_1, \ldots, a_n)$ schreiben wir hier $P(a_1, \ldots, a_n)$. Man beachte aber, dass $P(a_1, \ldots, a_n)$ keine Formel darstellt.

Beispiel 4.11

1. *Sei $S = (P)$ eine Signatur mit einem einstelligen Prädikatensymbol P.*
   Gegeben sei eine Kripke-Struktur $(W, R, \mathfrak{I})$ mit $W = \{p\}$, $R = \emptyset$ und $P^{\mathfrak{I}_p} = \emptyset$.
   Es gilt dann:

   (a) $M, p \not\Vdash_{ML} (\exists x)\, P\,(x)$,
   da für eine beliebige M-Belegung w gilt:
   $M, p, w \Vdash_{ML} (\exists x)\, P\,(x)$
   $\Leftrightarrow$ *es existiert $a \in |\mathfrak{I}_p|$ mit $M, p, w\,\{x/a\} \Vdash_{ML} P\,(x)$*
   $\Leftrightarrow$ *es existiert $a \in |\mathfrak{I}_p|$ mit $a \in P^{\mathfrak{I}_p}$.*
   Es gilt aber $P^{\mathfrak{I}_p} = \emptyset$.

   (b) $M, p \not\Vdash_{ML} \Diamond\,(\exists x)\, P\,(x)\,.$
   Sei w eine beliebige M-Belegung w. Es gilt dann:
   $M, p, w \Vdash_{ML} \Diamond\,(\exists x)\, P\,(x)$
   $\Leftrightarrow$ *es existiert $q \in W$ mit pRq und $M, q, w \Vdash_{ML} (\exists x)\, P\,(x)\,.$*
   Es existiert aber kein $q \in W$ mit pRq.

   (c) $M, p \Vdash_{ML} \Box\,(\exists x)\, P\,(x)\,.$
   Sei w eine beliebige M-Belegung w. Es gilt dann:
   $M, p, w \Vdash_{ML} \Box\,(\exists x)\, P\,(x)$
   $\Leftrightarrow$ *für alle $q \in W$: $pRq \;\Rightarrow\; M, q, w \Vdash_{ML} (\exists x)\, P\,(x)\,.$*
   Die Prämisse pRq ist aber nicht erfüllt, so dass also die rechte Seite der Äquivalenz wahr ist.

   (d) *Aus a), b) und c) folgt sofort*
   $M, p \not\Vdash_{ML} \Box\,(\exists x)\, P\,(x) \to (\exists x)\, P\,(x)\,.$
   $M, p \not\Vdash_{ML} \Box\,(\exists x)\, P\,(x) \to \Diamond\,(\exists x)\, P\,(x)\,.$

2. *Sei wiederum $S = (P)$ eine Signatur mit einem einstelligen Prädikatensymbol P.*
   Gegeben sei eine Kripke-Struktur $(W, R, \mathfrak{I})$ mit $W = \{p\}$, $R = \{(p, p)\}$ und $P^{\mathfrak{I}_p} = \emptyset$.
   Wegen $M, p, w \not\Vdash_{ML} (\exists x)\, P\,(x)$ und wegen
   $M, p, w \Vdash_{ML} \Box\,(\exists x)\, P\,(x)$
   $\Leftrightarrow$ *für alle $q \in W$: $pRq \;\Rightarrow\; M, q, w \Vdash_{ML} (\exists x)\, P\,(x)$*
   $\Leftrightarrow\; M, p, w \Vdash_{ML} (\exists x)\, P\,(x)$
   gilt dann für eine beliebige M-Belegung w :
   $M, p, w \not\Vdash_{ML} \Box\,(\exists x)\, P\,(x)$
   und somit
   $M, p \Vdash_{ML} \Box\,(\exists x)\, P\,(x) \to (\exists x)\, P\,(x)$ *und*
   $M, p \Vdash_{ML} \Box\,(\exists x)\, P\,(x) \to \Diamond\,(\exists x)\, P\,(x).$
   Damit erhalten wir
   $M, p, w \not\Vdash_{ML} \Box\,(\Box\,(\exists x)\, P\,(x) \to (\exists x)\, P\,(x)) \to \Box\,(\exists x)\, P\,(x)\,.$

3. *Es sei die Signatur* $S = (P; c, d)$ *mit einem einstelligen Prädikatensymbol* P *gegeben. Wir betrachten eine Kripke-Struktur* $(W, R, \mathfrak{J})$ *mit* $W = \{p, q, r, s\}$ *und* $R = W \times W \setminus \{(a, a) \mid a \in W\}$. *Für die Welteninterpretation* $\mathfrak{J}$ *gelte*

   $|\mathfrak{J}_p| = |\mathfrak{J}_q| = |\mathfrak{J}_r| = |\mathfrak{J}_s| = \{c, d\}$,

   $P^{\mathfrak{J}_p} = P^{\mathfrak{J}_q} = P^{\mathfrak{J}_r} = \{c\}$ *und* $P^{\mathfrak{J}_s} = \{c, d\}$.

   Graphisch veranschaulicht sieht die Kripke-Struktur wie folgt aus:

$$\begin{tikzcd} r\ \boxed{P(c)} \arrow[r, leftrightarrow] \arrow[d, leftrightarrow] \arrow[dr, leftrightarrow] & \boxed{P(c), P(d)}\ s \arrow[d, leftrightarrow] \arrow[dl, leftrightarrow] \\ p\ \boxed{P(c)} \arrow[r, leftrightarrow] & \boxed{P(c)}\ q \end{tikzcd}$$

   Es gilt dann

   $M, p \Vdash_{ML} \Box P(c)$ *und* $M, p \Vdash_{ML} \Diamond P(d)$, *aber nicht* $M, p \nVdash_{ML} \Box\Diamond P(d)$, *da* pRs *und* $M, s, w \nVdash_{ML} \Diamond P(d)$.

Bemerkung 4.12 *Für die Modallogik lassen sich leicht die folgenden Eigenschaften verifizieren.*

1. *Permanenz der Aussagenlogik in der Modallogik.*

   D.h. resultiert eine modallogische Formel aus einer aussagenlogischen Tautologie dadurch, dass Aussagenvariablen durch modallogische Formeln ersetzt worden sind, dann ist diese modallogische Formel ebenfalls allgemeingültig.

2. *Jede allgemeingültige Aussage* α *der Prädikatenlogik, in der keine Konstantensymbole vorkommen, ist auch modallogisch allgemeingültig, d.h. es gilt*

$$\models_{PL} \alpha \ \Rightarrow \Vdash_{ML} \alpha.$$

3. *Für die Modallogik gilt ebenfalls ein Koinzidenztheorem:*

   Sei $\alpha \in Fm_{ML}(S)$. *Ist* M *eine Kripke-Struktur,* $p \in M$, *und sind* w_1 *und* w_2 *zwei* M*-Belegungen, die in der Belegung der in* α *vorkommenden freien Variablen übereinstimmen, so gilt* $M, p, w_1 \Vdash_{ML} \alpha \Leftrightarrow M, p, w_2 \Vdash_{ML} \alpha$.

4. *Analog zur Prädikatenlogik können wir die Semantik für die Formeln* $(\forall x)\,\alpha$ *und* $(\exists x)\,\alpha$ *auch in der folgenden Weise angeben:*

   $M, p, w \Vdash_{ML} (\forall x)\,\alpha$ *genau dann, wenn für alle* $a \in |\mathfrak{J}_p|$ *gilt* $M, p, w\,\{x/a\} \Vdash_{ML} \alpha$ *bzw.*

   $M, p, w \Vdash_{ML} (\exists x)\,\alpha$ *genau dann, wenn ein* $a \in |\mathfrak{J}_p|$ *existiert mit* $M, p, w\,\{x/a\} \Vdash_{ML} \alpha$.

5. *Die Definition der Substitution wird für die Modallogik in kanonischer Weise durch die Festlegungen*

$$(\Box\alpha)\,\theta := \Box\,(\alpha\theta)$$

   und

$$(\Diamond\alpha)\,\theta := \Diamond\,(\alpha\theta)$$

   ergänzt.

6. *Es folgt leicht, dass für die Modallogik das Substitutionslemma in der folgenden Form zur Verfügung steht:*
   Sind $t_1^{M,p,w}, \ldots, t_n^{M,p,w}$ *definiert, dann gilt:*

$$M,p,w\left\{x_1/t_1^{M,p,w},\ldots,x_n/t_n^{M,p,w}\right\} \Vdash_{ML} \alpha \;\Leftrightarrow\; M,p,w \Vdash_{ML} \alpha\{x_1/t_1,\ldots,x_n/t_n\}\,.$$

Lemma 4.13 *Sei* S *eine Signatur, seien ferner* $\alpha, \beta \in Fm_{ML}(S)$.
Es gilt dann

1. *Für beliebige* S*-Kripke-Strukturen* M*, beliebige Welten* $p \in M$ *und* M*-Belegungen* w *sowie Terme* t *zur Signatur* S*, für die* $t^{M,p,w} \in |\mathfrak{I}_p|$ *und* $t^{M,p,w}$ *definiert sind*:
   $M,p,w \Vdash_{ML} (\forall x)\,\alpha \to \alpha\{x/t\}$
   $M,p,w \Vdash_{ML} \alpha\{x/t\} \to \exists x\alpha$
2. $\Vdash_{ML} (\forall x)(\alpha \to \beta) \to ((\forall x)\,\alpha \to (\forall x)\,\beta)$
3. $\Vdash_{ML} \Box\alpha \leftrightarrow \neg\Diamond\neg\alpha \qquad\qquad \Vdash_{ML} \Diamond\alpha \leftrightarrow \neg\Box\neg\alpha$
4. $\Vdash_{ML} \Box(\forall x)\,\alpha \to (\forall x)\,\Box\alpha$
5. $\nVdash_{ML} (\forall x)\,\Box\alpha \to \Box(\forall x)\,\alpha$ (Barcan'sche Formel)
6. $\Vdash_{ML} \Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$
7. $\Vdash_{ML} \alpha \quad\Rightarrow\quad \Vdash_{ML} \Box\alpha$

Beweis. Wir beweisen exemplarisch einige Fälle.

1. Es gelte $M,p,w \Vdash_{ML} (\forall x)\,\alpha$ und sei t ein Term, für den $t^{M,p,w}$ definiert ist. Auf Grund der Semantik-Definition gilt für alle $a \in |\mathfrak{I}_p|$

$$M,p,w\{x/a\} \Vdash_{ML} \alpha.$$

Damit gilt insbesondere $M,p,w\left\{x/t^{M,p,w}\right\} \Vdash_{ML} \alpha$, da $t^{M,p,w}$ definiert ist. Mit dem Substitutionslemma folgt $M,p,w \Vdash_{ML} \alpha\{x/t\}$. Es gilt somit $M,p,w \Vdash_{ML} (\forall x)\,\alpha \to \alpha\{x/t\}$.

Der Beweis für die Formel $\alpha\{x/t\} \to \exists x\alpha$ verläuft entsprechend.

6. M, p und w seien beliebig gewählt und es gelte $M,p,w \Vdash_{ML} \Box(\alpha \to \beta)$ sowie $M,p,w \Vdash_{ML} \Box\alpha$. Es gilt dann für alle $q \in M$ mit pRq sowohl $M,q,w \Vdash_{ML} \alpha \to \beta$ als auch $M,q,w \Vdash_{ML} \alpha$. Daher gilt dann für alle $q \in M$ mit pRq auch $M,q,w \Vdash_{ML} \beta$, folglich $M,p,w \Vdash_{ML} \Box\beta$. Wir erhalten also $M,p,w \Vdash_{ML} \Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$. Hieraus folgt die Behauptung.

7. Es gelte $\Vdash_{ML} \alpha$. Dann gilt auch für alle Kripke-Strukturen M, alle Welten $p \in M$ und alle M-Belegungen w : $M,p,w \Vdash_{ML} \alpha$. Insbesondere gilt dann für alle $q \in M$ mit pRq : $M,q,w \Vdash_{ML} \alpha$. Aufgrund der Semantik-Definition folgt $M,p,w \Vdash_{ML} \Box\alpha$.

■

Es ist anschaulich klar, dass der Wahrheitswert einer modallogischen Aussage α in einer Welt p einer Kripke-Struktur M nur von den Welten abhängen kann, die von p aus über die Erreichbarkeitsrelation tatsächlich erreicht werden können. Um diese Eigenschaft zeigen zu können, führen wir eine geeignete Terminologie ein.

Definition 4.14 *Für eine binäre Relation* $R \subseteq W \times W$ *seien* $R^n \subseteq W \times W$, $n \in \mathbb{N}$, *sowie die* reflexive und transitive Hülle R^* *von* R *folgendermaßen definiert:*
$(p,q) \in R^0$ *genau dann, wenn* $p = q$.
$(p,q) \in R^{n+1}$ *genau dann, wenn ein* $r \in W$ *existiert mit* $(p,r) \in R$ *und* $(r,q) \in R^n$
$R^* := \bigcup_{n \geq 0} R^n$.

Definition 4.15 *Sei $M = (W, R, \mathfrak{I})$ eine Kripke-Struktur und $p \in W$.*
Die Kripke-Struktur

$$M^p := (W^p, R^p, \mathfrak{I}^p)$$

mit

$$\begin{aligned} W^p &= \{q \mid (p,q) \in R^*\} \\ R^p &= R \cap (W^p \times W^p) \\ \mathfrak{I}^p &= \mathfrak{I} \restriction W^p \ , \end{aligned}$$

wobei $\mathfrak{I} \restriction W^p$ die Restriktion von $\mathfrak{I}$ auf die Menge W^p ist, heißt die von der Welt p in M generierte Kripke-Unterstruktur.

Lemma 4.16 *Sei $M = (W, R, \mathfrak{I})$ eine S-Kripke-Struktur, $\alpha \in Fm_{ML}(S)$ eine modallogische Formel. $p, q \in M$ seien Welten mit $q \in M^p$ (d.h. q ist von p aus erreichbar).*
Dann gilt

$$M, q \Vdash_{ML} \alpha \ \Leftrightarrow\ M^p, q \Vdash_{ML} \alpha \, .$$

Beweis. Durch strukturelle Induktion über den Aufbau der modallogischen Formeln α zeigen wir:
Für eine beliebige Welt $q \in M^p$ und eine beliebige M-Belegung w mit $w\,(Var) \subseteq |\mathfrak{I}_q|$ gilt

$$M, q, w \Vdash_{ML} \alpha \ \Leftrightarrow\ M^p, q, w \Vdash_{ML} \alpha \, .$$

(Man beachte, dass auf Grund der Voraussetzung $w\,(Var) \subseteq |\mathfrak{I}_q|$ w insbesondere auch eine M^p-Belegung ist.)

Ind.Basis: Für $\top$ und $\bot$ ist die Behauptung offensichtlich erfüllt.

Sei nun P ein n-stelliges Prädikatensymbol der Signatur S, $t_1, \ldots, t_n \in Tm_{ML}(S)$. Es gilt dann:

$M, q, w \Vdash_{ML} P\,(t_1, \ldots, t_n)$

$\Leftrightarrow$ $t_1^{M,q,w}, \ldots, t_n^{M,q,w}$ sind definiert und $\left(t_1^{M,q,w}, \ldots, t_n^{M,q,w}\right) \in P^{\mathfrak{I}_q}$

$\Leftrightarrow$ $t_1^{M,q,w}, \ldots, t_n^{M,q,w}$ sind definiert und $\left(t_1^{M^p,q,w}, \ldots, t_n^{M^p,q,w}\right) \in P^{\mathfrak{I}_q}$ nach Voraussetzung über w

$\Leftrightarrow$ $M^p, q, w \Vdash_{ML} P\,(t_1, \ldots, t_n)$.

Ind.Schritt: Wir betrachten exemplarisch zwei Fälle.

$(\forall x)\,\alpha$:

$M, q, w \Vdash_{ML} (\forall x)\,\alpha$

$\Leftrightarrow$ für alle $a \in |\mathfrak{I}_q|$ gilt: $M, q, w\,\{x \diagup a\} \Vdash_{ML} \alpha$

$\Leftrightarrow$ für alle $a \in \left|\mathfrak{I}_q^p\right|$ gilt: $M^p, q, w\,\{x \diagup a\} \Vdash_{ML} \alpha$ (nach IV und da $|\mathfrak{I}_q| = \left|\mathfrak{I}_q^p\right|$)

$\Leftrightarrow$ $M^p, q, w \Vdash_{ML} (\forall x)\,\alpha$.

$\Box\alpha$:

$M, q, w \Vdash_{ML} \Box\alpha$

$\Leftrightarrow$ für alle $r \in W$ mit qRr gilt: $M, r, w \Vdash_{ML} \alpha$

$\Leftrightarrow$ für alle $r \in W$ mit qRr gilt: $M^p, r, w \Vdash_{ML} \alpha$ (nach IV)

$\Leftrightarrow$ für alle $r \in W^p$ mit qR^pr gilt: $M^p, r, w \Vdash_{ML} \alpha$

$\Leftrightarrow$ $M^p, q, w \Vdash_{ML} \Box\alpha$.

■

Korollar 4.17 *Sei $M = (W, R, \mathfrak{I})$ eine S-Kripke-Struktur und $\alpha \in Fm_{ML}(S)$ eine modallogische Formel.*
$M \Vdash_{ML} \alpha \ \Leftrightarrow\ \alpha$ ist allgemeingültig in allen generierten Kripke-Unterstrukturen von M.

4.1.2 Der modallogische Folgerungsbegriff

In Analogie zum aussagenlogischen und prädikatenlogischen Folgerungsbegriff wollen wir auch einen modallogischen Folgerungsbegriff einführen. Wir definieren hierzu zunächst einen modallogischen Interpretationsbegriff.

Definition 4.18 *Sei M eine S-Kripke-Struktur, p eine Welt in M und w eine M-Belegung. Dann heißt (M, p, w) eine* modallogische Interpretation.

Ist $I = (M, p, w)$ eine modallogische Interpretation, und gilt für eine modallogische Formel α $M, p, w \Vdash_{ML} \alpha$, so notieren wir dies durch $I \Vdash_{ML} \alpha$. In diesem Fall sagen wir auch, dass I α erzwingt.

Definition 4.19 *Sei Σ eine Menge modallogischer Formeln.*
Eine modallogische Interpretation I heißt Modell *von Σ genau dann, wenn I jede Formel $\beta \in \Sigma$ erzwingt.*

Definition 4.20 *Sei S eine Signatur. Ferner seien $\Sigma \subseteq Fm_{ML}(S)$ und $\alpha \in Fm_{ML}(S)$.*
Aus Σ folgt *(semantisch) α (notiert als $\Sigma \Vdash_{ML} \alpha$) genau dann, wenn für jede modallogische Interpretation I gilt:*

ist I Modell von Σ, so erzwingt I α.

(Andere Sprechweise: α ist modallogische Konsequenz aus Σ)

Beispiel 4.21

1. $\{\Box\alpha, \Diamond(\alpha \to \beta)\} \Vdash_{ML} \Diamond\beta$.
2. *Hingegen gilt i.a. nicht* $\{\alpha\} \Vdash_{ML} \Box\alpha$.

Analog zur Aussagen- und Prädikatenlogik können wir wiederum den Begriff der Konsequenzenmenge definieren. Für diesen können wir dann dieselben Eigenschaften zeigen, die wir schon zuvor erhalten haben. Wir verzichten daher an dieser Stelle auf Beweise.

Definition 4.22 *Für $\Sigma \subseteq Fm_{ML}(S)$ ist*

$$Cn(\Sigma) := \{\alpha \in Fm_{PL}(S) \mid \Sigma \Vdash_{ML} \alpha\}$$

die (modallogische) Konsequenzenmenge von Σ.

Proposition 4.23 *Σ, Σ_1 und Σ_2 seien Mengen modallogischer Formeln zu einer Signatur S. Dann gilt:*

1. $\Sigma_1 \subseteq \Sigma_2 \;\Rightarrow\; Cn(\Sigma_1) \subseteq Cn(\Sigma_2)$.
2. $\Sigma \subseteq Cn(\Sigma)$.
3. *Die Menge der allgemeingültigen Formeln ist für jedes $\Sigma \subseteq Fm_{ML}$ in $Cn(\Sigma)$ enthalten.*
4. $Cn(Cn(\Sigma)) \subseteq Cn(\Sigma)$.

Lemma 4.24 *Seien $\Sigma \subseteq Fm_{ML}(S)$ und $\alpha, \beta \in Fm_{ML}(S)$.*
Dann gilt:

1. *Aus $\Sigma \Vdash_{ML} \alpha \to \beta$ und $\Sigma \Vdash_{ML} \alpha$ folgt $\Sigma \Vdash_{ML} \beta$.*
   (semantische Modus ponens Regel)
2. *Wenn $\Sigma \Vdash_{ML} \alpha \to \beta$, so $\Sigma \cup \{\alpha\} \Vdash_{ML} \beta$.*
   (Ableitungstheorem)
3. *Wenn $\Sigma \cup \{\alpha\} \Vdash_{ML} \beta$, so $\Sigma \Vdash_{ML} \alpha \to \beta$.*

4.1.3 Modallogische Signaturerweiterungen

Im Zusammenhang mit dem noch einzuführenden modallogischen Tableau-Kalkül werden wir eine gegebene Signatur um neue Konstantensymbole erweitern. Wir wollen daher für Konstantenerweiterungen im modallogischen Fall die später noch benötigten Eigenschaften zur Verfügung stellen.

Definition 4.25 *Sei S' eine Konstantenerweiterung der Signatur S.*
Eine S'-Erweiterung einer modallogischen S-Struktur $M = (W, R, \mathfrak{J})$ *ist eine modallogische S'-Struktur $(W, R, \mathfrak{J}')$, für die für jedes $p \in W$ $\mathfrak{J}'_p$ eine Erweiterung (um Konstanten) von $\mathfrak{J}_p$ ist.*

Lemma 4.26 *Sei S' eine Konstantenerweiterung der Signatur S.*

1. *Sei $M = (W, R, \mathfrak{J})$ eine modallogische S-Struktur, $p \in W$ und w eine M-Belegung. Ist M' eine Erweiterung der S-Struktur M zu einer modallogischen S'-Struktur, dann gilt für Formeln $\alpha \in Fm_{ML}(S)$*
$$M, p, w \Vdash_{ML} \alpha \;\Leftrightarrow\; M', p, w \Vdash_{ML} \alpha.$$
(Man beachte, dass w auch eine M'-Belegung ist.)

2. *Ist α eine allgemeingültige Formel der Signatur S, dann ist α auch eine allgemeingültige Formel der Signatur S'.*

Den sehr einfachen Beweis (mittels struktureller Induktion) führen wir nicht aus.

Lemma 4.27 *Sei die Signatur $S' = S(\{c\})$ eine Konstantenerweiterung der Signatur $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$. Ferner sei $\alpha \in Fm_{ML}(S)$ und $\Sigma \subseteq Fm_{ML}(S)$. Es ist dann $\alpha\{x/c\} \in Fm_{ML}(S')$.*
Es gilt:

1. $\Vdash_{ML} (\forall x)\,\alpha \Leftrightarrow\; \Vdash_{ML} \alpha\{x/c\}$
$\Vdash_{ML} \neg(\exists x)\,\alpha \Leftrightarrow\; \Vdash_{ML} \neg\alpha\{x/c\}$

2. *Hat $\Sigma \cup \{(\exists x)\,\alpha\}$ (bzw. $\Sigma \cup \{\neg(\forall x)\,\alpha\}$) ein Modell zur Signatur S, so hat $\Sigma \cup \{(\exists x)\,\alpha, \alpha\{x/c\}\}$ (bzw. $\Sigma \cup \{\neg(\forall x)\,\alpha, \neg\,\alpha\{x/c\}\}$) ein Modell zur Signatur S'.*

Der Beweis verläuft mit den notwendigen Anpassungen weitgehend analog zu Lemma 2.69, so dass wir auf die Ausführung verzichten.

4.2 Spezielle Modallogiken

Der sehr allgemein gehaltene Ansatz der Kripke-Strukturen bietet die Möglichkeit, durch zusätzliche Anforderungen an die Erreichbarkeitsrelation modallogische Formelklassen – und damit Modallogiken – auszuzeichnen. Dies ist einer der Gründe, warum es so vielfältige Anwendungen der Modallogik gibt. Wir werden im Folgenden die wichtigsten Modallogiken vorstellen.

4.2.1 Rahmeneigenschaften und die Definition klassischer Modallogiken

Definition 4.28 *Sei $\mathcal{F} = (W, R)$ ein Kripke-Rahmen.*
Die binäre Relation $R \subseteq W \times W$ heißt

1. reflexiv *genau dann, wenn für alle $p \in W$ gilt: pRp.*

2. irreflexiv *genau dann, wenn für kein $p \in W$ pRp gilt.*

3. symmetrisch *genau dann, wenn für alle $p, q \in W$ gilt: $pRq \Rightarrow qRp$.*

4. transitiv *genau dann, wenn für alle $p, q, r \in W$ gilt: wenn pRq und qRr, so pRr.*

5. seriell *genau dann, wenn für alle $p \in W$ ein $q \in W$ existiert mit pRq.*

6. euklidisch *genau dann, wenn für alle* $p, q, r \in W$ *gilt: wenn* pRq *und* pRr, *so* qRr.

7. schwach dicht *genau dann, wenn für alle* $p, q \in W$ *gilt: wenn* pRq, *so existiert* $r \in W$ *mit* pRr *und* rRq.

8. schwach linear (*engl.: weak connected*) *genau dann, wenn für alle* $p, q, r \in W$ *gilt: wenn* pRq *und* pRr, *so* qRr *oder* $q = r$ *oder* rRq.

Der Kripke-Rahmen $\mathfrak{F}$ *heißt reflexiv (irreflexiv, symmetrisch, transitiv, seriell, euklidisch, schwach dicht, schwach linear) genau dann, wenn die Relation* R *diese Eigenschaft hat.* $\mathfrak{F}$ *heißt* G-Rahmen *genau dann, wenn* R *eine transitive Relation ist und jede nichtleere Teilmenge* W' *von* W R*-maximale Elemente besitzt, d.h. es gibt* $p \in W'$, *so dass für kein* $q \in W'$ pRq *gilt. Ein G-Rahmen kann daher nie reflexiv sein.*

Bemerkung 4.29 *Aus der Definition geht hervor, dass - ausgenommen die G-Rahmen - die aufgeführten Rahmeneigenschaften mit den Ausdrucksmittel der Prädikatenlogik definierbar sind. G-Rahmen sind hingegen nicht mit Mitteln der Prädikatenlogik (erster Stufe) definierbar.*

Es zeigt sich, dass – ausgenommen die irreflexiven Rahmen – jeder dieser Rahmen über ein passendes Axiomenschema charakterisierbar ist. Es hat sich in der modallogischen Literatur eingebürgert, diese Axiomenschemata mit bestimmten Kürzeln zu bezeichnen.

Kürzel:	**Axiomenschema**
D	$\Box\alpha \to \Diamond\alpha$
T	$\Box\alpha \to \alpha$
4	$\Box\alpha \to \Box\Box\alpha$
B	$\alpha \to \Box\Diamond\alpha$
5	$\Diamond\alpha \to \Box\Diamond\alpha$
G	$\Box(\Box\alpha \to \alpha) \to \Box\alpha$
X	$\Box\Box\alpha \to \Box\alpha$
L	$\Box(\alpha \land \Box\alpha \to \beta) \lor \Box(\beta \land \Box\beta \to \alpha)$

(Statt G - nach Gödel - findet man auch die Kürzel GL für Gödel-Löb bzw. KW)

Definition 4.30 *Eine modallogische Formel* $\alpha \in Fm_{ML}(S)$ *heißt allgemeingültig im Rahmen*$\mathfrak{F} = (W, R)$ (*notiert als:* $\mathfrak{F} \Vdash_{ML} \alpha$) *genau dann, wenn für alle auf diesem Rahmen* $\mathfrak{F} = (W, R)$ *basierenden Kripke-Strukturen* $(W, R, \mathfrak{I})$ *gilt:* $(W, R, \mathfrak{I}) \Vdash_{ML} \alpha$.

Beispiel 4.31 *Nach Lemma 4.13 (6) ist*

$$\Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$$

in allen Rahmen allgemeingültig.

Theorem 4.32 *Sei* $\mathcal{F} = (W, R)$ *ein Kripke-Rahmen.*

1. D *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *seriell.*

2. T *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *reflexiv.*

3. 4 *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *transitiv.*

4. B *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *symmetrisch.*

5. 5 *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *euklidisch.*

6. G *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *ist ein G-Rahmen.*

7. X *ist allgemeingültig im Kripke-Rahmen* $\mathcal{F} \Leftrightarrow \mathcal{F}$ *ist schwach dicht.*

8. *L ist allgemeingültig im Kripke-Rahmen $\mathcal{F}$ $\Leftrightarrow$ $\mathcal{F}$ ist schwach linear.*

Beweis.

1. Übungsaufgabe.

2. „$\Leftarrow$":

   Es sei $\mathcal{F}$ ein reflexiver Rahmen.

   Es seien dann eine Welteninterpretation $\mathfrak{J}$, eine Welt $p \in W$ und eine Belegung w zur Kripke-Struktur $M = (W, R, \mathfrak{J})$ beliebig gewählt. Es gelte $M, p, w \Vdash_{ML} \Box\alpha$. Da R reflexiv ist, gilt somit insbesondere $M, p, w \Vdash_{ML} \alpha$. Damit gilt dann aber auch $M, p, w \Vdash_{ML} \Box\alpha \to \alpha$. Da $\mathfrak{J}$, p und w beliebig gewählt waren, gilt also $(W, R) \Vdash_{ML} \Box\alpha \to \alpha$.

   „$\Rightarrow$":

   Es sei $\alpha = (\exists x)\, P\,(x) \in Fm_{ML}\,(S)$. Falls in S nur ein n-stelliges Prädikatensymbol P mit $n > 1$ vorhanden ist, wähle $\alpha = (\exists x_1) \ldots (\exists x_n)\, P\,(x_1, \ldots, x_n)$.

   Sei $\mathfrak{F}$ nicht reflexiv. Es existiert dann ein $p \in W$, für das pRp nicht gilt. Sei $\mathfrak{J}$ eine Welteninterpretation mit $\mathfrak{J}_p \nvDash_{PL} \alpha$ sowie $\mathfrak{J}_q \models_{PL} \alpha$ für alle $q \in W$ mit pRq (man beachte, dass $\mathfrak{J}_p$ und $\mathfrak{J}_q$ prädikatenlogische Interpretationen sind und hier die prädikatenlogische Modelleigenschaft betrachtet wird). Eine solche Welteninterpretation $\mathfrak{J}$ mit dieser Eigenschaft existiert offensichtlich.Damit gilt aber für eine beliebige Belegung w $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\alpha$, aber nicht $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \alpha$. T kann daher im Rahmen $\mathfrak{F}$ nicht gelten.

3. „$\Leftarrow$":

   Sei (W, R) ein transitiver Rahmen. $\mathfrak{J}$, p und w seien wiederum beliebig gewählt.

   Es gelte $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\alpha$. Dann gilt

   $$\text{für alle } q \in W \text{ mit } pRq \quad (W, R, \mathfrak{J})\,, q, w \Vdash_{ML} \alpha. \tag{4.1}$$

   Sei r eine beliebig gewählte Welt in W mit qRr. Da R transitiv ist, gilt wegen pRq und qRr auch pRr. Wegen (1) gilt somit $(W, R, \mathfrak{J})\,, r, w \Vdash_{ML} \alpha$. Da r eine beliebig gewählte Welt mit der Eigenschaft qRr war, gilt damit $(W, R, \mathfrak{J})\,, q, w \Vdash_{ML} \Box\alpha$. Da dies für jede Welt q mit pRq gilt, folgt daher weiter $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\Box\alpha$. Damit ergibt sich dann $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\alpha \to \Box\Box\alpha$ und folglich $(W, R) \Vdash_{ML} \Box\alpha \to \Box\Box\alpha$.

   „$\Rightarrow$":

   Es sei α wiederum eine prädikatenlogische Aussage, z.B. $\alpha = (\exists x)\, P\,(x) \in Fm_{ML}\,(S)$.

   Sei $\mathfrak{F}$ nicht transitiv., d.h. es existieren $p, q, r \in W$ mit pRq und qRr, aber nicht pRr.

   Sei $\mathfrak{J}$ eine Welteninterpretation mit $\mathfrak{J}_p \models_{PL} \alpha$ sowie $\mathfrak{J}_t \models \alpha$ für alle $t \in W$ mit pRt. Ferner gelte $\mathfrak{J}_r \nvDash_{PL} \alpha$. Insbesondere gilt $\mathfrak{J}_q \models_{PL} \alpha$. Es ist wiederum klar, dass eine solche Welteninterpretation existiert. w sei eine beliebige Belegung. Damit gilt dann $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\alpha$, aber nicht $(W, R, \mathfrak{J})\,, q, w \Vdash_{ML} \Box\alpha$. Folglich gilt auch nicht $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\Box\alpha$. Damit ist 4 in $\mathfrak{F}$ nicht gültig.

4. „$\Leftarrow$":

   Sei (W, R) ein symmetrischer Rahmen. $\mathfrak{J}$, p und w seien beliebig gewählt. Da R symmetrisch ist, gilt für jedes $q \in W$ mit pRq auch qRp.

   Es gelte $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \alpha$. Sei $q \in W$ mit pRq beliebig gewählt. Dann gilt $(W, R, \mathfrak{J})\,, q, w \Vdash_{ML} \Diamond\alpha$ wegen qRp. Da q beliebig gewählt war, gilt somit $(W, R, \mathfrak{J})\,, p, w \Vdash_{ML} \Box\Diamond\alpha$.

Es folgt $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \alpha \to \Box\Diamond\alpha$ und damit auch $(W, R) \Vdash_{ML} \alpha \to \Box\Diamond\alpha$.

„$\Rightarrow$":

Es sei wiederum α eine prädikatenlogische Aussage, o.B.d.A. $\alpha = (\exists x) P(x) \in Fm_{ML}(S)$.

Sei $\mathfrak{F}$ nicht symmetrisch. Es existieren dann $p, q \in W$ mit pRq, aber nicht qRp.

Sei $\mathfrak{I}$ eine Welteninterpretation mit $\mathfrak{I}_p \models_{PL} \alpha$ und $\mathfrak{I}_r \not\models_{PL} \alpha$ für alle $r \in W$ mit qRr. Dann gilt für eine beliebige Belegung w $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \alpha$. Weiter haben wir $(W, R, \mathfrak{I}), q, w \nVdash_{ML} \Diamond\alpha$. Hieraus folgt $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \Box\Diamond\alpha$ und damit $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \alpha \to \Box\Diamond\alpha$. Damit gilt B nicht im Rahmen $\mathfrak{F}$.

5.

„$\Leftarrow$":

Sei (W, R) ein euklidischer Rahmen. $\mathfrak{I}$, p und w seien beliebig gewählt.

Es gelte $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Diamond\alpha$. Es existiert dann ein $r \in W$ mit pRr und $(W, R, \mathfrak{I}), r, w \Vdash_{ML} \alpha$. Sei $q \in W$ beliebig gewählt mit pRq. Da R euklidisch ist, gilt dann qRr und damit auch $(W, R, \mathfrak{I}), q, w \Vdash_{ML} \Diamond\alpha$.

Da q als beliebig mit der Eigenschaft pRq gewählt war, folgt dann $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Box\Diamond\alpha$. Daher gilt $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Diamond\alpha \to \Box\Diamond\alpha$. Es folgt $(W, R) \Vdash_{ML} \Diamond\alpha \to \Box\Diamond\alpha$.

„$\Rightarrow$":

Es sei $\alpha = (\exists x) P(x) \in Fm_{ML}(S)$.

Sei $\mathfrak{F}$ nicht euklidisch. Es existieren dann $p, q, r \in W$ mit pRq, pRr, aber nicht qRr.

Sei $\mathfrak{I}$ eine Welteninterpretation mit $\mathfrak{I}_q \models_{PL} \alpha$ und $\mathfrak{I}_t \not\models_{PL} \alpha$ für alle $t \in W$ mit qRt. Sei w wiederum eine beliebige Belegung. Damit gilt dann $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Diamond\alpha$ und $(W, R, \mathfrak{I}), q, w \nVdash_{ML} \Diamond\alpha$, also $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \Box\Diamond\alpha$. Daher gilt nicht $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Diamond\alpha \to \Box\Diamond\alpha$.

5 ist daher nicht gültig im Rahmen $\mathfrak{F}$.

6.

„$\Leftarrow$":

Sei $\mathfrak{F} = (W, R)$ ein G-Rahmen.

Sei $p \in W$ und $\alpha \in Fm_{ML}(S)$ beliebig. Wir wollen zeigen, dass $\Box(\Box\alpha \to \alpha) \to \Box\alpha$ im Rahmen $\mathfrak{F}$ allgemeingültig ist.

Angenommen, es gilt für eine Belegung w $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \Box\alpha$ für eine Welteninterpretation $\mathfrak{I}$. Sei

$$W' := \{q \in W \mid (p, q) \in R \text{ und } (W, R, \mathfrak{I}), q, w \nVdash_{ML} \alpha\}.$$

Wegen $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \Box\alpha$ gibt es ein $q \in W$ mit $(p, q) \in R$ und $(W, R, \mathfrak{I}), q, w \nVdash_{ML} \alpha$. Also $q \in W'$.

In W' existiert ein R-maximales Element r. Nach Definition von W' gilt $(p, r) \in R$ und $(W, R, \mathfrak{I}), r, w \nVdash_{ML} \alpha$.

Gibt es in W ein r' mit $(r, r') \in R$, so ist $r' \notin W'$. Wegen der Transitivität von R gilt für dieses r' $(p, r') \in R$ und, da $r' \notin W'$, also $(W, R, \mathfrak{I}), r', w \Vdash_{ML} \alpha$. Da dies für ein beliebiges r' mit $(r, r') \in R$ gilt, erhalten wir $(W, R, \mathfrak{I}), r, w \Vdash_{ML} \Box\alpha$, also $(W, R, \mathfrak{I}), r, w \nVdash_{ML} \Box\alpha \to \alpha$ und damit $(W, R, \mathfrak{I}), p, w \nVdash_{ML} \Box(\Box\alpha \to \alpha)$.

Somit gilt $(W, R, \mathfrak{I}), p, w \Vdash_{ML} \Box(\Box\alpha \to \alpha) \to \Box\alpha$. Es folgt $(W, R, \mathfrak{I}), p \Vdash_{ML} \Box(\Box\alpha \to \alpha) \to \Box\alpha$ für beliebiges $p \in W$ und beliebige Welteninterpretation $\mathfrak{I}$, d.h. $\Box(\Box\alpha \to \alpha) \to \Box\alpha$ ist allgemeingültig im Rahmen $\mathfrak{F}$.

„$\Rightarrow$":

Wir müssen zeigen, dass $\mathcal{F}$ transitiv ist und dass jede nichtleere Teilmenge W' von W R-maximale Elemente besitzt.

(a) Wir zeigen zunächst, dass aus „G allgemeingültig im Rahmen $\mathfrak{F} = (W, R)$" folgt, dass der Rahmen $\mathfrak{F}$ transitiv ist.
Wir zeigen hierzu die Kontraposition
$\mathfrak{F}$ nicht transitiv $\Rightarrow$ G nicht allgemeingültig im Rahmen $\mathfrak{F}$:
Sei also $\mathfrak{F}$ nicht transitiv. Es gibt dann $p, q, r \in W$ mit $(p, q), (q, r) \in R$ und $(p, r) \notin R$.
Sei $\alpha = (\exists x) P(x) \in Fm_{ML}(S)$. $\mathfrak{J}$ sei eine Welteninterpretation mit $\mathfrak{J}_q \not\models_{PL} \alpha$, $\mathfrak{J}_r \not\models_{PL} \alpha$ und $\mathfrak{J}_{p'} \models_{PL} \alpha$ für alle $p' \in W$, $p' \neq q$ und $p' \neq r$.
Es gilt dann $(W, R, \mathfrak{J}), q \not\Vdash_{ML} \alpha$ und damit $(W, R, \mathfrak{J}), p \not\Vdash_{ML} \Box\alpha$.
Weiter gilt für $p' \in W$ mit $(p, p') \in R$ und $p' \neq q$ $(W, R, \mathfrak{J}), p' \models_{ML} \alpha$, also $(W, R, \mathfrak{J}), p' \Vdash_{ML} \Box\alpha \to \alpha$.
Wegen $(W, R, \mathfrak{J}), r \not\Vdash_{ML} \alpha$ gilt $(W, R, \mathfrak{J}), q \not\Vdash_{ML} \Box\alpha$, also $(W, R, \mathfrak{J}), q \Vdash_{ML} \Box\alpha \to \alpha$.
Insgesamt ergibt sich damit $(W, R, \mathfrak{J}), p \Vdash_{ML} \Box(\Box\alpha \to \alpha)$, so dass gilt

$$(W, R, \mathfrak{J}), p \not\Vdash_{ML} \Box(\Box\alpha \to \alpha) \to \Box\alpha.$$

Damit ist also G nicht allgemeingültig im Rahmen $\mathfrak{F}$.

(b) Es ist weiter zu zeigen, dass jede nichtleere Teilmenge $W' \subseteq W$ R-maximale Elemente enthält.
Sei also $W' \subseteq W$ nichtleer. Angenommen, W' besitzt keine R-maximalen Elemente.
Sei $p \in W'$. Sei ferner $\alpha = (\exists x) P(x) \in Fm_{ML}(S)$ und $\mathfrak{J}$ eine Welteninterpretation mit der Eigenschaft, dass für Welten q aus W gilt: $\mathfrak{J}_q \models_{PL} \alpha \Leftrightarrow q \notin W'$.
Sei $q \in W$ beliebig gewählt mit $(p, q) \in R$.
Falls $q \notin W'$, dann gilt $(W, R, \mathfrak{J}), q \Vdash_{ML} \alpha$, also $(W, R, \mathfrak{J}), q \Vdash_{ML} \Box\alpha \to \alpha$.
Falls $q \in W'$, dann existiert ein $r \in W'$ mit $(q, r) \in R$, da W' keine R-maximalen Elemente besitzt.
Es gilt dann $(W, R, \mathfrak{J}), r \not\Vdash_{ML} \alpha$, also $(W, R, \mathfrak{J}), q \not\Vdash_{ML} \Box\alpha$ und damit $(W, R, \mathfrak{J}), q \Vdash_{ML} \Box\alpha \to \alpha$.

Da in beiden Fällen für ein beliebiges $q \in W$ mit $(p, q) \in R$ gilt $(W, R, \mathfrak{J}), q \Vdash_{ML} \Box\alpha \to \alpha$, gilt also auch $(W, R, \mathfrak{J}), p \Vdash_{ML} \Box(\Box\alpha \to \alpha)$. Da zu p ein $p' \in W'$ existiert mit $(p, p') \in R$, gilt dann $(W, R, \mathfrak{J}), p' \not\Vdash_{ML} \alpha$ und $(W, R, \mathfrak{J}), p \not\Vdash_{ML} \Box\alpha$.

Insgesamt folgt also $(W, R, \mathfrak{J}), p \not\Vdash_{ML} \Box(\Box\alpha \to \alpha) \to \Box\alpha$. $\mathfrak{F}$ ist somit kein G-Rahmen.

7. Übungsaufgabe.

8. „$\Leftarrow$":

Wir zeigen die Kontraposition. Sei L nicht allgemeingültig im Ramen (W, R). Es gibt dann Formeln $\alpha, \beta \in Fm_{ML}(S)$, eine Welteninterpretation $\mathfrak{J}$, eien Welt p und eine Belegung w mit $(W, R, \mathfrak{J}), p, w \not\Vdash_{ML} \Box(\alpha \wedge \Box\alpha \to \beta) \vee \Box(\beta \wedge \Box\beta \to \alpha)$. Dies ist äquivalent zu $(W, R, \mathfrak{J}), p, w \Vdash_{ML} \neg\Box(\alpha \wedge \Box\alpha \to \beta) \wedge \neg\Box(\beta \wedge \Box\beta \to \alpha)$ und damit auch äquivalent zu $(W, R, \mathfrak{J}), p, w \Vdash_{ML} \Diamond(\alpha \wedge \Box\alpha \wedge \neg\beta) \wedge \Diamond(\beta \wedge \Box\beta \wedge \neg\alpha)$.
Es gibt daher Welten q und r mit pRq und pRr sowie $(W, R, \mathfrak{J}), q, w \Vdash_{ML} \alpha \wedge \Box\alpha \wedge \neg\beta$ und $(W, R, \mathfrak{J}), r, w \Vdash_{ML} \beta \wedge \Box\beta \wedge \neg\alpha$. Damit folgt sofort $q \neq r$. Wegen $(W, R, \mathfrak{J}), q, w \Vdash_{ML} \Box\alpha$ und $(W, R, \mathfrak{J}), r, w \Vdash_{ML} \neg\alpha$ ergibt sich, dass $(q, r) \notin R$. Analog folgt $(r, q) \notin R$. Damit ist (W, R) kein schwach linearer Rahmen.

„$\Rightarrow$": Übungsaufgabe.

■

Auf der Grundlage von Theorem 4.32 werden die folgenden Modallogiken definiert. Statt von Kripke-Rahmen sprechen wir im Folgenden auch einfach von Rahmen.

Definition 4.33
K $:= \{\alpha \in Fm_{ML}(S) \mid \Vdash_{ML} \alpha\}$
D $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für serielle Rahmen*$\}$
T $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive Rahmen*$\}$
S4 $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive und transitive Rahmen*$\}$
S4.3 $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive, transitive und schwach lineare Rahmen*$\}$
B $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive und symmetrische Rahmen*$\}$
S5 $:= \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive und euklidische Rahmen*$\}$
G := $\{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für G-Rahmen*$\}$

Bemerkung 4.34 *Die Bezeichnung* $\boldsymbol{K}$ *(zu Ehren von S. Kripke) für die Menge aller modallogischen Formeln rührt daher, dass unter Verwendung des Axiomenschemas*

$$K: \qquad \Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$$

als einzigem modallogischen Axiomenschema eine vollständige Hilbert-Axiomatisierung der Menge der allgemeingültigen modallogischen Formeln möglich ist.
Unter Verwendung der weiteren Axiomenschemata können die restlichen Modallogiken axiomatisiert werden: z.B. $\boldsymbol{S4}$ *durch die Verwendung von* K, T, *und* 4, *so dass sich in der Literatur gelegentlich die Bezeichnung* $\boldsymbol{KT4}$ *für* $\boldsymbol{S4}$ *findet.* $\boldsymbol{S5}$ *entspricht dann* $\boldsymbol{KT5}$ *usw. .*

Bemerkung 4.35 *Es gilt*
$\boldsymbol{S5} = \{\alpha \in Fm_{ML}(S) \mid \alpha$ *allgemeingültig für reflexive, symmetrische und transitive Rahmen*$\}$, *d.h.* α *ist allgemeingültig in allen Rahmen* (W, R), *für die* R *eine Äquivalenzrelation ist. (Für* $\boldsymbol{S5}$ *findet sich daher auch die Bezeichnung* $\boldsymbol{KTB4}$*.)*

Bemerkung 4.36 *Die zuvor eingeführten modallogischen Systeme* L *mit* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$ *haben die folgende, leicht zu beweisende Abschlusseigenschaft.*
Für modallogische Formeln $\alpha \in Fm_{ML}(S)$ *gilt*

$$\alpha \in L \;\; \Rightarrow \;\; \Box\alpha \in L\,.$$

Modallogische Systeme mit dieser Eigenschaft werden als normale modallogische Systeme *bezeichnet.*

Notation 4.37 *Sei* $\alpha \in Fm_{ML}(S)$, $L \in \{\boldsymbol{K}, \boldsymbol{D}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$. *Für jedes dieser Logiksysteme* L *führen wir eine Gültigkeitsrelation* $\Vdash_L$ *ein durch die Definition*

$$\Vdash_L \alpha \;:\Leftrightarrow\; \alpha \in L\,.$$

Lemma 4.38 *Seien* $\alpha, \beta \in Fm_{ML}(S)$. *Es gilt dann*

1. $\Vdash_K \alpha \;\Rightarrow \Vdash_L \alpha$ *für* $L \in \{\boldsymbol{T}, \boldsymbol{D}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$
2. $\Vdash_T \alpha \;\Rightarrow \Vdash_L \alpha$ *für* $L \in \{\boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}\}$
3. $\Vdash_B \alpha \;\Rightarrow \Vdash_{S5} \alpha$
4. $\Vdash_{S4} \alpha \;\Rightarrow \Vdash_{S5} \alpha$
5. $\Vdash_L \alpha \to \beta \;\;\Rightarrow\;\; \Vdash_L \Box\alpha \to \Box\beta$ *für* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$
6. $\Vdash_L \alpha \to \beta \;\;\Rightarrow\;\; \Vdash_L \Diamond\alpha \to \Diamond\beta$ *für* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$
7. $\Vdash_{K4} \alpha \;\Rightarrow \Vdash_G \alpha$.

Beweis. Einfache Übungsaufgabe. ■

4.2.2 Modalitäten

Unserem Aufbau der formalen Sprache der Modallogik liegen die Modaloperatoren $\Box$ und $\Diamond$ zugrunde. Eine endliche Folge solcher Modaloperatoren kann aber selbst als eine Modalität angesehen werden, so dass sich die Frage stellt, ob diese Modalitäten alle verschieden sind. Von Bedeutung ist dies dann, wenn man Normalformen für modallogische Formeln benötigt wie z.B. bei der Untersuchung von modallogischen Resolutionskalkülen.

Definition 4.39 *Ist* O *eine endliche Folge von Operatoren aus* $\{\neg, \Box, \Diamond\}$, *so heißt* O *eine* Modalität. *Die leere Folge notieren wir mit* ε.

In analoger Weise zu dem Negationsoperator $\neg$ und den Modaloperatoren $\Box$ und $\Diamond$ wird mit einer Modalität O aus einer Formel α die modallogische Formel $O\alpha$ erzeugt, z.B. wird aus der Modalität $\Box\Diamond$ und der atomaren Formel $P(c)$ die Formel $\Box\Diamond P(c)$ gebildet. (Wir haben hier diese informelle Darstellung gewählt, um auf der formalen Ebene die Sprache der Modallogik nicht erweitern zu müssen.)

Definition 4.40 *Zwei Modalitäten O_1 und O_2 heißen* äquivalent bezüglich des modallogischen Systems $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{S4.3}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$ *genau dann, wenn für alle $\alpha \in Fm_{ML}(S)$ gilt*

$$\Vdash_L O_1\alpha \leftrightarrow O_2\alpha.$$

Sind O_1 und O_2 nicht äquivalent, so heißen sie verschieden.

Modalitäten in $S5$

Wir beweisen zunächst eine Reduktionseigenschaft für die Modaloperatoren in $\boldsymbol{S5}$:

Lemma 4.41 *Sei $\alpha \in Fm_{ML}(S)$. Es gilt dann*

1. $\Vdash_{S5} \Box\alpha \leftrightarrow \Box\Box\alpha$
2. $\Vdash_{S5} \Diamond\alpha \leftrightarrow \Diamond\Diamond\alpha$
3. $\Vdash_{S5} \Box\alpha \leftrightarrow \Diamond\Box\alpha$
4. $\Vdash_{S5} \Diamond\alpha \leftrightarrow \Box\Diamond\alpha$.

Beweis.

1. Da $\boldsymbol{S5} = \boldsymbol{KTB4}$, gilt T und 4 in $\boldsymbol{S5}$, d.h. insbesondere gilt $\Vdash_{S5} \Box\Box\alpha \rightarrow \Box\alpha$ und $\Vdash_{S5} \Box\alpha \rightarrow \Box\Box\alpha$.
2. Folgt unter Verwendung von $\Vdash_{ML} \Diamond\alpha \rightarrow \neg\Box\neg\alpha$.
3. Es gilt $\Vdash_{S5} \beta \rightarrow \Diamond\beta$ für eine beliebige Formel $\beta \in Fm_{ML}(S)$, da in $\boldsymbol{S5}$ reflexive Rahmen zugrundegelegt werden. Speziell für $\beta = \Box\alpha$ folgt $\Vdash_{S5} \Box\alpha \rightarrow \Diamond\Box\alpha$.

   In euklidischen Rahmen gilt zudem $\Vdash_{S5} \Diamond\neg\alpha \rightarrow \Box\Diamond\neg\alpha$, also auch die Kontraposition hierzu: $\Vdash_{S5} \neg\Box\Diamond\neg\alpha \rightarrow \neg\Diamond\neg\alpha$. Damit folgt $\Vdash_{S5} \Diamond\Box\alpha \rightarrow \Box\alpha$.
4. Folgt aus 3. unter Verwendung von $\Vdash_{ML} \Diamond\alpha \leftrightarrow \neg\Box\neg\alpha$.

■

Korollar 4.42 *In $\boldsymbol{S5}$ gibt es höchstens 6 verschiedene Modalitäten.*

Für diese 6 Modalitäten kann nun gezeigt werden, dass sie auch wirklich verschieden sind.

Theorem 4.43 *In $S5$ gibt es genau die folgenden 6 Modalitäten:*

$$\varepsilon,\ \neg,\ \Box,\ \neg\Box,\ \Diamond,\ \neg\Diamond.$$

Beweis. Nach dem vorangegangenen Korollar wissen wir, dass es höchstens diese 6 Modalitäten geben kann. Zu zeigen ist daher, dass es zu je zwei Modalitäten O_1, O_2 mit $O_1 \neq O_2$ unter diesen 6 Modalitäten eine Kripke-Struktur $(W, R, \mathfrak{J})$, eine Welt $p \in W$, eine Belegung w sowie einen Ausdruck α gibt mit

$$(W, R, \mathfrak{J}), p, w \Vdash_{S5} O_1\alpha$$

und

$$(W, R, \mathfrak{J}), p, w \nVdash_{S5} O_2\alpha\,.$$

Exemplarisch soll dies für $O_1 = \Diamond$ und $O_2 = \Box$ vorgeführt werden.
Es sei der Rahmen (W, R) mit $W = \{p, q\}$, $R = W \times W$ zugrundegelegt. Es sei $\alpha = (\exists x)\, P(x)$. $\mathfrak{J}$ sei eine Welteninterpretation mit $\mathfrak{J}_p \models_{PL} \alpha$ und $\mathfrak{J}_q \not\models_{PL} \alpha$. Eine solche Welteninterpretation $\mathfrak{J}$ existiert offensichtlich. Damit gilt $(W, R, \mathfrak{J}), p, w \Vdash_{S5} \Diamond\alpha$, aber nicht $(W, R, \mathfrak{J}), p, w \Vdash_{S5} \Box\alpha$. D.h. die beiden Modalitäten $\Box$ und $\Diamond$ sind verschieden. ■

Modalitäten in S4 und in S4.3

Auch hier können wir wiederum eine Reduktionseigenschaft für Modaloperatoren zeigen.

Lemma 4.44 *Sei $\alpha \in Fm_{ML}(S)$. Es gilt dann in **S4** (und in **S4.3**) :*

1. $\Vdash_{S4} \Box\alpha \leftrightarrow \Box\Box\alpha$
2. $\Vdash_{S4} \Diamond\alpha \leftrightarrow \Diamond\Diamond\alpha$
3. $\Vdash_{S4} \Diamond\Box\alpha \leftrightarrow \Diamond\Box\Diamond\Box\alpha$
4. $\Vdash_{S4} \Box\Diamond\alpha \leftrightarrow \Box\Diamond\Box\Diamond\alpha$

Beweis. Der Beweis von 1 und 2 ist identisch mit dem entsprechenden für ***S*5**, so dass wir nur noch die verbleibenden beiden Fälle zeigen müssen.

3. Aus der Reflexivität folgt zunächst

$\Vdash_{S4} \Box\alpha \to \Diamond\Box\alpha$

$\Rightarrow \Vdash_{S4} \Box\Box\alpha \to \Box\Diamond\Box\alpha$ nach Lemma 4.38 5.

$\Rightarrow \Vdash_{S4} \Box\alpha \to \Box\Diamond\Box\alpha$ nach 1.

$\Rightarrow \Vdash_{S4} \Diamond\Box\alpha \to \Diamond\Box\Diamond\Box\alpha$ nach Lemma 4.38 6.

Die umgekehrte Richtung ergibt sich wie folgt:

$\Vdash_{S4} \Box\Diamond\Box\alpha \to \Diamond\Box\alpha$ (Reflexivität)

$\Rightarrow \Vdash_{S4} \Diamond\Box\Diamond\Box\alpha \to \Diamond\Diamond\Box\alpha$ nach Lemma 4.38 6.

$\Rightarrow \Vdash_{S4} \Diamond\Box\Diamond\Box\alpha \to \Diamond\Box\alpha$ nach 2.

4. Folgt leicht aus 3.

■

Theorem 4.45 ***S*4** *hat genau die folgenden 14 verschiedenen Modalitäten:*

$$\varepsilon,\ \Box,\ \Diamond,\ \Box\Diamond,\ \Diamond\Box,\ \Box\Diamond\Box,\ \Diamond\Box\Diamond$$

sowie deren Negationen. Zwischen den nichtnegierten Modalitäten besteht folgende Beziehung

$$\begin{tikzcd}
 & & \Box \arrow[dl] \arrow[ddr] & \\
 & \Box\Diamond\Box \arrow[dl] \arrow[dr] & & \\
\Diamond\Box \arrow[dr] & & \Box\Diamond \arrow[dl] & \varepsilon \arrow[ddl] \\
 & \Diamond\Box\Diamond \arrow[dr] & & \\
 & & \Diamond &
\end{tikzcd}$$

wobei $O_1 \to O_2$ bedeutet: $\Vdash_{S4} O_1\alpha \to O_2\alpha$ für alle $\alpha \in Fm_{ML}(S)$.
Beweis. Aufgrund der Reduktionsbeziehungen aus dem vorangegangenen Lemma kann es höchstens diese Modalitäten geben. Bezüglich der Implikationsbeziehungen gilt:
$\Box \to \varepsilon$ und $\Box\Diamond\Box \to \Diamond\Box$, da $\Vdash_{S4} T$.
$\Box \to \Box\Diamond\Box$ (nach dem Beweis von Lemma 4.44 (3.) im vorangegangenen Lemma)
$\Box\Diamond\Box \to \Box\Diamond$ ergibt sich folgendermaßen:
Es gilt für eine beliebige Formel $\alpha \in Fm_{ML}(S) : \Vdash_{S4} \Box\alpha \to \alpha$.
$\Rightarrow \Vdash_{S4} \Diamond\Box\alpha \to \Diamond\alpha$
$\Rightarrow \Vdash_{S4} \Box\Diamond\Box\alpha \to \Box\Diamond\alpha$.
Für die verbleibenden Implikationsbeziehungen ist der Beweis analog.
Es ist noch zu zeigen, dass es mindestens diese Modalitäten gibt. Hierzu ist nachzuweisen, dass es zu je zwei Modalitäten O_1 und O_2, $O_1 \neq O_2$, unter diesen 14 Modalitäten eine Kripke-Struktur $(W, R, \mathfrak{J})$, eine Welt $p \in W$, eine Belegung w und eine Formel α gibt mit

$$(W, R, \mathfrak{J}), p, w \Vdash_{S4} O_1\alpha$$

und

$$(W, R, \mathfrak{J}), p, w \nVdash_{S4} O_2\alpha\,.$$

Dies ist aber eine einfache Übungsaufgabe. ■
In ***S*4.3** fallen darüberhinaus die Modalitäten $\Box\Diamond\Box$ und $\Diamond\Box$ sowie $\Diamond\Box\Diamond$ und $\Box\Diamond$ zusammen:

Lemma 4.46 *Sei $\alpha \in Fm_{ML}(S)$. Es gilt dann in* ***S*4.3** :

1. $\Vdash_{S4.3} \Diamond\Box\Diamond\alpha \leftrightarrow \Box\Diamond\alpha$.
2. $\Vdash_{S4.3} \Box\Diamond\Box\alpha \leftrightarrow \Diamond\Box\alpha$.

Beweis. Es genügt aus Dualitätsgründen 1. zu zeigen.
$\Vdash_{S4.3} \Box\Diamond\alpha \to \Diamond\Box\Diamond\alpha$ folgt sofort, da in ***T*** und damit auch in ***S*4.3** jede Formel der Gestalt $\beta \to \Diamond\beta$ allgemeingültig ist.
Wir zeigen $\Vdash_{S4.3} \Diamond\Box\Diamond\alpha \to \Box\Diamond\alpha$.
Wir beweisen hierzu die Kontraposition $\Vdash_{S4.3} \neg\Box\Diamond\alpha \to \neg\Diamond\Box\Diamond\alpha$.
Es gelte für eine beliebige Kripke-Struktur $M = (W, R, \mathfrak{J})$ mit reflexiver, transitiver und schwach linearer Erreichbarkeitsrelation R und für eine beliebige Welt $p \in M$ und eine beliebige M-Belegung w $M, p, w \Vdash_{S4.3} \neg\Box\Diamond\alpha$. Damit gilt auch $M, p, w \Vdash_{S4.3} \Diamond\Box\neg\alpha$. Es existiert somit ein $q \in M$ mit pRq, so dass

$$\text{für alle } r \in M \text{ mit } qRr \text{ gilt } M, r, w \Vdash_{S4.3} \neg\alpha. \tag{4.2}$$

Angenommen es gilt $M, p, w \Vdash_{S4.3} \Diamond\Box\Diamond\alpha$. Es existiert dann ein p' mit pRp' und $M, p', w \Vdash_{S4.3} \Box\Diamond\alpha$. Daher gilt

$$\text{für alle } p'' \text{ mit } p'Rp''M, p'', w \Vdash_{S4.3} \Diamond\alpha. \tag{4.3}$$

Da R schwach linear ist, gilt wegen pRq und pRp' entweder $p' = q$ oder $p'Rq$ oder qRp'.

$p' = q$: Wegen (4.3) und der Reflexivität von R erhalten wir $M, q, w \Vdash_{S4.3} \Diamond\alpha$. Da außerdem (4.2) gilt, folgt sofort ein Widerspruch.

$p'Rq$: Wegen (4.3) und $p'Rq$ gilt $M, q, w \Vdash_{S4.3} \Diamond\alpha$. Wegen (4.2) folgt wiederum sofort ein Widerspruch.

qRp' : Aufgrund der Reflexivität von R erhalten wir mit (4.3) $M, p', w \Vdash_{S4.3} \Diamond\alpha$. Es existiert somit ein $r \in M$ mit $M, r, w \Vdash_{S4.3} \alpha$. Mit der Transitivität von R folgt qRr. Damit steht aber $M, r, w \Vdash_{S4.3} \alpha$ im Widerspruch zu (4.2).

In allen drei möglichen Fällen erhalten wir also einen Widerspruch. Es muss daher $M, p, w \Vdash_{S4.3} \neg\Diamond\Box\Diamond\alpha$ gelten. Da M, p und w beliebig gewählt waren, folgt somit $\Vdash_{S4.3} \neg\Box\Diamond\alpha \to \neg\Diamond\Box\Diamond\alpha$. ■

Eine direkte Konsequenz ist dann das folgende

Theorem 4.47 ***S*4.3** *hat genau die folgenden 10 verschiedenen Modalitäten:*

$$\varepsilon,\ \Box,\ \Diamond,\ \Box\Diamond,\ \Diamond\Box$$

sowie deren Negationen. Zwischen den nichtnegierten Modalitäten besteht folgende Beziehung

$$\begin{tikzcd}
 & \Box \arrow[dl] \arrow[d] \\
\Diamond\Box \arrow[d] & \varepsilon \arrow[dd] \\
\Box\Diamond \arrow[dr] & \\
 & \Diamond
\end{tikzcd}$$

wobei $O_1 \to O_2$ bedeutet: $\Vdash_{S4.3} O_1\alpha \to O_2\alpha$ für alle $\alpha \in Fm_{ML}(S)$.

Modalitäten in *K*, *T*, *B* und *G*

Im Unterschied zu den beiden Modallogiken ***S*4** und ***S*5** hat jede dieser Modallogiken unendlich viele verschiedene Modalitäten.

Notation 4.48 $\Box^0 := \varepsilon$, $\Box^{n+1} := \Box\Box^n$ *für* $n \geq 0$.

Theorem 4.49 ***K****,* ***T****,* ***B*** *und* ***G*** *haben unendlich viele verschiedene Modalitäten.*

Beweis. Sei $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{B}\}$. Wir zeigen:
Für alle $m, n \in \mathbb{N}$ mit $n > m \geq 0$ und $\alpha \in Fm_{ML}(S)$ gilt $\nVdash_L \Box^m\alpha \to \Box^n\alpha$.
Hieraus folgt dann sofort die Behauptung des Theorems für die Modallogiken $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{B}\}$.
Seien $m, n \in \mathbb{N}$ mit $m < n$.
Wir betrachten den folgenden reflexiven und symmetrischen Rahmen (W, R) mit
$W = \{p_0, \ldots, p_n\}$
und für $i, j \in \{0, \ldots, n\}$:

$$(p_i, p_j) \in R :\Leftrightarrow i = j \text{ oder } i = j + 1 \text{ oder } j = i + 1.$$

Sei $\alpha = (\exists x)\, P(x)$ und $\mathfrak{J}$ eine Welteninterpretation mit $\mathfrak{J}_{p_k} \models_{PL} \alpha \Leftrightarrow 0 \leq k \leq m$, so dass also für l mit $m < l \leq n$ gilt: $\mathfrak{J}_{p_l} \not\models_{PL} \alpha$. Eine solche Welteninterpretation existiert offensichtlich. Es gilt dann

$$(W, R, \mathfrak{J}), p_0 \Vdash_B \Box^m\alpha,$$

aber nicht $(W, R, \mathfrak{J}), p_0 \Vdash_B \Box^n\alpha$. Wir erhalten somit $(W, R, \mathfrak{J}), p_0 \nVdash_B \Box^m\alpha \to \Box^n\alpha$ und damit auch $\nVdash_B \Box^m\alpha \to \Box^n\alpha$.
Da die Behauptung für ***B*** gilt, gilt sie folglich auch für ***K*** und ***T***.
Für ***G*** zeigt man mit einer sehr ähnlichen Konstruktion:
Für alle $m, n \in \mathbb{N}$ mit $n > m \geq 0$ und $\alpha \in Fm_{ML}(S)$ gilt $\nVdash_G \Box^n\alpha \to \Box^m\alpha$.
Damit gilt die Behauptung dann auch für die Modallogik ***G***. ■

4.2.3 Das $p-$Morphismus Lemma

Für die folgende Betrachtung müssen wir unsere Voraussetzungen hinsichtlich der Grundbereiche verschärfen. Wir legen für diesen Abschnitt „constant domains" zugrunde, d.h. wir haben in allen Welten einer Kripke-Struktur jeweils den gleichen Grundbereich, den wir mit $|\mathfrak{J}|$ bezeichnen.

Wir wollen nun mit dem $p-$Morphismus Lemma ($p-$Morphismus abgeleitet von Pseudo-Morphismus) uns ein Werkzeug zur Verfügung stellen, mit dem wir in einigen Fällen die Frage beantworten können, ob eine bestimmte Rahmeneigenschaft über ein modales Axiomenschema charakterisierbar ist. Wir definieren dazu zunächst den Begriff des $p-$Morphismus als eine spezielle Verallgemeinerung des Begriffs „Isomorphismus" für prädikatenlogische Strukturen.

Definition 4.50 *Seien $M = (W, R, \mathfrak{I})$ und $M' = (W', R', \mathfrak{I}')$ $S-$Kripke-Strukturen.*
*Ein $p-$*Morphismus *von M in M' ist ein Tupel (f, g), bestehend aus einer Abbildung*

$$f : W \to W'$$

und einer bijektiven Abbildung

$$g : |\mathfrak{I}| \to |\mathfrak{I}'|$$

für die Folgendes gilt:

1. *Für alle $p, q \in W : pRq \;\Rightarrow\; f(p) R' f(q)$.*
2. *Für alle $p \in W$ und $r \in W' : f(p) R' r \;\Rightarrow$ es existiert $q \in W$ mit $f(q) = r$ und pRq.*
3. *Für alle $p \in W$ weisen die prädikatenlogischen Strukturen $\mathfrak{I}_p$ und $\mathfrak{I}'_{f(p)}$ dieselbe Signatur auf.*
4. *Für alle $p \in W$ ist g ein (prädikatenlogischer) Isomorphismus von $\mathfrak{I}_p$ auf $\mathfrak{I}'_{f(p)}$.*

*Eine Abbildung $f : W \to W'$, die den beiden Bedingungen 1. und 2. genügt, heißt $p-$*Morphismus *vom Rahmen (W, R) in den Rahmen (W', R'). Ist f ein surjektiver $p-$*Morphismus *vom Rahmen (W, R) in den Rahmen (W', R'), dann heißt (W', R') $p-$*morphes Bild *von (W, R).*

Lemma 4.51 ($p-$Morphismus Lemma für Kripke-Strukturen)
Seien $M = (W, R, \mathfrak{I})$ und $M' = (W', R', \mathfrak{I}')$ $S-$Kripke-Strukturen. Ferner sei (f, g) ein $p-$Morphismus von M in M'. Es gilt dann für $p \in W$, $M-$Belegungen w und alle $\alpha \in Fm_{ML}(S)$:

$$M, p, w \Vdash_{ML} \alpha \Leftrightarrow M', f(p), g(w) \Vdash_{ML} \alpha.$$

Beweis. Durch strukturelle Induktion.

Ind.Basis: Sei α eine atomare Formel.

g ist nach Voraussetzung ein Isomorphismus von $\mathfrak{I}_p$ auf $\mathfrak{I}'_{f(p)}$. Damit gilt auf Grund von Lemma 2.56

$$\mathfrak{I}_p, w \models_{PL} \alpha \;\Leftrightarrow\; \mathfrak{I}_{f(p)}, g \circ w \models_{PL} \alpha.$$

Es folgt sofort

$$M', f(p), g \circ w \Vdash_{ML} \alpha \;\Leftrightarrow\; M, p, w \Vdash_{ML} \alpha.$$

Ind.Schritt: Wir betrachten exemplarisch 2 Fälle:

1. $(\exists x)\, \alpha$:
   Es gilt dann
   $M', f(p), g \circ w \Vdash_{ML} (\exists x)\, \alpha$
   $\Leftrightarrow$ es ex. $g(a) \in \left|\mathfrak{I}_{f(p)}\right|$, so dass $M', f(p), (g \circ w)\{x/g(a)\} \Vdash_{ML} \alpha$
   $\Leftrightarrow$ es ex. $a \in |\mathfrak{I}_p|$, so dass $M', f(p), g \circ w\{x/a\} \Vdash_{ML} \alpha$
   $\Leftrightarrow$ es ex. $a \in |\mathfrak{I}_p|$, so dass $M, p, w\{x/a\} \Vdash_{ML} \alpha$ nach Ind.Vor.
   $\Leftrightarrow\; M, p, w \Vdash_{ML} (\exists x)\, \alpha$.
2. $\Diamond\alpha$:
   Es gilt
   $M', f(p), g \circ w \Vdash_{ML} \Diamond\alpha$
   $\Leftrightarrow$ es gibt ein $r \in W'$ mit $f(p) R' r$ und $M', r, g \circ w \Vdash_{ML} \alpha$
   $\Rightarrow$ es gibt ein $q \in W$ mit pRq, $f(q) = r$ und $M', f(q), g \circ w \Vdash_{ML} \alpha$
   $\Rightarrow$ es gibt ein $q \in W$ mit pRq und $M, q, w \Vdash_{ML} \alpha$ nach Ind.Vor.
   $\Leftrightarrow\; M, p, w \Vdash_{ML} \Diamond\alpha$.

   Weiter gilt

$M, p, w \Vdash_{ML} \Diamond\alpha$
$\Leftrightarrow$ es gibt ein $q \in W$ mit pRq und $M, q, w \Vdash_{ML} \alpha$
$\Leftrightarrow$ es gibt ein $q \in W$ mit pRq und $M', f(q), g \circ w \Vdash_{ML} \alpha$ nach Ind.Vor.
$\Rightarrow$ es gibt ein $q \in W$ mit $f(p) R' f(q)$ und $M', f(q), g \circ w \Vdash_{ML} \alpha$ nach Vor. über f
$\Rightarrow$ $M', f(p), g \circ w \Vdash_{ML} \Diamond\alpha$.

Insgesamt folgt also auch in diesem Fall die behauptete Äquivalenz.

■

Lemma 4.52 ($p-$Morphismus Lemma für Kripke-Rahmen) *Sei S eine Signatur. Der Kripke-Rahmen $\mathfrak{F}' = (W', R')$ sei ein $p-$morphes Bild von $\mathfrak{F} = (W, R)$.*
Es gilt dann für beliebiges $\alpha \in Fm_{ML}(S)$:

$$\mathfrak{F} \Vdash_{ML} \alpha \Rightarrow \mathfrak{F}' \Vdash_{ML} \alpha.$$

Beweis.
Wir zeigen die Kontraposition.
Sei α eine Formel, die in dem Kripke-Rahmen $\mathfrak{F}'$ nicht allgemeingültig ist. Es gibt dann eine auf diesem Rahmen basierende Kripke-Struktur $M' = (W', R', \mathfrak{I}')$, eine Welt $p' \in W'$ und eine $M'-$Belegung w', so dass nicht gilt: $M', p', w' \Vdash_{ML} \alpha$.
Sei f ein surjektiver $p-$Morphismus vom Kripke-Rahmen $\mathfrak{F}$ in den Kripke-Rahmen $\mathfrak{F}'$. Wir definieren eine auf dem Kripke-Rahmen $\mathfrak{F}$ basierende Kripke-Struktur $M = (W, R, \mathfrak{I})$ wie folgt: für $p \in W$ sei $\mathfrak{I}_p = \mathfrak{I}_{f(p)}$.
Damit ist offensichtlich (f, id) ein $p-$Morphismus der Kripke-Struktur $M = (W, R, \mathfrak{I})$ in die Kripke-Struktur $M' = (W', R', \mathfrak{I}')$. Wähle nun ein $p \in W$ mit $f(p) = p'$. Dann gilt nach dem vorangegangenen Lemma 4.51 zu $p-$Morphismen von Kripke-Strukturen $M, p, w \not\Vdash_{ML} \alpha$. Damit ist α auch im Kripke-Rahmen $\mathfrak{F}$ nicht allgemeingültig.
■

Beispiel 4.53 *Wir haben in Theorem 4.32 gesehen dass eine Vielzahl von Rahmen mittels modaler Axiomenschemata charakterisiert werden kann. Es gibt jedoch kein modales Axiomenschema, das z.B. die irreflexiven Rahmen charakterisiert (Übungsaufgabe).*

4.2.4 Zur Anwendung der Modallogik

Wir haben gesehen, dass wir durch die Beschränkung auf spezielle Rahmen jeweils ausgezeichnete Logiken erhalten haben. Diese Tatsache, dass wir über die Rahmeneigenschaften Logiken definieren können, macht es uns möglich, den Modaloperatoren ganz unterschiedliche, auf den jeweiligen Anwendungsaspekt bezogene semantische Interpretationen zuzuordnen. So wird beispielsweise für die Logik ***S*5** der Box-Operator häufig als „es wird gewusst” interpretiert und ***S*5** demgemäß als Wissenslogik bezeichnet. (Statt des $\Box$-Symbols wird in ***S*5** meistens das Symbol K (von *knowledge* herrührend) verwendet.) Für ***S*4** kann der Operator $\Box$ z.B. räumlich interpretiert werden: „in allen direkt erreichbaren Orten”.
Viele sprachliche Formulierungen, die Modalitäten benutzen, können daher in einer der hier vorgestellten Modallogiken formalisiert und auf ihren Wahrheitsgehalt hin untersucht werden. Für praktisch alle Anwendungen in der Informatik sind jedoch die Ausdrucksmittel dieser Modallogiken nicht ausreichend, so dass immer geeignete Erweiterungen vorzunehmen sind. Wir wollen im folgenden 2 verschiedene Anwendungsbereiche von Modallogiken vorstellen, wobei wir allerdings die erforderlichen Erweiterungen der Modallogik nicht bzw. nicht vollständig formal einführen, sondern uns mit informellen Eräuterungen begnügen.

Von der Modallogik zur linearen Temporallogik

Zu den sicherlich bedeutendsten Gebieten der Modallogik gehört die Temporallogik. Bereits in den Logiken ***S*4** und ***S*4.3** kann den Modaloperatoren eine brauchbare zeitlogische Deutung gegeben werden. Betrachten wir dazu einen reflexiven und transitiven bzw. einen reflexiven, transitiven und schwach linearen Rahmen $\mathfrak{F} = (W, R)$ sowie eine ausgezeichnete Welt p_0 in diesem Rahmen. Aufgrund von Lemma 4.16 dürfen wir annehmen, dass dieser Rahmen von dieser Welt $p_0 \in W$ generiert worden ist, d.h. für dieses p_0 gilt $W =$

$\{q \in W \mid (p_0, q) \in R\}$ (man beachte, dass für den Rahmen $\mathfrak{F}$ $R^* = R$ gilt). Definiert man für beliebige $p, q \in W$

$$p \leq q \text{ genau dann, wenn } (p,q) \in R$$

und interpretiert man eine Welt p als einen Zeitpunkt, dann gilt für Aussagen α, für eine auf $\mathfrak{F}$ basierende Kripke-Struktur $(W, R, \mathfrak{I})$ und für einen Zeitpunkt $p \in W$:

$$(W,R,\mathfrak{I}), p \Vdash_L \Box\alpha \;\Leftrightarrow\; \text{zu allen Zeitpunkten } q \in W \text{ mit } p \leq q \text{ gilt } (W,R,\mathfrak{I}), q \Vdash_L \alpha$$

sowie

$$(W,R,\mathfrak{I}), p \Vdash_L \Diamond\alpha \;\Leftrightarrow\; \text{es gibt einen Zeitpunkt } q \in W \text{ mit } p \leq q \text{ und } (W,R,\mathfrak{I}), q \Vdash_L \alpha$$

für $L = \boldsymbol{S4}$ bzw. $L = \boldsymbol{S4.3}$.

Da für $\boldsymbol{S4}$ $\leq$ eine Halbordnung darstellt, erhalten wir somit im Falle von $\boldsymbol{S4}$ eine einfache Zeitlogik (oder wie wir im Folgenden sagen wollen: *Temporallogik*) mit verzweigender Zukunft, in der (z.B. aufgrund von Nichtdeterminismus) unterschiedliche zeitliche Entwicklungen berücksichtigt sind. Legt man die Logik $\boldsymbol{S4.3}$ zugrunde, so ist $\leq$ sogar eine lineare Ordnung. In diesem Fall spricht man von einer *linearen Temporallogik*. Durch weitere Anforderungen an die Erreichbarkeitsrelation, die sich durch geeignete modallogische Axiomenschemata charakterisieren lassen, können wir u.a. erreichen, dass die Rahmen dann isomorph zum Rahmen $(\mathbb{N}, \leq)$ (in der Literatur i.a. mit $(\omega, \leq)$ notiert) sind – wir sprechen dann von einer *diskreten Temporallogik* – oder aber die Welten (Zeitpunkte) dicht liegen:

Bezeichnung	Axiomenschema	Rahmeneigenschaft
X	$\Box\Box\alpha \rightarrow \Box\alpha$	„sichert Dichtheit”
Dum	$\Box(\Box(\alpha \rightarrow \Box\alpha) \rightarrow \alpha) \rightarrow (\Diamond\Box\alpha \rightarrow \alpha)$	„sichert Diskretheit”

Für konkrete Anwendungen sind die Ausdrucksmittel aber noch zu schwach, so dass man zum einen häufig weitere Modaloperatoren einführt und zum anderen die sprachlichen Ausdrucksmittel erweitert. In Abschnitt 4.4 werden wir darauf etwas näher eingehen.

Von der Modallogik zur (multimodalen) epistemischen Logik

Im Zusammenhang mit der Einführung der Modallogik $\boldsymbol{S5}$ hatten wir schon darauf hingewiesen, dass in dieser Logik der $\Box$-Operator i.a. als Wissensoperator interpretiert wird. Demgemäß wird $\boldsymbol{S5}$ oft auch als epistemische Logik bezeichnet (abgeleitet vom griechischen „episteme”, was soviel wie Erkenntnis bedeutet). Welche Eigenschaften der Modallogik $\boldsymbol{S5}$ sind es, die die Interpretation von $\Box$ als „es wird gewusst” rechtfertigen? Auf die philosphischen Aspekte, was denn Wissen nun genau ist („wahre Überzeugung” oder „gerechtfertigter wahrer Glaube” oder …?) wollen wir hier nicht eingehen, sondern in mehr pragmatischer Vorgehensweise uns vergewissern, dass wesentliche Aspekte, die wir mit dem Wissensbegriff für die jeweilige Anwendung verbinden, erfüllt sind. Sehen wir uns hierzu die $\boldsymbol{S5}$ charakterisierenden Axiomenschemata an:

T $\Box\alpha \rightarrow \alpha$: „Wird α gewusst, so muss α insbesondere gelten”.

Mit anderen Worten: was nicht gilt, kann man auch nicht wissen.

Diese Forderung an einen Wissensbegriff ist unstrittig. Von „Wissensstrukturen” wird man daher immer erwarten, dass sie eine reflexive Erreichbarkeitsrelation aufweisen.

5 $\Diamond\alpha \rightarrow \Box\Diamond\alpha$: Speziell für $\alpha = \neg\beta$ erhält man $\Diamond\neg\beta \rightarrow \Box\Diamond\neg\beta$. Unter Ausnutzung der Beziehung zwischen $\Box$ und $\Diamond$ ist die letzte Formel dann äquivalent zu $\neg\Box\beta \rightarrow \Box\neg\Box\beta$, so dass also das Axiomenschema 5 epistemisch interpretiert ausdrückt „man weiß, was man nicht weiß”. Von einem philosphischen Standpunkt wird man dies kaum als eine Eigenschaft von Wissen akzeptieren. Dieses Schema für eine Wissenslogik zu postulieren, ist daher recht fragwürdig. Wenn man es dennoch tut, dann aus dem Grund, dass die resultierende Logik schöne Eigenschaften aufweist (z.B. einfache Normalform), Die Transitivität der Erreichbarkeitsrelation, die durch das Schema 4 ausgedrückt wird, ist auf der Basis von T und 5 dann folgerbar.

4 $\Box\alpha \to \Box\Box\alpha$: „Man weiß, was man weiß”.

Diese Eigenschaft wird für den Wissensbegriff überwiegend akzeptiert. Es gibt aber durchaus kontroverse Standpunkte zu diesem Prinzip.

Diese kurze Erläuterung der epistemischen Interpretation dieser Axiomenschemata macht somit deutlich, dass eine „adäquate” Wissenslogik zwischen den Logiksystemen $\boldsymbol{S4}$ und $\boldsymbol{S5}$ liegen muss.
Wir wollen noch auf eine weitere Schwierigkeit des modallogischen Ansatzes im Zusammenhang mit dem Wissensbegriff hinweisen. Da z.B. für $\boldsymbol{S5}$ gilt

$$\Vdash_{S5} \alpha \ \Leftrightarrow \Vdash_{S5} \Box\alpha,$$

wird also jede allgemeingültige Aussage gewusst. Weiter impliziert dies, dass auch alle logischen Folgerungen des eigenen Wissens gewusst werden:
Aus $\Vdash_{S5} \Box\alpha$ und $\Vdash_{S5} \alpha \to \beta$ folgt $\Vdash_{S5} \Box\beta$.
Dieses „Prinzip der logischen Allwissenheit” (englisch: logical omniscience) ist für viele Anwendungen von Wissenslogiken in der Informatik nicht akzeptierbar, so dass man modale Wissenslogiken entwickelt hat, die dieses Prinzip vermeiden. (Siehe hierzu z.B. das ausgezeichnete Buch:
R.Fagin, J.H. Halpern, Y. Moses, M.Y. Vardi, Reasoning About Knowledge, MIT Press 1995)

Epistemische Logiken finden häufig Anwendung in Multi-Agentensystemen, wo man von außen betrachtet das Wissen der Agenten logisch beschreiben bzw. analysieren möchte. Da man in diesen Fällen in der Regel mehrere Agenten hat und man das Wissen jedes einzelnen Agenten beschreiben können will, benötigt man somit auch für jeden einzelnen Agenten einen eigenen Wissensoperator. Wir werden so zu einer multimodalen Wissenslogik geführt, in der im Alphabet der Sprache statt der Modaloperatoren $\Box$ und $\Diamond$ nun mehrere geeignet indizierte Modaloperatoren $\Box_1, \dots, \Box_k$ und $\Diamond_1, \dots, \Diamond_k$ vorhanden sind. (Im Falle der Wissenslogiken wird dann i.a. K_i statt $\Box_i$ verwendet.)
Der Aufbau der formalen Sprache einer solchen multimodalen Sprache passiert dann genau so, wie wir es schon für die Modallogik gesehen haben, wobei wir nun aber beliebige dieser $\Box_i$ und $\Diamond_j$ bei der Bildung von Formeln verwenden dürfen.
Die Semantik-Definition muss entsprechend angepasst werden. Wir haben für jeden Agenten i (bzw. Modaloperator $\Box_i$) eine eigene Erreichbarkeitsrelation, so dass also ein passender Kripke-Rahmen ein Tupel der Gestalt $(W, R_1, \dots, R_k)$ mit $R_i \subseteq W \times W$, $1 \leq i \leq k$, ist.
Wir legen dann fest:
$(W, R_1, \dots, R_k, \mathfrak{J}), p, w \Vdash_{ML} \Box_i\alpha$ $:\Leftrightarrow$ für alle q mit $(p, q) \in R_i$ gilt $(W, R_1, \dots, R_k, \mathfrak{J}), q, w \Vdash_{ML} \alpha$.
Im Zusammenhang mit Logiken für Multi-Agentensysteme werden häufig noch weitere Modaloperatoren verwendet, mit denen das Wissen einer Gruppe von Agenten, das gemeinsame Wissen sowie das verteilte Wissen, das in der Gruppe von Agenten vorhanden ist, behandelt werden kann.

Bemerkung 4.54 *Wir haben hier multimodale Logiken im Zusammenhang mit den epistemischen Logiken angesprochen. Es gibt in ganz anderen Bereichen viele natürliche multimodale Logiken, z.B. dann, wenn man Zeit- und Raumaspekte berücksichtigen muss wie in der Robotertechnologie. Wir werden später im Rahmen eines Beispiels eine weitere, ganz bedeutende multimodale Logik, die dynamische Logik, kurz vorstellen.*

Sehen wir uns zum Abschluß dieses Exkurses noch ein sehr einfaches Beispiel zur Modellierung (logischen Beschreibung) einer Problemstellung mit Mitteln der epistemischen Logik an.

Beispiel 4.55 Das Problem der 3 Weisen.
Ein König hatte 3 weise Männer, hier mit a,b, und c bezeichnet. Eines Tages beschloß er, herauszufinden, wer der weiseste von ihnen ist.Dazu setzte er jedem von ihnen einen weißen oder schwarzen Hut auf und stellte sie dann so in Form eines Dreiecks auf, dass jeder die Hüte der beiden anderen sehen konnte, nicht aber seinen eigenen. Er teilte ihnen dann mit, dass wenigstens einer von ihnen einen weißen Hut trägt.

Er fragt dann a, ob er die Farbe seines Hutes wüsste. a verneint. Anschließend fragt er b dasselbe. b verneint ebenfalls. Als er schließlich c fragt, antwortet dieser korrekterweise, dass er einen weißen Hut trägt. Wie konnte c das wissen?
Zur formalen Beschreibung legen wir das Prädikatensymbol W für „trägt weißen Hut" zugrunde. $\neg W(x)$ verwenden wir, um ausdrücken, dass x einen schwarzen Hut trägt, d.h. wir unterstellen, dass jeder entweder einen weißen oder einen schwarzen Hut trägt. Wir haben ferner die Modaloperatoren $\Box_a, \Box_b, \Box_c$ mit der Intention „a weiß ..." usw. zur Verfügung.
Der in der Geschichte formulierte Sachverhalt kann dann in der folgenden Weise logisch beschrieben werden.
Sei die Formelmenge Σ definiert durch

$$\begin{aligned}\Sigma := &\{\Box\Box_i (W(a) \vee W(b) \vee W(c)) \mid i \in \{a,b,c\}\} \\ &\cup \{\Box(\neg W(j) \rightarrow \Box_k \neg W(j)) \mid j,k \in \{a,b,c\},\ j \neq k\} \\ &\cup \{\Box\Box_c\Box_b\neg\Box_a W(a), \Box\Box_c\neg\Box_b W(b)\},\end{aligned}$$

wobei zur Abkürzung $\Box$ hier für eine beliebige endliche Folge von Modaloperatoren aus $\{\Box_a, \Box_b, \Box_c\}$ steht.
Es gilt dann $\Sigma \Vdash_{S5} \Box_c W(c)$. (Übungsaufgabe)
(*Wir benutzen hier als Index für die entsprechende multimodale Wissenslogik ebenfalls* $\boldsymbol{S5}$.)

4.2.5 $\boldsymbol{\nu}$- und π-Formeln

Wir hatten für die Aussagen- und Prädikatenlogik eine uniforme Notation für Formeln eingeführt. Insbesondere erlaubte diese uns eine kompakte Darstellung der Beweiskalküle. Wir wollen nun die für die Prädikatenlogik verwendete uniforme Notation der $\boldsymbol{\alpha}$-, $\boldsymbol{\beta}$-, $\boldsymbol{\gamma}$- und δ-Formeln übernehmen und um zwei weitere Kategorien ergänzen, die die modallogischen Operatoren berücksichtigen:

- $\boldsymbol{\nu}$–Formeln ($\boldsymbol{\nu}$ von **n**eccesity),

  das sind modallogische Formeln der Gestalt $\Box\alpha$ bzw. $\neg\Diamond\alpha$.

- π-Formeln ($\boldsymbol{\pi}$ von **p**ossibility),

  das sind modallogische Formeln der Gestalt $\Diamond\alpha$ bzw. $\neg\Box\alpha$.

ν-Formeln und ihre Komponenten $\boldsymbol{\nu}_0$:

$\boldsymbol{\nu}$	$\boldsymbol{\nu}_0$
$\Box\alpha$	α
$\neg\Diamond\alpha$	$\neg\alpha$

π-Formeln mit Komponenten $\boldsymbol{\pi}_0$ und der Konjugierten $\bar{\boldsymbol{\pi}}$:

$\boldsymbol{\pi}$	$\boldsymbol{\pi}_0$	$\bar{\boldsymbol{\pi}}$
$\neg\Box\alpha$	$\neg\alpha$	$\Box\alpha$
$\Diamond\alpha$	α	$\neg\Diamond\alpha$

Man könnte nun wiederum als Beweisprinzip strukturelle Induktion und als Definitionsprinzip strukturelle Rekursion über Formeln dieser Kategorien formulieren. Da wir dies für den Rest des Skripts nicht mehr benötigen, verzichten wir darauf.

4.3 Modallogische Tableau-Kalküle

Wir wollen im Folgenden aufbauend auf dem prädikatenlogischen Tableau-Kalkül Tableau-Kalküle für die verschiedenen Modallogiken angeben. Nun stellt der Aufbau eines Tableaus einen Versuch dar, ein Modell für die infrage stehende Formel(menge) zu konstruieren. Im Falle der Modallogik bedeutet dies, dass ein Tableau die Struktur eines Kripke-Rahmens widerspiegeln muss. Z.B. muss für $\Diamond\alpha$ eine i.a. von der aktuellen

Welt verschiedene Welt vorhanden sein, in der α gelten muss. Ein spezielles Problem stellen nun die $\boldsymbol{\delta}$- und $\boldsymbol{\gamma}$-Formeln dar. Im Falle von δ-Formeln werden in der Prädikatenlogik $\boldsymbol{\delta}$-Instanzen unter Verwendung von neuen Konstantensymbolen gebildet, die Objekte bezeichnen können, die in einer vorausgegangenen Welt nicht vorhanden waren. In Kripke-Strukturen werden aber bisher Konstantensymbole global gleich interpretiert. Die Problematik, dass beim Übergang von einer Welt in eine andere die Menge der Objekte zunehmen kann, bzw. dass für Welten, zwischen denen kein Pfad existiert, keine Inklusionsbeziehung zwischen den jeweiligen Grundbereichen existieren muss, betrifft auch die Bildung der $\boldsymbol{\gamma}$-Instanzen zu γ-Formeln.

4.3.1 Modallogische $\boldsymbol{K}$-Tableaus

Wir definieren zunächst allgemein $\boldsymbol{K}$-Tableaus. Über weitere Regeln werden wir hieraus die Tableaus für die weiteren Logiken erhalten.
$\boldsymbol{K}$-Tableaus sind wiederum endliche markierte Binärbäume, wobei die Markierungen jetzt aus Paaren (p,α) (notiert als $p \Vdash \alpha$) oder aber (p,q) (notiert als pRq) mit Welten p und q sowie modallogischen Formeln α bestehen. Liegt eine Folge $p_0Rp_1, p_1Rp_2, \ldots, p_{n-1}Rp_n$ von Markierungen vor, so wollen wir diesen Sachverhalt auch durch $p_0R^*p_n$ notieren.
Wir übernehmen - soweit möglich - die Terminologie aus der aussagenlogischen bzw. prädikatenlogischen Definition. Sei $W = \{p_i \mid i \in \mathbb{N}\}$ ein Vorrat von Welten und $\mathcal{K}_0 = \{c_i \mid i \in \mathbb{N}\}$ ein Vorrat von Konstantensymbolen, die in der Signatur S nicht vorhanden sind. Für die Signatur S nehmen wir o.B.d.A. an, dass sie wenigstens ein Konstantensymbol enthält.

Definition 4.56 (Σ-*Tableau für* $\{\alpha_0,\ldots,\alpha_n\}$)
Sei $\Sigma \cup \{\alpha_0,\ldots,\alpha_n\} \subseteq Fm_{ML}(S)$ *eine Menge von Aussagen.*

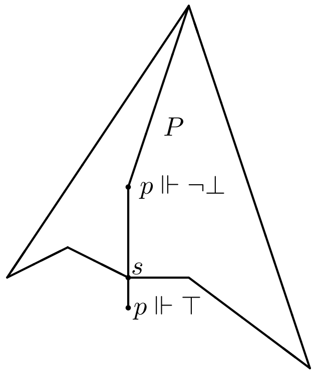
1. *Der mit* $p_0 \Vdash \alpha_0, \ldots, p_0 \Vdash \alpha_n$ *markierte Binärbaum*

$$\begin{array}{c} p_0 \Vdash \alpha_0 \\ | \\ p_0 \Vdash \alpha_1 \\ | \\ \vdots \\ | \\ p_0 \Vdash \alpha_n \end{array}$$

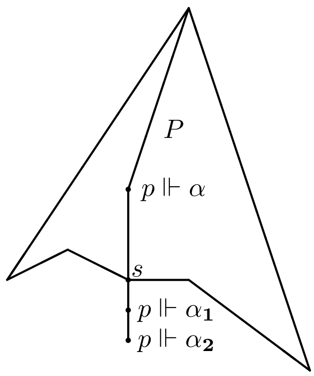
ist ein Σ-*Tableau(das* Starttableau) *für* $\{\alpha_0,\ldots,\alpha_n\}$.

2. *Ist* τ *ein* Σ-*Tableau für* $\{\alpha_0,\ldots,\alpha_n\}$ *und* P *ein Zweig in* τ *mit Blatt* s, *ist ferner ein Knoten in* P *mit* $p \Vdash N$ *markiert, dann ist der in Abhängigkeit von* N *nachfolgend definierte, markierte Binärbaum* τ' *ebenfalls ein* Σ-*Tableau für* $\{\alpha_0,\ldots,\alpha_n\}$:

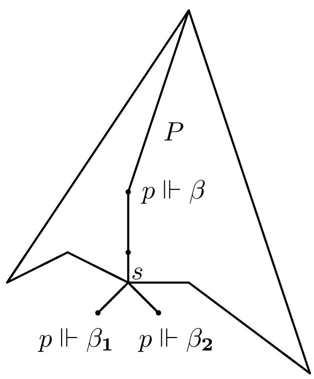
   (a) $N = \neg\bot \quad (\neg\top$ *resp.* $\neg\neg\beta)$:
   τ' *resultiert aus* τ, *indem an das Blatt* s *ein mit* $p \Vdash \top$ $(p \Vdash \bot$ *resp.* $p \Vdash \beta)$ *markierter Knoten als Sohn angehängt wird. (als Tableau-Regel notiert:* $\dfrac{p \Vdash \neg\bot}{p \Vdash \top}$)



(b) *N ist eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_1$ und $\boldsymbol{\alpha}_2$:*
τ' resultiert aus τ dadurch, dass das Blatt s von P einen Sohn und einen Enkel erhält, die mit $p \Vdash \boldsymbol{\alpha}_1$ bzw. $p \Vdash \boldsymbol{\alpha}_2$ markiert werden. (*als Tableau-Regel notiert:* $\dfrac{p \Vdash \boldsymbol{\alpha}}{\begin{array}{c} p \Vdash \boldsymbol{\alpha}_1 \\ p \Vdash \boldsymbol{\alpha}_2 \end{array}}$)

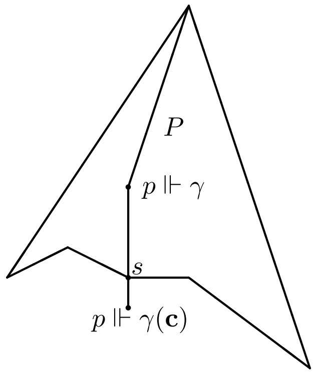


(c) *N ist eine β-Formel mit den Komponenten $\boldsymbol{\beta}_1$ und $\boldsymbol{\beta}_2$:*
τ' wird aus τ dadurch erhalten, dass an das Blatt s von P zwei Söhne, die mit $\boldsymbol{\beta}_1$ bzw. $\boldsymbol{\beta}_2$ markiert werden, angehängt werden. (*als Tableau-Regel notiert:* $\dfrac{p \Vdash \boldsymbol{\beta}}{p \Vdash \boldsymbol{\beta}_1 \mid p \Vdash \boldsymbol{\beta}_2}$)

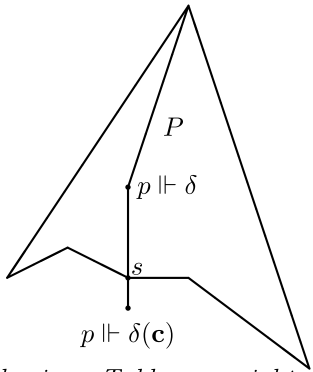


a) - c) sind somit die „aussagenlogischen Tableau-Regeln".

(d) *N ist eine γ-Formel :*
τ' resultiert aus τ, indem an das Blatt s ein mit $\boldsymbol{\gamma}\,(c)$ markierter Knoten als Sohn angehängt wird. (*als Tableau-Regel notiert:* $\dfrac{p \Vdash \boldsymbol{\gamma}}{p \Vdash \boldsymbol{\gamma}\,(c)} \quad (c\,!)$),
wobei $c \in S$ oder
$c \in \mathcal{K}_0$ und es gibt im Zweig P eine Welt q mit qR^*p, und $q \Vdash \beta$ und c kommt in β vor.
(d.h. c ist auf dem Pfad zu p zuvor über eine $\boldsymbol{\delta}$-Instanz eingeführt worden)

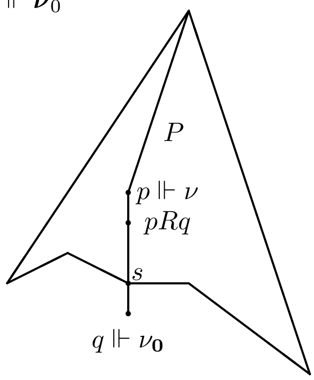


(e) *N ist eine δ-Formel mit einer $\boldsymbol{\delta}$-Instanz $\boldsymbol{\delta}(c)$ für ein $c \in \mathcal{K}_0$, das in τ nicht vorkommt*:
τ' resultiert aus τ, indem an das Blatt s ein mit $\boldsymbol{\delta}(c)$ markierter Knoten als Sohn angehängt wird (als Tableau-Regel notiert: $\dfrac{p \Vdash \boldsymbol{\delta}}{p \Vdash \boldsymbol{\delta}(c)}$ *(c neu!)),*



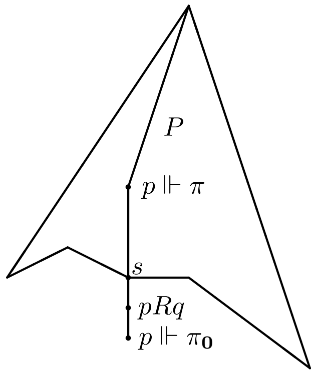
wobei $c \in \mathcal{K}_0$ und c im bisherigen Tableau τ nicht vorkommt.
d) und e) sind die „prädikatenlogischen Tableau-Regeln".

(f) *N ist eine ν-Formel :*
Ist im Zweig P für ein $q \in W$ eine Markierung pRq vorhanden, so resultiert τ' aus τ, indem an das Blatt s ein mit $q \Vdash \boldsymbol{\nu}_0$ markierter Knoten als Sohn angehängt wird. (als Tableau-Regel notiert: $\dfrac{p \Vdash \boldsymbol{\nu}}{q \Vdash \boldsymbol{\nu}_0}$ *(pRq !))*

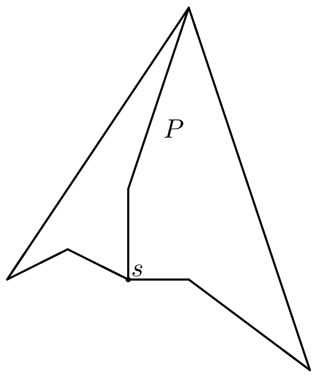


(g) *N ist eine π-Formel:*
τ' resultiert aus τ, indem an das Blatt s ein mit pRq markierter Knoten als Sohn und ein mit $q \Vdash \pi_0$ markierter Knoten als Enkel angehängt wird für ein $q \in W$, das im bisherigen Tableau τ nicht vorkommt.
(als Tableau-Regel notiert: $\dfrac{p \Vdash \pi}{\begin{array}{c} pRq \\ q \Vdash \pi_0 \end{array}}$ *(q neu!))*

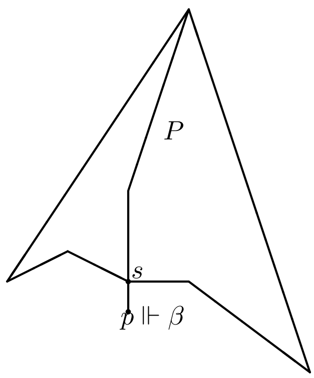
f) und g) sind die „modallogischen Tableau-Regeln".



(h) *Ist τ ein Σ-Tableau*



und P ein Zweig in τ mit Blatt s, dann ist für eine Aussage $\beta \in \Sigma$ auch der folgende markierte Binärbaum ein Σ-Tableau:



3. *Nur solche markierten Bäume, die nach 1. und 2. erhalten werden, sind Σ-Tableaus für $\{\alpha_0, \ldots, \alpha_n\}$.*

Definition 4.57

1. *Ein* Zweig P eines modallogischen Tableaus *heißt* geschlossen *genau dann, wenn für ein $p \in W$*
   (a) *ein Knoten in P mit $p \Vdash \bot$ markiert ist, oder wenn*
   (b) *es eine Aussage $\alpha \in Fm_{ML}(S(\mathcal{K}_0))$ und zwei Knoten in P gibt, von denen einer mit $p \Vdash \alpha$ und der andere mit $p \Vdash \neg\alpha$ markiert ist.*
2. *Ein* Σ-Tableau heißt geschlossen *genau dann, wenn jeder Zweig des Tableaus geschlossen ist.*
3. *Ein modallogischer Σ-Tableau-Beweis für α ist ein geschlossenes Σ-Tableau für $\{\neg\alpha\}$*
   (notiert als $\Sigma \vdash_{K-T} \alpha$).
4. *Ein* modallogischer Tableau-Beweis für α *ist ein modallogischer $\emptyset-$Tableau-Beweis für α (notiert als $\vdash_{K-T} \alpha$).*

4.3.2 Ergänzende Tableau-Regeln für die weiteren Modallogiken

Tableaus für die verschiedenen hier eingeführten Modallogiken erhalten wir nun leicht jeweils durch ergänzende Regeln. Es gelten dieselben Voraussetzungen wie in Definition 4.56.

Definition 4.58 (Spezielle modallogische Tableauregeln)

1. *Modallogik* $\boldsymbol{T}$:

   (*Σ-Tableaus werden in diesem Fall als $\boldsymbol{T}$-Tableaus (mit Prämissenmenge Σ) bezeichnet.*)

   Reflexive Tableau-Regel:

   $$\frac{}{pRp}$$

   Ist im Zweig P des $\boldsymbol{T}$-Tableaus τ ein Knoten mit $p \Vdash \alpha$ für ein $\alpha \in Fm_{ML}(S(\mathcal{K}_0))$ oder mit pRq für $p \neq q$ markiert, und resultiert τ' aus τ dadurch, dass an das Blatt von P ein mit pRp markierter Sohn angehängt wird, dann ist τ' ebenfalls ein $\boldsymbol{T}$-Tableau.

2. *Modallogik* $\boldsymbol{S}4$:

   (*Σ-Tableaus werden als $\boldsymbol{S}4$-Tableaus (mit Prämissenmenge Σ) bezeichnet.*)

   (a) *Reflexive Tableau-Regel und*

   (b) Transitive Tableau-Regel:

   $$\frac{\begin{array}{c}pRq\\ qRr\end{array}}{pRr}$$

   Ist im Zweig P des $\boldsymbol{S}4$-Tableaus τ ein Knoten mit pRq und ein weiterer Knoten mit qRr markiert, und resultiert τ' aus τ dadurch, dass an das Blatt von P ein mit pRr markierter Sohn angehängt wird, dann ist τ' ebenfalls ein $\boldsymbol{S}4$-Tableau.

3. *Modallogik* $\boldsymbol{B}$:

   (*Tableaus werden als $\boldsymbol{B}$-Tableaus (mit Prämissenmenge Σ) bezeichnet.*)

   (a) *Reflexive Tableau-Regel und*

   (b) Symmetrische Tableau-Regel:

   $$\frac{pRq}{qRp}$$

   Ist im Zweig P des $\boldsymbol{B}$-Tableaus τ ein Knoten pRq markiert, und resultiert τ' aus τ dadurch, dass an das Blatt von P ein mit qRp markierter Sohn angehängt wird, dann ist τ' ebenfalls ein $\boldsymbol{B}$-Tableau.

4. *Modallogik* $\boldsymbol{S}5$:

   (*Tableaus werden als $\boldsymbol{S}5$-Tableaus (mit Prämissenmenge Σ) bezeichnet.*)

   (a) *Reflexive Tableau-Regel und*

   (b) Euklidische Tableau-Regel:

   $$\frac{\begin{array}{c}pRq\\ pRr\end{array}}{qRr}$$

   Ist im Zweig P des $\boldsymbol{S}5$-Tableaus τ ein Knoten mit pRq und ein weiterer Knoten pRr markiert, und resultiert τ' aus τ dadurch, dass an das Blatt von P ein mit qRr markierter Sohn angehängt wird, dann ist τ' ebenfalls ein $\boldsymbol{S}5$-Tableau.

5. *Modallogik* $\boldsymbol{G}$:

   (*Tableaus werden als $\boldsymbol{G}$-Tableaus (mit Prämissenmenge Σ) bezeichnet.*)

(a) Transitive Tableau-Regel und

(b) $\boldsymbol{\pi}$-Regel für G :

$$\frac{p \Vdash \boldsymbol{\pi}}{\begin{array}{c} pRq \\ q \Vdash \boldsymbol{\pi}_0 \\ q \Vdash \bar{\boldsymbol{\pi}} \end{array}} \quad (q \text{ neu!})$$

(modifizierte $\boldsymbol{\pi}$*-Tableau-Regel)*
Ist im Zweig P des $\boldsymbol{G}$-Tableaus τ ein Knoten mit $p \Vdash \boldsymbol{\pi}$ markiert, und resultiert τ' aus τ dadurch, dass für ein $q \in W$, das in τ nicht vorkommt, an das Blatt von P ein Sohn, ein Enkel und ein Großenkel angehängt werden, die mit qRr, $q \Vdash \boldsymbol{\pi}_0$ und mit $q \Vdash \bar{\boldsymbol{\pi}}$ markiert sind, dann ist τ' ebenfalls ein $\boldsymbol{G}$-Tableau.

Es mag aufgefallen sein, dass wir für die Logik $\boldsymbol{S4.3}$ keine Tableau-Regeln angegeben haben. Eine dem Axiomenschema L entsprechende Tableau-Regel lässt sich ohne Schwierigkeiten formulieren, da sie jedoch von ihrer formalen Gestalt etwas aus dem Rahmen fällt, haben wir auf ihre Angabe verzichtet.

Beispiel 4.59

1. *Wir zeigen zunächst an einem einfachen Beispiel, wie der Tableau-Kalkül genutzt werden kann, um die Nicht-Allgemeingültigkeit einer modallogischen Formel nachzuweisen.*

   Sei $\alpha = P(c)$.

   Es gilt $\nVdash_K \Box\alpha \to \alpha$. *Dies kann man leicht durch die Angabe einer Kripke-Struktur zeigen, in der diese Formel nicht allgemeingültig ist. Zu solch einer Kripke-Struktur wird man aber auch durch den Versuch geführt, einen Tableau-Beweis für* $\Box\alpha \to \alpha$ *aufzubauen:*

$$\begin{array}{ll} p_0 \Vdash \neg(\Box\alpha \to \alpha) & \alpha\text{-}\mathit{Formel} \\ \quad | & \\ p_0 \Vdash \Box\alpha & \boldsymbol{\alpha}_1\text{-}\mathit{Komponente} \\ \quad | & \\ p_0 \Vdash \neg\alpha & \boldsymbol{\alpha}_2\text{-}\mathit{Komponente} \end{array}$$

   Danach hat man keine weiteren Möglichkeiten mehr, ohne schon durchgeführte Schritte zu wiederholen. Wir erhalten also in keinem Fall ein geschlossenes Tableau. Eine Kripke-Struktur M, in der $\Box\alpha \to \alpha$ nicht allgemeingültig ist, ist dann z.B. $(\{p_0\}, R, \mathfrak{I})$ mit $R = \emptyset$, wobei $\alpha^{\mathfrak{I}_{p_0}} = F$. Damit gilt dann $M, p_0 \Vdash_K \neg\alpha$. Da außerdem $M, p_0 \Vdash_K \Box\alpha$ wegen $R = \emptyset$ gilt, folgt sofort $\nVdash_K \Box\alpha \to \alpha$.

2. *Es gilt* $\Vdash_K (\exists x)\Box\neg P(x) \to \neg\Diamond(\forall x)P(x)$:

$$\begin{array}{lll} 1 & p_0 \Vdash \neg((\exists x)\Box\neg P(x) \to \neg\Diamond(\forall x)P(x)) & \alpha\text{-}\mathit{Formel} \\ & \quad | & \\ 2 & p_0 \Vdash (\exists x)\Box\neg P(x) & \boldsymbol{\alpha}_1\text{-}\mathit{Komponente\ von}\ 1 \\ & \quad | & \\ 3 & p_0 \Vdash \neg\neg\Diamond(\forall x)P(x) & \boldsymbol{\alpha}_2\text{-}\mathit{Komponente\ von}\ 1 \\ & \quad | & \\ 4 & p_0 \Vdash \Diamond(\forall x)P(x) & \mathit{aus}\ 3 \\ & \quad | & \\ 5 & p_0 \Vdash \Box\neg P(c) & \mathit{aus}\ \boldsymbol{\delta}\text{-}\mathit{Formel}\ 2;\ c\ \mathit{neu} \\ & \quad | & \\ 6 & p_0Rp_1 & \mathit{aus}\ \pi\text{-}\mathit{Formel}\ 4 \\ & \quad | & \\ 7 & p_1 \Vdash (\forall x)P(x) & \mathit{aus}\ \pi\text{-}\mathit{Formel}\ 4 \\ & \quad | & \\ 8 & p_1 \Vdash P(c) & \mathit{aus}\ \gamma\text{-}\mathit{Formel}\ 7 \\ & \quad | & \\ 9 & p_1 \Vdash \neg P(c) & \mathit{aus}\ \nu\text{-}\mathit{Formel}\ 5 \\ & \quad * & \mathit{geschlossen\ wegen}\ 8\ \mathit{und}\ 9 \end{array}$$

3. In $\boldsymbol{S5}$ *ist die Barcan'sche Formel beweisbar, d.h. es gilt* $\Vdash_{S5} (\forall x)\, \Box P(x) \to \Box (\forall x)\, P(x)$:

1	$p_0 \Vdash \neg ((\forall x)\, \Box P(x) \to \Box (\forall x)\, P(x))$	*α-Formel*
2	$p_0 \Vdash (\forall x)\, \Box P(x)$	*$\boldsymbol{\alpha}_1$-Komponente von 1*
3	$p_0 \Vdash \neg\Box (\forall x)\, P(x)$	*$\boldsymbol{\alpha}_2$-Komponente von 1*
4	$p_0 R p_1$	*aus π-Formel 3*
5	$p_1 \Vdash \neg (\forall x)\, P(x)$	*aus π-Formel 3*
6	$p_1 \Vdash \neg P(c)$	*aus δ-Formel 5; c neu*
7	$p_0 R p_0$	*refl. Tableau-Regel*
8	$p_1 R p_0$	*eukl. Tableau-Regel aus 4 und 7*
9	$p_0 \Vdash \Box P(c)$	*aus γ-Formel 2*
10	$p_1 \Vdash P(c)$	*aus ν-Formel 9*
	$*$	*geschlossen wegen 6 und 10*

4.3.3 Korrektheit der Tableau-Kalküle

Unser nächstes Ziel ist es, für die aufgeführten modallogischen Tableau-Kalküle die Korrektheit und mit Ausnahme von $\boldsymbol{G}$ auch die Vollständigkeit für diese Beweiskalküle zu zeigen. Wir beginnen zunächst mit dem einfacheren Teil, dem Nachweis der Korrektheit.
Im Falle der klassischen Aussagen- und Prädikatenlogik wurde gezeigt, dass die Anwendung einer Tableau-Regel auf ein erfüllbares Σ-Tableau wieder zu einem erfüllbaren Σ-Tableau führte. Im Falle der Modallogiken besteht die Schwierigkeit, dass ein „Modell" für eine Aussagenmenge nun durch eine Kripke-Struktur gegeben ist. Unterstellen wir, dass Erfüllbarkeit eines modallogischen Tableaus gegeben ist, dann muss das Resultat einer Tableau-Regel bzw. der Aufbau des Tableaus in gewissem Sinne konform zu einem vorhandenen „Modell" sein. Hier ergibt sich aber eine Schwierigkeit mit den im Tableau neu eingeführten Konstantensymbolen. Ähnlich wie im Falle der Prädikatenlogik, wo wir in diesem Fall Konstantenerweiterungen betrachtet haben, werden wir auch hier entsprechende Konstantenerweiterungen benutzen.

Notation 4.60 *Wir schreiben* $\Sigma \vdash_{L-T} \alpha$ *für* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$*, falls ein geschlossenes L-Tableau mit Prämissenmenge* Σ *für* $\{\neg\alpha\}$ *existiert.*

Notation 4.61 *Sei wie in der Definition der* Σ*-Tableaus* S *eine Signatur und* $\mathcal{K}_o$ *eine abzählbar unendliche Menge von Konstantensymbolen, die in der Signatur* S *nicht enthalten sind.*
Sei weiter $M = (W, R, \mathfrak{J})$ *eine* $S(\mathcal{K})$*-Kripke-Struktur,* $\mathcal{K} \subseteq \mathcal{K}_0$*, und* $\mathfrak{J}_p$ *für* $p \in W$ *eine prädikatenlogische* $S(\mathcal{K}_p)$*-Struktur mit* $\mathcal{K}_p \subseteq \mathcal{K}$.

1. *Hat* M *die Eigenschaft, dass*
$$\mathcal{K}_0 \setminus \mathcal{K}$$
abzählbar unendlich viele Konstantensymbole enthält, so wird M *als* $\mathcal{K}_0$-passend *bezeichnet.*

2. *Für eine* $S(\mathcal{K} \cup \{c\})$*-Kripke-Struktur* $M' = (W, R, \mathfrak{J}')$*, für die* $\mathfrak{J}'_p$ *eine prädikatenlogische* $S(\mathcal{K}_p \cup \{c\})$*-Struktur ist, die wir aus der* $\mathcal{K}_0$*-passenden Kripke-Struktur* M *dadurch erhalten, dass wir zunächst die Signatur* $\mathcal{K}$*, die* M *zugrundeliegt, um das* neue *Konstantensymbol* $c \in \mathcal{K}_0$ *erweitern und in der Welt* $p \in M$ *(und damit auch in allen von* p *aus erreichbaren Welten)*

dem Konstantensymbol c eine Konstante $c^{\mathfrak{I}'} := a$ *zuordnen, verwenden wir die Schreibweise* $M' = M\left(\mathfrak{I}_p\left\{c/a\right\}\right)$.

$\mathfrak{I}'_p$ *und damit auch alle* $\mathfrak{I}'_q$ *für Welten q, die von p aus erreichbar sind, sind also prädikatenlogische* $S\left(\mathcal{K}_p \cup \{c\}\right)$*-Strukturen bzw.* $S\left(\mathcal{K}_q \cup \{c\}\right)$*-Strukturen.* M' *werden wir auch als* Strukturerweiterung *von M bezeichnen. Generell werden wir den Begriff Strukturerweiterung in transitiver Weise verwenden, d.h. ist* M'' *Strukturerweiterung von* M' *und* M' *Strukturerweiterung von M, so ist* M'' *auch Strukturerweiterung von M.*

Für den Begriff der Strukturerweiterung ergibt sich die folgende, leicht zu beweisende Eigenschaft:

Lemma 4.62 *Sei* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$. *Sei ferner M eine* $S(\mathcal{K})$*-Kripke-Struktur und die* $S\left(\mathcal{K} \cup \mathcal{K}'\right)$*-Kripke-Struktur* M' *eine Strukturerweiterung von M.*
Für $p \in M$ *und Aussagen* $\alpha \in Fm_{ML}(S(\mathcal{K}))$ *gilt dann:*

$$M, p \Vdash_L \alpha \;\Leftrightarrow\; M', p \Vdash_L \alpha.$$

Lemma 4.63 *Sei* $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$. *Sei weiter* $M = (W, R, \mathfrak{I})$ *eine* $\mathcal{K}_0$*-passende* $S(\mathcal{K})$*Kripke-Struktur und* $p \in M$.
Es gilt dann:

1. *Für* α*-Formeln mit Komponenten* $\boldsymbol{\alpha}_1$ *und* $\boldsymbol{\alpha}_2$ *sowie für* β*-Formeln mit Komponenten* $\boldsymbol{\beta}_1$ *und* $\boldsymbol{\beta}_2$ *gilt:*

   $M, p \Vdash_L \boldsymbol{\alpha} \;\Rightarrow\; M, p \Vdash_L \boldsymbol{\alpha}_1$ *und* $M, p \Vdash_L \boldsymbol{\alpha}_2$

   $M, p \Vdash_L \boldsymbol{\beta} \;\Rightarrow\; M, p \Vdash_L \boldsymbol{\beta}_1$ *oder* $M, p \Vdash_L \boldsymbol{\beta}_2$

2. *Für* γ*-Formeln mit der* $\boldsymbol{\gamma}$*-Instanz* $\boldsymbol{\gamma}(c)$ *für ein* $c \in S(\mathcal{K}_p)$ *gilt:*

   $M, p \Vdash_L \boldsymbol{\gamma} \;\Rightarrow\; M, p \Vdash_L \boldsymbol{\gamma}(c)$.

3. *Für* δ*-Formeln haben wir:*

   Gilt $M, p \Vdash_L \boldsymbol{\delta}$, *so existiert eine Strukturerweiterung* $M\left(\mathfrak{I}_p\{c/a\}\right)$, $c \in \mathcal{K}_0$, *mit* $M\left(\mathfrak{I}_p\{c/a\}\right), p \Vdash_L \boldsymbol{\delta}(c)$.

4. $M, p \Vdash_L \boldsymbol{\nu} \;\Rightarrow\; M, q \Vdash_L \boldsymbol{\nu}_0$ *für alle* $q \in M$ *mit* $(p, q) \in R$.

   $M, p \Vdash_L \boldsymbol{\pi} \;\Rightarrow\; M, q \Vdash_L \boldsymbol{\pi}_0$ *für ein* $q \in M$ *mit* $(p, q) \in R$.

5. $\Vdash_G \pi \to \Diamond\left(\pi_0 \wedge \bar{\pi}\right)$

Beweis. Wir wollen die Teilaussagen 3. und 5. beweisen. Für die restlichen Aussagen ist der Beweis sehr einfach.

3. Wir betrachten der Einfachheit halber den Fall $\boldsymbol{\delta} = (\exists x)\,\alpha$. Sei w eine M-Belegung. Es gilt dann:

$M, p, w \Vdash_L (\exists x)\,\alpha$

$\Leftrightarrow$ es existiert ein $a \in |\mathfrak{I}_p|$ mit $M, p, w\{x/a\} \Vdash_L \alpha$.

Sei nun $(W, R, \mathfrak{I}') = M\left(\mathfrak{I}_p\{c/a\}\right)$ eine durch die Festlegung $c^{\mathfrak{I}'} := a$ bestimmte Strukturerweiterung von M zu einer $S(\mathcal{K} \cup \{c\})$-Kripke-Struktur. Es gilt dann

$$M\left(\mathfrak{I}_p\{c/a\}\right), p, w\left\{x/c^{\mathfrak{I}'}\right\} \Vdash_L \alpha$$

und damit nach dem Substitutionstheorem

$$M\left(\mathfrak{I}_p\{c/a\}\right), p, w \Vdash_L \alpha\{x/c\}.$$

Da dies für beliebige Belegungen w gilt, folgt die Behauptung.

5. Wegen $\Vdash_{ML} \neg\Box\alpha \leftrightarrow \Diamond\neg\alpha$ genügt es, den Fall $\pi = \neg\Box\alpha$ zu betrachten.

Es gilt:

$\Vdash_G \Box(\Box\alpha \rightarrow \alpha) \rightarrow \Box\alpha$

$\Rightarrow \Vdash_G \neg\Box\alpha \rightarrow \neg\Box(\Box\alpha \rightarrow \alpha)$

$\Rightarrow \Vdash_G \neg\Box\alpha \rightarrow \Diamond(\Box\alpha \wedge \neg\alpha)$

d.h. es gilt $\Vdash_G \pi \rightarrow \Diamond(\overline{\boldsymbol{\pi}} \wedge \boldsymbol{\pi}_0)$

■

Notation 4.64 *Sei Σ eine Menge von Markierungen der Form $p \Vdash \beta$ bzw. pRq. (Beachten Sie, dass dies informelle Schreibweisen für (p, β) bzw. für (p, q) sind).*

1. *R^*_Σ sei die Menge aller Paare (p, q), so dass gilt:*

   (a) $p = q$ und $p \Vdash \beta \in \Sigma$ für eine Aussage β oder

   (b) es gibt für ein $n \in \mathbb{N}$ $p_0, \ldots, p_n$ mit $p_0 = p$, $p_n = q$ und $p_iRp_{i+1} \in \Sigma$, $0 \leq i < n$.

2. *$\mathcal{K}_{\Sigma,p}$ sei die Menge aller Konstantensymbole aus $S(\mathcal{K}_0)$, die in einer Formel β mit $q \Vdash \beta \in \Sigma$ für ein q mit $(q, p) \in R^*_\Sigma$ vorkommen.*

3. *$W_\Sigma := \{p_i \mid i \in \mathbb{N},\ p_i$ kommt in Σ vor$\}$*

Definition 4.65 *Sei $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$ und $M' = (W', R', \mathfrak{J}')$ eine $S(\mathcal{K})$-Kripke-Struktur. Σ sei eine Menge von Markierungen der Form pRq bzw. $p \Vdash \beta$ für $\beta \in Fm_{ML}(S(\mathcal{K}))$.*

1. *Eine Abbildung $f : W_\Sigma \rightarrow W'$ heißt* L-Einbettung *von Σ in M' genau dann, wenn gilt*

$$pRq \in \Sigma_P \ \Rightarrow\ (f(p), f(q)) \in R'.$$

2. *Σ heißt* L-erfüllbar unter der L-Einbettung f *von Σ in M' genau dann, wenn für jede Markierung $p \Vdash \beta \in \Sigma$ gilt:*

$$M', f(p) \Vdash_L \beta.$$

3. *Σ heißt* L-erfüllbar *(in M') genau dann, wenn eine L-Einbettung f von Σ in M' existiert, so dass Σ L-erfüllbar unter der L-Einbettung f ist.*

Notation 4.66 *Die Menge der beim Aufbau eines Tableaus T neu eingeführten Konstantensymbole aus $\mathcal{K}_0$ sei mit $\mathcal{K}_T$ notiert.*

Definition 4.67 *Sei T ein L-Tableau für ein $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$.
Für einen Zweig P des L-Tableaus T sei Σ_P die Menge der im Zweig P vorkommenden Markierungen. Das Tableau T heißt* L-erfüllbar *(in M') genau dann, wenn für einen Zweig P im L-Tableau T Σ_P L-erfüllbar (in M') ist.*

Wir sind nun in der Lage, das für den Korrektheitsbeweis zentrale Lemma in kompakter Weise zu formulieren und anschließend zu beweisen.

Lemma 4.68 *Sei $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$ und T ein L-Tableau für $\{\alpha\} \subseteq Fm_{ML}(S)$. Sei ferner $M = (W', R', \mathfrak{J})$ eine $S(\mathcal{K}_T)$-Kripke-Struktur.
Es gilt dann:
Ist T L-erfüllbar (in M) und resultiert das Tableau T' aus T durch die Anwendung einer Tableau-Regel für die Logik L, so ist T' L-erfüllbar (in einer $S(\mathcal{K}_{T'})$-Kripke-Struktur M').*

Beweis. Sei T L-erfüllbar (in M). Es gibt dann einen Zweig P in T, so dass Σ_P L-erfüllbar (in M) ist.
Wir müssen nach der zuletzt angewandten Tableau-Expansionsregel unterscheiden. Wir werden exemplarisch die Fälle der $\boldsymbol{\delta}$-Formeln und der π-Formeln betrachten. In den übrigen Fällen ist die Argumentation sehr einfach (entweder weitestgehend analog dem aussagenlogischen bzw. dem prädikatenlogischen Fall oder aber trivial).
Wir können uns auf den Fall beschränken, dass beim Übergang von T zu T' der Zweig P bearbeitet worden ist, da im anderen Fall die Aussage sofort folgt (vgl. den Beweis im prädikatenlogischen Fall).
Sei $f : W_{\Sigma_P} \to W'$ eine Einbettung von Σ_P in M, die Σ_P L-erfüllt.

Fall δ-Formel: Der Zweig P ist dann aufgrund der Tableau-Expansionsregel $\dfrac{p \Vdash \boldsymbol{\delta}}{p \Vdash \boldsymbol{\delta}(c)}$ zu P' erweitert worden.

Nach Voraussetzung gilt $M, f(p) \Vdash_L \boldsymbol{\delta}$. Nach Lemma 4.63 3. existiert dann eine Strukturerweiterung M' mit $M\prime, f(p) \Vdash_L \boldsymbol{\delta}(c)$.

Da für alle Aussagen β, in denen c nicht vorkommt, und für alle $q \in W'$ gilt

$$M, q \Vdash_L \beta \Leftrightarrow M', q \Vdash_L \beta \ ,$$

folgt, dass f eine Einbettung von $\Sigma_{P'}$ in M' ist, die $\Sigma_{P'}$ L-erfüllt.

Fall π-Formel: In diesem Fall ist P aufgrund der Expansionsregel für eine π-Formel in einem Knoten $s \in P$ zu P' erweitert worden.

Sei die Markierung von s $p \Vdash \boldsymbol{\pi}$ für eine π-Formel.

Wir betrachten hier den Fall der Logik $\boldsymbol{G}$, da in diesem Fall der Zweig P außer um pRq (für ein neues q) und $q \Vdash \boldsymbol{\pi}_0$ zuätzlich verlängert wird um die „Markierung" $q \Vdash \overline{\boldsymbol{\pi}}$.

Nach Voraussetzung gilt $M, f(p) \Vdash_G \boldsymbol{\pi}$. In $\boldsymbol{G}$ gilt nach Lemma 4.63 5. $\Vdash_G \boldsymbol{\pi} \to \Diamond(\boldsymbol{\pi}_0 \wedge \overline{\boldsymbol{\pi}})$. Es gibt daher eine Welt $r \in W'$ mit $(f(p), r) \in R'$ und $M, r \Vdash_G \boldsymbol{\pi}_0 \wedge \overline{\boldsymbol{\pi}}$.

Wir definieren daher eine Fortsetzung f' von f auf $W_{\Sigma_{P'}} = W_{\Sigma_P} \cup \{q\}$ folgendermaßen:

$$f'(s) := \begin{cases} r & \text{falls } s = q \\ f(s) & \text{sonst.} \end{cases} \ .$$

f' ist dann eine Einbettung von $\Sigma_{P'}$ in M (leicht zu verifizieren), die $\Sigma_{P'}$ L-erfüllt.

In jedem Fall ergibt sich somit, dass das resultierende Tableau T' L-erfüllbar ist. ■

Mit diesem Lemma erhalten wir nun leicht die Korrektheit der Tableau-Kalküle.

Theorem 4.69 (*Korrektheit der Tableau-Kalküle*)
Für $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$ und $\alpha \in Fm_{ML}(S)$ Aussage gilt:

$$\vdash_{L-T} \alpha \Rightarrow \Vdash_L \alpha \ .$$

Beweis. Es gelte $\vdash_{L-T} \alpha$ für $L \in \{\boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}, \boldsymbol{G}\}$. Es existiert dann ein geschlossenes L-Tableau T für $\{\neg\alpha\}$. Sei $p_0 \Vdash \neg\alpha$ die Markierung des Startknotens von T.
Angenommen α ist nicht allgemeingültig für L. Es existiert dann eine Kripke-Struktur $M = (W, R, \mathfrak{J})$ und eine Welt $q \in M$ mit $M, q \Vdash_L \neg\alpha$. Damit ist die Abbildung $f : \{p_0\} \to W$ mit $f(p_0) = q$ eine L-Einbettung von Σ_{T_0} des Starttableaus T_0, die Σ_{T_0} und damit T_0 L-erfüllt. (Man beachte, dass T_0 in diesem Fall auch ein Zweig ist.)
Analog zum aussagenlogischen und prädikatenlogischen Korrektheitsbeweis folgt durch Induktion und unter Anwendung von Lemma 4.68, dass dann auch das Tableau T L-erfüllbar in einer Strukturerweiterung M' von M ist. Es gibt daher einen Zweig P in T sowie eine Einbettung f_P von Σ_P in M', die Σ_P L-erfüllt. Da T ein geschlossenes Tableau ist, gibt es in dem Zweig P entweder einen Knoten mit einer Markierung $r \Vdash \bot$ oder aber zwei Knoten mit Markierungen $r \Vdash \beta$ und $r \Vdash \neg\beta$ für ein r und eine Aussage β.
Aufgrund der Einbettung von Σ_P in M' gilt dann $M', f_P(r) \Vdash_L \bot$ bzw. $M', f_P(r) \Vdash_L \beta$ und $M', f_P(r) \Vdash_L \neg\beta$. Dies kann für keine Kripke-Struktur gelten. Die Annahme, dass α nicht allgemeingültig ist, ist somit falsch. ■

4.3.4 Vollständigkeit der Tableau-Kalküle

Mit Ausnahme der Modallogik $\boldsymbol{G}$ kann für die hier vorgestellten Modallogiken die Vollständigkeit der angegebenen Tableau-Kalküle gezeigt werden. Wir könnten dazu im Prinzip in ähnlicher Weise vorgehen wie im Falle der Prädikaten- bzw. Aussagenlogik.
Hier soll eine weitere Methode, die auf der systematischen Erzeugung von Tableaus basiert, vorgestellt werden. Diese Methode zur Erzeugung von Tableaus hat den Vorteil, dass sie für allgemeingültige Aussagen α *in jedem Fall* ein geschlossenes Tableau für $\{\neg\alpha\}$ liefert, im Gegensatz zu dem bisher betrachteten Tableau-Verfahren.
In einem ersten Schritt modifizieren wir die Tableau-Regeln für die γ-Formeln und für die ν-Formeln:

- Die Expansionsregeln für γ-Formeln und ν-Formeln werden im Aufbau eines Tableaus im allgemeinen mehrfach benötigt. Um später eine leichtere Übersicht darüber zu erhalten, welche Formeln in einem Tableau bereits bearbeitet worden sind, erweitern wir die Tableau-Regeln dahingehend, dass an einen Zweig, der aufgrund einer $\boldsymbol{\gamma}$-Expansionsregel $\dfrac{p \Vdash \boldsymbol{\gamma}}{p \Vdash \boldsymbol{\gamma}(c)}$ oder einer $\boldsymbol{\nu}$-Expansionsregel $\dfrac{p \Vdash \boldsymbol{\nu}}{q \Vdash \boldsymbol{\nu}_0}$ erweitert wird, zusätzlich ein Knoten mit Markierung $p \Vdash \boldsymbol{\gamma}$ bzw. $p \Vdash \boldsymbol{\nu}$ angehängt wird.

Systematisches Tableau-Verfahren

Sei wie schon in der Definition des Tableau-Kalküls $W = \{p_i \mid i \in \mathbb{N}\}$ ein Vorrat von „Welten" und $\mathcal{K}_0 = \{c_i \mid i \in \mathbb{N}\}$ eine Vorrat von Konstantensymbolen.
Wir wollen jetzt ein Tableau-Verfahren vorstellen, das zwar wenig intuitiv ist, aber den Vorteil hat, dass für allgemeingültige Aussagen das Verfahren in jedem Fall zum Abschluss kommt und somit einen Beweis liefert.

Es soll $\Vdash_L \alpha$ für eine Logik $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$ gezeigt werden.

Notation 4.70 *Für eine atomare modallogische Aussage α heiße eine Markierung $p \Vdash \alpha$* atomar.

Systematische Tableau-Methode:

Die Prämissenmenge Σ liege in Form einer Aufzählung vor.

Schritt 1**:** Erzeuge das Starttableau für $\{\neg\alpha\}$. Der Startknoten ist dann mit $p_0 \Vdash \neg\alpha$ markiert.
Seien nun bereits n Schritte ausgeführt und das Tableau T_n aufgebaut worden.

Schritt $n+1$:

$n+1$ **gerade:** Wähle im Tableau T_n nichtdeterministisch einen noch nicht bearbeiteten Knoten s geringster Tiefe aus. Sei $p \Vdash N$ die Markierung von s.
Markiere s als bearbeitet.
In Abhängigkeit von N wird T_n erweitert.
Für jeden nicht geschlossenen Zweig P mit $s \in P$ und Blatt u wird an u angefügt:

- falls $N = \neg\top$: $p \Vdash \bot$
- falls $N = \neg\bot$: $p \Vdash \top$
- falls $N = \neg\neg\beta$: $p \Vdash \beta$
- falls N α-Formel mit den Komponenten $\boldsymbol{\alpha}_0$ und $\boldsymbol{\alpha}_1$:

$$\begin{array}{c} p \Vdash \boldsymbol{\alpha}_0 \\ | \\ p \Vdash \boldsymbol{\alpha}_1 \end{array}$$

- falls N β-Formel mit den Komponenten $\boldsymbol{\beta}_0$ und $\boldsymbol{\beta}_1$:

$$\begin{array}{ccc} / & & \backslash \\ p \Vdash \boldsymbol{\beta}_0 & & p \Vdash \boldsymbol{\beta}_1 \end{array}$$

- falls N γ-Formel:
  K sei die Menge der Konstantensymbole c, die der folgenden Bedingung genügen:
  c kommt in einer Markierung $q \Vdash \beta \in \Sigma_P$ (β Aussage) des Zweiges P vor, wobei gilt:

$$\begin{array}{ll} (q,p) \in R^* \text{ und } q \neq p & \text{oder} \\ p = q \text{ und } \beta \neq \boldsymbol{\gamma}(c)\,. & \end{array}$$

  Falls $K = \emptyset$, setze $K := \{c\}$, wobei c das erste Konstantensymbol der Signatur S ist.
  Sei somit $K = \{c_1, \ldots, c_k\}$ mit $k \geq 1$.
  An u wird dann angehängt

$$\begin{array}{c} p \Vdash \boldsymbol{\gamma}(c_1) \\ | \\ \vdots \\ | \\ p \Vdash \boldsymbol{\gamma}(c_k) \\ | \\ p \Vdash \boldsymbol{\gamma} \end{array}$$

  Anschaulich: Für jedes Konstantensymbol c, das in einem Pfad zur „Welt" p auftritt, bilden wir eine $\boldsymbol{\gamma}$-Instanz $\boldsymbol{\gamma}(c)$ und hängen die Markierung $p \Vdash \boldsymbol{\gamma}(c)$ an den Zweig P an, sofern sie im Zweig noch nicht vorhanden ist. Kommt kein Konstantensymbol vor, so hängen wir für ein Konstantensymbol c aus der Signatur die Markierung $p \Vdash \boldsymbol{\gamma}(c)$ an den Zweig P an.
- falls N δ-Formel:

$$p \Vdash \boldsymbol{\delta}(c_j)$$

  für das kleinste $j \in \mathbb{N}$, so dass $c_j \in \mathcal{K}_0$ nicht in T_n vorkommt.
- falls N ν-Formel:
  Seien $q_1, \ldots, q_k$ sämtliche „Welten", für die in P eine Markierung pRq_i und eine Markierung $q_i \Vdash \beta$ für ein β vorhanden ist $(i = 1, \ldots, k)$:

$$\begin{array}{c} q_1 \Vdash \nu_0 \\ | \\ \vdots \\ | \\ q_k \Vdash \nu_0 \\ | \\ p \Vdash \nu \end{array}$$

- falls N π-Formel:

$$\begin{array}{c} pRp_i \\ p_i \Vdash \boldsymbol{\pi}_0 \end{array}$$

  für das kleinste $i \in \mathbb{N}$, so dass $p_i \in W$ nicht in T_n vorkommt.

$n + 1$ **ungerade:** In diesem Fall werden die Tableau-Erweiterungsregeln für die Prämissen sowie für die jeweilige Modallogik systematisch durchgeführt. Z.B. für:
***T* :** es wird an jeden Zweig von T_n, in dem eine Markierung der Form $p \Vdash \beta$ vorkommt,ein Knoten mit Markierung pRp angefügt, sofern diese im Zweig noch nicht vorhanden ist.
S4 :Wie für ***T***. Zusätzlich wird an jeden Zweig, in dem Markierungen pRq und qRr vorhanden sind, eine Markierung pRr angehängt.
Für die Prämissenmenge Σ : für die in der Aufzählung erste Aussage α, die in Σ_P nicht vorkommt, wird $p_0 \Vdash \alpha$ angehängt.

Das Verfahren endet, wenn alle Knoten, die mit Formeln markiert sind, bearbeitet sind oder wenn das Tableau geschlossen ist. In diesem Fall heißt das Tableau beendet (oder abgearbeitet). Ist das Tableau noch nicht geschlossen und sind auch nicht alle Knoten als bearbeitet markiert, dann wird Schritt $n+2$ ausgeführt.

Es sollte klar sein, dass dieses so beschriebene systematische Verfahren wiederum ein korrektes Verfahren ist. Wir wollen im folgenden den Nachweis erbringen, dass es auch ein vollständiges Beweisverfahren ist, und dass es im Falle von allgemeingültigen Aussagen nach endlich vielen Schritten abbricht.

Modallogische Hintikka-Mengen

Hintikka-Mengen spielten eine große Rolle in den Vollständigkeitsbeweisen für die Aussagen- und Prädikatenlogik. Im Beweis des jeweiligen Modellexistenztheorems wurde eine Aussagenmenge zu einer maximal widerspruchsfreien Hintikka-Menge erweitert. Eine Hintikka-Menge war in diesen Fällen aber immer erfüllbar.

Definition 4.71 (Modallogische Hintikka-Mengen) *Sei Σ eine Menge von Markierungen der Form pRq bzw $p \Vdash \beta$ und β modallogische Aussage.*
Σ heißt modallogische L-Hintikka-Menge
($L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$) genau dann, wenn gilt:

1. *Ist α eine atomare Aussage und ist $p \Vdash \alpha \in \Sigma$, so gilt $p \Vdash \neg\alpha \notin \Sigma$.*
2. *Für alle p gilt: $p \Vdash \neg\top \notin \Sigma, \quad p \Vdash \bot \notin \Sigma$.*
3. *$p \Vdash \neg\neg\beta \in \Sigma \;\Rightarrow\; p \Vdash \beta \in \Sigma$.*
4. *Für eine α-Formel mit den Komponenten $\boldsymbol{\alpha}_0$ und $\boldsymbol{\alpha}_1$:*
   $p \Vdash \boldsymbol{\alpha} \in \Sigma \;\Rightarrow\; p \Vdash \boldsymbol{\alpha}_0,\ p \Vdash \boldsymbol{\alpha}_l \in \Sigma$.
5. *Für eine β-Formel mit den Komponenten $\boldsymbol{\beta}_0$ und $\boldsymbol{\beta}_1$:*
   $p \Vdash \boldsymbol{\beta} \in \Sigma \;\Rightarrow\; p \Vdash \boldsymbol{\beta}_0 \in \Sigma$ *oder* $p \Vdash \boldsymbol{\beta}_l \in \Sigma$.
6. *Für eine γ-Formel:*
   $p \Vdash \boldsymbol{\gamma} \in \Sigma \;\Rightarrow$ *für alle* $c \in \mathcal{K}_{\Sigma,p}$ *gilt* $p \Vdash \boldsymbol{\gamma}(c) \in \Sigma$.
7. *Für eine δ-Formel:*
   $p \Vdash \boldsymbol{\delta} \in \Sigma \;\Rightarrow$ *es existiert ein* $c \in S(\mathcal{K}_0)$ *mit* $p \Vdash \boldsymbol{\delta}(c) \in \Sigma$.
8. *Für eine ν-Formel:*
   $p \Vdash \boldsymbol{\nu} \in \Sigma \;\Rightarrow$ *für alle q mit $pRq \in \Sigma$ gilt* $q \Vdash \boldsymbol{\nu}_0 \in \Sigma$.
9. *Für eine π-Formel:*
   $p \Vdash \boldsymbol{\pi} \in \Sigma \;\Rightarrow$ *es existiert ein q mit $pRq \in \Sigma$ und* $q \Vdash \boldsymbol{\pi}_0 \in \Sigma$
10. *Die L zugrundeliegenden Rahmeneigenschaften bedingen Abschlusseigenschaften für „R" :*
    z.B. für $\boldsymbol{S4}$:
    $p \Vdash \beta \in \Sigma \;\Rightarrow\; pRp \in \Sigma$
    $pRq,\ qRr \in \Sigma \;\Rightarrow\; pRr \in \Sigma$

Notation 4.72 *Eine Kripke-Struktur $M = (W', R', \mathfrak{J})$ wird als L-*Kripke-Struktur *bezeichnet, wenn der Rahmen (W', R') die für die Logik $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$ charakteristischen Rahmeneigenschaften aufweist (z.B. Reflexivität für $L = \boldsymbol{T}$).*

Theorem 4.73 *Sei Σ eine modallogische L-Hintikka-Menge, $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$.*
Dann gilt:
Σ ist L-erfüllbar in einer L-Kripke-Struktur $M = (W, R, \mathfrak{J})$ mit $W = W_\Sigma$.

Beweis. Sei $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$.
Wir definieren eine L-Kripke-Struktur $M = (W', R', \mathfrak{J})$ wie folgt:
$W' := W_\Sigma$
$R' := \{(p, q) \mid pRq \in \Sigma\}$
Offensichtlich gilt für die Relation R' :

$$R' \text{ ist } \begin{cases} \text{ohne Beschränkung} & & L = \boldsymbol{K} \\ \text{reflexiv} & & L = \boldsymbol{T} \\ \text{reflexiv, transitiv} & \text{falls} & L = \boldsymbol{S4} \\ \text{reflexiv, symmetrisch} & & L = \boldsymbol{B} \\ \text{reflexiv, euklidisch} & & L = \boldsymbol{S5} \end{cases}$$

Für jedes $p \in W'$ sei $\mathfrak{J}_p$ eine prädikatenlogische $S(\mathcal{K}_{\Sigma,p})$-Herbrand-Struktur mit Grundbereich $\mathcal{K}_{\Sigma,p}$, so dass für ein n-stelliges Prädikatensymbol $Q \in S$ gilt

$$Q^{\mathcal{J}_p}(c_1, \ldots, c_n) = T \;\Leftrightarrow\; p \Vdash Q(c_1, \ldots, c_n) \in \Sigma.$$

Die Welteninterpretation $\mathfrak{J}$ ist dann definiert durch $\mathfrak{J}(p) := \mathfrak{J}_p$ für alle $p \in W'$.
Es gilt für beliebige $p \in W'$ und $\beta \in Fm_{ML}(S(\mathcal{K}_0))$:

$$p \Vdash \beta \in \Sigma \;\Rightarrow\; M, p \Vdash_L \beta. \tag{4.4}$$

Der Beweis wird durch strukturelle Induktion über den Aufbau von Formeln geführt.
Für den Induktionsschritt wollen wir exemplarisch einige Fälle beweisen.

1. Fall: β ist eine γ-Formel:

   Nach Voraussetzung gilt $p \Vdash \boldsymbol{\gamma} \in \Sigma$. Da Σ eine Hintikka-Menge ist, gilt $p \Vdash \gamma(c) \in \Sigma$ für alle $c \in \mathcal{K}_{\Sigma,p}$.

   Nach Induktionsvoraussetzung gilt dann für alle $c \in \mathcal{K}_{\Sigma,p}$ $M, p \Vdash_L \boldsymbol{\gamma}(c)$, also aufgrund der Definition der Semantik $M, p \Vdash_L \boldsymbol{\gamma}$.

2. Fall: β ist eine ν-Formel:

   Nach Voraussetzung gilt $p \Vdash \boldsymbol{\nu} \in \Sigma$. Da Σ eine Hintikka-Menge ist, gilt für alle q mit $pRq \in \Sigma$ $p \Vdash \boldsymbol{\nu}_0 \in \Sigma$, also für alle $q \in W'$ mit $(p, q) \in R'$ $p \Vdash \boldsymbol{\nu}_0 \in \Sigma$. Nach Induktionsvoraussetzung gilt für alle $q \in W'$ mit $(p, q) \in R'$ $M, q \Vdash_L \boldsymbol{\nu}_0$ und damit $M, p \Vdash_L \boldsymbol{\nu}$.

3. Fall: β ist eine π-Formel:

   Nach Voraussetzung gilt $p \Vdash \boldsymbol{\pi} \in \Sigma$. Da Σ eine Hintikka-Menge ist, gibt es ein q mit $pRq \in \Sigma$ und $q \Vdash \boldsymbol{\pi}_0 \in \Sigma$. Also gibt es ein $q \in W'$ mit $(p, q) \in R'$ und $q \Vdash \boldsymbol{\pi}_0 \in \Sigma$. Nach Induktionsvoraussetzung gilt für dieses q $M, q \Vdash_L \boldsymbol{\pi}_0$. Somit gilt $M, p \Vdash_L \boldsymbol{\pi}$.

Aus 4.4 folgt aber sofort die L-Erfüllbarkeit von Σ. ■

Theorem 4.74 (Vollständigkeit der Tableau-Kalküle) *Sei $\alpha \in Fm_{ML}(S)$ eine modallogische Aussage.*
Es gilt dann:
$\Vdash_L \alpha \;\Rightarrow \vdash_{L-T} \alpha$ *für $L \in \{\boldsymbol{K}, \boldsymbol{T}, \boldsymbol{S4}, \boldsymbol{B}, \boldsymbol{S5}\}$.*

Beweis. (Skizze)
Wir beweisen die Kontraposition

$$\nvdash_{L-T} \alpha \;\Rightarrow \nVdash_L \alpha$$

$\vdash_{L-T} \alpha$ gelte nicht, d.h. es existiert kein geschlossenes L-Tableau für $\{\neg\alpha\}$.

Es existiert somit insbesondere kein geschlossenes systematisches L-Tableau.
Wir betrachten eine systematische Tableau-Konstruktion für $\{\neg\alpha\}$. Diese bricht entweder nach endlich vielen Schritten mit einem Tableau ab, in dem alle Knoten bearbeitet sind und das einen nicht geschlossenen Zweig P enthät, oder aber die Konstruktion terminiert nicht. Im letzteren Fall führt die Konstruktion zur Erzeugung eines unendlich langen, nicht geschlossenen Pfades, der ebenfalls mit P bezeichnet sei. In beiden Fällen ist die Menge der Markierungen von Knoten des Zweiges bzw. Pfades P in einer modallogischen L-Hintikka-Menge Σ enthalten (leicht zu beweisen unter Verwendung der Tatsache, dass die Tableau-Konstruktion systematisch erfolgt, also jeder Knoten des Zweiges irgendwann bearbeitet wird). Nach Theorem 4.73 ist Σ somit L-erfüllbar in einer L-Kripke-Struktur $M = (W, R, \mathfrak{J})$. Da $p_0 \Vdash \neg\alpha \in \Sigma$, gilt insbesondere dann $M, p_0 \Vdash_L \neg\alpha$. Da in α keine Konstantensymbole aus $\mathcal{K}_0$ enthalten sind, gilt daher auch $M', p_0 \Vdash_L \neg\alpha$ für die aus M resultierende S-Kripke-Struktur $M' = (W, R, \mathfrak{J}')$, die sich aus M dadurch ergibt, dass für jedes $p \in W$ $\mathfrak{J}'_p$ die Beschränkung von $\mathfrak{J}_p$ auf die Signatur S ist. α ist folglich nicht allgemeingültig.
Damit ist die Kontraposition und also auch die Vollständigkeit bewiesen. ■

4.3.5 Ausblick auf eine Verallgemeinerung

Der Modallogik, so wie wir sie eingeführt haben, liegen sehr starke Voraussetzungen zugrunde. Wir hatten gefordert, dass Konstantensymbole in allen Welten gleich interpretiert werden und dass bei einem Übergang von einer Welt p in eine Welt q alle Objekte, die in der Welt p vorhanden sind, auch in q enthalten sind, also eine Monotonie-Bedingung erfüllt sein muss. Darüberhinaus hatten wir auf die Verwendung von Funktionssymbolen verzichtet. Dies ist im Falle der Prädikatenlogik völlig unproblematisch, da eine Funktion imer durch ihren Graph repräsentiert werden kann. Im Falle der Modallogik ist die Sachlage jedoch etwas komplizierter, vor allen Dingen dann, wenn ein allgemeinerer Aufbau der Modallogik angestrebt wird. Wir wollen hier kurz auf einige Probleme und Möglichkeiten eingehen, die sich aus solch einem allgemeineren Aufbau ergeben.
Neben Funktionssymbolen, die wir immer noch nicht in voller Allgemeinheit berücksichtigen werden, und dem Gleichheitssymbol werden wir noch Prädikatsabstraktionen als neue Elemente zum Aufbau der formalen Sprache verwenden.

Wir wollen zunächst λ-Abstraktionen informell erläutern, wobei in dieser informellen Darstellung Objekt- und Metasprache nicht streng auseinandergehalten werden. Betrachten wir den Ausdruck $x + 3$, wobei x als Variable für eine natürliche Zahl stehen soll. Mit $\langle \lambda x.\, x + 3\rangle$ bezeichnen wir dann dasjenige Objekt (in diesem Fall eine Funktion), das angewandt auf eine natürliche Zahl n das Ergebnis $n + 3$ liefert, d.h. es gilt $\langle \lambda x.\, x + 3\rangle\,(n) = n+3$. (Die spitzen Klammern verwenden wir nur zur besseren Lesbarkeit und sind eigentlich nicht Bestandteil der λ-Notation.) Z.B. ist also $\langle \lambda x.\, x + 3\rangle\,(5) = 8$. Entsprechend: haben wir einen Ausdruck $x > 2 \wedge x < 16$ gegeben und sind die natürlichen Zahlen zugrundegelegt, so bezeichnet $\left\langle \lambda x.\, x^2 > 2 \wedge x^2 < 16\right\rangle$ dasjenige Prädikat $\mathfrak{P}$, das auf ein n genau dann zutrifft, wenn $n^2 > 2 \wedge n^2 < 16$ gilt. So trifft $\mathfrak{P}$ also nur auf die natürlichen Zahlen 2 und 3 zu.
Haben wir einen Ausdruck $x \cdot y > 24 \wedge x + y < 15$ mit den Variablen x und y gegeben, so erhalten wir durch die Bildung von $\langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle$ ein von dem „Parameter y" abhängiges Prädikat, das angewandt auf ein n das Resultat $n \cdot y > 24 \wedge n + y < 15$ hat, d.h. es gilt

$$\langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(n) \;\Leftrightarrow\; n \cdot y > 24 \wedge n + y < 15.$$

Zu $\langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(n)$ kann man nun wiederum auch $\langle \lambda y.\, \langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(n)\rangle$ bilden. Dies ist dann das Prädikat, das angewandt auf die natürliche Zahl m das Ergebnis $\langle \lambda x.\, x \cdot m > 24 \wedge x + m < 15\rangle\,(n)$ hat. Es gilt daher

$$\begin{aligned}\langle \lambda y.\, \langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(n)\rangle\,(m) &\Leftrightarrow \langle \lambda x.\, x \cdot m > 24 \wedge x + m < 15\rangle\,(n)\\ &\Leftrightarrow n \cdot m > 24 \wedge n + m < 15.\end{aligned}$$

Statt $\langle \lambda y.\, \langle \lambda x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(n)\rangle\,(m)$ wollen wir auch $\langle \lambda y, x.\, x \cdot y > 24 \wedge x + y < 15\rangle\,(m, n)$ schreiben.

Wir wollen nun die Verwendung der λ-Abstraktion im Aufbau einer modallogischen Sprache formalisieren. Wir lassen dabei zu, dass unsere Signatur Funktionssymbole enthält. Ferner soll das Gleichheitssymbol in der Signatur vorhanden sein.
Die Definition der modallogischen Formeln wird dann um die folgende Punkte ergänzt:

Definition 4.75 (*Fortsetzung von Definition 4.1 2.*)

f)

1. *Ist* $\alpha \in Fm_{ML}(S)$ *eine modallogische Formel, in der* x *nicht gebunden vorkommt, und* $x \in Var_{ML}$ *eine Variable, dann ist* $\langle \lambda x. \alpha \rangle$ *ein* Prädikats-Abstrakt.
   Die freien Variablen von $\langle \lambda x. \alpha \rangle$ *sind* $fvar(\alpha) \setminus \{x\}$.
2. *Ist* $\langle \lambda x. \alpha \rangle$ *ein Prädikats-Abstrakt und* $t \in Tm_{ML}(S)$ *ein Term, dann ist* $\langle \lambda x. \alpha \rangle (t) \in Fm_{ML}(S)$.
   Die freien Variablen von $\langle \lambda x. \alpha \rangle (t)$ *sind* $(fvar(\alpha) \setminus \{x\}) \cup var(t)$.

g) Sind $s, t \in Tm_{ML}(S)$, *so ist* $s \approx t \in Fm_{ML}(S)$.

Beispiel 4.76 *Folgende Zeichenreihen sind Formeln:*

1. $\langle \lambda x. P(x,y) \to (\exists z) P(z,y) \rangle (f(x,y))$
2. $(\forall y) \langle \lambda x. Q(x) \rangle (f(x,y))$
3. $\langle \lambda x. x \approx y \rangle (z)$

Die Semantik für Formeln dieser Art wird dann in der folgenden Weise festgelegt.

Definition 4.77 (*Ergänzung zu Definition 4.5 2. und 4.9 2.*)

1. *Eine* normale $(S-)$ Kripke-Struktur *ist ein Tripel* $(W, R, \mathfrak{J})$, *wobei*
   (W, R) *ein Kripke-Rahmen und* $\mathfrak{J}$ *eine Abbildung ist, die jeder Welt* $p \in W$ *eine normale prädikatenlogische* S*-Struktur* $\mathfrak{J}_p$ *zuordnet.*
2. *Sei* $M = (W, R, \mathfrak{J})$ *eine* S*-Kripke-Struktur,* $p \in M$ *und* w *eine* M*-Belegung.*
   (a) $M, p, w \Vdash_{ML} \langle \lambda x. \alpha \rangle (t)$ *genau dann, wenn* $M, p, w \left\{ x / t^{\mathfrak{J}_p} \right\} \Vdash_{ML} \alpha$.
   (b) $M, p, w \Vdash_{ML} s \approx t$ *genau dann, wenn* $s^{\mathfrak{J}_p, w} = t^{\mathfrak{J}_p, w}$.

Bemerkung 4.78

1. *Wir lassen nun zu, dass für ein Konstantensymbol* c *und Welten* p *und* q $c^{\mathfrak{J}_p} \neq c^{\mathfrak{J}_q}$ *gilt. (Im Englischen werden Konstantensymbole, die nur noch lokal für jede Welt als Konstanten anzusehen sind, als* flexible designators *bezeichnet.)*
2. *Im Falle, dass in der Signatur Funktionssymbole vorhanden sind, legen wir zur Vereinfachung der weiteren Darstellung constant domains zugrunde.*
3. *Die Identität* $=$ *ist unabhängig von der aktuellen Welt als globale Relation gegeben.*

Die Tatsache, dass wir nun lokale und nicht globale Konstanten haben, bringt nun aber Probleme für die Semantik mit sich, die wir in der Definition der Semantik bisher nicht angesprochen haben. Wie ist z.B. eine Formel $\Diamond P(c)$ für ein Konstantensymbol c zu interpretieren? Legen wir eine S-Kripke-Struktur $M = (W, R, \mathfrak{J})$, eine Welt $p \in M$ und eine M-Belegung w zugrunde, dann gibt es zwei Alternativen:

1. $M, p, w \Vdash_{ML} \Diamond P(c) \quad \Leftrightarrow \quad$ es existiert $q \in W$ mit pRq und $M, q, w \Vdash_{ML} P(c)$
$\Leftrightarrow \quad$ es existiert $q \in W$ mit pRq und $c^{\mathfrak{I}_q} \in P^{\mathfrak{I}_q}$

   D.h. der $\Diamond$-Operator hat Priorität. Wir nehmen erst den Übergang in eine Welt q vor und verwenden dann die Interpretation von c in der Welt q. Wir sprechen dann von einer *de dicto Interpretation*, da erst der Übergan vorgenommen wird.

2. $M, p, w \Vdash_{ML} \Diamond P(c) \quad \Leftrightarrow \quad M, p, w \left\{x/c^{\mathfrak{I}_p}\right\} \Vdash_{ML} \Diamond P(x)$
$\Leftrightarrow \quad$ es existiert $q \in W$ mit pRq und $c^{\mathfrak{I}_p} \in P^{\mathfrak{I}_q}$

   In diesem Fall wird vorrangig das Konstantensymbol c in der aktuellen Welt p interpretiert und dann erst der Übergang in eine Welt q vollzogen (*de re Interpretation*).

Da wir beide Alternativen ausdrücken können wollen, müssen diese Alternativen aber durch unterschiedliche Formeln wiedergegeben werden. Die Prädikats-Abstraktion bietet uns hier eine Möglichkeit. Mit ihrer Hilfe können wir diese Mehrdeutigkeit umgehen. $\Diamond P(c)$ im Sinne von 1. können wir ausdrücken durch $\Diamond \langle \lambda x. P(x) \rangle (c)$.
Im Sinne von 2. wird es ausgedrückt durch $\langle \lambda x. \Diamond P(x) \rangle (c)$.

Beispiel 4.79 (*Umgangssprachlicher Art*)

1. *Notwendigerweise ist (in unserem Sonnensystem) 9 die Anzahl unserer Planeten.*

   (a) *In der de dicto Lesweise: (c stehe für die Anzahl unserer Planeten)*
   $\Box \langle \lambda x. x \approx 9 \rangle (c)$
   oder mit anderen Worten:
   Notwendigerweise muss folgendes gelten: Die Anzahl unserer Planeten ist 9.
   In diesem de dicto Sinn ist diese Aussage falsch, da eine andere „Welt" denkbar ist, in der die Anzahl der Planeten nicht 9 beträgt.

   (b) *de re Lesweise:*
   $\langle \lambda x. \Box x \approx 9 \rangle (c)$
   oder in Worten:
   Für die Anzahl der Planeten ist der Fall, dass sie notwendigerweise 9 beträgt. In diesem Fall ist die Aussage wahr, da in unserer Welt die Zahl der Planeten 9 ist, und notwendigerweise $9 = 9$ *gilt.*

2. *Notwendigerweise ist der Bundeskanzler zuvor Ministerpräsident von Niedersachsen gewesen.*

   In der de dicto Lesweise ist diese Aussage wiederum falsch, hingegen wahr unter der de re Interpretation.

Zum Abschluß wollen wir noch ein Beispiel aus der Informatik behandeln, für das die Prädikats-Abstraktion recht nützlich ist.

Beispiel 4.80 (Dynamische Logik)
In der dynamischen Logik wird mit jedem „Programm" A einer gegebenen Programmiersprache ein Modaloperator $\Box_A$ *verbunden (im allgemeinen als* $[A]$ *notiert), so dass dann eine Formel* $[A]\,\alpha$ *zu lesen ist als „nach jeder terminierenden Ausführung des Programms A gilt* α*". Für den dualen Operator* $\Diamond_A$ *(notiert als* $\langle A \rangle$*) ist* $\langle A \rangle\, \alpha$ *zu lesen als „es gibt eine Ausführung von A, die terminiert in einem Zustand, in dem* α *gilt".*
Die dynamische Logik ist somit eine multimodale Logik, die wie eine modale Logik aber unter Verwendung dieser Modaloperatoren $\langle A \rangle$ *und* $[A]$ *augebaut ist. In ihrer einfachsten Fassung werden in der dynamischen Logik die Programme aufgrund der folgenden Regeln gebildet.*
Ausgehend von einer Menge Π_0 *von atomaren Programmen definiert man:*

1. *Jedes atomare Programm $a \in \Pi_0$ ist ein Programm.*

2. *Sind A und B Programme und α eine Formel, so sind Programme:*

$A; B$	*Führe erst A und dann B aus (Komposition)*
$A \cup B$	*Führe nichtdeterministisch A oder B aus (Alternation)*
A^*	*Wiederhole A endlich oft (Iteration)*
$\alpha?$	*Test auf α: falls α gilt, fahre fort, ansonsten „Abbruch"*

Hiermit sind dann z.B. auch Programme der folgenden Art definierbar:

if α then A else B	*definiert durch*	$(\alpha?; A) \cup (\neg\alpha?; B)$
while α do A	$''$	$(\alpha?; A)^*; \neg\alpha?$
repeat A until α	$''$	$A; (\neg\alpha?; A)^*$
skip	$''$	$\top?$
abort	$''$	$\bot?$

Wir bezeichnen die Menge aller Programme mit Π.
Eine passende Kripke-Struktur hat dann die Form $M = (W, \{R_A \mid A \in \Pi\}, \mathfrak{I})$.
Für die Erreichbarkeitsrelationen R_A verlangen wir zusätzlich die folgende Eigenschaft:

$$\begin{aligned} R_{A;B} &= R_A \circ R_B \\ R_{A\cup B} &= R_A \cup R_B \\ R_{A^*} &= R_A^* \\ R_{\alpha?} &= \{(p,p) \mid p \in M \text{ und } M, p \Vdash_{ML} \alpha\} \end{aligned}$$

Damit gilt dann z.B.

$$M, p, w \Vdash_{ML} [A]\,\alpha \;\Leftrightarrow\; \text{für alle } q \text{ (wenn } pR_Aq \text{ so } M, q, w \Vdash_{ML} \alpha),$$

also genau das, was wir mit $[A]\,\alpha$ intendiert hatten.
Interpretieren wir die möglichen Belegungen der Speichervariablen des Programms als Welten (Zustände), so kann jede einzelne Speichervariable als eine (nicht rigide) Konstante aufgefasst werden. Betrachten wir z.B. die atomare Anweisung $c := c + 1$. Befinden wir uns in einem Zustand p und wird in diesem Zustand p diese Anweisung $c := c + 1$ ausgeführt, dann soll in allen Zuständen, die wir nach der Ausführung erreichen können, der Wert der Speichervariablen c gegenüber dem Wert in der Welt p um eins erhöht worden sein. Wollen wir dies formal ausdrücken, sind wir vielleicht versucht, dies in der folgenden Weise zu tun:

$$(\forall x)\,((c \approx x) \to [c := c + 1]\,(c \approx x + 1))\,.$$

Mit dem Leibniz-Prinzip (siehe Prädikatenlogik mit Identität) folgt jedoch sofort, dass dann

$$[c := c + 1]\,(c \approx c + 1)$$

gelten müsste, was natürlich widersinnig ist. Mit Hilfe der Prädikats-Abstraktion können wir den gewünschten Sachverhalt korrekt ausdrücken:

$$\langle \lambda x.\ [c := c + 1]\,(c \approx x)\rangle\,(c + 1)\,.$$

Das c in $(c + 1)$ bezieht sich noch auf den alten Zstand vor Ausführung der Anweisung $[c := c + 1]$. Das c in $(c \approx x)$ bezieht sich schon auf den neuen Zustand, so dass damit in korrekter Weise ausgedrückt wird „$c_{neu} = c_{alt} + 1$".

4.4 Temporallogik

Wir hatten bereits in Abschnitt 4.2.4 gesehen, dass den Modaloperatoren $\Box$ und $\Diamond$ in den Logiken ***S*4.3** und ***S*4** in kanonischer Weise eine zeitlogische Interpretation gegeben werden konnte, wobei ***S*4.3** zu einer Linearzeitlogik führt, da die in der Abfolge der Zeit erreichbaren Welten linear geordnet sind (*linearer Zeitfluss*),

während ***S4*** zu einer verzweigenden Zeitlogik führt, da die in der Abfolge der Zeit erreichbaren Welten i.a. nur partiell geordnet sind (*verzweigender Zeitfluss*). Nun ist es zwar möglich, weitergehende Aspekte des Zeitflusses (*diskrete* oder *dichte* „*Zeit*") über Rahmeneigenschaften bzw. Formelschemata zu charakterisieren, für die Anwendungen insbesondere im Bereich der Spezifikation und Verifikation sind die Ausdrucksmittel dieser Logiken dennoch nicht ganz ausreichend. Wir wollen hier im Rahmen eines modallogischen Ansatzes eine Linearzeitlogik – wir sprechen im Folgenden auch von einer linearen Temporallogik – LTL einführen, die für solche Anwendungen geeignet ist.
In der Definition der Semantik der Modallogik waren wir so verfahren, dass wir die Prädikatensymbole in den einzelnen Welten unterschiedlich – also lokal – interpretieren konnten, hingegen erfolgte die Interpretation der Konstantensymbole und die Belegung der Variablen global in allen Welten gleich. Dies war im Grunde genommen bereits eine willkürliche Festlegung, die mehr durch mögliche Anwendungen im Bereich der Künstlichen Intelligenz motiviert war. Prinzipiell steht für die Definition der Semantik der beteiligten Größen (Prädikaten- und Funktionssymbole, Konstantensymbole und Variablen) das ganze Spektrum – lokale Interpretation, globale Interpretation oder eine Kombination von beidem – zur Auswahl. Auch hinsichtlich der Grundbereiche in den einzelnen Welten ist eine Einschränkung nur aus pragmatischen Gründen oder aber bei globaler Interpretation der Prädikaten- und Funktionssymbole erforderlich.
Mit der hier im Folgenden näher zu definierenden Temporallogik LTL wollen wir eine Logik vorstellen, die in erster Linie für die Spezifikation und Verifikation von Programmen gedacht ist. Hierbei haben wir keineswegs nur sequentielle Programme im Auge, sondern auch nichtdeterministische oder parallele Programme. Es liegt somit nahe, für den Aufbau dieser Logik von dem folgenden Modellierungsansatz auszugehen:

1. Wir benutzen eine diskrete lineare Zeitstruktur mit einem nicht endenden Zeitfluss, der nur für die Zukunft betrachtet wird (d.h. wir verwenden implizit einen Zeitrahmen $(\mathbb{N}; \leq)$).

2. Für die Objekt- oder Individuenbereiche wird vorausgesetzt, dass sie zu allen Zeitpunkten dieselben sind (d.h. wir verwenden „constant domains").

3. Prädikaten-, Funktionen- und Konstantensymbole werden global interpretiert.

4. Neben den globalen Variablen haben wir auch lokale Variablen.

5. Darüber hinaus erweitern wir die Ausdruckskraft durch einen „zusätzlichen" Zeitoperator

4.4.1 Formalisierung der Temporallogik LTL

Der formale Aufbau der Temporallogik LTL ist mit kleinen Anpassungen derselbe wie für die Prädikatenlogik (mit Identität).

Wir gehen aus von einer Signatur S. Das temporallogische Alphabet $\Sigma_{LTL}(S)$ zur Signatur S enthält gegenüber dem prädikatenlogischen Alphabet zusätzlich den zweistelligen *Temporaloperator* ***until***, den wir in Infix-Schreibweise verwenden. Ferner wird die Variablenmenge Var aufgeteilt in zwei unendliche, disjunkte Mengen Var_l und Var_g, die Menge der lokalen und der globalen Variablen. Für die globalen Variablen behalten wir die bisherigen Bezeichnungen bei, die lokalen Variablen notieren wir mit $u, u_0, u_1, \ldots$.
Im Termaufbau werden die lokalen Variablen wie die bisher vorhandenen (globalen) Variablen behandelt.
Die Definition der temporallogischen Terme ($Tm_{LTL}(S)$) erfolgt wie im Falle der Prädikatenlogik, wobei im Termaufbau globale und lokale Variablen benutzt werden dürfen.
Für die Definition der temporallogischen Formeln ($Fm_{LTL}(S)$) sind die Bedingungen in der Definition der prädikatenlogischen Formeln (mit Identität) zu übernehmen und durch die folgenden Bedingungen zu ergänzen:

- Falls $\alpha, \beta \in Fm_{LTL}(S)$, auch so $(\alpha\ \textbf{\textit{until}}\ \beta) \in Fm_{LTL}S$.

- Lokale Variablen werden nicht quantifiziert.

Beispiel 4.81

1. *Die folgenden Zeichenreihen sind temporallogische Formeln der Signatur* $(P, Q; c)$ *mit* $ar(P) = 1$ *und* $ar(Q) = 2$:

   (a) $(\forall x)(P(x)$ ***until*** $(\exists y)\, Q(y, c))$.

   (b) $(\forall x)((P(x)$ ***until*** $Q(x, c))$ ***until*** $(\exists y)\, P(y))\,.\land\neg(\exists x)\, P(x)$.

2. *Keine temporallogische Formel ist* $(\forall x)(P(x)$ ***until*** $(\exists u)\, Q(u, c))$.

In Formeln (α ***until*** β) werden wir im Folgenden die äußeren Klammern weglassen.

4.4.2 Semantik

Definition 4.82 *Sei* S *eine Signatur.*
Eine lineare temporallogische S-Struktur *ist ein Tupel* $M = (\mathcal{A}, w, \varrho)$ *mit*

1. $\mathcal{A}$ *ist eine prädikatenlogische* S*-Struktur.*

2. w *ist eine* A*-Belegung der globalen Variablen.*

3. ϱ *ist eine* ω*-Folge* $\varrho = \eta_0\eta_1\eta_2\ldots$ *von* Zuständen η_i, *wobei jeder Zustand* $\eta_i : Var_l \to |\mathcal{A}|$ *eine* $\mathcal{A}$*-Belegung der lokalen Variablen ist.*

   ϱ *wird auch als* (Zustands-) Pfad *bezeichnet.*

Notation 4.83 *Sei* $i \in N$.

1. *Für* $\varrho = \eta_0\eta_1\eta_2\ldots$ *bezeichne* $\varrho^i := \eta_i\eta_{i+1}\eta_{i+2}\ldots$.

2. *Für* $M = (\mathcal{A}, w, \varrho)$ *sei* $M^i := \left(\mathcal{A}, w, \varrho^i\right)$.

Definition 4.84 *Sei* $S = (\mathcal{R}; \mathcal{F}; \mathcal{K})$ *eine Signatur und* $M = (\mathcal{A}, w, \varrho)$ *eine lineare temporallogische* S*-Struktur mit* $\varrho = \eta_0\eta_1\eta_2\ldots$.
Der Wert $t^{(\mathcal{A},w,\varrho)}$ eines Terms $t \in Tm_{LTL}(S)$ *in der Struktur* $\mathcal{A}$, *unter der Belegung* w *im Pfad* ϱ *ist definiert durch:*

1. $c^{(\mathcal{A},w,\varrho)} := c^{\mathcal{A}}$ *für* $c \in \mathcal{K}$.

2. $x^{(\mathcal{A},w,\varrho)} := w(x)$ *für eine globale Variable* x.

3. $u^{(\mathcal{A},w,\varrho)} := \eta_0(u)$ *für eine lokale Variable* u.

4. $[f(t_1, \ldots, t_n)]^{(\mathcal{A},w,\varrho)} := f^{\mathcal{A}}\left(t_1^{(\mathcal{A},w,\varrho)}, \ldots, t_n^{(\mathcal{A},w,\varrho)}\right)$ *für* $f \in \mathcal{F}$ *und Terme* $t_1, \ldots, t_n \in Tm_{LTL}(S)$.

Wir ordnen nun jeder LTL-Formel α in der LTL-Struktur $M = (\mathcal{A}, w, \varrho)$ einen Wahrheitswert α^M zu, wobei wir den Fall $\alpha^M = T$ durch

$$\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha$$

bzw. auch durch

$$M \Vdash_{LTL} \alpha$$

notieren. (Sprechweise: M bzw. A, w, ϱ *erzwingt* α.)
Entsprechend drückt

$$\mathcal{A}, w, \varrho \nVdash_{LTL} \alpha$$

bzw.

$$M \nVdash_{LTL} \alpha$$

aus, dass $\alpha^M = F$.

Definition 4.85 *Sei S eine Signatur und $M = (\mathcal{A}, w, \varrho)$ eine lineare temporallogische S-Struktur mit $\varrho = \eta_0\eta_1\eta_2\dots$.*

1. *Für eine atomare Formel $P(t_1, \dots, t_n)$:*

   $A, w, \varrho \Vdash_{LTL} P(t_1, \dots, t_n)$ *genau dann, wenn* $\left(t_1^{(\mathcal{A},w,\varrho)}, \dots, t_n^{(\mathcal{A},w,\varrho)}\right) \in P^{\mathcal{A}}$.

   $A, w, \varrho \Vdash_{LTL} \top$.

   $A, w, \varrho \nVdash_{LTL} \bot$

2. *Aussagenlogische Verknüpfungen:*

   (a) $A, w, \varrho \Vdash_{LTL} \neg\alpha$ *genau dann, wenn* $A, w, \varrho \nVdash_{LTL} \alpha$.

   (b) $A, w, \varrho \Vdash_{LTL} \alpha \wedge \beta$ *genau dann, wenn* $A, w, \varrho \Vdash_{LTL} \alpha$ *und* $A, w, \varrho \Vdash_{LTL} \beta$.

   (c) $A, w, \varrho \Vdash_{LTL} \alpha \vee \beta$ *genau dann, wenn* $A, w, \varrho \Vdash_{LTL} \alpha$ *oder* $A, w, \varrho \Vdash_{LTL} \beta$.

   (d) $A, w, \varrho \Vdash_{LTL} \alpha \to \beta$ *genau dann, wenn gilt: wenn* $A, w, \varrho \Vdash_{LTL} \alpha$, *so* $A, w, \varrho \Vdash_{LTL} \beta$.

3. *Quantifikationen:*

   (a) $A, w, \varrho \Vdash_{LTL} (\forall x)\,\alpha$ *genau dann, wenn für alle x-Varianten w' von w gilt* $A, w', \varrho \Vdash_{LTL} \alpha$.

   (b) $A, w, \varrho \Vdash_{LTL} (\exists x)\,\alpha$ *genau dann, wenn es eine x-Variante w' von w gibt mit* $A, w', \varrho \Vdash_{LTL} \alpha$.

4. *Temporallogische Bedingung:*

   $A, w, \varrho \Vdash_{LTL} \alpha$ ***until*** β *genau dann, wenn ein $i > 0$ existiert, so dass* $A, w, \varrho^i \Vdash_{LTL} \beta$ *und für alle k mit $0 < k < i$ gilt* $A, w, \varrho^k \Vdash_{LTL} \alpha$.

α *ist* allgemeingültig (*oder* gültig) in der LTL-Struktur M (*in Zeichen:* $\Vdash_{LTL}^{M} \alpha$) *genau dann, wenn für alle $i \in \mathbb{N}$ gilt* $M^i \Vdash_{LTL} \alpha$.
α *ist* allgemeingültig (*oder* gültig) (*in Zeichen:* $\Vdash_{LTL} \alpha$) *genau dann, wenn für jede LTL-Struktur M gilt* $\Vdash_{LTL}^{M} \alpha$.

Definition 4.86 *Wir führen als Abkürzungen ein:*

1. $\Diamond\alpha := \alpha \vee (\top\ \textbf{\textit{until}}\ \alpha)$

   Es gilt:

   $M \Vdash_{LTL} \Diamond\alpha$ *genau dann, wenn ein $i \in \mathbb{N}$ existiert, so dass* $M \Vdash_{LTL} \alpha$.

   Die informelle Bedeutung von $\Diamond\alpha$ ist somit „irgendwann gilt α".

2. $\Box\alpha := \neg\Diamond\neg\alpha$ *oder gleichwertig* $\Box\alpha := \alpha \wedge \neg(\top\ \textbf{\textit{until}}\ \neg\alpha)$.

   Es gilt:

   $M \Vdash_{LTL} \Box\alpha$ *genau dann, wenn für alle $i \in \mathbb{N}$ gilt* $M^i \Vdash_{LTL} \alpha$.

   Die informelle Bedeutung von $\Box\alpha$ ist daher „α gilt immer".

3. $\bigcirc\alpha := (\bot\ \textbf{\textit{until}}\ \alpha)$

   Es gilt:

   $M \Vdash_{LTL} \bigcirc\alpha$ *genau dann, wenn* $M^1 \Vdash_{LTL} \alpha$.

   Informelle Bedeutung von $\bigcirc\alpha$: „im nächsten Zeitpunkt gilt α".

Für den Umgang mit temporallogischen Formen ist die folgende informelle Deutung, die sich au der Definition der Semantik ergibt, recht nützlich:

- $\Box\Diamond\alpha$: „unendlich oft gilt α"

- $\Diamond\Box\alpha$: „fast immer gilt α".

Beispiel 4.87

1. *Sei $S = (P; c)$ eine Signatur, wobei P zweistellig ist.*

   Ferner sei die (prädikatenlogische) S-Struktur $\mathcal{A} = \left(\mathbb{N}; P^{\mathcal{A}}; c^{\mathcal{A}}\right)$ mit $P^{\mathcal{A}} =<$ und $c^{\mathcal{A}} = 4$ gegeben. w sei eine $\mathcal{A}$-Belegung mit $w(x) = 5$. Wir legen außerdem den Zustandspfad $\varrho = \eta_0\eta_1\eta_2\ldots$ zugrunde mit

   $$\eta_i(u) := i, \text{ für alle } i \in \mathbb{N}.$$

   Es gilt dann:

   (a) $\mathcal{A}, w, \varrho \Vdash_{LTL} P(c, x)$, *da*
   $$\begin{aligned} \mathcal{A}, w, \varrho \Vdash_{LTL} P(c, x) &\Leftrightarrow c^{\mathcal{A}} < w(x) \\ &\Leftrightarrow 4 < 5. \end{aligned}$$

   (b) $\mathcal{A}, w, \varrho \Vdash_{LTL} \bigcirc P(u, x)$, *da*
   $$\begin{aligned} \mathcal{A}, w, \varrho \Vdash_{LTL} \bigcirc P(u, x) &\Leftrightarrow \mathcal{A}, w, \varrho^1 \Vdash_{LTL} P(u, x) \\ &\Leftrightarrow \eta_1(u) < w(x) \\ &\Leftrightarrow 1 < 5. \end{aligned}$$

   (c) $\mathcal{A}, w, \varrho \nVdash_{LTL} \Box P(u, x)$, *da*
   $$\begin{aligned} \mathcal{A}, w, \varrho \Vdash_{LTL} \Box P(u, x) &\Leftrightarrow \textit{für alle } i \geq 0 : \mathcal{A}, w, \varrho^i \Vdash_{LTL} P(u, x) \\ &\Leftrightarrow \textit{für alle } i \geq 0 : \eta_i(u) < w(x) \\ &\Leftrightarrow \textit{für alle } i \geq 0 : i < 5. \end{aligned}$$

   (d) $\mathcal{A}, w, \varrho \Vdash_{LTL} \Box (\exists x) P(u, x)$, *da*
   $$\begin{aligned} \mathcal{A}, w, \varrho \Vdash_{LTL} \Box P(u, x) &\Leftrightarrow \textit{für alle } i \geq 0 : \mathcal{A}, w, \varrho^i \Vdash_{LTL} (\exists x) P(u, x) \\ &\Leftrightarrow \textit{für alle } i \geq 0 \textit{ ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho^i \Vdash_{LTL} P(u, x) \\ &\Leftrightarrow \textit{für alle } i \geq 0 \textit{ ex. } a \in |\mathcal{A}| : \eta_i(u) < w\{x/a\}(x) \\ &\Leftrightarrow \textit{für alle } i \geq 0 \textit{ ex. } a \in |\mathcal{A}| : i < a. \end{aligned}$$

   (e) *Wir verwenden hier, dass $\Box\Diamond$ „unendlich oft" bedeutet (siehe 4.4.2).*
   $\mathcal{A}, w, \varrho \nVdash_{LTL} (\exists x) \Box\Diamond P(u, x)$, *da*
   $$\begin{aligned} &\mathcal{A}, w, \varrho \Vdash (\exists x) \Box\Diamond P(u, x) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho \Vdash_{LTL} \Box\Diamond P(u, x) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für unendlich viele } i : \mathcal{A}, w\{x/a\}, \varrho^i \Vdash_{LTL} P(u, x) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für unendlich viele } i : \eta_i(u) < w\{x/a\}(x) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für unendlich viele } i : i < a. \end{aligned}$$

2. *Sei die Signatur $S = (P; f)$, wobei P und f zweistellig sind, gegeben.*

   Wir legen die (prädikatenlogische) S-Struktur $\mathcal{A} = \left(\mathbb{N}; P^{\mathcal{A}}; f^{\mathcal{A}}\right)$ mit $P^{\mathcal{A}} =<$ und $f^{\mathcal{A}} = +$ zugrunde. w sei eine beliebige $\mathcal{A}$-Belegung, $\varrho = \eta_0\eta_1\eta_2\ldots$ sei ein Zustandspfad mit $\eta_0(u) = \eta_1(u) = \eta_2(u) = 0$ und $\eta_3(u) = 3$.

   Es gilt dann

   (a) $\mathcal{A}, w, \varrho \nVdash_{LTL} (\exists x) \bigcirc (\forall y) P(x, f(y, u))$, *da*
   $$\begin{aligned} &\mathcal{A}, w, \varrho \Vdash_{LTL} (\exists x) \bigcirc (\forall y) P(x, f(y, u)) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho \Vdash_{LTL} \bigcirc (\forall y) P(x, f(y, u)) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho^1 \Vdash_{LTL} (\forall y) P(x, f(y, u)) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : \mathcal{A}, w\{x/a, y/b\}, \varrho^1 \Vdash_{LTL} P(x, f(y, u)) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : a < b + \eta_1(u) \\ &\quad\Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : a < b + 0. \end{aligned}$$

(b) $\mathcal{A}, w, \varrho^1 \Vdash_{LTL} (\exists x) \bigcirc (\forall y) \bigcirc P(x, f(y, u))$, *da*

$$\begin{aligned}
&\mathcal{A}, w, \varrho^1 \Vdash_{LTL} (\exists x) \bigcirc (\forall y) \bigcirc P(x, f(y, u)) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho^1 \Vdash_{LTL} \bigcirc (\forall y) \bigcirc P(x, f(y, u)) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}| : \mathcal{A}, w\{x/a\}, \varrho^2 \Vdash_{LTL} (\forall y) \bigcirc P(x, f(y, u)) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : \mathcal{A}, w\{x/a, y/b\}, \varrho^2 \Vdash_{LTL} \bigcirc P(x, f(y, u)) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : \mathcal{A}, w\{x/a, y/b\}, \varrho^3 \Vdash_{LTL} P(x, f(y, u)) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : a < b + \eta_3(u) \\
&\quad \Leftrightarrow \textit{ex. } a \in |\mathcal{A}|, \textit{ so dass für alle } b \in |\mathcal{A}| : a < b + 3.
\end{aligned}$$

Bemerkung 4.88 *Wie im Falle der Modallogik (siehe Bemerkung 4.12) erhalten wir für die Temporallogik einige wichtige Eigenschaften, die sich leicht beweisen lassen:*

1. *Permanenz der Aussagenlogik in der Linearzeitlogik LTL.*

2. *Jede allgemeingültige Aussage α der Prädikatenlogik ist auch temporallogisch allgemeingültig, d.h. es gilt*
$$\models_{PL} \alpha \ \Rightarrow \Vdash_{LTL} \alpha.$$
(Im Falle von LTL fällt gegenüber der Modallogik die Einschränkung hinsichtlich der Konstanten weg, da wir hier „constant domains" zugrunde legen.)

3. *Für die Temporallogik gilt ebenfalls ein Koinzidenztheorem in der folgenden Form:*
Sei $\alpha \in Fm_{ML}(S)$ und $\mathcal{A}$ eine prädikatenlogische S-Struktur. w und w' seien zwei $\mathcal{A}$-Belegungen, die in der Belegung der in α vorkommenden freien globalen Variablen übereinstimmen, und $\varrho = \eta_0\eta_1\eta_2\ldots$ sowie $\varrho' = \eta'_0\eta'_1\eta'_2\ldots$ seien zwei Zustandspfade, für die gilt, dass für alle $i \in \mathbb{N}$ η_i und η'_i für die in α vorkommenden lokalen Variablen übereinstimmen.
Dann gilt:
$$A, w, \varrho \Vdash_{LTL} \alpha \ \Leftrightarrow \ A, w', \varrho' \Vdash_{LTL} \alpha.$$

4. *Analog zur Prädikatenlogik können wir die Semantik für die Formeln $(\forall x)\alpha$ und $(\exists x)\alpha$ auch in der folgenden Weise angeben:*
$\mathcal{A}, w, \varrho \Vdash_{LTL} (\forall x)\alpha$ genau dann, wenn für alle $a \in |\mathcal{A}|$ gilt $\mathcal{A}, w\{x/a\}, \varrho \Vdash_{LTL} \alpha$ bzw.
$\mathcal{A}, w, \varrho \Vdash_{LTL} (\exists x)\alpha$ genau dann, wenn ein $a \in |\mathcal{A}|$ existiert mit $\mathcal{A}, w\{x/a\}, \varrho \Vdash_{LTL} \alpha$.

5. *Wie schon für die Modallogik ergänzen wir die Definition der Substitution für die Temporallogik durch die Festlegung*
$$(\alpha \textbf{ until } \beta)\theta := \alpha\theta \textbf{ until } \beta\theta.$$
Für die Temporallogik steht uns dann das Substitutionslemma in der folgenden Form zur Verfügung:
Sind $t_1, \ldots, t_n$ Terme, die keine lokalen Variablen enthalten, dann gilt:
$$\mathcal{A}, w\left\{x_1/t_1^{(\mathcal{A},w,\varrho)}, \ldots, x_n/t_n^{(\mathcal{A},w,\varrho)}\right\}, \varrho \Vdash_{LTL} \alpha \ \Leftrightarrow \ \mathcal{A}, w, \varrho \Vdash_{LTL} \alpha\{x_1/t_1, \ldots, x_n/t_n\}.$$

Wir geben zunächst einige Eigenschaften an, die den Gebrauch der Quantoren betreffen.

Lemma 4.89 *Sei S eine Signatur, seien ferner $\alpha, \beta \in Fm_{ML}(S)$.*
Es gilt dann

1. *Für Terme t, die keine lokalen Variablen enthalten, gilt:*
   $\Vdash_{LTL} (\forall x)\,\alpha \to \alpha\,\{x/t\}$
   $\Vdash_{LTL} \alpha\,\{x/t\} \to \exists x\alpha$
2. $\Vdash_{LTL} (\forall x)\,(\alpha \to \beta) \to ((\forall x)\,\alpha \to (\forall x)\,\beta)$
3. $\Vdash_{LTL} \Box\,(\forall x)\,\alpha \to (\forall x)\,\Box\alpha$
4. $\Vdash_{LTL} (\forall x)\,\Box\alpha \to \Box\,(\forall x)\,\alpha$ (Barcan'sche Formel)

Bemerkung 4.90 *Die Formel $(\forall x)\,\alpha \to \alpha\,\{x/t\}$ ist i.a. nicht mehr allgemeingültig, wenn t lokale Variablen enthält:*
Sei dazu S eine Signatur, die außer dem Gleichheitssymbol $\approx$ keine weiteren Symbole enthält. Sei weiter $\mathcal{A}$ eine normale S-Struktur und w eine $\mathcal{A}$-Belegung.
$\varrho = \eta_0\eta_1\eta_2\ldots$ *sei ein Zustandspfad mit $\eta_0\,(u) = \eta_0\,(v) = \eta_1\,(u) \neq \eta_1\,(v)$ für die lokalen Variablen u und v.*
Betrachte die Formel $\alpha := x \approx u \to \bigcirc\,(x \approx u)$.
Es gilt:
$\mathcal{A}, w, \varrho \Vdash_{LTL} (\forall x)\,\alpha$, *da*
$\mathcal{A}, w, \varrho \Vdash_{LTL} (\forall x)\,(x \approx u \to \bigcirc\,(x \approx u))$
$\Leftrightarrow$ *für alle* $a \in |\mathcal{A}| : \mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} x \approx u \to \bigcirc\,(x \approx u)$
$\Leftrightarrow$ *für alle* $a \in |\mathcal{A}|$: *wenn* $\mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} x \approx u$, *so* $\mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} \bigcirc\,(x \approx u)$
$\Leftrightarrow$ *für alle* $a \in |\mathcal{A}|$: *wenn* $\mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} x \approx u$, *so* $\mathcal{A}, w\,\{x/a\}\,, \varrho^1 \Vdash_{LTL} x \approx u$
$\Leftrightarrow$ *für alle* $a \in |\mathcal{A}| : a = \eta_0\,(u) \;\Rightarrow\; a = \eta_1\,(u)\,.$
Hingegen erhalten wir für die Formel $\alpha\,\{x/v\}$:
$\mathcal{A}, w, \varrho \nVdash_{LTL} v \approx u \to \bigcirc\,(v \approx u)$, *da*
$\mathcal{A}, w, \varrho \Vdash_{LTL} v \approx u \to \bigcirc\,(v \approx u)$
$\Leftrightarrow$ *wenn* $\mathcal{A}, w, \varrho \Vdash_{LTL} v \approx u$, *so* $\mathcal{A}, w, \varrho^1 \Vdash_{LTL} v \approx u$
$\Leftrightarrow$ *wenn* $\eta_0\,(v) = \eta_0\,(u)$, *so* $\eta_1\,(v) = \eta_1\,(u)\,.$
Wir erhalten somit:
$\mathcal{A}, w, \varrho \nVdash_{LTL} (\forall x)\,(x \approx u \to \bigcirc\,(x \approx u)) \to (v \approx u \to \bigcirc\,(v \approx u))\,.$

Im folgenden wollen wir einige einfache, vielfach jedoch grundlegende Eigenschaften für den ***until***-Operator und die definierten Operatoren $\bigcirc$, $\Box$ und $\Diamond$ angeben.
Aufgrund der Tatsache, dass der der Linearzeitlogik zugrundeliegende Rahmen $(\mathbb{N}; \leq)$ reflexiv, transitiv und linear ist, der zusätzlich die Diskretheitsbedingung erfüllt, übertragen sich alle Eigenschaften der Logik ***S*4.3**, die sich aufgrund der Rahmeneigenschaften ergeben, auf die Linearzeitlogik LTL.

Lemma 4.91 *Für $\alpha, \beta \in Fm_{LTL}\,(S)$ gilt:*

1. $\Vdash_{LTL} \Box\,(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$
2. $\Vdash_{LTL} \Box\alpha \leftrightarrow \Box\Box\alpha$
3. $\Vdash_{LTL} \Diamond\alpha \leftrightarrow \Diamond\Diamond\alpha$
4. $\Vdash_{LTL} \Diamond\Box\Diamond\alpha \leftrightarrow \Box\Diamond\alpha$
5. $\Vdash_{LTL} \Box\Diamond\Box\alpha \leftrightarrow \Diamond\Box\alpha$
6. $\Vdash_{LTL} \Box\,(\alpha \wedge \beta) \leftrightarrow \Box\alpha \wedge \Box\beta$
7. $\Vdash_{LTL} \Diamond\,(\alpha \vee \beta) \leftrightarrow \Diamond\alpha \vee \Diamond\beta$

Lemma 4.92 (Kommutativgesetze)
Für $\alpha \in Fm_{LTL}(S)$ *gilt:*

1. $\Vdash_{LTL} \neg \bigcirc \alpha \leftrightarrow \bigcirc \neg \alpha$
2. $\Vdash_{LTL} \Box \bigcirc \alpha \leftrightarrow \bigcirc \Box \alpha$
3. $\Vdash_{LTL} \Diamond \bigcirc \alpha \leftrightarrow \bigcirc \Diamond \alpha$

Beweis. Exemplarisch beweisen wir 1.:
Sei $\mathcal{A}$ eine S-Struktur, w eine $\mathcal{A}$-Belegung und ϱ ein Zustandspfad.
Es gilt dann

$$\begin{array}{rcl} \mathcal{A}, w, \varrho \Vdash_{LTL} \neg \bigcirc \alpha & \Leftrightarrow & \mathcal{A}, w, \varrho \nVdash_{LTL} \bigcirc \alpha \\ & \Leftrightarrow & \mathcal{A}, w, \varrho^1 \nVdash_{LTL} \alpha \\ & \Leftrightarrow & \mathcal{A}, w, \varrho^1 \Vdash_{LTL} \neg \alpha \\ & \Leftrightarrow & \mathcal{A}, w, \varrho \Vdash_{LTL} \bigcirc \neg \alpha\,. \end{array}$$

■

Lemma 4.93 (Distributivgesetze für $\bigcirc$)
Für $\alpha, \beta \in Fm_{LTL}(S)$ *und* $\circ \in \{\wedge, \vee, \rightarrow, \boldsymbol{until}\}$ *gilt:*
$\Vdash_{LTL} \bigcirc (\alpha \circ \beta) \leftrightarrow \bigcirc \alpha \circ \bigcirc \beta$.

Beweis. Exemplarisch beweisen wir $\Vdash_{LTL} \bigcirc (\alpha\ \boldsymbol{until}\ \beta) \leftrightarrow (\bigcirc \alpha\ \boldsymbol{until}\ \bigcirc \beta)$:
Sei $\mathcal{A}$ eine S-Struktur, w eine $\mathcal{A}$-Belegung und ϱ ein Zustandspfad.
Es gilt dann

$$\begin{array}{rcl} \mathcal{A}, w, \varrho \Vdash_{LTL} \bigcirc (\alpha\ \boldsymbol{until}\ \beta) & \Leftrightarrow & \mathcal{A}, w, \varrho^1 \Vdash_{LTL} (\alpha\ \boldsymbol{until}\ \beta) \\ & \Leftrightarrow & \text{ex. } i > 0 \text{ mit} \mathcal{A}, w, \varrho^{i+1} \Vdash_{LTL} \beta \\ & & \text{und für alle } k \text{ mit } 0 < k < i \text{ gilt } \mathcal{A}, w, \varrho^{k+1} \Vdash_{LTL} \alpha \\ & \Leftrightarrow & \text{ex. } i > 0 \text{ mit} \mathcal{A}, w, \varrho^{i} \Vdash_{LTL} \bigcirc \beta \\ & & \text{und für alle } k \text{ mit } 0 < k < i \text{ gilt } \mathcal{A}, w, \varrho^{k} \Vdash_{LTL} \bigcirc \alpha \\ & \Leftrightarrow & \mathcal{A}, w, \varrho \Vdash_{LTL} \bigcirc \alpha\ \boldsymbol{until}\ \bigcirc \beta. \end{array}$$

■

Lemma 4.94 (Rekursionsgleichungen) *Für* $\alpha, \beta \in Fm_{LTL}(S)$ *gilt:*

1. $\Vdash_{LTL} (\alpha\ \boldsymbol{until}\ \beta) \leftrightarrow \bigcirc \beta \vee \bigcirc (\alpha \wedge (\alpha\ \boldsymbol{until}\ \beta))$
2. $\Vdash_{LTL} \Box \alpha \leftrightarrow \alpha \wedge \bigcirc \Box \alpha$
3. $\Vdash_{LTL} \Diamond \alpha \leftrightarrow \alpha \vee \bigcirc \Diamond \alpha$

Beweis. (Übungsaufgabe). ■

Lemma 4.95 *Für* $\alpha, \beta \in Fm_{LTL}(S)$ *gilt:*

1. $\Vdash_{LTL} \alpha \ \Rightarrow \Vdash_{LTL} \Box \alpha$
2. $\Vdash_{LTL} \alpha \ \Rightarrow \Vdash_{LTL} \bigcirc \alpha$
3. $\Vdash_{LTL} \alpha \rightarrow \beta$ *und* $\Vdash_{LTL} \alpha \rightarrow \bigcirc \alpha \ \Rightarrow \Vdash_{LTL} \alpha \rightarrow \Box \beta$

Beweis. (Übungsaufgabe). ■

4.4.3 Zur Anwendung von LTL

Wir wollen uns als nächstes ein Beispiel ansehen, das zeigt, in welcher Weise die Temporallogik LTL zur Spezifikation und Verifikation herangezogen werden kann. Um den Aufwand gering zu halten, verzichten wir auf die Angabe der Signatur und auch auf die Angabe der semantischen Struktur, die hierbei zugrundegelegt ist.

Beispiel 4.96 *Betrachten wir das folgende nichtterminierende Programm, das sämtliche Primzahlen ausgibt:*

$$\begin{array}{ll}
\lambda_0: & u := 2 \\
\lambda_1: & \textbf{\textit{print}}\ (u) \\
\lambda_2: & u := u+1 \\
\lambda_3: & v := 2 \\
\lambda_4: & \textbf{\textit{if}}\ (v)^2 > u\ \textbf{\textit{then goto}}\ \lambda_1 \\
\lambda_5: & \textbf{\textit{if}}\ (u \bmod v) = 0\ \textbf{\textit{then goto}}\ \lambda_2 \\
\lambda_6: & v := v+1 \\
\lambda_7: & \textbf{\textit{goto}}\ \lambda_4
\end{array}$$

Ein Programmzustand (λ_i, r_u, r_v) ist dann gegeben durch

1. *die Marke λ_i der Anweisung, die als nächste auszuführen ist, und*
2. *durch die aktuellen Werte r_u und r_v von u und v.*

Da mit der Ausführung des Programms sich die Programmzustände permanent ändern, ist es naheliegend, die aktuelle Marke λ und die aktuellen Werte von u und v über lokale Variablen zu erfassen. Für die temporallogische Beschreibung des Programms benutzen wir diese Symbole als lokale Variablen.
Eine passende Signatur muss neben dem Gleichheitssymbol $\approx$ u.a. das einstellige Prädikatensymbol prime (für „ist Primzahl") enthalten. Weiter verwenden wir die Konstantensymbole $\lambda_0, \ldots, \lambda_7$.
Wir können damit unser Programm in der folgenden Weise durch eine Formel α beschreiben:

$$\begin{array}{l}
\alpha : (\forall x)\,(\forall y) \\
(\lambda \approx \lambda_0 \wedge \bigcirc\lambda \approx \lambda_1 \wedge \bigcirc u = 2 \qquad\qquad (\textit{Startsituation}) \\
\wedge \\
\Box(\\
\quad (\lambda \approx \lambda_1 \to \bigcirc\lambda \approx \lambda_2) \\
\quad \wedge (\lambda \approx \lambda_2 \wedge x \approx u \wedge y \approx v \to \bigcirc\,(u \approx x+1 \wedge v \approx y \wedge \lambda \approx \lambda_3)) \\
\quad \wedge (\lambda \approx \lambda_3 \wedge x \approx u \to \bigcirc\,(u \approx x \wedge v \approx 2 \wedge \lambda \approx \lambda_4)) \\
\quad \wedge (\lambda \approx \lambda_4 \wedge (v * v > u) \wedge x \approx u \to \bigcirc\,(u \approx x \wedge \lambda \approx \lambda_1)) \\
\quad \wedge (\lambda \approx \lambda_5 \wedge (u \bmod v \approx 0) \wedge x \approx u \to \bigcirc\,(u \approx x \wedge \lambda \approx \lambda_2)) \\
\quad \wedge (\lambda \approx \lambda_6 \wedge x \approx u \wedge y \approx v \to \bigcirc\,(u \approx x \wedge v \approx y+1 \wedge \lambda \approx \lambda_7)) \\
\quad \wedge (\lambda \approx \lambda_7 \wedge x \approx u \wedge y \approx v \to \bigcirc\,(u \approx x \wedge v \approx y \wedge \lambda \approx \lambda_4))\ )\ ).
\end{array}$$

Mittels der globalen Variablen können wir die lokalen Werte von u und v von einem Zustand in den Folgezustand übertragen. Dort, wo für u bzw. v der alte Wert nicht benötigt wird, haben wir auf die Übertragung dieser Werte verzichtet.
Damit können wir dann z.B. die folgenden „Korrektheits"-Aussagen formulieren:

- *„Es werden nur Primzahlen ausgegeben"*

$$\Box\,(\lambda \approx \lambda_1 \to \textit{prime}\,(u))$$

- *„Jede Primzahl wird ausgegeben"*

$$(\forall z)\,(\textit{prime}\,(z) \to \Diamond\,(\lambda \approx \lambda_1 \wedge u \approx z))$$

- *„Die Primzahlen werden in der korrekten Reihenfolge ausgegeben"*

$$(\forall z)\,(\lambda \approx \lambda_1 \wedge u \approx z \rightarrow \bigcirc\Box\,(\lambda \approx \lambda_1 \rightarrow u > z))\,.$$

Für einen formalen Beweis, dass das Programm eine solche Eigenschaft β besitzt, wäre also in einem geeigneten Beweiskalkül die Formel $\alpha \wedge \alpha' \rightarrow \beta$ zu herzuleiten, wobei α' eine Formel ist, die eine Axiomatisierung der zugrundeliegenden Struktur der natürlichen Zahlen darstellt.

4.4.4 Die Unvollständigkeit der linearen Temporallogik LTL

Im Falle der Aussagenlogik, der Prädikatenlogik und auch der verschiedenen Modallogiken (ausgenommen $\boldsymbol{G}$) hatten die angegeben Beweiskalküle die Eigenschaft, vollständige Beweiskalküle zu ein, d.h. jede allgemeingültige Aussage war mit diesen Beweiskalkülen auch beweisbar. Im Falle der hier betrachteten linearen Temporallogik werden wir sehen, dass es prinzipiell kein algorithmisches Beweisverfahren geben kann, mit dem alle in LTL allgemeingültigen Aussagen beweisbar sind. Unser Beweis wird sich dabei auf das folgende Theorem von Trakhtenbrot stützen.

Theorem 4.97 (Trakhtenbrot 1953) *Sei S eine Signatur, die ein Prädikatensymbol P mit $ar\,(P) \geq 2$ enthält.*
Es gilt dann:
Die Menge der prädikatenlogischen Aussagen in $Fm_{PL}\,(S)$, die genau in den prädikatenlogischen S-Strukturen mit endlichem Grundbereich allgemeingültig sind, ist nicht rekursiv aufzählbar.

(Für einen Beweis des Theorems von Trakhtenbrot siehe z.B. C. Smoryński: Logical Number Theory I, Springer Verlag 1991.)

Bezeichne FIN_S die Menge der Aussagen in $Fm_{PL}\,(S)$, die genau in den prädikatenlogischen S-Strukturen mit endlichem Grundbereich allgemeingültig sind. Es gibt also keinen Algorithmus, der die Menge der Aussagen in FIN_S aufzählen kann.

Im folgenden sei S eine Signatur, wie sie im Theorem von Trakhtenbrot vorausgesetzt wird.

Lemma 4.98 *Es gibt eine Formel $\alpha_{fin} \in Fm_{LTL}\,(S)$, so dass für jede prädikatenlogische S-Struktur $\mathcal{A}$, jede $\mathcal{A}$-Belegung w und jeden Zustandspfad ϱ gilt:*

$$\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin} \;\Rightarrow\; |\mathcal{A}| \text{ ist endlich.}$$

Beweis. Wir definieren zwei Formeln $\alpha\,(x)$ und $\beta\,(x)$, in denen die Variable x frei vorkommt:
$\alpha\,(x) := \Box\Diamond u \approx x$ („in unendlich vielen Zuständen nimmt u den Wert von x an")
$\beta\,(x) := \Box\,(u \approx x \rightarrow (\forall y)\,(\neg u \approx x\ \textbf{\textit{until}}\ u \approx y))$
(„Wann immer in einem Zustand die lokale Variable u den Wert von x erhält, nimmt u jeden anderen möglichen Wert in einem späteren Zustand an, ohne zwischenzeitlich den Wert von x wieder angenommen zu haben")
Es sei $\alpha_{fin} := (\exists x)\,(\alpha\,(x) \wedge \beta\,(x))$. In α_{fin} kommt also an lokalen Variablen nur u vor.
Seien nun $\mathcal{A}, w$ und $\varrho = \eta_0\eta_1\ldots$ beliebig gegeben und es gelte $\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin}$.
Dann gibt es ein $a \in |\mathcal{A}|$, so dass

$$\mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} \alpha\,(x) \tag{4.5}$$

und

$$\mathcal{A}, w\,\{x/a\}\,, \varrho \Vdash_{LTL} \beta\,(x)\,. \tag{4.6}$$

Wegen 4.5 folgt, dass ein $i \in \mathbb{N}$ existiert mit

$$\mathcal{A}, w\,\{x/a\}\,, \varrho^i \Vdash_{LTL} u \approx x. \tag{4.7}$$

Wegen 4.6 gilt für dieses i :

$$\mathcal{A}, w\,\{x/a\}\,, \varrho^i \Vdash_{LTL} u \approx x \rightarrow (\forall y)\,(\neg u \approx x\ \textbf{\textit{until}}\ u \approx y)$$

und damit wegen 4.7

$$\mathcal{A}, w\{x/a\}, \varrho^i \Vdash_{LTL} (\forall y)(\neg u \approx x \;\textbf{\textit{until}}\; u \approx y).$$

Damit gilt dann für alle $b \in |\mathcal{A}|$:

$$\mathcal{A}, w\{x/a, y/b\}, \varrho^i \Vdash_{LTL} \neg u \approx x \;\textbf{\textit{until}}\; u \approx y. \tag{4.8}$$

Aufgrund von 4.8 gilt dann für alle $b \in |\mathcal{A}|$:

$$\text{es existiert ein } k > i \text{ mit } \mathcal{A}, w\{x/a, y/b\}, \varrho^k \Vdash_{LTL} u \approx y$$

und

$$\text{für alle } j \text{ mit } i < j < k \text{ gilt } \mathcal{A}, w\{x/a, y/b\}, \varrho^j \Vdash_{LTL} \neg u \approx x.$$

Mit anderen Worten:
für das angegebene $a \in |\mathcal{A}|$ und für alle $b \in |\mathcal{A}|$ gilt:

$$\text{es existiert ein } k > i \text{ mit } \eta_k(u) = b \text{ und für alle } j \text{ mit } i < j < k \text{ ist } \eta_j(u) \neq a \tag{4.9}$$

Wegen 4.5 gilt aber weiter, dass ein $m > i$ existiert mit

$$\mathcal{A}, w\{x/a\}, \varrho^m \Vdash_{LTL} u \approx x,$$

also mit

$$\eta_m(u) = a. \tag{4.10}$$

Aus 4.9 und 4.10 folgt nun, dass $|\mathcal{A}|$ endlich sein muss. Denn wäre $|\mathcal{A}|$ nicht endlich, so müsste zu jedem $b \in |\mathcal{A}|$ (und damit für unendlich viele Elemente b) ein Zeitpunkt $k_b > i$ existieren mit $\eta_{k_b}(u) = b$ und zwischen dem Zeitpunkt i und k_b dürfte die lokale Variable u nie den Wert a annehmen. Damit kann es keinen auf i folgenden endlichen Zeitabschnitt geben, in dem u den Wert a erhält. ■

Wir zeigen als nächstes:

Lemma 4.99 *Zu jeder prädikatenlogischen S-Struktur $\mathcal{A}$ mit endlichem Grundbereich gibt es einen Zustandspfad ϱ, so dass für alle $\mathcal{A}$-Belegungen w gilt:*

$$\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin}.$$

Beweis. Sei $|\mathcal{A}| = \{a_0, a_1, \ldots, a_{n-1}\}$. Wir betrachten einen Zustandspfad $\varrho = \eta_0\eta_1\eta_2\ldots$ mit der folgenden Eigenschaft:

$$\eta_{k\cdot n+i}(u) = a_i$$

für $k \in \mathbb{N}$ und $i \in \{0, \ldots, n-1\}$.
Eine einfache Auswertung der Semantikdefinition zeigt dann, dass für eine beliebige $\mathcal{A}$-Belegung w gilt

$$\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin}.$$

■

Lemma 4.100 *Für jede prädikatenlogische (und damit insbesondere temporallogische) Aussage $\alpha \in Fm_{LTL}(S)$ gilt:*
$\Vdash_{LTL} \alpha_{fin} \to \alpha$ genau dann, wenn für jede S-Struktur $\mathcal{A}$ mit endlichem Grundbereich $|\mathcal{A}|$ gilt $\mathcal{A} \models_{PL} \alpha$.

Beweis.

„⇒": Wir zeigen die Kontraposition:

$$\text{Gibt es eine } S\text{-Struktur } \mathcal{A} \text{ mit } \mathcal{A} \not\models_{PL} \alpha \text{, so } \not\Vdash_{LTL} \alpha_{fin} \to \alpha.$$

Sei $\mathcal{A}$ eine S-Struktur mit endlichem Grundbereich, in der α nicht allgemeingültig ist. Es gibt dann eine $\mathcal{A}$-Belegung w, so dass $\mathcal{A}, w \not\models_{PL} \alpha$. Nach Lemma 4.99 gibt es einen Zustandspfad ϱ mit $\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin}$. Nach Bemerkung 4.88 2. gilt wegen $\mathcal{A}, w \not\models_{PL} \alpha$ auch $\mathcal{A}, w, \varrho \not\Vdash_{LTL} \alpha$. Damit folgt $\mathcal{A}, w, \varrho \not\Vdash_{LTL} \alpha_{fin} \to \alpha$.

„⇐": Sei $\mathcal{A}, w, \varrho$ beliebig gewählt mit $\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin}$. Nach Lemma 4.98 ist $|\mathcal{A}|$ dann endlich. Nach Voraussetzung gilt aber $\mathcal{A} \models_{PL} \alpha$ und damit auch $\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha$. Damit gilt dann $\mathcal{A}, w, \varrho \Vdash_{LTL} \alpha_{fin} \rightarrow \alpha$, und da $\mathcal{A}, w, \varrho$ beliebig gewählt waren, schließlich $\Vdash_{LTL} \alpha_{fin} \rightarrow \alpha$.

■

Theorem 4.101 *Die lineare Temporallogik LTL ist unvollständig.*
Genauer:
$TAUT_{LTL} := \{\alpha \in Fm_{LTL}(S) \mid \alpha \text{ Aussage}, \ \Vdash_{LTL} \alpha\}$ *ist nicht rekursiv aufzählbar.*

Beweis. Für die Abbildung F, die jeder Formel $\alpha \in Fm_{PL}(S)$ die Formel $F(\alpha) := \alpha_{fin} \rightarrow \alpha \in Fm_{LTL}(S)$ zuordnet, gilt, dass sie berechenbar ist und dass

$$\alpha \in FIN_S \ \Leftrightarrow \ F(\alpha) \in TAUT_{LTL}.$$

Es gilt also $FIN_S \leq TAUT_{LTL}$. Da FIN_S nicht rekursiv aufzählbar ist, ist somit auch die Menge der allgemeingültigen temporallogischen Aussagen zur Signatur S nicht rekursiv aufzählbar. ■