Einführung in die Logik

Jürgen Koslowski
unter Verwendung von Materialien von
Roland Meyer und
Thorsten Palm

Theoretical Computer Science
Technische Universität Braunschweig

SS 2022

https://www.tcs.cs.tu-bs.de/teaching/Logik_SS_2022.html

Teil 0
Ziele und Voraussetzungen

Inhaltsverzeichnis Teil 0

Kapitel 0
Überblick

Teil 1: Aussagenlogik

Ziel ist die Formalisierung des korrekten Schließens. Um festzustellen, ob eine Aussage A eine korrekte Schlussfolgerung aus einer gegebenen, potentiell unendlichen Menge Γ von Aussagen, den sog. Prämissen, ist,

- kann man prüfen, ob A logisch aus Γ folgt, Schreibweise $\Gamma \models A$;
- oder man kann versuchen, in einem geeigneten Kalkül A aus Γ herzuleiten (zu „beweisen“), Schreibweise $\Gamma \vdash A$;
- oder man kann versuchen mit geeigneten Algorithmen zu zeigen, dass $\Gamma \cup \{\neg A\}$ inkonsistent oder widersprüchlich ist.

Vorher ist die Syntax klären, d.h., der korrekte Aufbau von Formeln, die Aussagen repräsentieren sollen. Diese bilden eine sog. Termalgebra über gewissen logischen Junktoren und unspezifischen atomaren Formeln, denen in der Semantik Wahrheitswerte 0 (falsch) bzw. 1 (wahr) zugeordnet werden können. Da Termalgebren auch im 2. Teil eine wichtige Rolle spielen werden, stellen wir die nötigen mathematischen Grundlagen bereit, speziell den Rekursionssatz, den wir nachfolgend häufig anwenden werden.

Die Begriffsbildungen der Semantik werden erheblich nachvollziehbarer, wenn man über gewisse Grundkenntnisse der Ordnungstheorie verfügt.

Speziell liefert die Erfüllungsrelation (ob eine Formel A von einer Belegung φ der Atome erfüllt wird) einen Hüllenoperator auf der Menge aller Formeln, der der Folge-Relation $\models$ zugrunde liegt.

Der wichtige Kompaktheitssatz ist das erste beweistechnisch nicht-triviale Ergebnis. Er ermöglicht, sich auf endliche Formelmengen zu beschränken.

Einer Einführung zum Thema Deduktion folgt dann eine aufgrund des Deduktionstheorems praktikable Version des Hilbert-Kalküls. Ein (optionales?) Kapitel behandelt die (weiter verbreitete) natürliche Deduktion sowie Genzens Sequenzen-Kalkül.

Teil 1 schließt mit einem ausführlichen Kapitel zu drei verbreiteten Algorithmen zum häufig schnellen Nachweis der Nichterfüllbarkeit von Formel(menge)n: semantische Tableaus, dem Davis-Putnam Algorithmus und der Resolutionsmethode.

Teil 2: Prädikatenlogik

Genauere Analyse spezifischer Datenbereiche, die mittels einer Signatur $\boldsymbol{Fun} + \boldsymbol{Rel} \xrightarrow{\mathcal{S}} \mathbb{N}$ aus formalen Funktions- und Relationssymbolen (Operatoren und Prädikaten) beschrieben werden können.

- Über einer Variablenmenge $\mathcal{V}$ liefert $\boldsymbol{Fun}$ wieder eine Termalgebren.
- Die neuen atomaren Formeln sind nun Aussagen darüber, ob passende Tupel solcher Terme formal in eines der Prädikate erfüllen.
- Im Fall $\boldsymbol{Rel} = \emptyset$ genügt das nicht. Daher betrachtet man immer ein spezielles binären Prädikat $\doteq$ für formale Gleichungen. Dieses soll auch immer als Gleichheit interpretiert werden.
- Neben den aussagenlogischen Junktoren kommen nun weitere einstellige Junktoren aus mit sog. Quantoren $\forall$ (für alle) und $\exists$ (es gibt) versehenen Variablen zum Einsatz, die ihre jeweiligen Variablen in ihrem Geltungsbereich binden.

Die Semantik spielt sich dann in $\mathcal{S}$-Strukturen ab, d.h., Mengen mit einer Interpretation der Funktions- und Relationssymbole aus $\mathcal{S}$.

In dieser VL wird versucht [Kap. 10], wesentliche semantische Begriffsbildungen der PL so eng wie möglich an diejenigen der AL anzulehnen. In der Tat bestehen hier überraschende Ähnlichkeiten.

In [Kap. 11] dient die Skolemisierung zur Vorbereitung der Sätze von Herbrand und von Löwenheim-Skolem zur Charakterisierung erfüllbarer Formelmengen. Weiter impliziert er die Existenz von Nichtstandardmodellen z.B. der natürlichen Zahlen. Der Satz von Gödel-Herbrand-Skolem schlägt die Brücke zur AL und eröffnet mit dem Algorithmus von Gilbert die Möglichkeit, semantische Tableaus und die Resolutionsmethode an die Gegebenheiten der PL_1 anzupassen [Kap. 12], und andererseits, einen Kompaktheitssatz für die PL_1 mit Hilfe der Version für die AL zu beweisen.

Kapitel 1
Mathematische Voraussetzungen

Mengen naiv betrachtet

Wir betrachten Mengen zunächst naiv als Zusammenfassungen konkreter oder abstrakter Objekte; die grundlegende Relation $\in$ sagt aus, ob ein Objekt x zur Menge A gehört oder nicht:

$$x \in A \quad \text{bzw.} \quad x \notin A$$

Solange die Beschreibung der zusammengefassten Objekte nicht selbstreferentiell ist, treten keine Problem auf:

Beispiel

Die Mengen A, die sich nicht selber als Element enthalten, also $A \notin A$ erfüllen, kann man nicht zu einer Menge zusammenfassen.

Wir unterscheiden endliche Mengen, wie die leere Menge $\emptyset$ oder die drei-elementige Menge $\{a, b, c\}$, von unendlichen Mengen, wie den natürlichen Zahlen $\mathbb{N} := \{0, 1, 2, 3, \dots\}$ oder den reellen Zahlen $\mathbb{R}$.

Operationen auf Mengen

Definition

Eine Menge A heißt Teilmenge einer Menge B (und entsprechend B Obermenge von A), wenn aus $x \in A$ folgt $x \in B$; Schreibweise $A \subseteq B$.

Definition

Für je zwei Objekte a und b bezeichnet $\langle a, b \rangle$ deren geordnetes Paar; im Gegensatz zur Menge {a,b} genügt es der Spezifikation

$$\langle a, b \rangle = \langle c, d \rangle \qquad \text{gdw.} \qquad a = c \quad \text{und} \quad b = d$$

Definition

Für zwei Mengen A und B definiert man

- den Durchschnitt $A \cap B := \{ x : x \in A \quad \text{und} \quad x \in B \}$;
- die Vereinigung $A \cup B := \{ x : x \in A \quad \text{oder} \quad x \in B \}$;
- das cartesisches Produkt $A \times B := \{ \langle a, b \rangle : a \in A \quad \text{und} \quad b \in B \}$.

B^n bezeichnet das n-fache cartesische Produkt einer Menge B mit sich selbst, die Elemente heißen n-Tupel oder Wörter der Länge n.
Beachte: es gibt jeweils genau ein 0-Tupel $\varepsilon_B \in B^0$, das sog. leere Wort.

Definition

Für eine Menge B definiert man

- die Potenzmengen $\boldsymbol{P}(B) := \{ A : A \subseteq B \}$ und $\boldsymbol{P}_\omega(B) := \{ A : A \subseteq B \text{ endlich} \}$;
- die Menge aller Wörter $B^* := \bigcup_{n \in \mathbb{N}} B^n$ (automatisch disjunkt!).

Definition

Sind A und B Mengen, so heißen Teilmengen $R \subseteq A \times B$ auch (binäre) Relation von A nach B eine; Schreibweise in dieser VL: $A \xrightarrow{R} B$.
Die duale Relation $B \xrightarrow{R^{\mathrm{op}}} A$ besteht genau aus den geordneten Paaren $\langle b, a \rangle \in B \times A$, die $\langle a, b \rangle \in R$ erfüllen.

Statt $\langle a, b \rangle \in R$ schreibt man auch aRb (Infix-Notation), z.B. für $\leq$.

Wichtige binäre Relationen auf einer Menge B

Definition

Eine Relation $B \xrightarrow{R} B$ heißt

- reflexiv, falls xRx für alle $x \in B$;
- transitiv, falls aus $xRyRz$ folgt xRz für alle $x, y, z \in B$;
- symmetrisch, falls aus xRy folgt yRx für alle $x, y \in B$;
- anti-symmetrisch, falls aus $xRyRx$ folgt $x = y$ für alle $x, y \in B$;
- linear, falls xRy oder yRx gilt für alle $x, y \in B$;
- Äquivalenzrelation, falls R reflexiv, transitiv und symmetrisch ist;
- Quasiordnung, falls R reflexiv und transitiv ist;
- Halbordnung, falls R reflexiv, transitiv und anti-symmetrisch ist.

HA.

Welche dieser Eigenschaften bleiben unter Durchschnitten erhalten?

Funktionen als spezielle Relationen und Komposition

Definition

Eine Relation $A \xrightarrow{f} B$ heißt

(a) total, wenn zu jedem $a \in A$ ein $b \in B$ existiert mit afb;

(b) partielle Funktion oder einwertig, wenn aus afb und afc folgt $b = c$;

(c) Funktion, wenn sie total und einwertig ist.

Schreibweise: $A \xrightarrow{f} B$ und $f(a)$ für das eindeutige $b \in B$ mit afb. Die Menge aller Funktionen von A nach B wird mit B^A bezeichnet.

Identifiziert man $n \in \mathbb{N}$ mit $\{0, 1, \ldots, n-1\}$, so sind die obigen n-Tupel in B nichts weiter als Funktionen von n nach B.

Definition

Die Komposition $A \xrightarrow{R;S} C$ (alternativ: $S \circ R$) für $A \xrightarrow{R} B \xrightarrow{S} C$ ist gegeben durch: $a(R;S)c$ gdw. ein $b \in B$ existiert mit $aRbSc$. Hier spielt die Funktion $\mathbf{id}_B := \{\langle b, b\rangle : b \in B\}$ eine besondere Rolle.

Elementfreie Charakterisierung relationaler Eigenschaften

Lemma

Eine Relation $B \xrightarrow{R} B$ *ist genau dann*

- *reflexiv, wenn* $\mathbf{id}_B \subseteq R$;
- *transitiv, wenn* $R;R \subseteq R$;
- *symmetrisch, wenn* $R \subseteq R^{\mathrm{op}}$;
- *linear, wenn* $R \cup R^{\mathrm{op}} = B \times B$.

Wir schreiben oft R^n für die n-fache Komposition einer Relation $B \xrightarrow{R} B$.

Satz

- $R^+ = \bigcup\{ R^n : n > 0 \}$ *die kleinste transitive Relation, die* R *enthält, die sog. transitive Hülle;*
- $R^* = \bigcup\{ R^n : n \in \mathbb{N} \}$ *die kleinste reflexive und transitive Relation, die* R *enthält, die sog. reflexive transitive Hülle;*
- $R^{\ddot{a}} = \bigcup\{ (R \cup R^{\mathrm{op}})^n : n \in \mathbb{N} \}$ *die kleinste Äquivalenzrelation, die* R *enthält, die sog. Äquivalenz-Hülle.*

Eigenschaften von Funktionen

Satz

1. *Die Komposition von Relationen ist assoziativ mit neutralen Elementen der Form* $\boldsymbol{id}_B$, B *eine Menge.*
2. *Die Funktionen sind unter Komposition abgeschlossen.* □

Definition

Eine Funktion $A \xrightarrow{f} B$ heißt

- ▷ injektiv, falls f^{op} einwertig ist (aus $f(a) = f(c)$ folgt $a = c$);
- ▷ surjektiv, falls f^{op} total ist (zu $b \in B$ ex. $a \in A$ mit $f(a) = b$);
- ▷ bijektiv, wenn f injektiv und surjektiv ist.

Satz

Ist $A \xrightarrow{f} B$ *bijektiv, so gilt für die sog. Umkehrfunktion* $B \xrightarrow{f^{\mathrm{op}}} A$*:*
$f; f^{\mathrm{op}} = f^{\mathrm{op}} \circ f = \boldsymbol{id}_A$ *und* $f^{\mathrm{op}}; f = f \circ f^{\mathrm{op}} = \boldsymbol{id}_B$. □

Anmerkungen

- Es besteht eine bijektive Beziehung zwischen Teilmengen $B \in A$ und charakteristischen Funktionen $A \xrightarrow{\chi_B} \mathbb{B} := \{0,1\} = 2$, die genau die Elemente von B auf 1 abbilden.[1] Das führt auch zu der alternativen Bezeichnung 2^A für die Potenzmenge $\boldsymbol{P}(A)$.
  Insbesondere stehen Relationen $A \longrightarrow B$, also Teilmengen von $A \times B$, in bijektiver Beziehung zu Funktionen $A \times B \longrightarrow \mathbb{B}$. Wir werden je nach Bedarf zwischen diesen Sichtweisen wechseln.
- Manche Autoren mögen keine Relationen $A \xrightarrow{R} B$ und verwenden stattdessen Funktionen $A \xrightarrow{R} \boldsymbol{P}(B)$, die jedem $a \in A$ die Menge $\{\, b \in B \,:\, aRb \,\}$ zuordnen; diese kann auch leer sein. Das scheint weitgehend eine Geschmacksfrage zu sein.

[1] $\mathbb{B}$ ist ein subobject classifyer in der Kategorie ***set*** der Mengen und Funktionen.

Beispiele für Quasi-/Halb-Ordnungen

Beispiel

- Die Gleichheit $=$ ist sowohl eine Äquivalenzrelation als auch eine Halbordnung auf jeder Menge X;
- $\leq$ ist eine lineare Halb-Ordnung auf $\mathbb{B}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$;
- $\subseteq$ ist eine Halb-Ordnung auf der Potenzmenge $P(X)$ einer Menge X, i.A. nicht linear;
- die Teilbarkeitsrelation $\mid$ ist eine nicht lineare Halbordnung auf $\mathbb{N}$;
- Jede Abbildung $X \xrightarrow{f} Y$ induziert eine Äquivalenzrelation auf X vermöge $x \sim x'$ gdw. $f(x) = f(x')$;
- Jede Abbildung $X \xrightarrow{f} \langle Y, \leq \rangle$ in eine quasi-geordnete Menge induziert eine Quasi-Ordnung auf X vermöge $x \sqsubseteq x'$ gdw. $f(x) \leq f(x')$, was f dadurch zu einer monotonen Abbildung (s.u.) macht.

back to $\sqsubseteq$

Polaritäten

Definition

Eine Abbildung $\langle X, \leq \rangle \xrightarrow{f} \langle Y, \sqsubseteq \rangle$ zwischen quasi-geordneten Mengen heißt monoton, bzw. antiton, falls für alle $a, b \in X$

$$a \leq b \quad \text{impliziert} \quad f(a) \sqsubseteq f(b) \quad \text{bzw.} \quad f(b) \sqsubseteq f(a)$$

Satz

Jedes $R \subseteq X \times Y$ induziert ein Paar antitoner Abbildungen (Polarität)

$$P(X) \underset{R^{\triangleright}}{\overset{R^{\triangleleft}}{\leftrightarrows}} P(Y) \qquad \textit{via} \qquad \begin{aligned} R^{\triangleleft}(V) &:= \{\, x \in X \,:\, xRy \text{ für jedes } y \in V \,\} \\ R^{\triangleright}(U) &:= \{\, y \in Y \,:\, xRy \text{ für jedes } x \in U \,\} \end{aligned}$$

mit $U \subseteq R^{\triangleleft}(V)$ gdw. $V \subseteq R^{\triangleright}(U)$ für alle $U \subseteq X$ und alle $V \subseteq Y$, sowie

$$R^{\triangleleft} \circ R^{\triangleright} \circ R^{\triangleleft} = R^{\triangleleft} \quad \textit{und} \quad R^{\triangleright} \circ R^{\triangleleft} \circ R^{\triangleright} = R^{\triangleright}.$$

Hüllenoperatoren

Definition

Eine monotone Abbildung $\langle X, \leq \rangle \xrightarrow{H} \langle X, \leq \rangle$ heißt Hüllenoperator, falls

- H extensiv ist, d.h. $x \leq H(x)$ für alle $x \in X$;
- H idempotent ist, d.h., $H(H(x)) \leq H(x)$ für alle $x \in X$.

Im halb-geordneten Fall gilt wegen der Monotonie $H(H(x)) = H(x)$.

Satz

Für jedes $R \subseteq X \times Y$ *ist* $R^{\triangleleft} \circ R^{\triangleright}$ *und* $R^{\triangleright} \circ R^{\triangleleft}$ *ein Hüllenoperator auf* $P(X)$ *bzw.* $P(Y)$.

Beweis.

Wegen $R^{\triangleleft}(V) \subseteq R^{\triangleleft}(V)$ gdw. $V \subseteq R^{\triangleright}(R^{\triangleleft}(V))$ ist $R^{\triangleright} \circ R^{\triangleleft}$ extensiv; analog für $R^{\triangleleft} \circ R^{\triangleright}$. Daher gilt $R^{\triangleleft}(V) = R^{\triangleleft}(R^{\triangleright}(R^{\triangleleft}(V)))$, denn $R^{\triangleleft}$ ist antiton. Also ist $R^{\triangleright} \circ R^{\triangleleft}$ idempotent; analog für $R^{\triangleleft} \circ R^{\triangleright}$. □

Ordnungstheoretische Grundbegriffe

Definition

In einer quasi-geordneten Menge $\langle X, \sqsubseteq \rangle$ heißt $x \in X$ untere Schranke von $Q \subseteq X$ falls $x \sqsubseteq q$ für alle $q \in Q$ gilt.
Hat jede (jede endliche) Teilmenge $Q \subseteq X$ eine größte untere Schranke $\bigsqcap Q$ ($\sqcap Q$), das Infimum von Q, so heißt $\langle X, \sqsubseteq \rangle$ $\bigsqcap$-($\sqcap$-)Halbverband.
Duale Begriffe: obere Schranke, $\bigsqcup$-($\sqcup$-)Halbverband

In einem (vollständigen) Verband hat jede endliche (jede) Teilmenge ein Infimum und ein Supremum.

Beispiel

- Jede linear geordnete Menge ist ein Verband; $\sqcap = \min, \sqcup = \max$.
- Jede Potenzmenge ist ein vollständiger Verband; $\bigsqcap = \bigcap$, $\bigsqcup = \bigcup$.
- $\mathbb{N}$ ist ein Verband bzgl. $|$; $\sqcap =$ **ggT**, $\sqcup =$ **kgV**.

Satz

Ist $\langle X, \sqsubseteq \rangle$ *ein (vollständiger) Verband, und ist* $\mathcal{H} = \langle H_i : i \in I \rangle$ *eine Familie von Hüllenoperatoren auf* $\langle X, \sqsubseteq \rangle$ *mit* I *endlich (beliebig), so ist auch* $\bigsqcap \mathcal{H}$ *ein Hüllenoperator mit*

$$\bigsqcap \mathcal{H}(x) = \bigsqcap \{ H_i(x) : i \in I \}$$

Beweis

Monotonie: Für $x \sqsubseteq y$ in X gilt $H_i(x) \sqsubseteq H_i(y)$ für alle $i \in I$ und daher $\bigsqcap \{ H_i(x) : i \in I \} \sqsubseteq \bigsqcap \{ H_i(y) : i \in I \}$.

Extensivität: Aus $x \sqsubseteq H_i(x)$ für $i \in I$ folgt $x \sqsubseteq \bigsqcap \mathcal{H}$.

Idempotenz: $\bigsqcap \mathcal{H}(x) \sqsubseteq H_i(x)$ für $x \in X$ und $i \in I$ impliziert

$$H_i\big(\bigsqcap \mathcal{H}(x)\big) \sqsubseteq H_i(H_i(x)) = H_i(x) \text{ also } \bigsqcap \mathcal{H}\big(\bigsqcap \mathcal{H}(x)\big) \sqsubseteq \bigsqcap \mathcal{H}(x) \quad \square$$

Beachte: da es nur eine Menge von Hüllenoperatoren auf $\langle X, \sqsubseteq \rangle$ gibt, darf der Index i sogar eine echte Klasse durchlaufen. Diese Tatsache nutzen wir bei der Definition der logischen Folgerung $\models$ der Prädikatenlogik aus.

Teil 1
Aussagenlogik (AL)

Inhaltsverzeichnis, Teil 1

Kapitel 2
Syntax

Was soll abstrahiert werden?

- Ziel ist es, Aussagen zu abstrahieren, die wahr oder falsch sein können. Die Abstraktionen nennen wir Formeln.
- In der natürlichen Sprache lassen sich Aussagen mittels Bindewörtern („und“, „oder“, „nicht“ *etc.*) zu komplexeren Aussagen verknüpfen.
- Zum Start braucht man nicht weiter zerlegbare oder atomare Aussagen. Die AL ignoriert, dass diese eine innere Struktur (nicht-logischer Art) haben können (dafür ist die PL_1 zuständig).
- Die oben erwähnten Bindewörter sind ebenfalls zu abstrahieren, mit Hilfe sog. Junktoren. **Achtung:** mit dieselben Junktoren lassen sich Aussagen zu verknüpfen, *zu denen bisher konstruktive Beweise vorliegen oder nicht*, statt Aussagen, die wahr oder falsch sein können, was zu unterschiedlicher Semantik führt. Man spricht dann von intuitionistischer oder konstruktivistischer statt klassischer Logik.

Drei Sorten von Symbolen:

- Ein abzählbar unendlicher Vorrat $\mathcal{A}$ von Aussagen-Variablen oder atomaren Formeln p, q, p_i, q_j *etc.*, um immer von den relevanten konkreten Aussagen abstrahieren zu können;
- Abstraktionen der Bindewörter „und“, „oder“, „nicht“ etc. der natürlichen Sprache mittels einer endlichen Menge $\mathcal{J}$ sog. Junktoren, deren Namen die intendierte Semantik (s.u) andeuten, unter anderem

$\wedge$	Konjunktion	für „und“
$\vee$	Disjunktion	für „oder“
$\rightarrow$	Implikation	für „[wenn...], dann“
$\leftrightarrow$	Äquivalenz	für „genau dann wenn“
$\neg$	Negation	für „nicht“
$\top$	Gewissheit	für „wahr“
$\bot$	Absurdität	für „falsch“

- ggf. Klammern (und) zum Auflösen eventueller Mehrdeutigkeiten.

Stelligkeit der Junktoren und Rolle der Klammern

Die Anzahl der zu verbindenden Aussagen durch die Bindewörter weist den sie abstrahierenden Junktoren eine endliche Stelligkeit zu:

$$\text{nullär:}\quad \bot \text{ und } \top \quad ; \qquad \text{unär:}\quad \neg \quad ; \qquad \text{binär:}\quad \wedge,\ \vee,\ \to,\ \leftrightarrow\ .$$

ar (engl. "arity") bezeichnet die entsprechende Abbildung von der Menge $\mathcal{J} = \{\bot, \top, \neg, \wedge, \vee, \to, \leftrightarrow\}$ in die Menge $\mathbb{N} = \{0, 1, 2 \dots\}$ der natürlichen Zahlen; dies ist eine Signatur im Sinne von ▸ Anhang A auf $\mathcal{J}$.

Die zugehörigen Terme, d.h., bestimmte wohlgeformte Wörter über der disjunkten (!) Vereinigung $\mathcal{J}[\mathcal{A}] := \mathcal{J} + \mathcal{A}$, dem Alphabet der AL, heißen hier Formeln. Die Infix-Notation für binäre Junktoren erfordert zwingend die Klammern (und) im Alphabet, also $\mathcal{J}^{()}[\mathcal{A}] := \mathcal{J}[\mathcal{A}] + \{(,)\}$, derer man bald mit Ersparnis-Regeln Herr zu werden versucht sein wird.

[Da die "(reverse) polish notation" ((R)PN) für Terme ohne Klammern auskommt, sind diese kein echter Bestandteil der Sprache der AL, sondern nur ein darstellungstechnisches Artefakt ohne inhärent logische Bedeutung.]

Die formale Sprache der AL: Formeln

Definition (induktiv, Infix-Version)

Die formale Sprache $\mathcal{F}[\mathcal{A}]$ der AL ist die kleinste Menge von Wörtern über dem Alphabet $\mathcal{J}^{()}[\mathcal{A}]$ mit folgenden Abschlusseigenschaften:

- $\mathcal{A} \cup \{\bot, \top\} \subseteq \mathcal{F}[\mathcal{A}]$;
- wenn $A \in \mathcal{F}[\mathcal{A}]$, dann $\neg A \in \mathcal{F}[\mathcal{A}]$;
- wenn $A, B \in \mathcal{F}[\mathcal{A}]$, dann $(A \star B) \in \mathcal{F}[\mathcal{A}]$ mit $\star$ binär.

Die Elemente von $\mathcal{F}[\mathcal{A}]$ heißen Formeln; diejenigen mit mindestens einem Junktor heißen molekular (im Unterschied zu den atomaren Formeln in $\mathcal{A}$).

Dies ist im Wesentlichen die Definition der Terme aus [Anhang A] für die Signatur $\mathcal{J} \xrightarrow{ar} \mathbb{N}$, wobei für binäre Junktoren statt der polnischen Notation $\star A\, B$ oder, lesbarer, $\star(A, B)$, die zwingend Klammer-bewehrte Infix-Schreibweise $(A \star B)$ verwendet wird.

Alternative Beschreibung: Formeln via BNF

Informatiker drücken den Aufbau generischer Formeln $\boldsymbol{F}$ gern rekursiv mit Hilfe einer „Grammatik“ in Backus-Naur-Form, kurz BNF, aus:

$$\boldsymbol{F} ::= \mathcal{A} \mid \bot \mid \top \mid \neg \boldsymbol{F} \mid (\boldsymbol{F} \star \boldsymbol{F}) \quad \text{für } \star \text{ binär}$$

Interpretiere dies als rekursive (Un-)Gleichung für die Mengen-Variable $\boldsymbol{F}$:

- $::=$ steht für Mengeninklusion $\subseteq$;
- die senkrechten Striche $\mid$ bedeuten Vereinigung $\cup$;
- $\bot$ und $\top$ stehen für die Singletons $\{\bot\}$ bzw. $\{\top\}$;
- $\neg \boldsymbol{F}$ steht abkürzend für $\{ \neg A : A \in \boldsymbol{F} \}$;
- $(\boldsymbol{F} \star \boldsymbol{F})$ steht abkürzend für $\{ (A \star B) : A, B \in \boldsymbol{F} \}$.

Iteration ausgehend von $\boldsymbol{F}_0 = \emptyset$ liefert $\mathcal{F}[\mathcal{A}]$ als kleinsten Fixpunkt, also

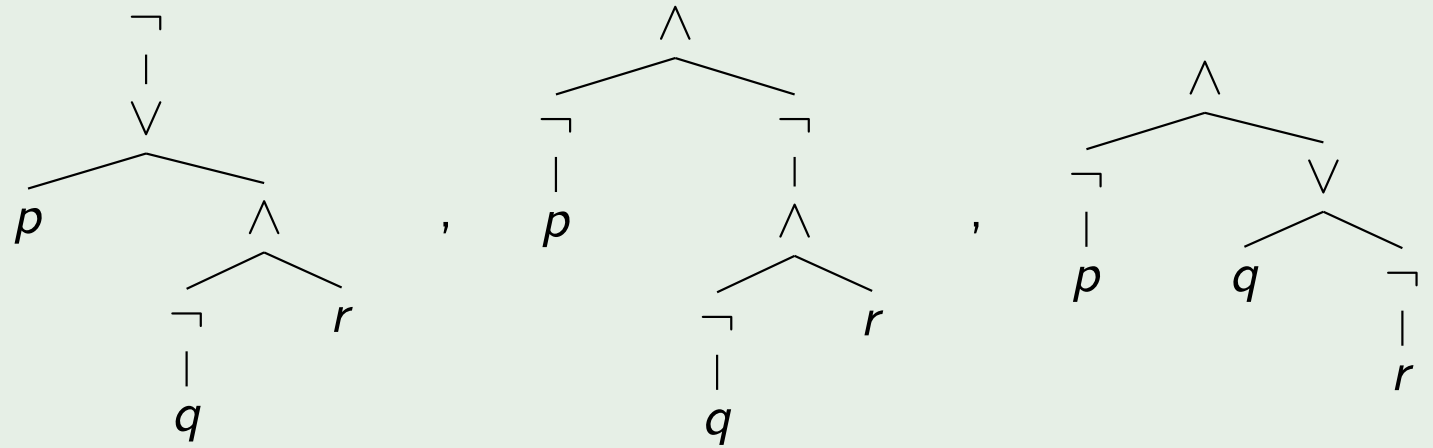
$$\mathcal{F}[\mathcal{A}] = \mathcal{A} \cup \{\bot\} \cup \{\top\} \cup \neg \mathcal{F}[\mathcal{A}] \cup (\mathcal{F}[\mathcal{A}] \star \mathcal{F}[\mathcal{A}]) \text{ für } \star \text{ binär}$$

Formeln der AL als Syntaxbäume

Beispiel

$$\neg(p \vee (\neg q \wedge r)) \quad , \quad (\neg p \wedge \neg(\neg q \wedge r)) \quad , \quad (\neg p \wedge (q \vee \neg r))$$

haben folgende Baum-Darstellung:



Aufgrund der Redundanz in Form vertikaler Ausdehnung erübrigen sich hier Klammern. Für ein Alphabet ist eine saubere Definition gelabelter „Bäume“ leider viel komplexer als die von Wörtern und wird hier unterschlagen.

Beobachtung

1. Die obige Signatur $\mathcal{J} \xrightarrow{ar} \mathbb{N}$ wird oft durch Indizierung der aufgelisteten Junktoren in $\mathcal{J}$ ausgedrückt:

$$\bot_{/0} \ , \ \top_{/0} \ , \ \neg_{/1} \ , \ \wedge_{/2} \ , \ \vee_{/2} \ , \ \rightarrow_{/2} \ , \ \leftrightarrow_{/2}$$

2. Die kanonische $\mathcal{J}$-Algebrastruktur auf $\mathcal{F}[\mathcal{A}]$ ist gegeben durch

$$\begin{array}{lll}
(\mathcal{F}[\mathcal{A}])^0 \xrightarrow{\Phi(\bot)} \mathcal{F}[\mathcal{A}]; & & \bullet \longmapsto \bot \\
(\mathcal{F}[\mathcal{A}])^0 \xrightarrow{\Phi(\top)} \mathcal{F}[\mathcal{A}]; & & \bullet \longmapsto \top \\
(\mathcal{F}[\mathcal{A}])^1 \xrightarrow{\Phi(\neg)} \mathcal{F}[\mathcal{A}]; & & A \longmapsto \neg A \\
(\mathcal{F}[\mathcal{A}])^2 \xrightarrow{\Phi(\star)} \mathcal{F}[\mathcal{A}]; & & \langle A, B\rangle \longmapsto (A \star B)
\end{array}$$

Wie üblich lässt man Φ meist weg.

Strukturelle Induktion

Das allgemeine Beweisprinzip nimmt hier folgende Form an:

Satz

Sei E eine Eigenschaft, die Formeln haben können oder nicht. Falls

- *jede atomare Formel Eigenschaft E hat;*
- *$\bot$ und $\top$ Eigenschaft E haben;*
- *mit A auch $\neg A$ Eigenschaft E hat;*
- *mit A und B auch $(A \star B)$ Eigenschaft E hat ($\star$ binär)*

dann hat jede Formel in $\mathcal{F}[\mathcal{A}]$ die Eigenschaft E. □

Andere Beweisprinzipien stehen weiterhin zur Verfügung, wie z.B. Induktion über die Länge von Formeln, oder über die Tiefe des Syntaxbaumes...

Beispiel für strukturelle Induktion

Lemma (spezifisch für Infix-Schreibweise)

Jede Formel in $\mathcal{F}[\mathcal{A}]$ hat gleich viele öffnende wie schließende Klammern.

Beweis.

Induktionsanfang: Die Behauptung ist korrekt für Formeln in $\mathcal{A} \cup \{\bot, \top\}$.
Induktionsannahme: Jede Formel mit weniger Junktoren als die molekulare Formel A erfülle die Behauptung.
Induktionsschluss: A ist durch genau eine der 5 Abschlussoperationen aus Formeln mit weniger Junktoren entstanden. Die Negation fügt keine Klammern hinzu, während die binären Junktoren je eine öffnende und eine schließende Klammer hinzufügen, was die Gleichheit bewahrt. □

(Anstelle von Formeln mit weniger Junktoren könnte man auch Formeln mit niedrigerer Baumdarstellung betrachten.)

Alternativer Beweis.

Idee: Buch führen über öffnende bzw. schließende Klammern einer Formel.

Wir setzen die Abbildung

$$\mathcal{J}^{()}[\mathcal{A}] \xrightarrow{\delta} \mathbb{Z} \quad , \quad x \longmapsto \begin{cases} 1 & \text{falls } x = (\\ -1 & \text{falls } x =) \\ 0 & \text{sonst} \end{cases},$$

auf ganz $(\mathcal{J}^{()}[\mathcal{A}])^*$ fort (nicht nur $\mathcal{F}[\mathcal{A}]$), indem wir definieren

$$\bar{\delta}(a_0 a_1 \dots a_{n-1}) := \textstyle\sum_{i<n} \delta(a_i)$$ [a]

$\bar{\delta}$ bestimmt also die Differenz zwischen der Anzahl der öffnenden bzw. schließenden Klammern in einem Wort.

Zu zeigen ist nun, dass für jede Formel $A \in \mathcal{F}[\mathcal{A}]$ gilt $\bar{\delta}(A) = 0$.

---

[a] Dem liegt die strukturelle Induktion für Monoide zugrunde, hier $\langle \mathbb{Z}, +, 0 \rangle$, wobei uns die Eindeutigkeit der Fortsetzung an dieser Stelle nicht interessiert.

Fortsetzung

Induktionsanfang: Beachte, dass $\bar{\delta}$ auf $\mathcal{A} + \{(,)\}$ mit δ übereinstimmt.
Induktionsannahme: Jede Formel mit weniger Junktoren als die molekulare Formel A erfülle die Behauptung.
Induktionsschluss: A ist durch genau eine der 5 Abschlussoperationen aus Formeln mit weniger Junktoren entstanden.

- Falls $A = \neg B$ gilt

$$\bar{\delta}(\neg B) = \delta(\neg) + \bar{\delta}(B) = 0 + \bar{\delta}(B) = 0$$

- Falls $A = (B \star C)$ gilt

$$\begin{aligned}\bar{\delta}\big((B \star C)\big) &= \delta\big(() + \bar{\delta}(B) + \delta(\star) + \bar{\delta}(C) + \delta()\big)\\ &= 1 + 0 + 0 + 0 + (-1) = 0\end{aligned}$$

Hier haben wir die Assoziativität der Addition ausgenutzt. □

Natürlich bildet $\bar{\delta}$ auch viele andere Wörter auf 0 ab, z.B. $w =)p\neg(\wedge$.

$\mathcal{J}$-Algebra

Alle wesentlichen Konstruktionen auf der $\mathcal{J}$-Algebra $\mathcal{F}[\mathcal{A}]$ werden nun mit Hilfe des ▸ Rekursionssatz vorgenommen.

Beispiel

Länge $|A|$ = Anzahl aller $\mathcal{J}^{()}[\mathcal{A}]$-Symbole in A: Verwende

- die konstante Abbildung $\mathcal{A} \xrightarrow{1} \mathbb{N}$;
- die folgenden Interpretationen von $\mathcal{J}$ in $\mathbb{N}$:

$$\begin{aligned}
\mathbb{N}^0 &\xrightarrow{I(\bot)} \mathbb{N}; & \bullet &\longmapsto 1\\
\mathbb{N}^0 &\xrightarrow{I(\top)} \mathbb{N}; & \bullet &\longmapsto 1\\
\mathbb{N}^1 &\xrightarrow{I(\neg)} \mathbb{N}; & n &\longmapsto n+1\\
\mathbb{N}^2 &\xrightarrow{I(\star)} \mathbb{N}; & \langle n, m\rangle &\longmapsto n+m+3
\end{aligned}$$

Nun ist $\mathcal{F}[\mathcal{A}] \xrightarrow{||} \mathbb{N}$ die eindeutige Fortsetzung von $\mathcal{A} \xrightarrow{\delta} \mathbb{N}$ zu einem $\mathcal{J}$-Homomorphismus nach $\langle \mathbb{N}, I\rangle$.

Beispiel

$\#_p(A)$ bezeichne die Anzahl der Vorkommen des Atoms $p \in \mathcal{A}$ in A.
Eine induktive Definition verwendet etwa

$$\#_p(q) = \begin{cases} 1 & \text{falls } q = p\ ; \\ 0 & \text{sonst} \end{cases} \quad \text{für} \quad q \in \mathcal{A}$$

$$\#_p(\bot) = \#_p(\top) = 0$$
$$\#_p(\neg B) = \#_p(B)$$
$$\#_p\big((A \star B)\big) = \#_p(A) + \#_p(B) \quad \text{für } \star \text{ binär}$$

Die letzten drei Zeilen spezifizieren eine weitere $\mathcal{J}$-Stuktur K auf $\mathbb{N}$:

$$K(\bot) = K(\top) = 0 \text{ (konstant)} \quad , \quad K(\neg) = \boldsymbol{id}_{\mathbb{N}} \quad , \quad K(\star) = +$$

bzgl. der $\mathcal{F}[\mathcal{A}] \xrightarrow{\#_p} \mathbb{N}$ dann ein $\mathcal{J}$-Homomorphismus ist.

Analog kann man z.B. die Menge $@(A)$ der in A vorkommenden Atome definieren (HA).

Teilformeln (T. Palm)

Die Überprüfung, ob auf der Zielmenge wirklich eine $\mathcal{J}$-Struktur vorliegt, darf man nicht vergessen:

Beispiel

Die endliche Menge $T(A)$ der Teilformeln von A enthält neben A die Teilformeln derjenigen Formel(n), aus denen A mit Hilfe des letzten Junktors (Wurzel des Syntaxbaumes) konstruiert wurde. Etwa

$$T(q) = \{q\} \text{ für } q \in \mathcal{A} \ , \ T(\bot) = \{\bot\} \ , \ T(\top) = \{\top\}$$
$$T(\neg B) = \{\neg B\} \cup T(B) = L(\neg)\big(T(B)\big)$$
$$T\big((A \star B)\big) = \{(A \star B)\} \cup T(A) \cup T(B) = T(A)\, L(\star)\, T(B) \quad (\star \text{ binär})$$

Aber wenn $\mathcal{F}[\mathcal{A}] \xrightarrow{T} \boldsymbol{P}_\omega(\mathcal{F}[\mathcal{A}])$ ein $\mathcal{J}$-Homomorphismus sein soll, welche $\mathcal{J}$-Struktur L auf $\boldsymbol{P}_\omega(\mathcal{F}[\mathcal{A}])$ liegt ihm zugrunde? Um welche negierte Formel soll etwa Γ ergänzt werden um $L(\neg)(\Gamma)$ zu liefern?

Beispiel (Fortsetzung)

Was sich hingegen als $\mathcal{J}$-Homomorphismus identifizieren lässt ist die Abbildung $\mathcal{F}[\mathcal{A}] \xrightarrow{\langle \mathbf{id}, T\rangle} \mathcal{F}[\mathcal{A}] \times \boldsymbol{P}_\omega(\mathcal{F}[\mathcal{A}])$. Die Spezifikation

$$
\begin{aligned}
\langle \mathbf{id}, T\rangle(q) &= \langle q, \{q\}\rangle \quad \text{für} \quad q \in \mathcal{A} \\
\langle \mathbf{id}, T\rangle(\bot) &= \langle \bot, \{\bot\}\rangle \\
\langle \mathbf{id}, T\rangle(\top) &= \langle \top, \{\top\}\rangle \\
\langle \mathbf{id}, T\rangle(\neg B) &= \langle \neg B, \{\neg B\} \cup T(B)\rangle \\
\langle \mathbf{id}, T\rangle\big((A \star B)\big) &= \langle (A \star B), \{(A \star B)\} \cup T(A) \cup T(B)\rangle \quad \text{für } \star \text{ binär}
\end{aligned}
$$

liefert folgende Interpretation I der Junktoren auf $\mathcal{F}[\mathcal{A}] \times \boldsymbol{P}_\omega(\mathcal{F}[\mathcal{A}])$:

- $I(\bot) = \langle \bot, \{\bot\}\rangle$ und $I(\top) = \langle \top, \{\top\}\rangle$;
- $I(\neg)\langle A, \Gamma\rangle = \langle \neg A, \{\neg A\} \cup \Gamma\rangle$;
- $\langle A, \Gamma\rangle\, I(\star)\, \langle B, \Delta\rangle = \langle (A \star B), \{(A \star B)\} \cup \Gamma \cup \Delta\rangle$.

Substitution

Für Formeln $A, B \in \mathcal{F}[\mathcal{A}]$ und ein Atom $q \in \mathcal{A}$ bezeichne $A[q/B]$ die Formel, die durch simultanes Ersetzen jedes Auftretens von q in A durch B entsteht. Genauer:

Definition

$\mathcal{F}[\mathcal{A}] \xrightarrow{[p/B]} \mathcal{F}[\mathcal{A}]$ ist die durch den ▸ Rekursionssatz bestimmte homomorphe Fortsetzung der Abbildung

$$\mathcal{A} \xrightarrow{p/B} \mathcal{F}[\mathcal{A}], \qquad q \longmapsto \begin{cases} B & \text{falls } q = p \\ q & \text{falls } q \neq p \end{cases}$$

bzgl. der kanonischen $\mathcal{J}$-Algebra-Struktur auf $\mathcal{F}[\mathcal{A}]$.

- Falls A kein Atom q enthält, stimmt $A[q/B]$ mit A überein.
- Falls B das Atom q enthält, kann q in $A[q/B]$ weiterhin auftreten; daher ist $[q/B]$ nicht notwendig idempotent!

Endlich: Klammerersparnisregeln!

Um Klammern einzusparen, verabreden wir folgende

Bindungskonventionen

- Junktoren niedrigerer Stelligkeit binden stärker als solche höherer Stelligkeit.
- $\wedge$ und $\vee$ binden gleich stark, aber stärker als $\rightarrow$ und $\leftrightarrow$.
- $\rightarrow$ und $\leftrightarrow$ binden gleich stark.
- binäre Junktoren assoziieren nach rechts, d.h., $A \star B \star C$ ist als $A \star (B \star C)$ zu interpretieren (evtl. anders als auf alten Folien!). Für $(A \star B) \star C$ gib es keine Vereinfachung.

Achtung: Im Rahmen der Semantik werden wir später sehen, dass die binären Operatoren $\wedge$, $\vee$ und $\leftrightarrow$ auf Formeln modulo Äquivalenz tatsächlich assoziativ sein werden; für $\rightarrow$ ist dies aber nicht der Fall!

Kapitel 3
Semantik

Die $\mathcal{J}$-Algebra $\mathbb{B} = \{0, 1\}$ der Wahrheitswerte

Definition

Die kanonische Interpretation von $\mathcal{J}$ in $\mathbb{B}$ (wir lassen I weg) ist

$$\mathbb{B}^0 \xrightarrow{\bot} \mathbb{B};\quad \bullet \longmapsto 0 \qquad \mathbb{B}^2 \xrightarrow{\wedge} \mathbb{B};\quad \langle x, y\rangle \longmapsto \inf\{x, y\}$$
$$\mathbb{B}^0 \xrightarrow{\top} \mathbb{B};\quad \bullet \longmapsto 1 \qquad \mathbb{B}^2 \xrightarrow{\vee} \mathbb{B};\quad \langle x, y\rangle \longmapsto \sup\{x, y\}$$
$$\mathbb{B}^1 \xrightarrow{\neg} \mathbb{B};\quad x \longmapsto 1 - x \qquad \mathbb{B}^2 \xrightarrow{\rightarrow} \mathbb{B};\quad \langle x, y\rangle \longmapsto \leq \langle x, y\rangle$$
$$\mathbb{B}^2 \xrightarrow{\leftrightarrow} \mathbb{B};\quad \langle x, y\rangle \longmapsto = \langle x, y\rangle$$

$\mathcal{J}$-Homomorphismen $\mathcal{F}[\mathcal{A}] \longrightarrow \mathbb{B}$ heißen Bewertungen.

Nach dem ▸ Rekursionssatz erlaubt jede Belegung der Variablen $\varphi \in \mathbb{B}^{\mathcal{A}}$ eine eindeutige homomorphe Fortsetzung zu einer Bewertung $\mathcal{F}[\mathcal{A}] \xrightarrow{\hat{\varphi}} \mathbb{B}$.

Definition

$\langle A, \varphi\rangle \longmapsto \hat{\varphi}(A)$ liefert eine Auswertungsfunktion $\mathcal{F}[\mathcal{A}] \times \mathbb{B}^{\mathcal{A}} \xrightarrow{\boldsymbol{E}} \mathbb{B}$, die wir auch als Erfüllungsrelation $\boldsymbol{E} \subseteq \mathcal{F}[\mathcal{A}] \times \mathbb{B}^{\mathcal{A}}$ auffassen können.

Wahrheitstabellen und Auswertung von Formeln

Die kanonische Interpretation von $\mathcal{J}$ in $\mathbb{B}$ lässt sich alternativ mittels Wahrheitstabellen darstellen:

$\bot$
0

$\top$
1

	$\neg$
0	1
1	0

		$\wedge$	$\vee$	$\rightarrow$	$\leftrightarrow$
0	0	0	0	1	1
0	1	0	1	1	0
1	0	0	1	0	0
1	1	1	1	1	1

Mit deren Hilfe lassen sich Bewertungen von Formeln aufgrund der Belegung der Variablen bestimmen.

Beispiel (Bestimme $\hat{\varphi}(\neg(p \wedge q) \rightarrow r)$ falls $\varphi(p) = 1$, $\varphi(q) = \varphi(r) = 0$)

$$\underbrace{\underbrace{\underbrace{\neg(\underbrace{p}_{1} \wedge \underbrace{q}_{0})}_{0} \rightarrow \underbrace{r}_{0}}_{1}}_{\mathbf{0}}$$

oder in „flacher“ Notation (verbirgt die Reihenfolge)

$$\underset{1}{\neg}\,(\,\underset{1}{p}\,\underset{0}{\wedge}\,\underset{0}{q}\,)\,\underset{\mathbf{0}}{\rightarrow}\,\underset{0}{r}$$

Boole'sche Funktionen

Definition

Die obige Auswertungsstrategie, angewandt auf alle Belegungen der $n = |@(A)|$ in A vorkommenden Atome liefert die von A induzierte Boole'sche Funktion $\mathbb{B}^n \xrightarrow{e_A} \mathbb{B}$.

Häufig wird e_A in Tabellenform angegeben:

Beispiel

p	q	r	p	$\vee$	q	$\rightarrow$	$(q$	$\leftrightarrow$	$r)$	$\wedge$	$\neg$	$(p$	$\vee$	$\neg$	$r)$
0	0	0		0		**1**		1		0	0		1	1	
0	0	1		0		**1**		0		0	1		0	0	
0	1	0		1		**0**		0		0	0		1	1	
0	1	1		1		**1**		1		1	1		0	0	
1	0	0		1		**0**		1		0	0		1	1	
1	0	1		1		**0**		0		0	0		1	0	
1	1	0		1		**0**		0		0	0		1	1	
1	1	1		1		**0**		1		0	0		1	0	

Offenbar gibt es höchstens 2^n viele Boole'sche Funktionen auf $\mathbb{B}^n$.

Polarität und Hüllenoperator für die Erfüllungsrelation $\boldsymbol{E}$

Die Relation $\boldsymbol{E} \subseteq \mathcal{F}[\mathcal{A}] \times \mathbb{B}^{\mathcal{A}}$ liefert eine sog. ▸ Polarität, d.h., ein Paar antitoner Abbildungen zwischen den entsprechenden Potenzmengen:

$$\boldsymbol{P}(\mathcal{F}[\mathcal{A}]) \underset{\boldsymbol{E}^{\triangleright}}{\overset{\boldsymbol{E}^{\triangleleft}}{\rightleftarrows}} \boldsymbol{P}\left(\mathbb{B}^{\mathcal{A}}\right) \quad \text{mit} \quad \begin{aligned} \Phi^{\triangleleft} &:= \boldsymbol{E}^{\triangleleft}(\Phi) = \{\, A \in \mathcal{F}[\mathcal{A}] \,:\, \varphi(A) = 1 \text{ für alle } \varphi \in \Phi \,\} \\ \Gamma^{\triangleright} &:= \boldsymbol{E}^{\triangleright}(\Gamma) = \{\, \varphi \in \mathbb{B}^{\mathcal{A}} \,:\, \varphi(A) = 1 \text{ für alle } A \in \Gamma \,\} \end{aligned}$$

die zudem beide beliebige Vereinigungen in Durchschnitte abbilden.

Für alle Formelmengen $\Gamma \subseteq \mathcal{F}[\mathcal{A}]$ und alle Belegungsmengen $\Phi \subseteq \mathbb{B}^{\mathcal{A}}$ gilt

$$\Gamma \subseteq \Phi^{\triangleleft} \quad \text{gdw.} \quad \Gamma^{\triangleright} \supseteq \Phi \qquad \text{und folglich} \qquad \Gamma \subseteq \Gamma^{\triangleright\triangleleft} \quad \text{sowie} \quad \Phi^{\triangleleft\triangleright} \supseteq \Phi$$

was ▸ Hüllenoperatoren auf den jeweiligen Potenzmengen ergibt.

Bemerkung.

Der Hüllenoperator $(\ )^{\triangleright\triangleleft}$ auf der Potenzmenge der Formelmenge $\mathcal{F}[\mathcal{A}]$ ist die Basis für den logischen Folgerungsbegriff.

Logische Folgerung $\models$, Erfüllbar- und Allgemeingültigkeit

Definition

Betrachte $\Gamma \subseteq \mathcal{F}[\mathcal{A}]$ (aus sog. Prämissen) und $A, B \in \mathcal{F}[\mathcal{A}]$.

- Falls $A \in \Gamma^{\triangleright\triangleleft}$, d.h., jeder Erfüller der Prämissen erfüllt auch A, schreibt man $\Gamma \models A$ und sagt „A folgt logisch aus Γ“.
- A heißt allgemeingültig/Tautologie, wenn jede Belegung A erfüllt. Schreibweise: $\models A$ statt $\emptyset \models A$ bzw. $A \in \emptyset^{\triangleright\triangleleft} = \left(\mathbb{B}^{\mathcal{A}}\right)^{\triangleleft}$.
- B (bzw. Γ) heißt erfüllbar, falls $\varphi \in \mathbb{B}^{\mathcal{A}}$ existiert mit $\hat{\varphi}(B) = 1$ (für jedes $B \in \Gamma$), d.h., falls $\{B\}^{\triangleright} \neq \emptyset$ (bzw. $\Gamma^{\triangleright} \neq \emptyset$).

Formeln $A \in \mathcal{F}[\mathcal{A}]$ und Belegungen $\varphi \in \mathbb{B}^{\mathcal{A}}$ erfüllen $\varphi(A) \neq \varphi(\neg A)$, daher zerlegt jedes A die Belegungsmenge disjunkt: $\mathbb{B}^{\mathcal{A}} = \{A\}^{\triangleright} + \{\neg A\}^{\triangleright}$.

Beobachtung.

A ist genau dann allgemeingültig, wenn $\neg A$ nicht erfüllbar ist. (HA) □

Charakterisierung der Unerfüllbarkeit von Formelmengen

Lemma

Folgende Bedingungen sind für eine Formelmenge $\Gamma \subseteq \mathcal{F}[\mathcal{A}]$ *äquivalent:*

(0) Γ *ist unerfüllbar, bzw.* $\Gamma^{\triangleright} = \emptyset$.

(1) $\Gamma \models A$ *für alle Formeln* A , *bzw.* $\Gamma^{\triangleright\triangleleft} = \mathcal{F}[\mathcal{A}]$.

(2) $\Gamma \models \bot$ *bzw.* $\bot \in \Gamma^{\triangleright\triangleleft}$.

(3) $\Gamma \models B$ *und* $\Gamma \models \neg B$ *bzw.* $\{B, \neg B\} \subseteq \Gamma^{\triangleright\triangleleft}$ *für ein* $B \in \mathcal{F}[\mathcal{A}]$.

Beweis.

(0) $\Rightarrow$ (1): Klar, da die Formeln keine Bedingungen erfüllen müssen.
(1) $\Rightarrow$ (2), (3): Trivial.
(2), (3) $\Rightarrow$ (0): $\Gamma^{\triangleright} = (\Gamma^{\triangleright})^{\triangleleft\triangleright} = (\Gamma^{\triangleright\triangleleft})^{\triangleright} \subseteq \{\bot\}^{\triangleright} = \{B, \neg B\}^{\triangleright} = \emptyset$ aufgrund der Eigenschaften von $(\ )^{\triangleright\triangleleft}$ und der Antitonie von $(\ )^{\triangleright}$. □

Logisches Schließen in der Praxis

Beispiel

Familie Z möchte im kommenden Jahr gern eine Waschmaschine, ein Auto und ein Moped anschaffen. Aber falls Herr Z seinen üblichen Bonus nicht bekommt, können sie sich nicht alles leisten. Die Waschmaschine ist aber unverzichtbar. Und die Familie braucht mindestens ein Fortbewegungmittel. Wenn sie auch in den Urlaub fahren will, kann sie sich kein Auto leisten. Wenn sie nicht in den Urlaub fährt, muss sie aber ein Moped kaufen, um den Sohn zu besänftigen.

- Zeigen Sie, dass Familie Z ein Moped und kein Auto anschafft, sofern Herr Z seinen üblichen Bonus nicht bekommt.

Beispiel (Formalisierung, Teil 1: Ereignisse und Junktoren)

Zunächst wollen wir die relevanten Informationen extrahieren und miteinander in Beziehung setzen:

- Dazu führen wir für die diversen Ereignisse, die eintreten können oder nicht, Abkürzungen ein:
  - p – Erhalt eines Bonusses;
  - q / r / s – Kauf einer Waschmaschine/eines Autos/eines Mopeds;
  - t – Antritt einer Urlaubsreise.
- Diese verknüpfen wir dann mit passenden logischen Junktoren
  - $\neg$ – nicht (Negation) (unär)
  - $\wedge$ / $\vee$ – und / oder (Konjunktion / Disjunktion) (binär)
  - $\rightarrow$ – wenn,dann (Implikation) (binär)

Beispiel (Formalisierung, Teil 2: die Prämissen)

- „Aber falls Herr Z seinen üblichen Bonus nicht bekommt, können sie sich nicht alles leisten.“

$$B_0 = \neg p \to \neg(q \wedge r \wedge s)$$

- „Die Waschmaschine ist aber unverzichtbar.“

$$B_1 = q$$

- „Die Familie braucht mindestens ein Fortbewegungsmittel.“

$$B_2 = r \vee s$$

- „Wenn sie in den Urlaub fahren will, kann sie sich kein Auto leisten.“

$$B_3 = t \to \neg r$$

- „Wenn sie nicht in den Urlaub fährt, muss sie ein Moped kaufen.“

$$B_4 = \neg t \to s$$

Beispiel (Formalisierung, Teil 3: die Schlußfolgerung)

Fraglich ist, ob aus der Wahrheit der Prämissen B_i, $i < 5$, auch die von

– „Sofern Herr Z seinen üblichen Bonus nicht bekommt, schafft Familie Z ein Moped und kein Auto an.“

$$A = \neg p \to (\neg r \wedge s)$$

folgt, formal geschrieben als

$$\{B_0, B_1, B_2, B_3, B_4\} \models A \qquad (*)$$

Anstatt jetzt alle $2^5 = 32$ Belegungen der Atome p, q, r, s und t darauf zu überprüfen, ob sie A erfüllen, sofern die Prämissen B_i, $i > 5$, erfüllt werden, kommen wir später auf dieses Problem zurück.

Das semantische Deduktionstheorem (sDT)

Lemma (sDT; Prämissen lassen sich zwischen $\models$ und $\to$ verschieben)

Für $\Gamma \subseteq \mathcal{F}[\mathcal{A}] \ni A, B$ *gilt:* $\Gamma \cup \{A\} \models B$ *gdw.* $\Gamma \models A \to B$.

Beweis.

($\Rightarrow$): Erfüllt φ die Menge Γ, folgt aus $\hat{\varphi}(A) = 1$ nach Voraussetzung $\hat{\varphi}(B) = 1$, oder es gilt $\hat{\varphi}(A) \leq \hat{\varphi}(B)$. Beides impliziert $\hat{\varphi}(A \to B) = 1$.

($\Leftarrow$): $\psi \in (\Gamma \cup \{A\})^{\triangleright} = \Gamma^{\triangleright} \cap \{A\}^{\triangleright}$ erfüllt Γ und A, nach Voraussetzung also auch $A \to B$, und wegen $\hat{\psi}(A) \leq \hat{\psi}(B)$ ebenfalls B. □

Korollar

$\{A, A \to B\} \models B$ bzw. $\Gamma \cup \{A\}$ unerfüllbar gdw. $\Gamma \models \neg A$.

Beweis.

Der 2. Teil verwendet, dass $\varphi(A \to \bot) = \varphi(\neg A)$ für alle $\varphi \in \mathbb{B}^{\mathcal{A}}$. □

Der Kompaktheitssatz (KPS)

Alle bisher eingeführten Begriffe sind entscheidbar, solange die Formelmengen endlich sind: dann kommen nur endlich viele Atome vor und man kann eine endliche Wahrheitstabelle aufstellen. Aber Γ darf unendlich sein.

Definition

Γ heißt endlich erfüllbar, wenn jede endliche Teilmenge von Γ erfüllbar ist.

Der folgende zentrale Satz besagt, dass man sich im Wesentlichen auf endliche Formelmengen beschränken kann:

Satz (Kompaktheitssatz)

Γ ist genau dann erfüllbar, wenn Γ endlich erfüllbar ist.

Die Notwendigkeit ist trivial, ebenso die Hinlänglichkeit, falls $\mathcal{A}$ endlich ist (Kontraposition). Vor dem Beweis der Hinlänglichkeit für unendliches $\mathcal{A}$ betrachten wir erst eine „Anwendung“ und einige Konsequenzen des KPS.

Ein Matching-Problem

Beispiel (Hochzeit im Mittelalter, nach T. Palm)

Alle Junggesellinnen eines Dorfes (Menge F) sollen verheiratet werden. Dazu liefern sie alle Listen $T(f) \subseteq M$ der ihnen genehmen Junggesellen ab. Gesucht ist als Lösung dieses speziellen Matchingproblems eine injektive Auswahlfunktion $F \xrightarrow{e} M$ mit $e(f) \in T(f)$ für alle $f \in F$.

Nun stellen wir uns das Dorf unendlich groß vor. Die Junggesellinnen bleiben aber bodenständig und liefern nur endliche Listen $T(f)$ ab.

Behauptung: Das Matching-Problem $F \xrightarrow{T} \boldsymbol{P}_\omega(M)$ lösbar, wenn es für jede endliche Teilmenge von F lösbar ist. (Die Umkehrung ist trivial.)

Der folgende Beweis nimmt einige Aspekte der PL aus Teil 2 vorweg.

Um den KPS anwenden zu können verwenden wir als Atome Symbole

$$p_{f,m} \quad \text{mit} \quad f \in F \quad \text{und} \quad m \in T(f) \qquad \text{„}f \text{ und } m \text{ heiraten.“}$$

Beispiel (Fortsetzung)

Γ besteht aus drei Komponenten, die verschiedene Aspekt der Problemstellung umsetzen ($f, f' \in F$, $m, m' \in M$):

- Auswahleigenschaft: $H_f := \bigvee_{m \in T(f)} p_{f,m}$ [a]
- Funktionalität: $I_{f;m,m'} := p_{f,m} \to \neg p_{f,m'}$ mit $m \neq m'$;
- Injektivität: $J_{f,f';m} := p_{f,m} \to \neg p_{f',m}$ mit $f \neq f'$.

Voraussetzung: Jede Einschränkung des Matchingproblems T auf eine endliche Teilmenge $F_0 \subseteq F$ ist lösbar.

Wir zeigen: Jede endliche Teilmenge $\Gamma_0 \subseteq \Gamma$ hat eine erfüllenden Belegung.

Für endliches $\Gamma_0 \subseteq \Gamma$ ist auch die Menge

$$F_* := \{ f \in F : f \text{ tritt als Index in einer Formel aus } \Gamma_0 \text{ auf } \}$$

endlich, denn jede Formel in Γ_0 hat nur endlich viele Atome.

[a] Endliche Disjunktion von Atomen; nimmt die Assoziativität von $\vee$ vorweg ▸ Folie 66.

Beispiel (Fortsetzung)

- Nach Voraussetzung lässt F_* mindestens eine Lösung e_* des eingeschränkten Matching-Problems T zu.
- Wiederholt man die Konstruktion von Γ mit F_* anstelle von F, liefert dies eine endliche Teilmenge $\Gamma_* \subseteq \Gamma$, die Γ_0 umfasst.
- Auswahleigenschaft, Funktionalität und Injektivität der Lösungen e_* entsprechen genau Belegungen, die Γ_* und damit auch Γ_0 erfüllen.

Somit ist die Voraussetzung des KPS gegeben, und Γ ist daher erfüllbar, was wiederum eine Lösung des Matchingproblems T bedeutet. □

Man beachte

- wie für ein konkretes Problem maßgeschneiderte atomare Formeln eingeführt wurden, die eine Formalisierung (noch) in der AL erlaubten;
- das geschickte Zusammenspiel von endlichen Formelmengen und endlichen Probleminstanzen.

Varianten des Kompaktheitssatzes

Korollar (wird auch oft als KPS bezeichnet).

$\Gamma \models A$ genau dann wenn $\Gamma_0 \models A$ für eine endliche Teilmenge $\Gamma_0 \subseteq \Gamma$.

Beweis.

$\Gamma \models A$ gdw. $\Gamma \cup \{\neg A\}$ unerfüllbar

gdw. $\Gamma \cup \{\neg A\}$ hat eine endliche unerfüllbare Teilmenge Γ_1

gdw. Γ hat eine endliche Teilmenge Γ_0 mit $\Gamma_0 \cup \{\neg A\}$ unerfüllbar

gdw. $\Gamma_0 \models A$ für eine endliche Teilmenge $\Gamma_0 \subseteq \Gamma$ □

Korollar (im Wesentlichen KPS für $\neg\Gamma$).

Für eine Formelmenge Γ sind folgende Aussagen äquivalent:

(a) Jede Belegung erfüllt mindestens eine Formel $B \in \Gamma$.

(b) Es gibt $n \in \mathbb{N}$ und $\boldsymbol{B} \in \Gamma^n$ mit $B_0 \vee \cdots \vee B_{n-1}$ allgemeingültig. □

Beweis der Hinlänglichkeit für den KPS

Nachfolgend sei p_i, $i \in \mathbb{N}$, eine feste Aufzählung von $\mathcal{A}$.

Definition

Eine Formelmenge Γ ist erfüllbar mit Präfix $\boldsymbol{b} \in \mathbb{B}^*$, sofern eine erfüllende Belegung $\varphi \in \mathbb{B}^{\mathcal{A}}$ mit $\varphi(p_i) = b_i$ für alle $i < |\boldsymbol{b}|$ existiert.

Proposition

Wenn jede endliche Teilmenge von Γ mit Präfix $\boldsymbol{b} \in \mathbb{B}^*$ erfüllbar ist, dann auch mit Präfix $\boldsymbol{b}0$ oder $\boldsymbol{b}1$.

Beweis.

Annahme: Es gibt endlichen Teilmengen $\Gamma_i \subseteq \Gamma$, die nicht mit Präfix $\boldsymbol{b}\,i$ erfüllbar sind, $i < 2$.

Dann ist $\Gamma_0 \cup \Gamma_1$ nicht mit Präfix $\boldsymbol{b}$ erfüllbar, Widerspruch. □

Beweis (KPS, Hinlänglichkeit).

$\mathcal{A}$ sei abzählbar[a] unendlich und Γ endlich erfüllbar.

Behauptung: Es existiert eine Folge $b_i \in \mathbb{B}$, $i \in \mathbb{N}$, so dass Γ für jedes $n \in \mathbb{N}$ endlich mit Präfix $\langle b_i : i < n \rangle$ erfüllbar ist.

Die Konstruktion von b_i erfolgt induktiv durch Anwendung obiger Proposition, wobei der Induktionsanfang aufgrund der endlichen Erfüllbarkeit von Γ funktioniert (Präfix ε).

Behauptung: Die Belegung β mit $\beta(p_i) = b_i$ erfüllt Γ.

Jede Formel $A \in \Gamma$ hat einen maximalen Atom-Index m und wird folglich mit Präfix $\langle b_i : i < m \rangle$ erfüllt. □

[a] Im überabzählbaren Fall braucht man eine transfinite Konstruktion für β.

Die kanonische Quasiordnung auf $\mathcal{F}[\mathcal{A}]$

Definition (vergl. letztes Beispiel auf ▸ Folie 17)

Für jede Belegung $\varphi \in \mathbb{B}^{\mathcal{A}}$ induziert $\mathcal{F}[\mathcal{A}] \xrightarrow{\hat{\varphi}} \mathbb{B}$ eine Quasiordnung $\sqsubseteq_\varphi$ auf $\mathcal{F}[\mathcal{A}]$; $\sqsubseteq$ [a] bezeichnet deren Durchschnitt (nicht anti-symm., HA).

[a] Meist wird leider $\models$ verwendet, ein ohnehin viel zu überladenes Symbol.

Lemma

Die folgenden Aussagen sind äquivalent für $A, B \in \mathcal{F}[\mathcal{A}]$*:*

1. $B \sqsubseteq A$
2. $\{B\} \models A$
3. $\hat{\varphi}(B \to A) = 1$ *für alle* $\varphi \in \mathbb{B}^{\mathcal{A}}$

Daher kann man $\sqsubseteq$ als Externalisierung des Junktors $\to$ auffassen.
Achtung: zwar gilt $\{p, q\} \models p \wedge q$, aber weder $p \sqsubseteq p \wedge q$ noch $q \sqsubseteq p \wedge q$.

Die Monotonie der logischen Operationen auf $\mathcal{F}[\mathcal{A}]$

Satz

Die logischen Operationen sind monoton als Abbildungen

$$\mathcal{F}[\mathcal{A}]^{\mathrm{op}} \xrightarrow{\neg} \mathcal{F}[\mathcal{A}] \ , \ \mathcal{F}[\mathcal{A}] \times \mathcal{F}[\mathcal{A}] \xrightarrow{\star} \mathcal{F}[\mathcal{A}] \ , \ \mathcal{F}[\mathcal{A}]^{\mathrm{op}} \times \mathcal{F}[\mathcal{A}] \xrightarrow{\rightarrow} \mathcal{F}[\mathcal{A}]$$

(mit $\star \in \{\wedge, \vee\}$ aber $\star \neq \ \leftrightarrow$) zwischen quasigeordneten Mengen.

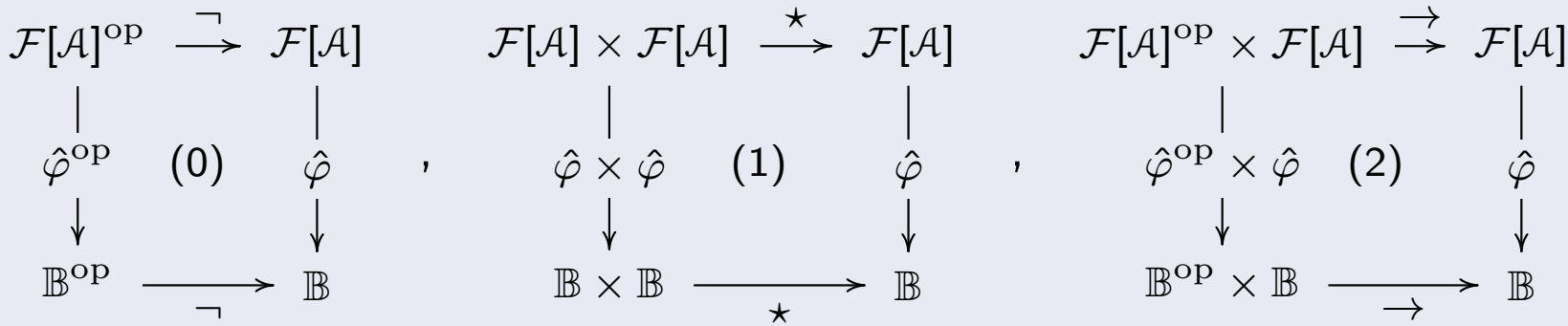
Beweis

Die kanonischen Interpretationen der Junktoren aus $\mathcal{J}$

$$\mathbb{B}^{\mathrm{op}} \xrightarrow{\neg} \mathbb{B} \ , \ \mathbb{B} \times \mathbb{B} \xrightarrow{\star} \mathbb{B} \ , \ \mathbb{B}^{\mathrm{op}} \times \mathbb{B} \xrightarrow{\rightarrow} \mathbb{B}$$

sind offenbar monoton bzgl. $\leq$. Für jedes $\varphi \in \mathbb{B}^{\mathcal{A}}$ ist $\mathcal{F}[\mathcal{A}] \xrightarrow{\hat{\varphi}} \mathbb{B}$ per Definition nicht nur monoton bzgl. $\sqsubseteq_{\varphi}$ und damit bzgl. des Durchschnitts $\sqsubseteq$ aller $\sqsubseteq_{\psi}$, $\psi \in \mathbb{B}^{\mathcal{A}}$, sondern sogar ein $\mathcal{J}$-Homomorphismus, d.h.

Fortsetzung



Wäre eine der log. Op'n auf $\mathcal{F}[\mathcal{A}]$ nicht monoton, so existierten

(0) $A \sqsubseteq B \in \mathcal{F}[\mathcal{A}]$ und $\varphi_0 \in \mathbb{B}^{\mathcal{A}}$ mit $\hat{\varphi}_0(B) < \hat{\varphi}_0(A)$

(1) $A \sqsubseteq B, C \sqsubseteq D \in \mathcal{F}[\mathcal{A}]$ und $\varphi_1 \in \mathbb{B}^{\mathcal{A}}$ mit $\hat{\varphi}_1(B \star D) < \hat{\varphi}_1(A \star C)$

(2) $A \sqsubseteq B, C \sqsubseteq D \in \mathcal{F}[\mathcal{A}]$ und $\varphi_2 \in \mathbb{B}^{\mathcal{A}}$ mit $\hat{\varphi}_2(A \to D) < \hat{\varphi}_2(B \to C)$

im Widerspruch zu der Tatsache, dass die andere Komposition via $\mathbb{B}^{\mathrm{op}}$, $\mathbb{B} \times \mathbb{B}$ bzw. $\mathbb{B}^{\mathrm{op}} \times \mathbb{B}$ nach Konstruktion monoton ist. Folglich sind die logischen Operationen auf $\mathcal{F}[\mathcal{A}]$ ebenfalls monoton. □

Im Folgenden wichtiger als die zugegebenermaßen recht seltsame Quasiordnung $\sqsubseteq$ auf $\mathcal{F}[\mathcal{A}]$ ist die daraus resultierende Äquivalenzrelation:

Die kanonische Kongruenz auf $\mathcal{F}[\mathcal{A}]$

Satz

$\models\!\!\!\models := \sqsubseteq \cap \sqsubseteq^{\mathrm{op}}$ *ist eine Kongruenz:* $A \models\!\!\!\models B$ *und* $C \models\!\!\!\models D$ *impliziert*

$$\neg A \models\!\!\!\models \neg B \quad \textit{und} \quad A \star C \models\!\!\!\models B \star D \ \textit{für} \ \star \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$$

Beweis.

- $A \models\!\!\!\models B$ gdw. $A \sqsubseteq B$ und $B \sqsubseteq A$
  gdw. $\neg B \sqsubseteq \neg A$ und $\neg A \sqsubseteq \neg B$
  gdw. $\neg A \models\!\!\!\models \neg B$
- $A \models\!\!\!\models B$ und $C \models\!\!\!\models D$ gdw. $A \sqsubseteq B$, $B \sqsubseteq A$, $C \sqsubseteq D$, $D \sqsubseteq C$
  gdw. $A \star C \sqsubseteq B \star D$ und $B \star D \sqsubseteq A \star C$
  gdw. $A \star C \models\!\!\!\models B \star D$ □

Satz (HA)

Die Menge $\mathcal{F}[\mathcal{A}]/\models\!\!\!\models$ *der Äquivalenzklassen ist kanonisch eine* $\mathcal{J}$*-Algebra.*

Künftig interessieren Formeln nur noch bis auf Äquivalenz, d.h., modulo $\models\!\!\!\models$.

Folgende Rechenregeln reflektieren das Verhalten der kanonischen Interpretation der Junktoren in $\mathbb{B}$, siehe Folien 18, 19.
Man kann sie leicht z.B. mittels Wahrheitstabellen nachweisen:

Satz (HA)

- $\neg$ *ist selbstinvers, d.h.,* $\neg\neg A \models\!\!\!\models A$.
- $\wedge$ *und* $\vee$ *sind*
  - *idempotent, d.h.,* $A \wedge A \models\!\!\!\models A \models\!\!\!\models A \vee A$.
  - *assoziativ, d.h.,* $(A \star B) \star C \models\!\!\!\models A \star B \star C$ *für* $\star \in \{\wedge, \vee\}$.
  - *kommutativ, d.h.,* $A \star B \models\!\!\!\models B \star A$ *für* $\star \in \{\wedge, \vee\}$.
  - *mit neutralem Element* $\top$ *bzw.* $\bot$*, d.h.,* $A \wedge \top \models\!\!\!\models A \models\!\!\!\models A \vee \bot$.
- $\bot$ *und* $\top$ *absorbieren bzgl.* $\wedge$ *bzw.* $\vee$*, d.h.,* $A \wedge \bot \models\!\!\!\models \bot$ *und* $A \vee \top \models\!\!\!\models \top$.
- *es gelten die Absorbtionsregeln*

$$A \wedge (A \vee B) \models\!\!\!\models A \quad \text{und} \quad A \vee (A \wedge B) \models\!\!\!\models A$$ □

Notationelle Konvention + weitere Rechenregeln

Die Assoziativität von $\wedge$ und $\vee$ rechtfertigt folgende Schreibweise:

Notation

$$\bigwedge_{i<0} A_i := \top \ , \quad \bigwedge_{i<n+1} A_i := \Big(\bigwedge_{i<n} A_i\Big) \wedge A_n \quad \text{, und dual mit } \bigvee \text{ für } \vee$$

Satz (HA)

1. *Es gelten die De Morgan'schen Regeln:*

$$\neg \bigwedge_{i<n} A_i \dashv\vdash \bigvee_{i<n} \neg A_i \qquad \textit{sowie} \qquad \neg \bigvee_{i<n} A_i \dashv\vdash \bigwedge_{i<n} \neg A_i$$

2. *Es gelten die Distributivgesetze:*

$$A \wedge \bigvee_{i<n} B_i \dashv\vdash \bigvee_{i<n} (A \wedge B_i) \qquad \textit{sowie} \qquad A \vee \bigwedge_{i<n} B_i \dashv\vdash \bigwedge_{i<n} (A \vee B_i)$$

Funktional vollständige Junktorenmengen

Offenbar gibt es 2^{2^n} n-stellige Boole'sche Funktionen.

Wie die De Morganschen Regeln zeigen, sind die durch $\mathcal{J}$ spezifizierten Funktionen redundant: neben $\vee$ können z.B. auch $\rightarrow$ und $\leftrightarrow$ eliminiert werden, ohne die Ausdrucksfähigkeit zu mindern.

Definition

Eine Menge $\mathcal{I}$ von Junktorsymbolen vorgegebener Semantik (= Wahrheitstabelle) heißt funktional vollständig, falls jede Boole'sche Funktion $\mathbb{B}^n \longrightarrow \mathbb{B}$ als $\boldsymbol{e}_A$ für geeignetes $A \in \boldsymbol{Term}(\mathcal{I}, \mathcal{A})$ auftritt.

Satz (HA)

$\mathcal{J}$ ist funktional vollständig. □

Ist $\mathcal{I}$ funktional vollständig, so auch jede Junktor-Menge $\mathcal{I}'$, deren Junktoren alle $\mathcal{I}$-Junktoren simulieren können.

Post'scher Vollständigkeitssatz (ohne Beweis)

Für $\mathcal{A}_0 \subseteq \mathcal{A}$ endlich heißt eine Formel $A \in \mathcal{F}[\mathcal{A}_0]$

- konstant, falls $\hat{\varphi}(A)$ für beide konstanten Belegungen $\varphi \in \mathbb{B}^{\mathcal{A}_0}$ denselben Wert wie φ annimmt;
- monoton, falls $\hat{\varphi}(A) \leq \hat{\psi}(A)$ für alle Belegungen $\varphi \leq \psi \in \mathbb{B}^{\mathcal{A}_0}$
- affin, falls für jedes Atom $p \in \mathcal{A}_0$ die Werte $\hat{\varphi}(A)$ und $\hat{\varphi}(A[p/\neg p])$
  - entweder für jedes $\varphi \in \mathbb{B}^{\mathcal{A}_0}$ komplementär sind;
  - oder für jedes $\varphi \in \mathbb{B}^{\mathcal{A}_0}$ übereinstimmen (p heißt Dummy-Variable);
- selbst-dual, falls $\hat{\varphi}(A) = 1 - \hat{\varphi}(A[p/\neg p \,:\, p \in \mathcal{A}_0])$ für alle $\varphi \in \mathbb{B}^{\mathcal{A}_0}$.

Satz (Post'scher Vollständigkeitssatz)

Eine Junktormenge $\mathcal{K}$ ist vollständig, wenn ***Term***$(\mathcal{K}, \mathcal{A})$ *je eine Formel enthält, die eine der obigen Eigenschaften nicht hat.* □

Insbesondere sind $\{\neg, \wedge\}$, $\{\neg, \vee\}$ und $\{\neg, \rightarrow\}$ vollständig (HA).

Kapitel 4
Deduktion allgemein

Wozu Deduktion?

Bisher konnten wir die Erfüllbarkeit einer Formel oder einer Formelmenge nur mit semantischen Mitteln überprüfen. Wegen $\Gamma \models A$ gdw. $\Gamma \cup \{\neg A\}$ unerfüllbar schließt das den Nachweis korrekter Schlussfolgerungen mit ein.

Mit wachsender Atom-Zahl nimmt der Aufwand, Wahrheitstabellen zu erstellen, exponentiell zu, stößt also bald an Grenzen.

Nach hinreichend vielen Beispielen wird man feststellen, dass sich bestimmte korrekte Schlüsse allein anhand der Syntax ihrer Prämissen und der Schlussfolgerung erkennen lassen, und kann so den Aufwand einer semantischen Überprüfung vermeiden.

Auf diese Weise erhält man ein „Kalkül“ zur Herleitung von Formeln, auch ▸ deduktives System genannt. Je nach Vorliebe des Autors oder intendierter Anwendung sind viele verschiedene Varianten möglich.

Idealerweise sollte hier die Herleitungsrelation $\vdash$ eines deduktiven Systems mit der logische Folgerelation $\models$ übereinstimmen.

Schlussregeln und Theoreme

Definition

Ein deduktives System $\mathcal{K} = \langle \mathcal{F}, \mathcal{R} \rangle$ ($\mathcal{K}$ für Kalkül) besteht aus

- einer Menge $\mathcal{F}$ von sog. Formeln;
- einer Menge $\mathcal{R} \subseteq \mathcal{F}^* \times \mathcal{F}$ sog. Schlussregeln, oder kurz Regeln.

Alternative Schreibweise für Regeln (mit Prämissen A_i und Konklusion B):

$$\frac{A_0, A_1, \ldots, A_{k-1}}{B} \qquad \text{anstatt} \qquad \langle A_0, A_1, \ldots, A_{k-1}; B \rangle \in \mathcal{R}$$

Die Menge $\boldsymbol{Ax}(\mathcal{K}) \subseteq \mathcal{F}$ der Axiome besteht aus den Konklusionen von Regeln ohne Prämissen, also mit $k = 0$. Ihr Abschluss unter Anwendung von Regeln liefert die Menge $\boldsymbol{Thm}(\mathcal{K}) \subseteq \mathcal{F}$ der Theoreme.

Regeln werden oft in Schemata zusammengefasst, die Formel-Variablen verwenden und zwecks leichterer Referenz mit kurzen Namen versehen sind.

Erste Beispiele

Beispiel (Ein deduktives System für partielle Syntax)

Mit der eingeschränkten Junktormenge $\mathcal{J}_0 = \{\neg, \to\}$ betrachten wir das Alphabet $\mathcal{J}_0^{()}[\mathcal{A}]$. Das deduktive System $\mathcal{K}_{\text{syn}} = \langle \mathcal{F}_{\text{syn}}, \mathcal{R}_{\text{syn}} \rangle$ verwendet als Formelmenge ganz $\left(\mathcal{J}_0^{()}[\mathcal{A}]\right)^*$ und als Regelmenge

$$\begin{aligned}\mathcal{R}_{\text{syn}} := \mathcal{A} &\cup \{ \langle A, \neg A \rangle \,:\, A \in \mathcal{J}_0^{()}[\mathcal{A}]^* \} \\ &\cup \{ \langle A, B; (A \to B) \rangle \,:\, A, B \in \mathcal{J}_0^{()}[\mathcal{A}]^* \}\end{aligned}$$

oder übersichtlicher in Form von Regelschemata

$$\frac{}{v}\ (\nu) \quad (v \in \mathcal{A}) \quad , \quad \frac{A}{\neg A}\ (\neg) \quad \text{und} \quad \frac{A \quad B}{(A \to B)}\ (\to)$$

Die Atome (aus $\mathcal{A}$) liefern die Axiome, während die Theoreme die aus dem Syntax-Kapitel bekannten Formeln in den Junktoren $\neg$ und $\to$ in Infix-Notation sind.

Beispiel (Ein deduktives System $\mathcal{K}_{\mathrm{Ar}} = \langle \mathcal{F}_{\mathrm{Ar}}, \mathcal{R}_{\mathrm{Ar}} \rangle$ für die Arithmetik)

Setze $\mathcal{F}_{\mathrm{Ar}} := \mathbb{Q}$, die Menge der rationalen Zahlen, und

$$\mathcal{R}_{\mathrm{Ar}}: \qquad \frac{}{1}\ (1) \quad , \quad \frac{x \quad y}{xy}\ (\times) \quad \text{und} \quad \frac{x \quad y}{x-y}\ (-)$$

Hierbei durchlaufen x und y alle rationalen Zahlen. Man mache sich klar, dass die Theoreme genau die ganzen Zahlen sind. Z.B. -2 erhält man mit

$$\dfrac{\dfrac{\dfrac{\dfrac{}{1}\,(1) \quad \dfrac{}{1}\,(1)}{0}\,(-) \quad \dfrac{}{1}\,(1)}{-1}\,(-) \quad \dfrac{}{1}\,(1)}{-2}\,(-)$$

Man kann diese Kombination von Regeln als gerichteten „partiellen Baum“ verstehen, mit Regeln als Knoten („Bruchstrich“ + Namen) und rationalen Zahlen als Kantenlabel. Knoten haben $k \in \mathbb{N}$ Kanten als Eingabe und genau eine als Ausgabe. Alternativ (und besser) versteht man die Regeln als Kanten eines sog. „Multigraphen“ mit rationalen Zahlen als Knoten.

Ab- und Herleitungen (auch Beweise genannt)

Definition

Eine Ableitung in $\mathcal{K} = \langle \mathcal{F}, \mathcal{R} \rangle$ von $A \in \mathcal{F}$ aus $\Gamma \subseteq \mathcal{F}$ ist ein nichtleeres Wort $\langle A_i : i \leq n \rangle \in \mathcal{F}^*$ mit $A_n = A$, so dass für jedes $k \leq n$ gilt

- ▷ entweder $A_k \in \Gamma$, oder es existiert $\mathbb{N} \ni j \xrightarrow{f} k$ (sog. Begründung) mit $\langle\langle A_{f(i)} : i < j \rangle; A_k \rangle \in \mathcal{R}$.

$A \in \mathcal{F}$ heißt aus Γ ableitbar, $\Gamma \vdash_{\mathcal{K}} A$, wenn es eine solche Ableitung gibt. Falls $\Gamma = \emptyset$ spricht man von einer Herleitung oder einem Beweis von A.

Beispiel (Keine HA!)

Eine mögliche Herleitung von -42 im obigen deduktiven System $\mathcal{K}_{\mathrm{Ar}}$ ist

$$\langle 1, 0, -1, -2, 4, -6, 36, -42 \rangle$$

Der entsprechende Baum, der -42 im Regelabschluss des Axioms 1 verortet, hat 71 Knoten und der Teilbaum für -2 tritt neunmal darin auf.

Satz

Eine Formel ist genau dann ein Theorem, wenn sie herleitbar ist.

Beweis.

Die Hinlänglichkeit folgt mittels Induktion über die Länge der Herleitung, die Notwendigkeit mittels Induktion über den Aufbau von Theoremen. □

Obiges Beispiel zeigt, wie viel kürzer Ableitungen sein können, verglichen mit der expliziten Angabe des Regel-Abschlusses ausgehend von bestimmten Axiomen. Sie enthalten viel weniger Redundanz.

Definition

Unter einer expliziten Ableitung versteht man eine vertikale Auflistung der Folgenglieder mit ihrem Index und ihrer Begründung, samt Namen der verwendeten Regel, bzw. der Angabe "Ann.", falls es sich um ein Element von Γ handelt.

Beispiel

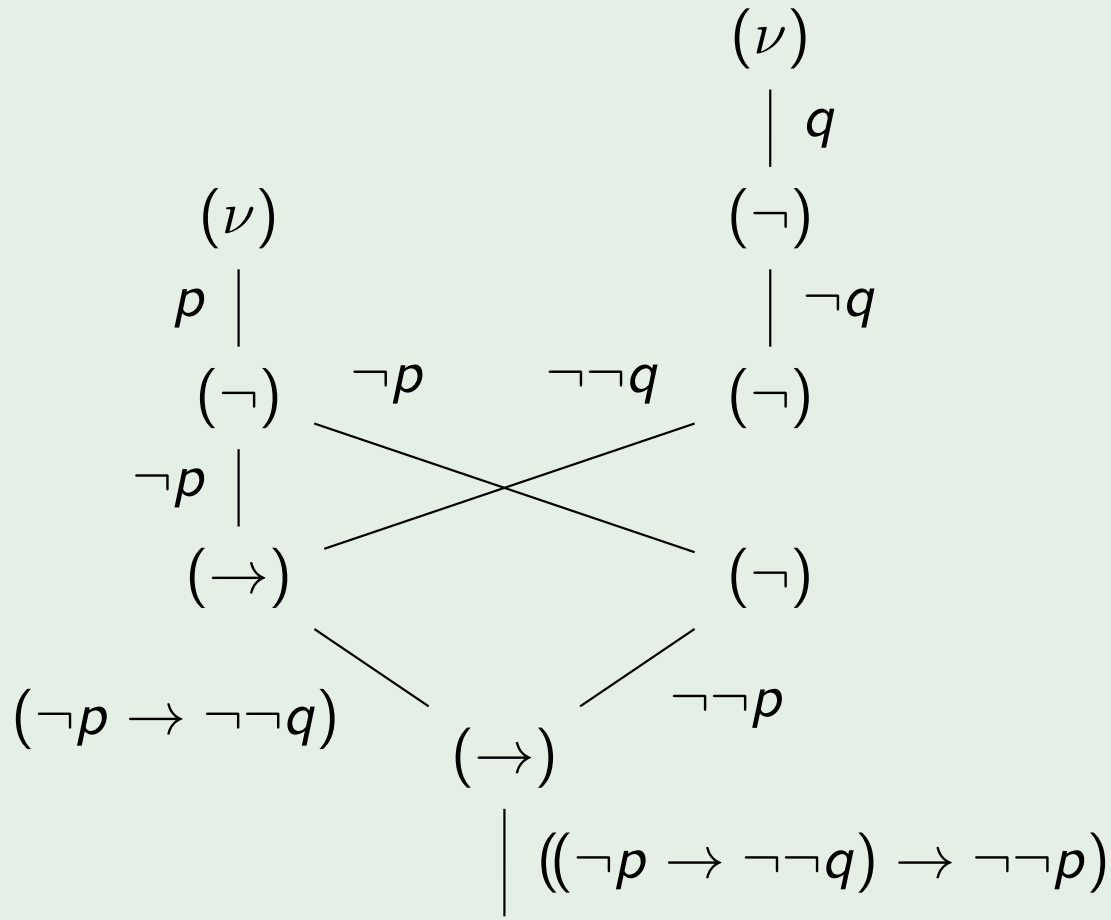
Explizite Herleitung von $((\neg p \to \neg\neg q) \to \neg\neg p)$ in $\mathcal{K}_{\mathrm{syn}}$:

0.	p	(ν)
1.	$\neg p$	$(\neg), 0$
2.	q	(ν)
3.	$\neg q$	$(\neg), 2$
4.	$\neg\neg q$	$(\neg), 3$
5.	$(\neg p \to \neg\neg q)$	$(\to), 1, 4$
6.	$\neg\neg p$	$(\neg), 1$
7.	$((\neg p \to \neg\neg q) \to \neg\neg p)$	$(\to), 5, 6$

Man beachte, dass die Reihenfolge der Schritte nicht vollständig festgelegt ist. So könnte etwa die doppelte Negation von p schon in Schritt 3 erfolgen, statt erst in Schritt 6.

Dies lässt sich graphisch aufbereiten mit Regelnamen als Knoten und Formeln als Ein- und Ausgaben; die Schrittnummern können dann entfallen.

Beispiel (Herleitungsgraph; Richtung von oben nach unten)



Ein Baum entsteht nur, wenn jede Formel außer der Wurzel genau einmal in einer Begründung auftritt. Sonst sind Knoten aufzuspalten:

Beispiel (Herleitungsbaum, „Bruchstriche“ als Knoten, oder Multikanten)

$$\dfrac{\dfrac{\dfrac{\dfrac{}{p}\,(\nu)}{\neg p}\,(\neg) \qquad \dfrac{\dfrac{\dfrac{}{q}\,(\nu)}{\neg q}\,(\neg)}{\neg\neg q}\,(\neg)}{(\neg p \to \neg\neg p)}\,(\to) \qquad \dfrac{\dfrac{\dfrac{}{p}\,(\nu)}{\neg p}\,(\neg)}{\neg\neg p}\,(\neg)}{((\neg p \to \neg\neg q) \to \neg\neg p)}\,(\to)$$

Ein Ableitungsbaum lässt die Struktur der Ableitung deutlicher erkennen und ist vertikal kompakter, aber zum Preis von deutlich mehr Knoten, wie das Beispiel der Herleitung von -42 in $\mathcal{K}_{\mathrm{Ar}}$ oben zeigt.

Satz (HA)

Ein deduktives System $\mathcal{K} = \langle \mathcal{F}, \mathcal{R} \rangle$ induziert einen Hüllenoperator auf $\boldsymbol{P}(\mathcal{F})$

$$\Gamma^{\vdash_{\mathcal{K}}} := \{ A \in \mathcal{F} : \Gamma \vdash_{\mathcal{K}} A \} \quad \textit{für} \quad \Gamma \subseteq \mathcal{F}$$

Dessen Fixpunkte heißen deduktiv abgeschlossene Mengen.

Der Beweis des folgenden Satzes ist viel einfacher als der Beweis seines semantischen Gegenstücks:

Satz (Kompkatheitssatz, syntaktisch (HA))

Eine Formel $A \in \mathcal{F}$ ist aus $\Gamma \subseteq \mathcal{F}$ genau dann ableitbar, wenn sie aus einer endlichen Teilmenge $\Gamma_0 \subseteq \Gamma$ ableitbar ist. □

Kapitel 5
Hilbert-Kalkül

Das Deduktive System $\mathcal{K}_0$

Wir wollen ein deduktives System $\mathcal{K}_0 = \langle \mathcal{F}_0, \mathcal{R}_0 \rangle$ einführen, für das die Relation $\vdash_{\mathcal{K}_0}$ mit $\models$ übereinstimmt. Vereinfachend setzen wir $\vdash := \vdash_{\mathcal{K}_0}$.

$\mathcal{F}_0$ besteht aus den Infix-Formeln in $\mathcal{A}$ über der vollständigen Junktorenmenge $\mathcal{J}_0 = \{\neg, \to\}$, während $\mathcal{R}_0$ durch vier Schemata gegeben ist (beachte die Klammervereinfachungen!):

- $$\frac{A \quad A \to B}{B}\ (\mathrm{MP})$$ Abtrennungsregel oder Modus Ponens
- $$\frac{}{B \to A \to B}\ (\mathrm{Ax1})$$
- $$\frac{}{(A \to B \to C) \to (A \to B) \to A \to C}\ (\mathrm{Ax2})$$
- $$\frac{}{(\neg B \to \neg A) \to A \to B}\ (\mathrm{Ax3})$$

Lemma

Für jede Formel $A \in \mathcal{F}_0$ gilt:

$$\vdash A \to A \qquad \text{(Th1)}$$

Beweis.

Wir geben eine explizite Herleitung:

0.	$A \to A \to A$	Ax1
1.	$A \to (A \to A) \to A$	Ax1
2.	$(A \to (A \to A) \to A) \to (A \to A \to A) \to A \to A$	Ax2
3.	$(A \to A \to A) \to A \to A$	MP, 1,2
4.	$A \to A$	MP, 0,3

□

Achtung: Die Klammerersparnisregeln für $\to$ können die Lesbarkeit zunächst beeinträchtigen. Dagegen hilft nur Routine.

Das syntaktische Deduktionstheorem

Der Nutzen von $\mathcal{K}_0$ beruht wesentlich auf dem

Satz (Herbrand, Tarski, veröffentlicht ca. 1930)

$\Gamma \cup \{A\} \vdash B$ *genau dann wenn* $\Gamma \vdash A \to B$.

Beweis.

($\Leftarrow$) Aus $\Gamma \cup \{A\}$ lassen sich A und $A \to B$ ableiten, also auch B .

($\Rightarrow$) Wende (MP) an auf B und das Axiom $B \to A \to B$. □

Hybride Ableitungen in $\mathcal{K}_0$

Das Deduktionstheorem erlaubt es, Ableitungen rekursiv zu verschachteln:

Definition (rekursiv!)

Eine hybride Ableitung von $A \in \mathcal{F}_0$ aus $\Gamma \subseteq \mathcal{F}_0$ ist eine Liste $\langle A_i : i \leq n \rangle$ aus Formeln und hybriden Ableitungen endlicher Schachtelungstiefe mit $A_n = A$, die folgende Bedingungen erfüllt: für $k \leq n$ ist

(0) entweder $A_k \in \Gamma$;

(1) oder es existiert $\mathbb{N} \ni j \xrightarrow{f} k$ mit $\langle \langle A_{f(i)} : i < j \rangle; A_k \rangle \in \mathcal{R}_0$;

(2) oder $A_k = B \to C$ und es existiert $j < k$ so dass A_j eine hybride Ableitung von C aus $\Gamma + \{B\}$ ist.

(3) oder A_k ist eine hybride Ableitung von C aus einer Prämissenmenge der Form $\Gamma + \{B\}$, und $A_j = B \to C$ für ein $k < j \leq n$.

Bedingungen (2) und (3) entsprechen Anwendungen von (DT), und (3) vermeidet unbenutzte hybride Ableitungen als Komponenten.

Definition

Eine explizite hybride Ableitung ist eine durchnummerierte vertikale Auflistung der Formeln einer entfalteten hybriden Ableitung mit Begründungen in Form des Namens der relevanten Regel bzw. "DT" und Angabe der relevanten Zeilennummern.

Wir markieren die Formeln einer echten expliziten hybriden Unter-Ableitung links mit einer eckigen Klammer. Sie muss mit einer neuen Prämisse B als Annahme beginnen, und wenn sie mit einer Formel C endet, sollte die Zeile $B \to C$ aufgrund des Deduktionstheorems unmittelbar folgen.

Achtung: Die Begründung einer Formel A_j aufgrund einer Regel in $\mathcal{R}_0$ darf nur Zeilen auf derselben Klammer-Ebene wie A_j oder darüber referenzieren; tiefere Klammer-Ebenen sind unzugänglich (out-of-scope).

Strategie: Prämissen mittels geeigneter Annahmen soweit wie möglich zerlegen, Axiome oder schon bewiesene Tautologie-Schemata (s.u.) geschickt(!) einsetzen, um die gewünschte Schlußfolgerung zu erhalten.

Sieben weitere Tautologie-Schemata

Lemma

$$\vdash \neg\neg A \to A \qquad \text{(Th2)}$$

Beweis (hybrid!).

0.	$\neg\neg A$	Ann.
1.	$\neg\neg A \to \neg\neg\neg\neg A \to \neg\neg A$	Ax1
2.	$\neg\neg\neg\neg A \to \neg\neg A$	MP, 0,1
3.	$(\neg\neg\neg\neg A \to \neg\neg A) \to \neg A \to \neg\neg\neg A$	Ax3
4.	$\neg A \to \neg\neg\neg A$	MP, 2,3
5.	$(\neg A \to \neg\neg\neg A) \to \neg\neg A \to A$	Ax3
6.	$\neg\neg A \to A$	MP, 4,5
7.	A	MP, 0,6
8.	$\neg\neg A \to A$	DT, 0–7

□

Lemma

(a) $\vdash \neg B \to B \to A$ (Th3)

(b) $\vdash B \to \neg\neg B$ (Th4)

Beweis

(a)

0.	$\neg B$	Ann.
1.	$\neg B \to \neg A \to \neg B$	Ax1
2.	$\neg A \to \neg B$	MP, 0,1
3.	$(\neg A \to \neg B) \to B \to A$	Ax3
4.	$B \to A$	MP, 2,3
5.	$\neg B \to B \to A$	DT, 0–4

(b)

0.	$\neg\neg\neg B \to \neg B$	Th2
1.	$(\neg\neg\neg B \to \neg B) \to B \to \neg\neg B$	Ax3
2.	$B \to \neg\neg B$	MP, 0,1

□

Lemma

$$\vdash (A \to B) \to \neg B \to \neg A \qquad \text{(Th5)}$$

Beweis.

0.	$A \to B$	Ann.
1.	$\neg\neg A$	Ann.
2.	$\neg\neg A \to A$	Th2
3.	A	MP, 1,2
4.	B	MP, 3,0
5.	$B \to \neg\neg B$	Th4
6.	$\neg\neg B$	MP, 4,5
7.	$\neg\neg A \to \neg\neg B$	DT, 2–7
8.	$(\neg\neg A \to \neg\neg B) \to \neg B \to \neg A$	Ax3
9.	$\neg B \to \neg A$	MP, 7,8
10.	$(A \to B) \to \neg B \to \neg A$	DT, 0–9

□

Lemma

$$\vdash A \to \neg B \to \neg(A \to B) \qquad \text{(Th6)}$$

Beweis.

0.	A	Ann.
1.	$A \to B$	Ann.
2.	B	MP, 0,1
3.	$(A \to B) \to B$	DT, 2–3
4.	$((A \to B) \to B) \to \neg B \to \neg(A \to B)$	Th5
5.	$\neg B \to \neg(A \to B)$	MP, 3,4
6.	$A \to \neg B \to \neg(A \to B)$	DT, 0–5

□

Lemma

$$\vdash (A \to B) \to (A \to \neg B) \to \neg A \qquad \text{(Th7)}$$

Beweis von Th7.

0.	$A \to B$	Ann.
1.	$A \to \neg B$	Ann.
2.	A	Ann.
3.	B	MP, 2,0
4.	$\neg B$	MP, 2,1
5.	$\neg B \to B \to X$	Th3
6.	$B \to X$	MP, 4,5
7.	$X = \neg Y$	MP, 3,7
8.	$A \to X$	DT, 2–7
9.	$(A \to X) \to (\neg X \to \neg A)$	Th5
10.	$\neg X \to \neg A$	MP, 8,9
11.	Y z.B. $A \to A$	irgendeine Tautologie z.B. Th1
12.	$Y \to \neg\neg Y$	Th4
n-4.	$\neg X = \neg\neg Y$	MP, 11,12
n-3.	$\neg A$	MP (n-4),10
n-2.	$(A \to \neg B) \to \neg A$	DT, 1–(n-3)
n-1.	$(A \to B) \to (A \to \neg B) \to \neg A$	DT, 0–(n-2)

Lemma

$$\vdash (B \to A) \to (\neg B \to A) \to A \qquad \text{(Th8)}$$

Beweis.

0.	$B \to A$	Ann.
1.	$(B \to A) \to \neg A \to \neg B$	Th5
2.	$\neg A \to \neg B$	MP, 0,1
3.	$\neg B \to A$	Ann.
4.	$(\neg B \to A) \to \neg\, A \to \neg\neg B$	Th5
5.	$\neg\, A \to \neg\neg B$	MP, 3,4
6.	$(\neg A \to \neg B) \to (\neg A \to \neg\neg B) \to \neg\neg A$	Th7
7.	$(\neg A \to \neg\neg B) \to \neg\neg A$	MP, 2,6
8.	$\neg\neg A$	MP, 5,7
9.	$\neg\neg A \to A$	Th2
10.	A	MP, 8,9
11.	$(\neg B \to A) \to A$	DT, 3–10
12.	$(B \to A) \to (\neg B \to A) \to A$	DT, 0–11

□

Zusammenfassung

Künftig dürfen wir neben den ursprünglichen Regeln für $\mathcal{K}_0$ (MP, Ax1, Ax2, Ax3) und dem syntaktischen Deduktionstheorem (DT) auch die Tautologie-Schemata (Th1) – (Th8) in Ableitungen verwenden:

- (Th1) $\vdash A \to A$
- (Th2) $\vdash \neg\neg A \to A$
- (Th3) $\vdash \neg B \to B \to A$
- (Th4) $\vdash B \to \neg\neg B$
- (Th5) $\vdash (A \to B) \to (\neg B \to \neg A)$
- (Th6) $\vdash A \to \neg B \to \neg(A \to B)$
- (Th7) $\vdash (A \to B) \to (A \to \neg B) \to \neg A$
- (Th8) $\vdash (B \to A) \to (\neg B \to A) \to A$

Definition

Eine Menge $\Gamma \subseteq \mathcal{F}_0$ heißt widersprüchlich oder inkonsistent, wenn eine Formel $B \in \mathcal{F}_0$ mit $\Gamma \vdash B$ und $\Gamma \vdash \neg B$ existiert.

In Analogie des Korollars zum semantischen Deduktionstheorem gilt nun

Korollar (Inkonsistenzregel)

$\Gamma \cup \{A\}$ inkonsistent gdw. $\Gamma \vdash \neg A$.

Beweis.

($\Rightarrow$): Wähle B mit $\Gamma \cup \{A\} \vdash B$ und $\Gamma \cup \{A\} \vdash \neg B$, nach (DT) also $\Gamma \vdash A \to B$ und $\Gamma \vdash A \to \neg B$. Mit (Th7) folgt dann $\Gamma \vdash \neg A$.

($\Leftarrow$): $\Gamma \vdash \neg A$ impliziert $\Gamma \cup \{A\} \vdash \neg A$, während $\Gamma \cup \{A\} \vdash A$ trivial ist. □

Ziel: $\mathcal{K}_0$ ist korrekt, d.h., $\vdash \subseteq \models$, und vollständig, d.h., $\vdash \supseteq \models$; also stimmt $\vdash$ mit $\models$ überein. Zunächst betrachten wir den Fall $\Gamma = \emptyset$.

0-Korrektheit und 0-Vollständigkeit von $\mathcal{K}_0$, Vorarbeiten

Satz

Eine Formel ist genau dann Theorem von $\mathcal{K}_0$, wenn sie eine Tautologie ist.

Die 0-Korrektheit ist leicht. Die 0-Vollständigkeit erfordert etwas Vorarbeit.

Definition

Für $\varphi \in \mathbb{B}^{\mathcal{A}}$ und $A \in \mathcal{F}[\mathcal{A}]$ setze

$$ {}^{\varphi}A := \begin{cases} A & \text{falls } \hat{\varphi}(A) = 1 \\ \neg A & \text{sonst} \end{cases} \qquad \text{sowie} \qquad {}^{\varphi}\Gamma := \{ {}^{\varphi}A \,:\, A \in \Gamma \} $$

Lemma

Für eine Teilmenge $\mathcal{A}_0$ von Atomen, eine Formel $A \in \mathcal{F}_0[\mathcal{A}_0]$ (in den Junktoren $\neg$ und $\rightarrow$) und eine Belegung $\varphi \in \mathbb{B}^{\mathcal{A}}$ gilt ${}^{\varphi}\mathcal{A}_0 \vdash {}^{\varphi}A$.

Beweis. (Strukturelle Induktion über A)

- $A = p \in \mathcal{A}_0$ impliziert ${}^{\varphi}A \in {}^{\varphi}\mathcal{A}_0$.
- $A = \neg B$ und ${}^{\varphi}\mathcal{A}_0 \vdash {}^{\varphi}B$: Dann gilt ${}^{\varphi}(\neg B) = \neg B$ gdw. $\hat{\varphi}(B) = 0$.
  Z.z.: $\neg B \vdash \neg B$ bzw. $B \vdash \neg\neg B$. Wende (DT) auf (Th1) bzw. (Th4) an.
- $A = B \to C$ und ${}^{\varphi}\mathcal{A}_0 \vdash {}^{\varphi}B$ sowie ${}^{\varphi}\mathcal{A}_0 \vdash {}^{\varphi}C$: Dann gilt
  ${}^{\varphi}(B \to C) = B \to C$ gdw. $\hat{\varphi}(B) = 0$ oder $\hat{\varphi}(C) = 1$. Z.z.:
  - $\neg B \vdash B \to C$: wende (DT) auf (Th3) an;
  - $C \vdash B \to C$: wende (DT) auf (Ax1) an;
  - $\{B, \neg C\} \vdash \neg(B \to C)$: wende (DT) zweimal auf (Th6) an. □

Lemma

Mit $\Gamma \cup \{B\} \vdash A$ *und* $\Gamma \cup \{\neg B\} \vdash A$ *gilt auch* $\Gamma \vdash A$.

Beweis.

(DT) liefert $\Gamma \vdash B \to A$ sowie $\Gamma \vdash \neg B \to A$. Aber $\{B \to A, \neg B \to A\} \vdash A$ folgt nach zweimaliger Anwendung von (MP) aus (Th8). □

Beweis der 0-Vollständigkeit von $\mathcal{K}_0$

Beweis.

A sei eine Tautologie, in der n Atome auftreten, etwa $\{ p_i : i < n \}$.
Nach dem ersten Lemma gilt für jede Belegung φ

$$ {}^{\varphi}\{ p_i : i < n \} \vdash {}^{\varphi}A = A \qquad (\star) $$

Wähle nun φ beliebig und setze

$$ \varphi_k(p_i) = \begin{cases} \varphi(p_i) & \text{falls } i < k \\ 1 - \varphi(p_i) & \text{falls } k \leq i \end{cases} \quad \text{für} \quad k \leq n $$

All diese $n+1$ Belegungen erfüllen $(\star)$. Da sich ${}^{\varphi_k}\{ p_i : i < n \}$ und ${}^{\varphi_{k+1}}\{ p_i : i < n \}$ an genau bei p_k unterscheiden, folgt nach dem zweiten Lemma ${}^{\varphi}\{ p_i : i < k \} \vdash A$ für alle $k < n$, speziell $k = 0$. □

Korrektheit und Vollständigkeit von $\mathcal{K}_0$

Satz

A ist genau dann in $\mathcal{K}_0$ aus Γ ableitbar, wenn A logisch aus Γ folgt.

Beweis.

Im Folgenden sei $\Gamma_0 = \{ B_i : i < n \}$ eine endliche Teilmenge von Γ.

$$
\begin{aligned}
\Gamma \vdash A \quad & \text{gdw.} \quad \text{es gibt } \Gamma_0 \subseteq \Gamma \text{ mit } \Gamma_0 \vdash A \\
& \text{gdw.} \quad \vdash B_0 \to B_1 \to \ldots \to B_{n-1} \to A \\
& \text{gdw.} \quad \models B_0 \to B_1 \to \ldots \to B_{n-1} \to A \\
& \text{gdw.} \quad \text{es gibt } \Gamma_0 \subseteq \Gamma \text{ mit } \Gamma_0 \models A \\
& \text{gdw.} \quad \Gamma \models A
\end{aligned}
$$

□

Hier wurden auf der semantischen und der syntaktischen Seite jeweils die entsprechenden Kompaktheitssätze und Deduktionstheoreme angewendet, und „in der Mitte“ die obige prämissenfreie Variante des Satzes.

Bemerkungen

- Der wesentliche Punkt des obigen Beweises ist die Verfügbarkeit des syntaktischen Deduktionstheorems (DT) für den Kalkül $\mathcal{K}_0$.
- Für andere Kalküle $\mathcal{K}$ mit dem Junktor $\rightarrow$ stellt sich die Frage, ob sie ein Deduktionstheorem $\Gamma \vdash_{\mathcal{K}} A \rightarrow B$ gdw. $\Gamma \cup \{A\} \vdash_{\mathcal{K}} B$ erfüllen.
- Man kann zeigen, dass dann die Schemata (MP), (Ax1) und (Ax2) herleitbar sind. Das Schema (Ax3), das nicht im Beweis des (DT) für $\mathcal{K}_0$ verwendet wurde, kann hingegen auch durch andere Schemata ersetzt werden (z.B. Varianten des Modus Tollens $\frac{A \rightarrow B \quad \neg B}{\neg A}$ (MT)).
- Deduktive Systeme der Logik, die wie $\mathcal{K}_0$ Axiome aber möglichst wenige Schlussregeln verwenden, werden Hilbert[2]-Kalküle genannt. $\mathcal{K}_0$ selbst geht auf Łukasiewicz[3] zurück, und stellt eine Vereinfachung des bahnbrechenden Systems von Frege[4] dar.

[2] David Hilbert (1862–1943)

[3] Jan Łukasiewicz (1878–1956), auch Erfinder der sog. „polnischen Notation“

[4] Gottlob Frege (1848–1925)

Nachteile des Hilbert-Kalküls

Die axiomatische Methode des Hilbert-Kalküls $\mathcal{K}_0$, mit nur einer Schlussregeln und beschränkter Junktorenmenge, hat einige Nachteile:

- Das Verfahren ist unhandlich; erst in Verbindung mit dem externen(!) Deduktionstheorem scheint effizientes Arbeiten möglich zu sein.
- Die herausgehobene Bedeutung des Junktors $\rightarrow$ und die Auswahl der Axiomschemata gelten nicht universell als Vorteile: was etwa zeichnet (Ax2) aus, wenn $A \rightarrow A$ aufwändig hergeleitet werden muss?
- Modus Ponens als einzige Schlussregeln erschwert die mechanische, etwa Computer-gestützte, Suche nach Beweisen, bei der man von der gewünschten Konklusion rückwärts auf geeignete Axiome hinarbeitet, denn Modus Ponens reduziert die Komplexität von Formeln.
- Zumindest subjektiv spiegeln Beweise in $\mathcal{K}_0$ die tatsächliche Arbeitsweise von Mathematikern nur sehr eingeschränkt wider; diese untersucht häufig die Konsequenzen bestimmter Annahmen (benötigt (DT)) und verwendet uneingeschränkt Konjunktion wie Disjunktion.

Kapitel 6
Natürliche Deduktion und Sequenzen-Kalkül (optional)

Geschichte und Überblick

1934 erschienen unabhängig voneinander zwei Vorschläge für ein deduktives System $\mathcal{K}_{\mathrm{nat}}$, das die tatsächliche mathematische Arbeitsweise besser widerspiegeln sollte: die Arbeiten von Jaśkowski[5] (mit Vorarbeiten seit mindestens 1929) und Gentzen[6] (basierend auf seiner Dissertation von 1933) begründeten den Kalkül des sog. ▸ natürlichen Schließens; zu Hintergrund und Einordnung siehe auch ▸ Pelletier-Hazen. Auch für Computer-unterstützte Beweissysteme ist $\mathcal{K}_{\mathrm{nat}}$ (wie auch der noch zu besprechende Sequenzen-Kalkül Gentzens) dem Hilbert-Kalkül $\mathcal{K}_0$ deutlich überlegen.

Während Gentzen für Ableitungen zur Baumnotation griff, nutzte Jaśkowski bereits eine vertikale Auflistung mit Blockstruktur (Kästen statt linker Klammern), verwendete später aber die Label zur Buchführung.

---

[5] Stanisław Jaśkowski (1906–1965): On the rules of suppositions in formal logic, Studia Logica 1, 5–32, 1934

[6] Gerhard Gentzen (1909–1945): Untersuchungen über das logische Schließen. I, Mathematische Zeitschrift, 39(2), 176–210, 1934

Grundlagen

- $\mathcal{F}_{\mathrm{nat}}$ enthält alle Formeln in $\mathcal{A}$ über $\mathcal{J}_{\mathrm{nat}} = \{\bot, \neg, \wedge, \vee, \to\}$.
- $\mathcal{K}_{\mathrm{nat}}$ hat keine Axiome, $\mathcal{R}^{-}_{\mathrm{nat}}$ enthält 11 Schemata, und zwar Einführungs- und Eliminationsregeln für $\neg$, $\wedge$, $\vee$, und $\to$, sowie eine Eliminationsregel für $\bot$. Dies deckt bereits die konstruktivistische oder intuitionistische Logik ab. Hinzufügen von (DNE) "double negation elimination", oder (LEM) ("law of excluded middle"), oder (PBC) ("proof by contradiction") liefert $\mathcal{K}_{\mathrm{nat}}$ für klassische Logik. Wir schreiben $\Vdash$ statt $\vdash_{\mathrm{nat}}$ (zur Unterscheidung von $\vdash$).
- **Achtung**: In Ableitungen kann jeder Einführungsregel eine entsprechende Eliminationsregel folgen, sofern deren Hauptteile (unten rot markiert) übereinstimmen (nicht zielführende Macke des Kalküls).
- Drei der Regeln entfalten eine Fernwirkung: $(\to i)$ simuliert das Deduktionstheorem in den hybriden Ableitungen des Hilbert-Kalküls direkt, und $(\neg i)$ entspricht dem Korollar zum ▸ (sDT). Dagegen hat $(\vee e)$ kein Vorbild in $\mathcal{R}_0$ und bedarf einer separaten Überlegung.

Die Regeln für Konjunktion und Implikation

	Introduktion	Elimination
$\wedge$	$\dfrac{G \quad H}{G \wedge H}\ (\wedge i)$	$\dfrac{G \wedge H}{G}\ (\wedge e)$ sowie $\dfrac{G \wedge H}{H}\ (\wedge e)$
$\rightarrow$	$\dfrac{\begin{bmatrix} G \\ \vdots \\ H \end{bmatrix}}{G \rightarrow H}\ (\rightarrow i)$	$\dfrac{G \quad G \rightarrow H}{H}\ (\rightarrow e)$

- $(\rightarrow e)$ entspricht Modus Ponens (MP) aus dem Hilbert-Kalkül $\mathcal{K}_0$.
- $(\rightarrow i)$ ist ein neuer Typ von Regel, die es ermöglichen soll, den Effekt von DT in hybriden $\mathcal{K}_0$-Ableitungen in $\mathcal{K}_{\mathrm{nat}}$ direkt zu realisieren. Sie erlaubt, eine Teilableitung von H aus der um G vergrößerten aktuellen Prämissenmenge zugunsten von $G \rightarrow H$ unzugänglich zu machen.

Die Regeln für Disjunktion, Negation und Absurdität

	Introduktion	Elimination
$\vee$	$\dfrac{G}{G \vee H}\,(\vee i)$ sowie $\dfrac{H}{G \vee H}\,(\vee i)$	$\dfrac{G \vee H \qquad \begin{bmatrix} G \\ \vdots \\ K \end{bmatrix} \qquad \begin{bmatrix} H \\ \vdots \\ K \end{bmatrix}}{K}\,(\vee e)$
$\neg$	$\dfrac{\begin{bmatrix} G \\ \vdots \\ \bot \end{bmatrix}}{\neg G}\,(\neg i)$	$\dfrac{G \qquad \neg G}{\bot}\,(\neg e)$
$\bot$	$\dfrac{}{\neg\bot}\,(\bot i)$ herleitbar! (HA)	$\dfrac{\bot}{G}\,(\bot e)$
$\neg\neg$	$\dfrac{G}{\neg\neg G}\,(\text{DNI})$ herleitbar! (HA)	$\dfrac{\neg\neg G}{G}\,(\text{DNE})$

- $(\neg\, i)$ kann angesichts $(\to i)$ ohne Fernwirkung ersetzt werden durch

$$\frac{G \to \bot}{\neg G}$$

  was Ableitungen an dieser Stelle um eine Zeile verlängert.
- Die Regel $(\vee\, e)$ ist am kompliziertesten anzuwenden: kann $G \vee H$ aus den aktuellen Prämissen abgeleitet werden, sind zwei Teilableitungen von K aus der um G bzw. H vergrößerten Prämissenmenge einzuschachteln, zugunsten von K.
- Alle Regeln entsprechen wahren semantischen Sachverhalten bzw. korrekten logischen Folgerungen, damit ist $\mathcal{K}_{\mathrm{nat}}$ korrekt.

Definition

Explizite Herleitungen sind nun vertikale nummerierte Auflistungen von Formeln mit Blockstruktur (linke eckige Klammern), deren Begründungen gemäß der Regeln auf vorangehende Formeln bzw. Blöcke verweisen. Auf einzelne Formeln in Blöcken kann von außerhalb nicht verwiesen werden.

Satz

$\vdash F \vee \neg F$

Beweis

0.	$\neg(F \vee \neg F)$	Ann.
1.	$\quad F$	Ann.
2.	$\quad F \vee \neg F$	$(\vee i)$, 1
3.	$\quad \bot$	$(\neg e)$, 0,2
4.	$\neg F$	$(\neg i)$, 1–3
5.	$F \vee \neg F$	$(\vee i)$, 4
6.	$\bot$	$(\neg e)$, 0,5
7.	$\neg\neg(F \vee \neg F)$	$(\neg i)$, 0–6
8.	$F \vee \neg F$	(DNE) , 7

□

Beispiel (obiger Beweis als Baum)

$$\dfrac{\dfrac{\dfrac{[\neg(F \vee \neg F)] \quad \dfrac{[F]}{F \vee \neg F}\,(\vee i)}{\dfrac{\bot}{\neg F}\,(\neg i)}\,(\neg e)}{F \vee \neg F}\,(\vee i) \qquad [\neg(F \vee \neg F)]}{\dfrac{\dfrac{\bot}{\neg\neg(F \vee \neg F)}\,(\neg i)}{F \vee \neg F}\,(\mathrm{DNE})}\,(\neg e)$$

Der Gültigkeitsbereich (Scope) der jeweiligen geschachtelten Teil-Ableiungen ist nur schwer auszumachen. Daher haben wir deren Anfang und Ende hier farblich markiert.

In dieser VL werden wir die Baum-Darstellung erst bei Gentzens Sequenzen-Kalkül wieder aufgreifen.

Satz

In $\mathcal{K}_{\text{nat}}$ lässt sich (DNE) äquivalent ersetzen durch

$$\frac{}{F \vee \neg F}\ (\text{LEM})$$

Beweis

Es genügt, im neuen Kalkül $\neg\neg F \vdash F$ nachzuweisen.

0.		$\neg\neg F$	Prämisse
1.	⌈	$\neg F$	Ann.
2.	|	$\bot$	$(\neg\, e)$, 0,1
3.	⌊	F	$(\bot\, e)$, 3
4.	⌈	F	Ann.
5.		$F \vee \neg F$	LEM
6.		F	$(\vee\, e)$, 5, 1–3, 4–4

□

Satz

$\Gamma \cup \{\neg G\} \vdash \bot$ *gdw* $\Gamma \vdash \neg\neg G$ *(immer) gdw* $\Gamma \vdash G$ *(klassisch).*

Beweis.

($\Rightarrow$) Falls $\Gamma \vdash \bot$, wende $(\bot e)$ an. Andernfalls gibt es eine Ableitung für $\bot$ aus $\Gamma \cup \{\neg G\}$ mit erster Zeile $\neg G$. Anwendung von $(\neg i)$ liefert $\Gamma \vdash \neg\neg G$, und mit (DNE) folgt $\Gamma \vdash G$.
($\Leftarrow$) Wende $(\neg e)$ auf $\Gamma \cup \{\neg G\} \vdash G$ bzw. $\Gamma \cup \{\neg G\} \vdash \neg\neg G$ an. □

Satz (HA)

In $\mathcal{K}_{\text{nat}}$ *lässt sich* (DNE) *äquivalent ersetzen durch*

$$\frac{\begin{array}{|c} \neg G \\ \vdots \\ \bot \end{array}}{G}\ (\text{PBC})$$

„Normalformen“ und die Teilformel-Eigenschaft

Einer Introduktion die entsprechende Elimination folgen zu lassen, so dass die Hauptteile (rot markiert) übereinstimmen, verlängert einen Beweis unnötig ohne Fortschritt.

Im Kalkül $\mathcal{K}^{-}_{\mathrm{nat}}$ ohne die Regel (DNE) hat das folgenden positiven Effekt:

Satz

In $\mathcal{K}^{-}_{\mathrm{nat}}$ lässt sich jeder Beweis durch Eliminierung der oben genannten Introduktions-Eliminations-Paare in eine „Normalform“ mit der sog. Teilformel-Eigenschaft bringen: jede auftretende Formel ist Teilformel der Konklusion oder einer der Prämissen/Annahmen. □

Man zerlegt einfach alle Prämissen durch Eliminationsregeln, und die gewünschte Konklusion rückwärts durch Introduktionsregeln, bis Übereinstimmung eintritt, oder nicht. Leider lassen sich in der klassischen Logik mit (DNE) (oder (LEM) oder (PBC)) Beweise nicht so einfach finden.

Ausdrucksstärke im Vergleich zum Hilbert-Kalkül

Satz

Die Axiom-Schemata (Ax1), (Ax2) und (Ax3) des Hilbert-Kalküls sind Theoreme in $\mathcal{K}_{\mathrm{nat}}$.

Beweis

(Ax1):

0.	B	Ann.
1.	A	Ann.
2.	$A \wedge B$	$(\wedge\, i)$, 1,0
3.	B	$(\wedge\, e)$, 2
4.	$A \to B$	$(\to i)$, 1–3
5.	$B \to A \to B$	$(\to i)$, 0–4

Beweis (Fortsetzung)

(Ax2):

0.	$A \to B \to C$	Ann.
1.	$\quad A \to B$	Ann.
2.	$\quad\quad A$	Ann.
3.	$\quad\quad B$	$(\to e)$, 2,1
4.	$\quad\quad B \to C$	$(\to e)$, 2,0
5.	$\quad\quad C$	$(\to e)$, 3,4
6.	$\quad A \to C$	$(\to i)$, 2–5
7.	$(A \to B) \to A \to C$	$(\to i)$, 1–6
8.	$(A \to B \to C) \to (A \to B) \to A \to C$	$(\to i)$, 0–7

Beweis (Fortsetzung)

(Ax3):

0.	$\neg B \to \neg A$	Ann.
1.	A	Ann.
2.	$\neg B$	Ann.
3.	$\neg A$	$(\to e)$, 2,0
4.	$\bot$	$(\neg e)$, 1,3
5.	$\neg\neg B$	$(\neg i)$, 2–4
6.	B	(DNE) , 5
7.	$A \to B$	$(\to i)$, 1–6
8.	$(\neg B \to \neg A) \to A \to B$	$(\to i)$, 0–7

□

Corollar

Die Theoreme (Th1)–(Th8) des Hilbert-Kalküls sind Theoreme in $\mathcal{K}_{\mathrm{nat}}$. □

▸ Folie 120

Einige Hilfssätze

Die obigen Theoreme liefern einige nützliche Rechenregeln:

Corollar

- $\neg B \nvdash B \to A$ (tH3)
- $B \nvdash \neg\neg B$ (tH4)
- $A \to B \nvdash \neg B \to \neg A$ (tH5)
- $A \nvdash \neg B \to \neg(A \to B)$ (tH6)
- $A \to B, A \to \neg B \nvdash \neg A$ (tH7)
- $B \to A, \neg B \to A \nvdash A$ (tH8)

(tH4) lässt sich aber schneller direkt zeigen als via $\mathcal{K}_0$, HA!

Künftig werden wir diese abgeleiteten Regeln wie auch (LEM) und (PBC) zur Verkürzung von Beweisen einsetzen.

Lemma (Contraposition)

$F \to G \vdash \neg G \to \neg F$ und $\neg G \to \neg F \vdash F \to G$. (CP)+(PC)

Beweis.

0.	$F \to G$	Prämisse
1.	$\quad \neg G$	Ann.
2.	$\quad\quad F$	Ann.
3.	$\quad\quad G$	$(\to e)$, 2,0
4.	$\quad\quad \bot$	$(\neg e)$, 3,1
5.	$\quad \neg F$	$(\neg i)$, 2–4
6.	$\neg F \to \neg G$	$(\to i)$, 1–5

0.	$\neg G \to \neg F$	Prämisse
1.	$\quad F$	Ann.
2.	$\quad\quad \neg G$	Ann.
3.	$\quad\quad \neg F$	$(\to e)$, 1,0
4.	$\quad\quad \bot$	$(\neg e)$, 3,1
5.	$\quad G$	(PBC) , 2–4
6.	$F \to G$	$(\to i)$, 1–5

□

(CP) ist zwar schon als (tH5) bekannt, aber dieser $\mathcal{K}_{\mathrm{nat}}$ -Beweis ist kürzer als der obige ▸ Beweis in $\mathcal{K}_0$.

Lemma (1. De Morgansche Regel)

$\neg(G \vee H) \vdash \neg G \wedge \neg H$ und $\neg G \wedge \neg H \vdash \neg(G \vee H)$. (DM1)

Beweis.

0.	$\neg(G \vee H)$	Prämisse
1.	$\quad G$	Ann.
2.	$\quad G \vee H$	$(\vee\, i)$, 1
3.	$\quad \bot$	$(\neg\, e)$, 2,0
4.	$\neg G$	$(\neg\, i)$, 1–3
5.	$\quad H$	Ann.
6.	$\quad G \vee H$	$(\vee\, i)$, 5
7.	$\quad \bot$	$(\neg\, e)$, 7,0
8.	$\neg H$	$(\neg\, i)$, 5–7
9.	$\neg G \wedge \neg H$	$(\wedge\, i)$, 4,8

0.	$\neg G \wedge \neg H$	Prämisse
1.	$\quad G \vee H$	Ann.
2.	$\quad\quad G$	Ann.
3.	$\quad\quad \neg G$	$(\wedge\, e)$, 0
4.	$\quad\quad \bot$	$(\neg\, e)$, 2,3
5.	$\quad\quad H$	Ann.
6.	$\quad\quad \neg H$	$(\wedge\, e)$, 0
7.	$\quad\quad \bot$	$(\neg\, e)$, 5,6
8.	$\quad \bot$	$(\vee\, e)$, 1,2–4,5–7
9.	$\neg(G \vee H)$	$(\neg\, i)$, 1–8

□

Lemma (2. De Morgansche Regel)

$\neg G \vee \neg H \vdash \neg(G \wedge H)$ und $\neg(G \wedge H) \vdash \neg G \vee \neg H$. (DM2)

Beweis.

0.	$\neg G \vee \neg H$	Prämisse
1.	$\quad G \wedge H$	Ann.
2.	$\quad G$	$(\wedge e)$, 1
3.	$\quad H$	$(\wedge e)$, 1
4.	$\quad \neg\neg G$	(tH4), 2
5.	$\quad \neg\neg H$	(tH4), 3
6.	$\quad \neg\neg G \wedge \neg\neg H$	$(\wedge i)$, 5,6
7.	$\quad \neg(\neg G \vee \neg H)$	DM1, 7
8.	$\quad \bot$	$(\neg e)$, 0,8
9.	$\neg(G \wedge H)$	$(\neg i)$, 1–9

0.	$\neg(G \wedge H)$	Prämisse
1.	$\quad \neg(\neg G \vee \neg H)$	Ann.
2.	$\quad \neg\neg G \wedge \neg\neg H$	DM1
3.	$\quad \neg\neg G$	$(\wedge e)$, 2
4.	$\quad \neg\neg H$	$(\wedge e)$, 2
5.	$\quad G$	(DNE) , 3
6.	$\quad H$	(DNE) , 4
7.	$\quad G \wedge H$	$(\wedge i)$, 5,6
8.	$\quad \bot$	$(\neg e)$, 7,0
9.	$\neg G \vee \neg H$	(PBC) , 1–8

Lemma (Ersetzung der Junktoren $\vee$ und $\wedge$)

$G \vee H \Vdash \neg G \to H$ und $G \wedge H \Vdash \neg(H \to \neg G)$. $(\vee x) + (\wedge x)$

Beweis.

0.	$G \vee H$	Prämisse
1.	$\quad G$	Ann.
2.	$\quad\quad \neg G$	Ann.
3.	$\quad\quad \bot$	$(\neg\, e)$, 1,2
4.	$\quad\quad H$	$(\bot\, e)$, 3
5.	$\quad \neg G \to H$	$(\to i)$, 2–4
6.	$\quad H$	Ann.
7.	$\quad\quad \neg G$	Ann.
8.	$\quad\quad \neg G \wedge H$	$(\wedge\, i)$, 7,6
9.	$\quad\quad H$	$(\wedge\, e)$, 8
10.	$\quad \neg G \to H$	$(\to i)$, 7–9
11.	$\neg G \to H$	$(\vee\, e)$, 0,1–5,6–10

0.	$G \wedge H$	Prämisse
1.	G	$(\wedge\, e)$, 0
2.	H	$(\wedge\, e)$, 0
3.	$\quad H \to \neg G$	Ann.
4.	$\quad\quad \neg H$	$(\to e)$, 2,3
5.	$\quad\quad \bot$	$(\neg\, e)$, 1,4
6.	$\neg(H \to \neg G)$	$(\neg\, i)$, 3–5

$\square$

Lemma (Rückersetzung der Junktoren $\vee$ und $\wedge$)

$\neg G \to H \vdash G \vee H$ und $\neg(H \to \neg G) \vdash G \wedge H$ $\qquad (\vee z) + (\wedge z)$

Beweis.

0.	$\neg G \to H$	Prämisse
1.	$\neg H \to \neg\neg G$	(CP), 0
2.	$\quad \neg(F \vee G)$	Ann.
3.	$\quad \neg G \wedge \neg H$	DM2, 2
4.	$\quad \neg(\neg H \to \neg\neg G)$	$(\vee x)$, 3
5.	$\quad \bot$	$(\neg e)$, 1,4
6.	$G \vee H$	(PBC) , 2–5

0.	$\neg(H \to \neg G)$	Prämisse
1.	$\quad \neg(G \wedge H)$	Ann.
2.	$\quad \neg G \vee \neg H$	DM1, 1
3.	$\quad \neg\neg G \to \neg H$	$(\vee x)$, 1,0
4.	$\quad H \to \neg G$	(CP), 3
5.	$\quad \bot$	$(\neg e)$, 0,4
6.	$F \wedge G$	(PBC) , 1–5

□

Der Junktor $\bot$ kann durch $\neg(p \to p)$ mit einem frischen Atom p ersetzt werden. Wir sind nun in der Lage, $\mathcal{K}_{\text{nat}}$ mit $\mathcal{K}_0$ in Beziehung zu setzen.

Die Vollständigkeit von $\mathcal{K}_{\text{nat}}$

Satz

Jede Tautologie ist ein $\mathcal{K}_{\text{nat}}$-Theorem, d.h., $\mathcal{K}_{\text{nat}}$ ist 0-vollständig.

Beweis.

Jede Tautologie F bzgl. $\mathcal{J}_{\text{nat}} = \{\neg, \to, \wedge, \vee, \bot\}$ kann mittels $G \vee H \dashv\vdash \neg G \to H$ und $G \wedge H \dashv\vdash \neg(H \to \neg G)$ sowie $\bot \dashv\vdash \neg(p \to p)$ in eine Tautologie F' bzgl. $\mathcal{J}_0 = \{\neg, \to\}$ transformiert werden. Diese ist ein $\mathcal{K}_0$-Theorem, im Hinblick auf die $\mathcal{K}_0$-Axiome also auch ein $\mathcal{K}_{\text{nat}}$-Theorem. Aufgrund von $\neg G \to H \vdash G \vee H$ und $\neg(H \to \neg G) \vdash G \wedge H$ sowie $\neg(p \to p) \vdash \bot$ lassen sich die obigen Transformationsschritte in $\mathcal{K}_{\text{nat}}$ rückgängig machen. Damit ist F ebenfalls ein $\mathcal{K}_{\text{nat}}$-Theorem. □

▸ Folie 124

Auch ein direkter Beweis ohne Rückgriff auf das $\mathcal{K}_0$-Ergebnis ist möglich. Dazu sind die vorbereitenden Lemmata ▸ Lemma und ▸ Lemma anzupassen.

Lemma

Für eine Teilmenge $\mathcal{A}_0$ *von Atomen, eine Formel* $A \in \mathcal{F}_{\mathrm{nat}}[\mathcal{A}_0]$ *über* $\mathcal{J}_{\mathrm{nat}}$ *und eine Belegung* $\varphi \in \mathbb{B}^{\mathcal{A}}$ *gilt* ${}^{\varphi}\mathcal{A}_0 \Vdash {}^{\varphi}A$.

Beweis. (Strukturelle Induktion über A)

- $A = \bot$ trivial wegen $(\bot\, i)$.
- $A = p \in \mathcal{A}_0$ impliziert ${}^{\varphi}A \in {}^{\varphi}\mathcal{A}_0$.
- $A = \neg B$ und ${}^{\varphi}\mathcal{A}_0 \Vdash {}^{\varphi}B$: ${}^{\varphi}(\neg B) = \neg B$ gdw. $\hat{\varphi}(B) = 0$.
  Z.z. $\neg B \Vdash \neg B$ (trivial) und $B \Vdash \neg\neg B$ (umseitig).
- $A = B \star C$ und ${}^{\varphi}\mathcal{A}_0 \Vdash {}^{\varphi}B$ sowie ${}^{\varphi}\mathcal{A}_0 \Vdash {}^{\varphi}C$.
  - ${}^{\varphi}(B \wedge C) = B \wedge C$ gdw. $\hat{\varphi}(B) = 1 = \hat{\varphi}(C)$.
    Z.z. $\{B, C\} \Vdash B \wedge C$ sowie $\neg B \Vdash \neg(B \wedge C)$ und $\neg C \Vdash \neg(B \wedge C)$.
  - ${}^{\varphi}(B \vee C) = B \vee C$ gdw. $\hat{\varphi}(B) = 1 = 1$ oder $\hat{\varphi}(C) = 1$.
    Z.z. $B \Vdash B \vee C$ und $C \Vdash B \vee C$ sowie $\{\neg B, \neg C\} \Vdash \neg(B \vee C)$.
  - ${}^{\varphi}(B \to C) = B \to C$ gdw. $\hat{\varphi}(B) = 0$ oder $\hat{\varphi}(C) = 1$.
    Z.z. $\neg B \Vdash B \to C$ und $C \Vdash B \to C$ sowie $\{B, \neg C\} \Vdash \neg(B \to C)$.

  $\star \in \{\vee, \wedge\}$: trivial bzw. De Morgan; Rest umseitig.

Beweis (Fortsetzung)

0.	B	Prämisse
1.	$\neg B$	Ann.
2.	$\bot$	$(\neg\, e)$, 0,1
3.	$\neg\neg B$	$(\neg\, i)$, 1–2

0.	C	Prämisse
1.	B	Ann.
2.	$B \wedge C$	$(\wedge\, i)$, 1,0
3.	C	$(\wedge\, e)$, 3
4.	$B \to C$	$(\to i)$, 1–3

0.	B	Prämisse
1.	$\neg C$	Prämisse
2.	$B \to C$	Ann.
3.	C	$(\to e)$, 0,2
4.	$\bot$	$(\neg\, e)$, 3,1
5.	$B \to C$	$(\to i)$, 2–4

0.	$\neg B$	Prämisse
1.	B	Ann.
2.	$\bot$	$(\neg\, e)$, 1,0
3.	C	$(\bot\, e)$, 2
4.	$B \to C$	$(\to i)$, 1–3

Lemma

Mit $\Gamma \cup \{B\} \vdash A$ *und* $\Gamma \cup \{\neg B\} \vdash A$ *gilt auch* $\Gamma \vdash A$.

Beweis.

0.	$B \vee \neg B$	LEM
1.	$\quad B$	Ann.
2.	$\quad A$	$(\rightarrow e)$, 3,0
3.	$\quad \neg B$	Ann.
4.	$\quad A$	$(\rightarrow e)$, 5,1
5.	A	$(\vee e)$, 2,3–4,5–6

□

Der ▸ Beweis der 0-Vollständigkeit von $\mathcal{K}_0$ funktioniert nun auch für $\mathcal{K}_{\mathrm{nat}}$.
Mit Hilfe des Kompaktheitssatzes folgt sofort die Vollständigkeit:

Satz

Aus $\Gamma \models A$ *folgt* $\Gamma \vdash A$.

□

Die Korrektheit von $\mathcal{K}_{\mathrm{nat}}$

- In expliziten nicht hybriden $\mathcal{K}_0$-Ableitungen ist aufgrund der Axiome des Hilbert-Kalküls jede Zeile eine Tautologie.
- Aufgrund der eingebauten Möglichkeit zur Schachtelung ist die Situation in $\mathcal{K}_{\mathrm{nat}}$ komplizierter. Dafür braucht der Fall $\Gamma = 0$ nicht gesondert behandelt zu werden.
- Wir wollen eine $\mathcal{K}_{\mathrm{nat}}$-Ableitung für $\Gamma \Vdash G$, Γ endlich, so umschreiben, dass Zeile i die Form $\Gamma_i \Vdash A_i$ erhält, wobei Γ_i die Menge der in Position i aktuellen Prämissen und Annahmen ist.
  - Auf dem äußeren Level gilt immer $\Gamma_i = \Gamma$.
  - Mit Beginn/Ende jeder linken Klammer nimmt die Anzahl der aktuellen Annahmen um eins zu- bzw. ab. Insbesondere bestimmt $|\Gamma_i - \Gamma|$ die Schachtelungstiefe der Zeile i. Die eckigen Klammern entfallen.

Als Beispiel betrachten wir den obigen Beweis für (Ax3):

Beispiel

0.	$\{\neg B \to \neg A\} \vdash \neg B \to \neg A$	Ann.
1.	$\{A, \neg B \to \neg A\} \vdash A$	Ann.
2.	$\{\neg B, A, \neg B \to \neg A\} \vdash \neg B$	Ann.
3.	$\{\neg B, A, \neg B \to \neg A\} \vdash \neg A$	$(\to e)$, 2,0
4.	$\{\neg B, A, \neg B \to \neg A\} \vdash \bot$	$(\neg e)$, 1,3
5.	$\{A, \neg B \to \neg A\} \vdash \neg\neg B$	$(\neg i)$, 2–4
6.	$\{A, \neg B \to \neg A\} \vdash B$	(DNE), 5
7.	$\{\neg B \to \neg A\} \vdash A \to B$	$(\to i)$, 1–6
8.	$\vdash (\neg B \to \neg A) \to A \to B$	$(\to i)$, 0–7

Jede Zeile enthält eine korrekte Aussage über die Ableitbarkeit $\vdash$ in $\mathcal{K}_{\mathrm{nat}}$.
In mit "Ann." markierten Zeilen gehört die rechte Formel zur linken Menge.
Sonst liegt eine logische Regel oder eine frühere Ableitung zugrunde.

Satz

Aus $\Gamma \Vdash G$ *folgt* $\Gamma \models G$.

Beweis

Mit Hilfe der Ergebnisse aus dem Semantik-Kapitel lässt sich leicht zeigen:

- Aus $\Gamma \models G$ und $\Gamma \models H$ folgt $\Gamma \models G \wedge H$.
- Aus $\Gamma \models G \wedge H$ folgt $\Gamma \models G$ und $\Gamma \models H$.
- Aus $\Gamma \cup \{G\} \models H$ folgt $\Gamma \models G \to H$.
- Aus $\Gamma \models G$ und $\Gamma \models G \to H$ folgt $\Gamma \models H$.
- Aus $\Gamma \models G$ oder $\Gamma \models H$ folgt $\Gamma \models G \vee H$.
- Aus $\Gamma \models G \vee H$ und $\Gamma \cup \{G\} \models K$ sowie $\Gamma \cup H \models K$ folgt $\Gamma \models K$.
- Aus $\Gamma \cup \{G\} \models \bot$ folgt $\Gamma \models \neg G$.
- Aus $\Gamma \models G$ und $\Gamma \models \neg G$ folgt $\Gamma \models \bot$

Beweis, Fortsetzung

- Aus $\Gamma \models \bot$ folgt $\Gamma \models G$.
- Aus $\Gamma \models \neg\neg G$ folgt $\Gamma \models G$

Dies sind die semantischen Versionen der Regeln für $\mathcal{K}_{\mathrm{nat}}$.

Für $\Gamma \Vdash G$ betrachte die wie oben umgeschriebene Ableitung.

Induktion über die Anzahl k der Schritte (=Regel-Anwendungen, keine Abkürzungen aufgrund früherer Ergebnisse) in $\mathcal{K}_{\mathrm{nat}}$-Ableitungen zeigt nun, dass wir in jeder Zeile die Relation $\Vdash$ durch $\models$ ersetzen können. □

Dies legt nahe ein Kalkül zu entwickeln, das Ausdrücke der Form $\Gamma \Vdash G$ direkt manipuliert.

Gentzens Sequenzen-Kalkül, Vorbereitung

Gentzens ▸ zweites System war ursprünglich als technisches Hilfmittel zum Beweis der Konsistenz der Prädikatenlogik gedacht, hat sich aber in vielen Bereichen der Logik als äußerst nützlich erwiesen.

Abstraktion des oben beschriebenen Prozesses mutiert das Relationssymbol $\Vdash$ zu einem Trennsymbol der Syntax; wir verwenden nun wieder $\vdash$.

Syntaktische Bausteine sind zunächst(!) Ausdrücke der Form $\Gamma \vdash \varphi$ mit $\Gamma \in (\mathcal{F}_{\mathrm{nat}})^*$, sog. einseitige Sequenzen (engl. sequents, auch judgements genannt). Den Zusammenhang mit den oben verwendeten endlichen Mengen auf der linken Seite vermitteln sog. strukturellen Regeln, informell

- Vertauschung (engl. “Exchange”): Prämissen-Reihenfolge irrelevant;
- Kontraktion (engl. “Contraction”): Prämissen mehrfach verwendbar;
- Verdünnung (engl. “Weakening” oder “Thinning”): Vergrößerung um überflüssige Prämissen zulässig.

Strukturelle Regeln des ND-Sequenzen-Kalküls

Mittels der bekannten Bruch-Notation erhält man nun:

$$\frac{\Gamma, G, F \vdash H}{\Gamma, F, G \vdash H}\ (xch) \quad , \quad \frac{\Gamma, F, F \vdash H}{\Gamma, F \vdash H}\ (ctr) \quad , \quad \frac{\Gamma \vdash H}{\Gamma, F \vdash H}\ (wkg)$$

- (xch) erlaubt es, die aktuellen Prämissen in Multi-Mengen anstelle von (geordneten) Folgen zusammenzufassen;
- (ctr) erlaubt es, die aktuellen Prämissen in Mengen anstelle von Multi-Mengen zusammenzufassen;
- (wkg) erlaubt es, Prämissenmengen zu vergrößern (und damit zu verdünnen). Wie oben gesehen, lassen sich Prämissenmengen von Zeilen verschiedener Schachtelungstiefe bei Bedarf vereinheitlichen.

Übertragung der $\mathcal{K}^{-}_{\mathrm{nat}}$-Regeln eliminiert die Sonderstllung der Regeln mit Fernwirkung. Dafür kommt ein Axiom-Schema hinzu:

Logische Regeln des ND-Sequenzen-Kalküls

$$\frac{}{\Gamma, F, \Delta \vdash F}\ (ax)$$

	Introduktion	Elimination
$\wedge$	$\frac{\Gamma \vdash G \quad \Gamma \vdash H}{\Gamma \vdash G \wedge H}\ (\wedge i)$	$\frac{\Gamma \vdash G \wedge H}{\Gamma \vdash G}\ (\wedge e)$, $\frac{\Gamma \vdash G \wedge H}{\Gamma \vdash H}\ (\wedge e)$
$\rightarrow$	$\frac{\Gamma, F \vdash G}{\Gamma \vdash G \rightarrow H}\ (\rightarrow i)$	$\frac{\Gamma \vdash G \quad \Gamma \vdash G \rightarrow H}{\Gamma \vdash H}\ (\rightarrow e)$
$\vee$	$\frac{\Gamma \vdash G}{\Gamma \vdash G \vee H}\ (\vee i)$, $\frac{\Gamma \vdash H}{\Gamma \vdash G \vee H}\ (\vee i)$	$\frac{\Gamma \vdash G \vee H \quad \Gamma, G \vdash K \quad \Gamma, H \vdash K}{\Gamma \vdash K}\ (\vee e)$
$\neg$	$\frac{\Gamma, F \vdash \bot}{\Gamma \vdash \neg F}\ (\neg i)$	$\frac{\Gamma \vdash F \quad \Gamma \vdash \neg F}{\Gamma \vdash \bot}\ (\neg e)$
$\bot$		$\frac{\Gamma \vdash \bot}{\Gamma \vdash G}\ (\bot e)$

Ein Analogon zur Regel (DNE), die bei der Suche nach Normalformen beliebiger $\mathcal{K}_{\mathrm{nat}}$-Beweise gestört hatte, bzw. zu (LEM) oder (CTR), fehlt, also ist der bisher skizzierte Kalkül intuitionistisch.

Gentzens geniale Idee zur Behandlung klassischer Logik mittels Sequenzen kommt ohne weiteres Regel-Schema des obigen Typs aus. Sein Kalkül $\mathcal{K}_{\mathrm{seq}}$ verwendet zweiseitige Sequenzen der Form $\Gamma \vdash \Delta$, Γ, Δ Formelmengen. Eine solche Sequenz ist korrekt, wenn $\bigwedge \Gamma \to \bigvee \Delta$ allgemeingültig ist.

Die strukturellen Regeln werden zu links- und rechtsseitigen Varianten verdoppelt. Bei den obigen logischen Regeln treten alle Junktoren rechts von $\vdash$ auf. Bei zweiseitigen Sequenzen würde man aber auch Regeln mit Junktoren links von $\vdash$ erwarten.

Die bisherigen Einführungs-Regeln lassen sich leicht in zweiseitige Versionen überführen, nun Rechts-Introduktionen genannt (beachte auch $\bot = \bigvee \emptyset$). Beim Ersetzen der bisherigen Eliminierungen durch Links-Introduktionen beachte man die Dualität zwischen $\wedge$ und $\vee$ und nutze die klassische Ersetzung von $\to$ durch $\neg$ und $\vee$ als Motivation:

Strukturelle Regeln für $\mathcal{K}_{\text{seq}}$

- Exchange:

$$\frac{\Gamma \vdash \Delta_0, A, B, \Delta_1}{\Gamma \vdash \Delta_0, B, A, \Delta_1}\ (\text{XCH-}R) \qquad \frac{\Gamma_0, A, B, \Gamma_1 \vdash \Delta}{\Gamma_0, B, A, \Gamma_1 \vdash \Delta}\ (\text{XCH-}L)$$

- Contraction:

$$\frac{\Gamma \vdash \Delta_0, A, A, \Delta_1}{\Gamma \vdash \Delta_0, A, \Delta_1}\ (\text{CTR-}R) \qquad \frac{\Gamma_0, A, A, \Gamma_1 \vdash \Delta}{\Gamma_0, A, \Gamma_1 \vdash \Delta}\ (\text{CTR-}L)$$

- Weakening:

$$\frac{\Gamma \vdash \Delta_0, \Delta_1}{\Gamma \vdash \Delta_0, A, \Delta_1}\ (\text{WKG-}R) \qquad \frac{\Gamma_0, \Gamma_1 \vdash \Delta}{\Gamma_0, A, \Gamma_1 \vdash \Delta}\ (\text{WKG-}L)$$

Erste logische Regeln für $\mathcal{K}_{\text{seq}}$

$$\frac{}{\Gamma, F \vdash \Delta, F}\ (\text{AX})$$

	R	L
$\wedge$	$\dfrac{\Gamma \vdash \Delta, G \quad \Gamma \vdash \Delta, H}{\Gamma \vdash \Delta, G \wedge H}\ (\wedge R)$	$\dfrac{\Gamma, G, H \vdash \Delta}{\Gamma, G \wedge H \vdash \Delta}\ (\wedge L)$
$\rightarrow$	$\dfrac{\Gamma, G \vdash \Delta, H}{\Gamma \vdash \Delta, G \rightarrow H}\ (\rightarrow R)$	$\dfrac{\Gamma \vdash \Delta, G \quad \Gamma, H \vdash \Delta}{\Gamma, G \rightarrow H \vdash \Delta}\ (\rightarrow L)$
$\vee$	$\dfrac{\Gamma \vdash \Delta, G, H}{\Gamma \vdash \Delta, G \vee H}\ (\vee R)$	$\dfrac{\Gamma, G \vdash \Delta \quad \Gamma, H \vdash \Delta}{\Gamma, G \vee H \vdash \Delta}\ (\vee L)$
$\neg$	$\dfrac{\Gamma, F \vdash \Delta}{\Gamma \vdash \Delta, \neg F}\ (\neg R)$	$\dfrac{\Gamma \vdash \Delta, F}{\Gamma, \neg F \vdash \Delta}\ (\neg L)$

Warum $\mathcal{K}_{\text{seq}}$ die klassische Logik beschreibt

Beispiel (DNE)

0.	$A \vdash A$	Axiom
1.	$\vdash A, \neg A$	$(\neg R)$, 0
2.	$\neg\neg A \vdash A$	$(\neg L)$, 1

Beispiel (LEM)

0.	$A \vdash A$	Axiom
1.	$\vdash A, \neg A$	$(\neg R)$, 0
2.	$\vdash A \vee \neg A$	$(\vee R)$, 1

Was fehlt noch in $\mathcal{K}_{\text{seq}}$?

Analog zu den Regeln aus $\mathcal{R}^{-}_{\text{nat}}$ lassen sich eine oder mehrere Sequenzen zu einer neuen transformieren, wobei die Blockstruktur entfällt.

Es fällt auf, dass die logischen Regeln sämtlich zu komplexeren logischen Formeln im „Nenner“ führen. Damit lassen sich nicht alle Ableitungen von $\mathcal{K}_{\text{nat}}$ nachspielen, etwa die (ziemlich sinnlose) Ableitung:

$\mathcal{K}_{\text{nat}}$	.	$\vdots$	
	n .	G	
$n+1$.		H	
$n+2$.		$G \wedge H$	$(\wedge i)$, n,n+1
$n+3$.		G	$(\wedge e)$, n+2

$\mathcal{K}_{\text{seq}}$	.	$\vdots$	
	n .	$\Gamma \vdash G$	
$n+1$.		$\Gamma \vdash H$	
$n+2$.		$\Gamma \vdash G \wedge H$	$(\wedge R)$, n,n+1
	.	???	
	m .	$G \vdash G$	Axiom
$m+1$.		$G \wedge H \vdash G$	$(\wedge L)$, m
$m+\lvert\Gamma\rvert$.		$\Gamma, G \wedge H \vdash G$	(WKG-L) , $\lvert\Gamma\rvert$ times

Wie will man in $\mathcal{K}_{\text{seq}}$ die Sequenzen $n+2$ und $m+|\Gamma|$ verbinden?

Die Schnittregel

Gentzen behalf sich, indem er $\mathcal{K}_{\text{seq}}$ duch Einführung einer neuen strukturellen Regel, der Schnittregel, zu $\mathcal{K}^{+}_{\text{seq}}$ erweiterte:

$$\frac{\Gamma_0 \vdash \Delta_0, A \qquad A, \Gamma_1 \vdash \Delta_1}{\Gamma_0, \Gamma_1 \vdash \Delta_0, \Delta_1} \ (\text{CUT})$$

Mit ihrer Hilfe lässt sich oben aus den Sequenzen $\Gamma \vdash G \wedge H$ und $G \wedge H \vdash G$ wie gewünscht $\Gamma \vdash G$ erhalten.

Als Gegenstück zur Existenz von $\mathcal{K}^{-}_{\text{nat}}$-Beweisen in Normalform mit der Teilformel-Eigenschaft für den intuitionistischen Teil der Aussagenlogik konnte Gentzen für den zweiseitigen Sequenzen-Kalkül zeigen

Satz (Gentzens Hauptsatz, ohne Beweis)

Zu jedem $\mathcal{K}^{+}_{\text{seq}}$-Beweis $\Gamma \vdash \Delta$ existiert ein (CUT)-freier $\mathcal{K}_{\text{seq}}$-Beweis von Δ aus Γ, und dieser kann algorithmisch aus dem ursprünglichen Beweis konstruiert werden. □

Kapitel 7
Algorithmen für die Aussagenlogik

Übersicht

Zum Nachweis von $\Gamma \models A$ kann man anstelle von $\Gamma \vdash A$ natürlich auch die Nichterfüllbarkeit von $\Gamma \cup \{\neg A\}$ überprüfen. Aufgrund des KPS lässt sich dies selbst bei unendlichem Γ (z.B. bei einer ▸ Herbrand Expansion) in endlich vielen Schritten feststellen, im Gegensatz zur Erfüllbarkeit. Daher ist das Problem der Unerfüllbarkeit semi-entscheidbar.

Wir stellen drei Syntax-basierte Algorithmen vor, mit deren Hilfe die Nichterfüllbarkeit von Formel(menge)n häufig schneller gezeigt werden kann, als mittels Wahrheitstabellen (brute force) oder deduktiver Ableitungen. Im ungünstigsten Fall ist ihre Laufzeit aber immer noch exponentiell. Bei erfüllbaren endlichen Mengen werden erfüllende Belegungen gefunden.

Während semantische Tableaus auf beliebige Formelmengen anwendbar ist, erfordert der Davis-Putnam Algorithmus Formeln in Negations-Normalform (NNF), während die Resolutionsmethode Formeln in konjunktiver Normalform (KNF) benötigt. Insofern werden wir kurz auf diese Normalformen und ihre Konstruktion eingehen.

Unerfüllbarkeit von Γ via Tableaus: grobe Strategie

- Iterative Zerlegung der Formeln in Γ nach bestinnten Kriterien (s.u.);
- Einsortieren der Bestandteile in einen dynamisch wachsenden subbinären Baum (Tableau) mit Formel(menge)n als Knoten;
- Knoten, in denen ein Widerspruch zu einer Formel in einem Vorgängerknoten auftritt, werden samt Folgeknoten von der weiteren Bearbeitung ausgeschlossen. Das liefert sog. geschlossene Äste.
- Γ ist genau dann unerfüllbar, wenn alle Äste geschlossen sind.

Wir klassifizieren Formeln über $\mathcal{J}_1 = \{\neg, \wedge, \vee, \rightarrow\}$ als

- ▷ Atome in $\mathcal{A}$ oder deren Negationen, genannt Literale;
- ▷ doppelte Negationen – überflüssig, werden sofort eliminiert;
- ▷ bis auf Äquivalenz binäre Konjunktionen, genannt α-Formeln;
- ▷ bis auf Äquivalenz binäre Disjunktionen, genannt β-Formeln.

Zerlegungsregeln

Wir notieren die syntaktisch(!) definierten Zerlegungsschritte analog zu den Schlussregeln deduktiver Systeme. Zwei Typen sind zu unterscheiden:

- α-Formeln (und manchmal in der Literatur doppelte Negation):

$$\frac{A \wedge B}{A, B} \quad , \quad \frac{\neg(A \vee B)}{\neg A, \neg B} \quad , \quad \frac{\neg(A \to B)}{A, \neg B} \quad , \quad \left(\frac{\neg\neg A}{A}\right)$$

  Sind die α-Teilchen im „Nenner“ wahr, so auch der „Zähler“.

- β-Formeln:

$$\frac{A \vee B}{A \mid B} \quad , \quad \frac{\neg(A \wedge B)}{\neg A \mid \neg B} \quad , \quad \frac{A \to B}{\neg A \mid B}$$

  Ist ein β-Teilchen im „Nenner“ wahr, so auch der „Zähler“.

In diesem Abschnitt werden wir führende doppelte Negationen immer sofort eliminieren, d.h., wir werden „modulo“ $\neg\neg(-)$ rechnen.

Definition

1. $\mathcal{F}^{\neg\neg}[\mathcal{A}] \subseteq \mathcal{F}[\mathcal{A}]$ besteht aus genau den Formeln, die nicht mit einer doppelten Negation beginnen.
2. $\Gamma \subseteq \mathcal{F}^{\neg\neg}[\mathcal{A}]$ heißt α- bzw. β-gesättigt, wenn mit jeder Formel aus Γ auch beide α- bzw. β-Teilchen modulo $\neg\neg(-)$ zu Γ gehören.
3. Die kleinste α- bzw. β-gesättigte Obermenge von Γ in $\mathcal{F}^{\neg\neg}[\mathcal{A}]$, genannt α- bzw. β-Hülle, bezeichnen wir mit $\Gamma^{<\alpha>}$ bzw. $\Gamma^{<\beta>}$.
4. Für die kleinste sowohl α- als auch β-gesättigte Obermenge von Γ, genannt Teilchen-Hülle, bezeichnen wir mit $\Gamma^{<\alpha,\beta>}$.

Bemerkung

Die Teilchen-Hülle von $G \in \mathcal{F}^{\neg\neg}[\mathcal{A}]$ besteht zwar nicht notwendig aus ▸ Teilformeln von G, aber aus solchen von von $\mathrm{NNF}(G)$, ▸ Folie 154 unten.

Die Tableau-Methode bestimmt die Teilchen-Hülle $\Gamma^{<\alpha\beta>}$ iterativ und arrangiert ihre Elemente so in einer Baumstruktur, dass erfüllende Belegungen bzw. deren Nichtexistenz ablesbar werden.

Vorüberlegung zur praktischen Tableau-Konstruktion

Ein Algorithmus zur Konstruktion eines Tableaus für potentiell unendliches Γ kann zu jedem Zeitpunkt immer nur endlich viele Formeln im Blick haben.

Die Reihenfolge der Abarbeitung wird man nur schwer vorschreiben können, dafür lassen sich bei Bedarf später Strategien formulieren.

Im folgenden Beispiel haben wir die Formeln bis zu zweifach markiert:

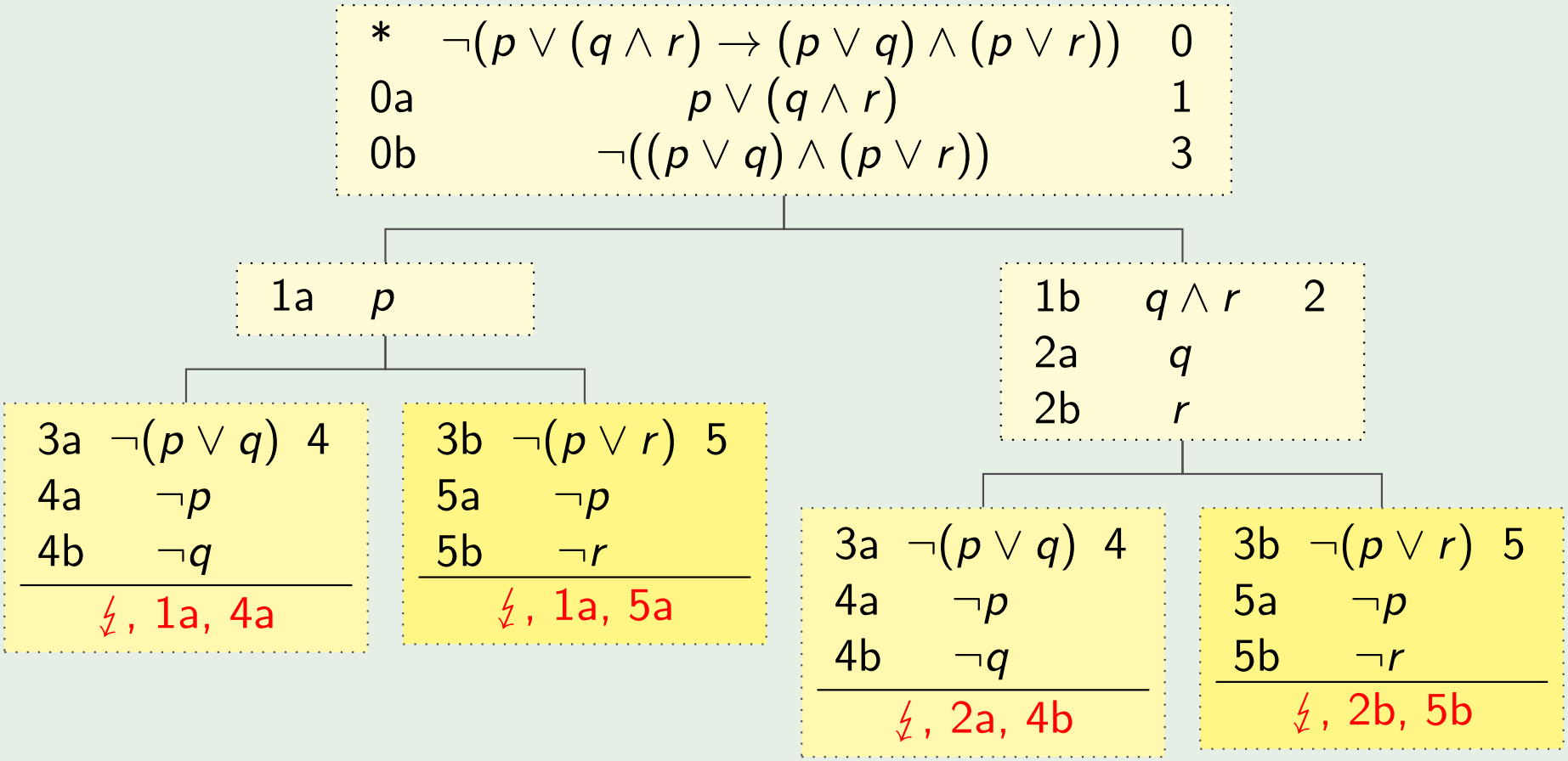
- rechts, sofern es sich nicht um ein Literal handelt, mit der fortlaufenden Nummer k des Arbeitsschritts, in dem sie zerlegt werden;
- links mit einem Verweis auf den Arbeitsschritt ihrer Entstehung;

Weiter wollen wir vereinbaren, dass im Fall gleicher α- oder β-Teilchen die Ursprungsformel direkt durch das einfachere Teilchen ersetzt wird.

Widersprüche suchen wir nur unter den auftretenden Literalen.

Beispiel

$\Gamma = \{\neg(p \vee (q \wedge r) \to (p \vee q) \wedge (p \vee r))\}$ ist nicht erfüllbar:



Ob die α-Teilchen im aktuellen Knoten bleiben oder in spätere Blätter wandern macht hier keinen Unterschied.

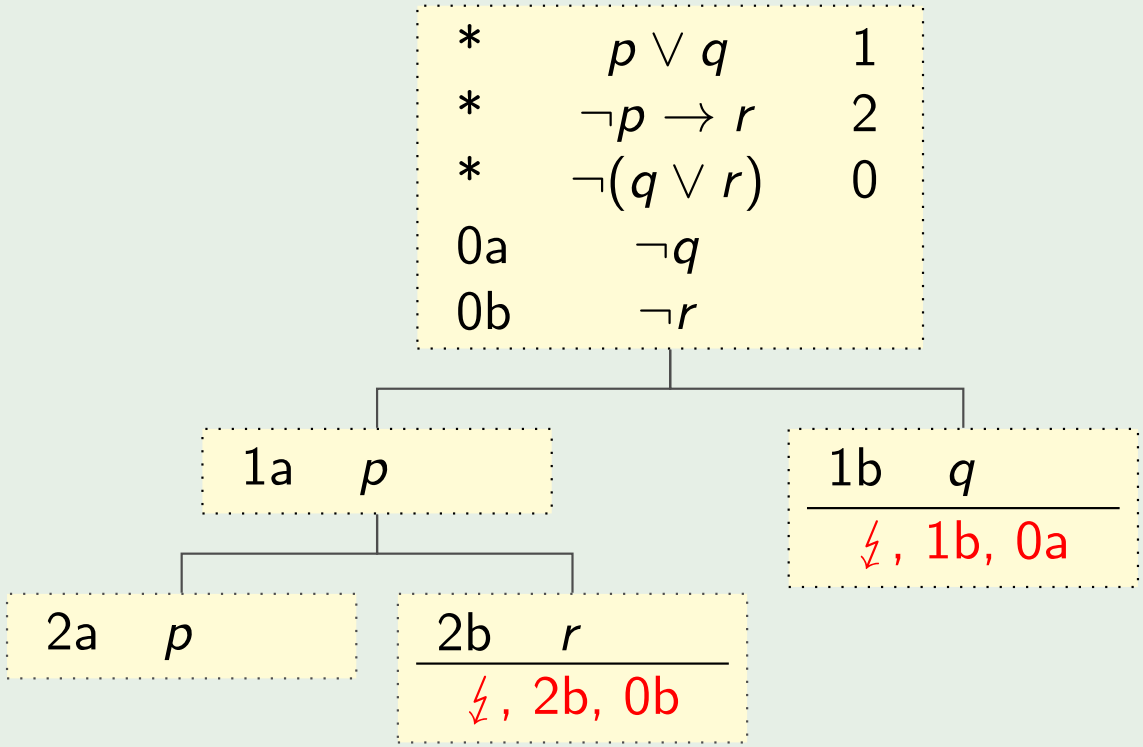
Praktische Konstruktion des Tableaus für $\Gamma \subseteq \mathcal{F}[\mathcal{A}]$

1. Γ bildet zunächst die potentiell unendliche Wurzelmenge;
2. Auswahl einer unverarbeiteten Formel A; ggf. Markierung mit $*$;
3. Ist A in Schritt k kein Literal, wird A regel-konform zerlegt und die Teilchen werden mit ka bzw. kb als „erfaßt“ markiert:
   - α-Teilchen landen je nach Strategie entweder in allen Blättern ab dem aktuellen Knoten, oder direkt im Knoten von A;
   - für β-Formeln werden bei allen Blättern ab dem aktuellen Knoten zwei neue Knoten mit je einem der β-Teilchen eröffnet.
4. Treten innerhalb der markierten(!) Formeln entlang eines Astes widersprüchliche Literale auf, so werden der tiefere der betroffenen Knoten und seine Nachfolger von der weiteren Bearbeitung ausgeschlossen, man erhält einen geschlossenen Ast.

Um evtl. widersprüchliche Literale schnell zu finden, wollen wir jede Formel, deren Bearbeitung begonnen wurde, bis zum Ende zerlegen (depth first).

Beispiel

$\Gamma = \{A_0 = p \vee q, A_1 = \neg p \to r, A_2 = \neg(q \vee r)\}$ ist erfüllbar.

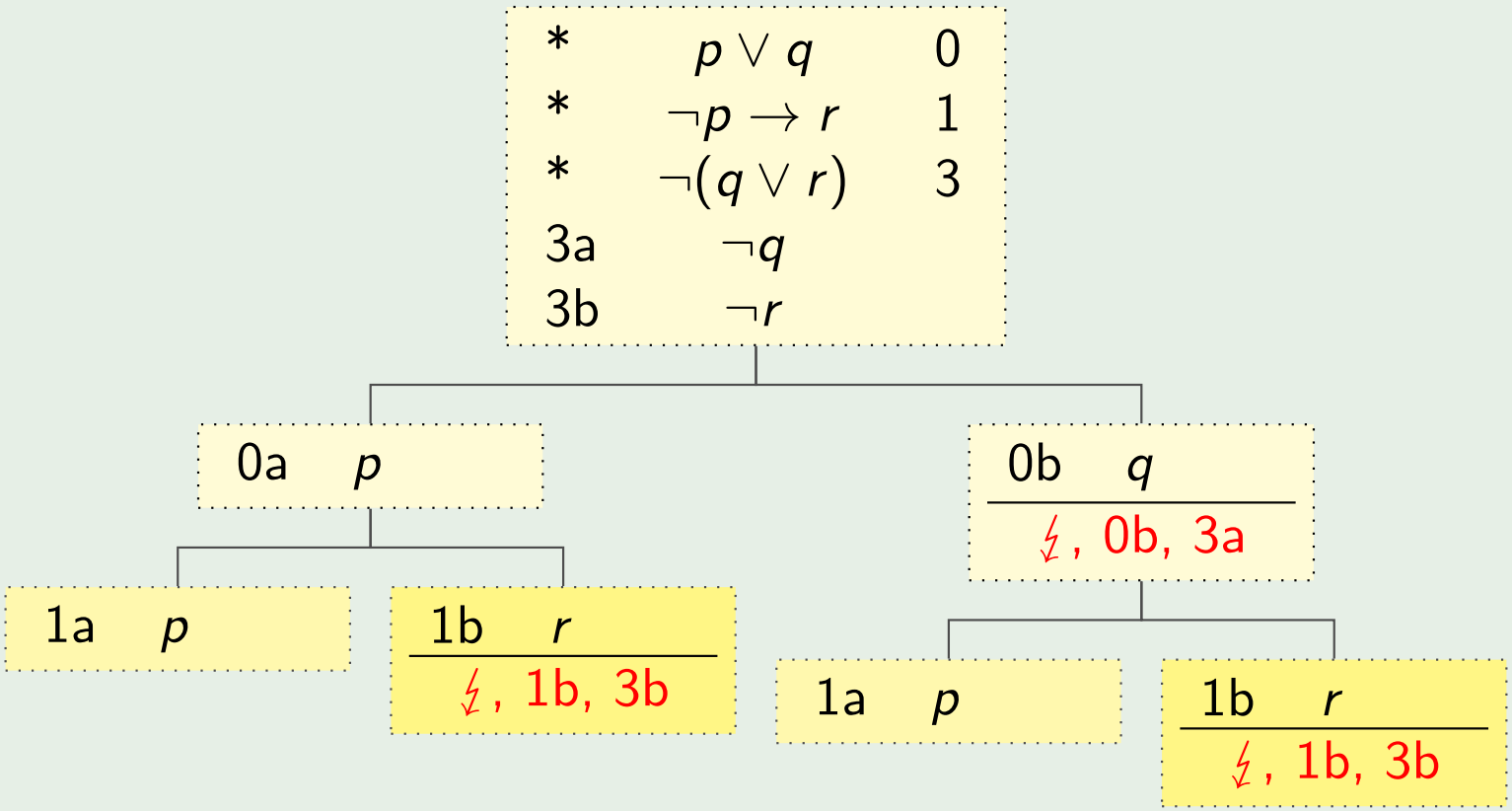


In Position (2a) wurde eine doppelte Negation sofort eliminiert.

Entlang des linken Astes macht die Belegung $p \longmapsto 1$, $q \longmapsto 0$, $r \longmapsto 0$ die Literale in Position (2a und 1a), (0a) bzw. (0b), und damit Γ wahr.

Beispiel

$\Gamma = \{p \vee q, \neg p \to r, \neg(q \vee r)\}$ ist erfüllbar, andere Tableau-Reihenfolge:



Warum funktioniert das obige Verfahren?

Definition

Ein Tableau der AL ist Abbildung $\mathbb{B}^* \xrightarrow{\tau} \boldsymbol{P}(\mathcal{F}^{\neg\neg}[\mathcal{A}])$, so dass

- das Urbild $\mathcal{B}(\tau)$ der nichtleeren Formelmengen ist Präfix-abgeschlossen und erfüllt $w0 \in \mathcal{B}(\tau)$ gdw. $w1 \in \mathcal{B}(\tau)$ für alle $w \in \mathbb{B}^*$.
- $\Gamma \subseteq \tau(\varepsilon) \subseteq \Gamma^{<\alpha>}$.
- für jedes $\varepsilon \neq w \in \mathcal{B}(\tau)$ existiert $B \in \tau(w)$ mit $\tau(w) \subseteq \{B\}^{<\alpha>}$ und B ist β-Teilchen einer Formel aus einem echten Vorgängerknoten.

Ein Ast Θ von $\mathcal{B}(\tau)$ (maximale lineare Präfix-geordnete Teilmenge) heißt

- vollständig, falls $\bigcup \tau[\Theta]$ unter α-Teilchen abgeschlossen ist und mit jeder β-Formel mindestens ein β-Teilchen enthält.
- abgeschlossen, falls $\bigcup \tau[\Theta]$ eine Formel und ihre Negation enthält (es genügt, sich auf Literale zu beschränken); andernfalls heißt Θ offen.

Ein Tableau heißt abgeschlossen, falls jeder Ast abgeschlossen ist.

Die Existenz vollständiger Tableaus

Lemma

Jedes Tableau τ kann vervollständigt werden.

Beweis.

$\mathcal{F}^{\neg\neg}[\mathcal{A}]$ ist abzählbar! Setze $\tau_0 := \tau$ und definiere τ_{k+1} durch Hinzufügen der Teilchen für jedes Nicht-Literal in τ_k, wobei α-Teilchen in ihrem Knoten verbleiben, während β-Teilchen neue Blätter im entsprechenden Zweig eröffnen. Dann gilt $\tau_k(w) \subseteq \tau_{k+1}(w)$ für alle $w \in \mathbb{B}^*$ und $k \in \mathbb{N}$. Vereinigung liefert ein vollständiges Tableau τ_∞:

$$\tau_\infty(w) := \bigcup \{ \tau_k(w) : k \in \mathbb{N} \}$$ □

Offenbar gilt immer $\tau_\infty(\varepsilon) = \Gamma^{<\alpha>}$. Speziell kann man für Γ mit dem Minimaltableau starten, also $\tau(\varepsilon) = \Gamma$ und $\tau(w) = \emptyset$ falls $w \neq \varepsilon$.

Lemma (Hintikka)

Für vollständige Tableau-Astmengen gilt: „erfüllbar“=„offen“.

Beweis.

„$\Rightarrow$“ ist klar. Umgekehrt sei Θ offen. φ möge genau die Atome p mit $\neg p \in \bigcup \tau[\Theta]$ auf 0 abbilden. Induktion über den Aufbau der Formeln in $\bigcup \tau[\Theta]$ (vollständig!) und die Abhängigkeit des Wahrheitswerts einer Formel von denen ihrer Teilchen zeigt $\varphi \in (\bigcup \tau[\Theta])^{\triangleright}$. $\square$

Die ursprüngliche Tableau-Methode (um 1955) markiert Knoten sub-binärer Bäume mit einzelnen Formeln statt Formelmengen. Sie geht auf Beth[7] und (unabhängig) Hintikka[8] zurück. Smullyans[9] Vereinfachung (1968, 1995) machte sie populär. Die maximale linearen Ketten bis zu den Verzweigungsknoten entsprechen im Wesentlichen den hier verwendeten Mengen.

[7] Evert Willem Beth (1908–1964)
[8] Jaakko Hintikka (1929–2015)
[9] Raymond Merrill Smullyan (1919–2017)

Unerfüllbarkeit via Tableaus

Satz

Γ *ist unerfüllbar gdw.* Γ *hat ein abgeschlossenes Tableau.*

Beweis.

Jedes Tableau τ für Γ enthält Γ in allen Astmengen. Unerfüllbarkeit überträgt sich auf diese, und nach Hintikkas Lemma sind sie abgeschlossen.

Umgekehrt folgt aus $\varphi \in \Gamma^{\triangleright}$ per Induktion über die Wörter in $\mathbb{B}^*$ die Existenz eines offenen Asts in jedem Tableau τ für Γ:
Anfang: φ erfüllt $\Gamma^{<\alpha>} = \tau_\infty(\varepsilon)$ für die Vervollständigung τ_∞ von τ.
Annahme: $\tau_\infty(w) \neq \emptyset$ und φ erfüllt $\bigcup\{\, \tau_\infty(u) : u \leq w \,\}$.
Schluss: Ist $\tau_\infty(wi)$ leer, $i < 2$, bestimmt w einen offenen Ast für τ_∞.

Sonst existiert ein Präfix $u \leq w$ und eine β-Formel $B \in \tau_\infty(u)$ mit $\tau_\infty(wi) = \{B_i\}^{<\alpha>}$, $i < 2$, für ihre β-Teilchen. Wegen $\varphi(B) = 1$ erfüllt φ nun B_i für ein $i < 2$, und damit ganz $\tau_\infty(wi) \supseteq \tau(wi)$. □

Rechtfertigung der Tableau-Konstruktion

Lemma

1. *Für endliches* $\Gamma \subseteq \mathcal{F}^{\neg\neg}[\mathcal{A}]$ *terminiert obige Konstruktion.*
2. *Ist* Γ *unendlich und unerfüllbar, so terminiert die Konstruktion mit einem endlichen abgeschlossenen Tableau.*

Beweis.

Der endliche Fall ist klar; im unendlichen Fall kann die Konstruktion nur mit einem abgeschlossenen Tableau terminieren.

Annahme: Wenn ausgehend vom Minimaltableau τ die Konstruktion nicht terminiert, erhalten wir eine streng monoton wachsende Folge τ_i, $i \in \mathbb{N}$, von Tableaus, deren Vereinigung $\bar{\tau}$ in allen Positionen, wo Äste noch nicht abgeschlossen sind, mit τ_∞ übereinstimmt. Da der Baum für $\bar{\tau}$ binär ist, existiert nach Königs Lemma (Graphtheorie) eine unendlicher Ast Θ.
Wegen $\Gamma \subseteq \tau_\infty(\varepsilon) \subseteq \bigcup \tau_\infty[\Theta]$ ist diese vollständige und offene Menge nach Hintikkas Lemma erfüllbar, also auch Γ, Widerspruch. $\square$

Normalformen und Erfüllbarkeitsäquivalenz

Manche Algorithmen erfordern Eingaben in einer bestimmten einfachen Gestalt. Der hierbei verwendet Begriff Normalform ist immer im Hinblick auf den spezifischen Algorithmus zu verstehen; leider gibt es keine universell beste/kleinste/einfachste Darstellung einer jeden Formel über $\mathcal{J}_2[\mathcal{A}]$ (s.u.).

Die Effizienz von Algorithmen beinhaltet die Umwandlung der Eingabe(n) in die nötige Normalform. Sie darf bzgl. Zeit/Platz nicht zu teuer sein. Statt Äquivalenz zur Ausgangsformel genügt oft folgende schwächere Bedingung:

Definition (strenger als in anderen Quellen!)

Formeln A und B mit $@(A) \subseteq @(B)$ genügen der Relation $A \sqsubseteq_e B$, falls

- zu jedem $\varphi \in \{A\}^\triangleright$ existiert ein $\psi \in \{B\}^\triangleright$, das auf $@(A)$ mit φ übereinstimmt;
- $\{B\}^\triangleright \subseteq \{A\}^\triangleright$.

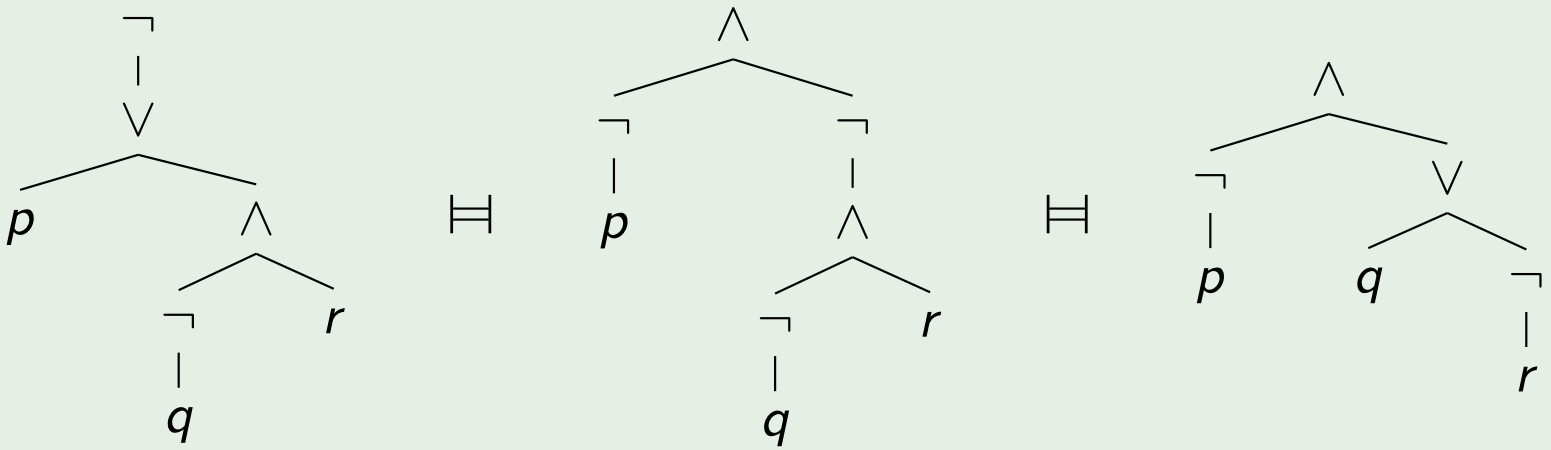
Erfüllbarkeitsäquivalenz $\models\!\!\!\models_e$ ist nun die kleinste ÄR, die $\sqsubseteq_e \cup (\sqsubseteq_e)^{\mathrm{op}}$ enthält, d.h., die transitive Hülle von $\sqsubseteq_e \cup (\sqsubseteq_e)^{\mathrm{op}}$.

Beispiel

Folgende Formeln lassen sich durch de Morgan'sche Regeln ineinander umformen (die Distributivgesetze vergrößern unnötig die Blätterzahl):

$$\neg(p \vee (\neg q \wedge r)) \quad , \quad (\neg p \wedge \neg(\neg q \wedge r)) \quad , \quad (\neg p \wedge (q \vee \neg r))$$

Die äußeren Formeln sind gleich lang, ihre Syntaxbäume haben 7 Knoten und 6 Kanten, unterscheiden sich aber in der Tiefe:



Die rechte Formel in KNF (s.u.) wird am Nützlichsten sein.

NNF für $\mathcal{J}_2 = \{\neg, \wedge, \vee\}$

Definition (NNF)

NNF $\subseteq \mathcal{F}_2[\mathcal{A}]$ ist der Abschluß der Menge $\mathcal{A} + \neg\mathcal{A}$ der Literale unter Kon- und Disjunktion. Ihre Elemente liegen in Negationsnormalform vor.

Jede Formel $A \in \mathcal{F}[\mathcal{A}]$ lässt sich durch Elimination der Junktoren $\top$, $\bot$, $\rightarrow$ und $\leftrightarrow$ und durch Iteration der de Morgan'schen Regeln in eine äquivalente Formel $B \in$ NNF umwandeln, die wir mit NNF(A) bezeichen.

Definition

Für $A \in \mathcal{F}[\mathcal{A}]$ bezeichne $\|A\|$ die Anzahl der Blätter des Syntaxbaums.

Lemma (HA!)

Zu jeder Formel $A \in \mathcal{F}[\mathcal{A}]$ ohne $\leftrightarrow$ existiert eine äquivalente Formel $B \in$ NNF, so dass $\|B\|$ in $\mathcal{O}(\|A\|)$ liegt, d.h., bis auf eine Konstante linear von $\|A\|$ abhängt. □

KNF und DNF für $\mathcal{J}_2 = \{\neg, \wedge, \vee\}$

Definition

- Eine Disjunktion von Literalen heißt Klausel. Man unterscheidet
  - positive/negative Klauseln, in denen alle Literale positiv/negativ sind;
  - Horn-Klauseln mit maximal einem positiven Literal;
  - k-Klauseln, wenn maximal k Literale auftreten; im Fall $k = 1$ spricht man auch von Unit-Klauseln.
- Eine Konjunktion von Klauseln heißt konjunktive Normalform (KNF), speziell k-KNF im Fall von k-Klauseln. Die entsprechenden Teilmengen von $\mathcal{F}_2[\mathcal{A}]$ bezeichnen wir ebenso, also k-KNF $\subseteq$ KNF $\subseteq$ NNF.

Definition

Vertauscht man in der obigen Definition die Rollen von Konjunktion ($\wedge$) und Disjunktion($\vee$), so erhält man die Begriffe der co-Klausel und der disjunktiven Normalform (DNF). Es gilt k-DNF $\subseteq$ DNF $\subseteq$ NNF.

Bemerkung (alternative Mengendarstellung für KNF)

Klauseln lassen sich alternativ als endliche Mengen von Literalen auffassen, und Formeln in KNF als endliche Mengen von Klauseln, z.B.

$$\{\{p, q, r\}, \{p, \neg q\}, \{\neg p, q\}\} \quad \text{statt} \quad (p \vee q \vee r) \wedge (p \vee \neg q) \wedge (\neg p \vee q)$$

Für Mengeninklusion $\subseteq$ („subsumiert“) und Vereinigung $\cup$ gilt nun:

Klauseln: $K \subseteq L$ impliziert $K \sqsubseteq L$, und $K \cup L = K \vee L$;

KNF-Formeln: $F \subseteq G$ impliziert $G \sqsubseteq F$, und $F \cup G = F \wedge G$.

Lemma (KNF-Reduktion)

- *Für Klauseln mit $K \subseteq L$ gilt $K \models\!\dashv K \wedge L$.*
- *Enthält die Klausel K komplementäre Literale, so gilt $K \models\!\dashv \top$.* □

Daher kann man subsumierte und tautologische Klauseln aus einer KNF entfernen, bzw. von weiterer Bearbeitung ausschließen.

Lemma

Zu jeder Formel $A \in \mathcal{F}_2[\mathcal{A}]$ existiert eine äquivalente Formel $B \in KNF$, so dass $\|B\|$ in $\mathcal{O}\left(2^{\|A\|}\right)$ liegt.

Beweis.

Der Syntaxbaum von $\mathrm{NNF}(A)$ hat maximal $\|A\|$ Blätter und ist, bis auf eventuelle Negationen von Blättern, streng binär. Daher ist seine Höhe durch $\|A\| - 1$ beschränkt, und die Summe der Entfernungen zweier Blätter von der Wurzel (Astlängen) durch $\|A\|$. Zieht man mit dem 2. Distributibutivgesetz Disjunktionen „nach innen“, kann die Höhe des Baums höchstens auf die Summe zweier Astlängen wachsen, also $\|B\| \leq 2^{\|A\|}$. $\square$

Lemma (HA!)

Es gibt eine Folge von NNF-Formeln A_n, $n \in \mathbb{N}$, mit $\|A_n\| = 2n$, so dass jede logisch äquivalente Formel B_n in KNF mindestens $2^n = (\sqrt{2})^{\|A\|}$ Positionen mit Literalen hat.

Duale Formeln und einfache Zusammenhänge

Definition

Die duale Formel zu $A \in \mathcal{F}_1[\mathcal{A}]$ ist gegeben durch

$$d(p) := p \text{ für } p \in \mathcal{A} \qquad d(B \wedge C) = d(B) \vee d(C)$$
$$d(\neg A) = \neg d(A) \qquad d(B \vee C) = d(B) \wedge d(C)$$

Lemma

- $A \in KNF$ *impliziert* $NNF(\neg A) \in DNF$
- $A \in KNF$ *genau dann wenn* $d(A) \in DNF$. □

Lemma

- $(1 - \varphi)(d(A)) = 1 - \varphi(A)$ *für jede Belegung* φ.
- A *ist eine Tautologie gdw.* $d(A)$ *ist widersprüchlich.*
- A *ist erfüllbar gdw.* $d(A)$ *ist keine Tautologie.* □

Davis-Putnam-Verfahren (für NNF)

Die Erfüllbarkeit einer Formel lässt sich statt mittels einer Wahrheitstabelle gezielter überprüfen, indem man bottom-up $\top$ oder $\bot$ für einzelne Variablen substituiert (Homomorphismus!, siehe ▸ Folie 16), was bei NNF-Formeln aufgrund der Rechenregeln für $\top$ und $\bot$ leicht zu vereinfachen ist und einen binären Formel-Baum mit Blättern aus $\{\top, \bot\}$ liefert.

Lemma (HA)

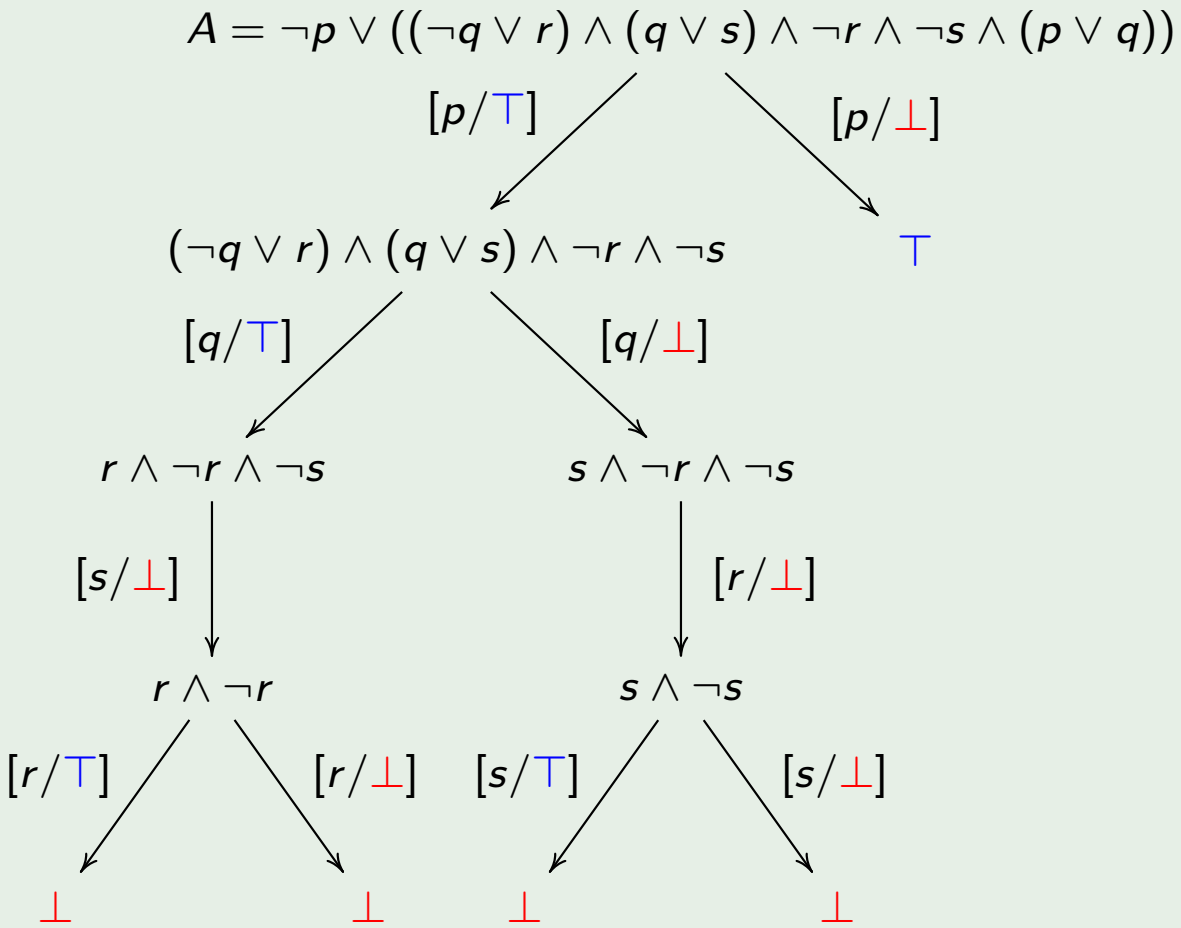
$A \in NNF$ ist genau dann erfüllbar, wenn eine der Formeln $A[p/\top]$ und $A[p/\bot]$ äquivalent zu einer erfüllbaren Formel ist. □

Algorithmen, die diese Idee iterativ umsetzen und mit Heuristiken weiter verfeinern, sind als Davis[10]-Putnam[11]-Verfahren bekannt. Das ursprüngliche Verfahren wurde 1960 veröffentlicht, und noch heute kommen seine Varianten in den schnellsten SAT-Solvern zur Anwendung.

[10] Martin Davis (*1928)

[11] Hilary Whitehall Putnam (1926-2016)

Beispiel



Damit erfüllt jede Belegung mit $\varphi(p) = 1$ die Formel A.

Regel-baserte Definition von Davis-Putnam

(0) Keine Verzweigungen sind erforderlich im Fall der
- Unit-Regel: A hat für ein Atom p die Form $p \wedge B$ oder $\neg p \wedge B$, oder
- Pure-Literal Regel: ein Atom $p \in \mathcal{A}$ tritt in A nur positiv/negativ auf.

Wende nur $[p/\top]$ bzw. $[p/\bot]$ an; dann ist A erfüllbarkeitsäquivalent im obigen strengen Sinn zum Substitut.

(1) Kann man den Erfolg keiner Substitutionen ausschließen, liegt die
- Splitting-Regel vor

Heuristische Auswahl des zu substituierenden Atoms (umseitig).

Im Fall von KNF-Formeln greifen weitere Vereinfachungen: subsumierte bzw. tautologische Klauseln dürfen sofort entfernt werden ▸ Folie 156.

Auswahlkriterien für die Splitting-Regel

Bei der Splitting-Regel ist der „Preis“ für den Prozess der Auswahl des Atoms abzuwägen gegen die zu erwartende Vereinfachung beim restlichen Verfahren. Mögliche Auswahlkriterien sind etwa

- das erste vorkommende Atom (keine Auswahlkosten);
- ein am häufigsten vorkommendes Atom;
- ein Atom p mit $\sum_{p \in K_i} |K_i|$ minimal;
- ein Atom p, das in den „kurzen“ Klauseln am häufigsten vorkommt; (braucht eine explizite Schranke, ab wann eine Klausel als „kurz“ gilt);
- ein Atom, bei dem die Differenz zwischen positiven und negativen Auftreten in den „kurzen“ Klauseln maximal ist.

Konkrete Implementierungen können weitere Heuristiken verwenden.

Resolutionslemma

Satz

Für alle Formeln A, B, C *gilt* $\{A \vee B, \neg B \vee C\} \models A \vee C$.

Beweis.

Die Behauptung ist äquivalent zu $\{\neg A \to B, B \to C\} \models \neg A \to C$ und damit auch zu $\{\neg A \to B, B \to C\} \vdash \neg A \to C$:

0.	$\neg A \to B$	Ann.
1.	$B \to C$	Ann.
2.	$\neg A$	Ann.
3.	B	MP, 2,0
4.	C	MP, 3,1
5.	$\neg A \to C$	DT, 2–4

□

Dies liefert die Basis für den folgenden Resolventen-Kalkül:

Resolution für Klauseln in Mengenschreibweise

Definition

Das deduktive System $\mathcal{R}$ auf der Menge $\mathcal{C}$ aller Klauseln hat nur die Regel

$$\frac{K + \{\ell\} \quad \{\neg\ell\} + M}{K \cup M}$$ mit der Resolvente $K \cup M$

für Klauseln K, M und ein Literal ℓ, das nicht in K, und dessen Negation nicht in M vorkommt („$+$“ steht für disjunkte Vereinigung).

Eine $\mathcal{R}$-Herleitung einer Klausel K aus $F \in \mathrm{KNF}$, geschrieben $F \vdash_{\mathrm{res}} K$, ist eine Klausel-Folge $\langle K_i : i \leq n \rangle$ mit $K_n = K$ und

- für $k < n$ gilt entweder $K_k \in F$, oder K_k ist Resolvente von K_i und K_j mit $i, j < k$.

Beachte die Ähnlichkeit der Resolventen-Regel mit der Schnittregel des Sequenzen-Kalküls unter Berücksichtigung von $(\neg R)$, wobei $\Gamma_0 = \Gamma_1 = \emptyset$.

Aussagenlogik versus Klausel-Logik

In Analogie zur bisherigen Aussagenlogik erhalten wir nun eine „Klausel-Logik“ mit folgenden Entsprechungen:

	Aussagenlogik	Klausel-Logik
Grundbausteine	Formeln $A \in \mathcal{F}[\mathcal{A}]$	Klauseln K
Prämissenmengen	$\Gamma \subseteq \mathcal{F}[\mathcal{A}]$	Formeln $F \in$ KNF
Folgerung	$\Gamma \models A$	$F \models_{\mathrm{res}} K$
Herleitung	$\Gamma \vdash A$, $\Gamma \Vdash A$, etc.	$F \vdash_{\mathrm{res}} K$

Fasst man $F \in \mathrm{KNF}$ als Menge von Klauseln auf, kann man für beliebiges $A \in \mathcal{F}[\mathcal{A}]$ fragen, ob in der Aussagenlogik $F \models A$ gilt. Falls A keine Klausel ist, gibt es dafür keine Entsprechung in der Klausel-Logik.

Trotzdem wird auch die Klausel-Logik die Unerfüllbarkeit von $F \cup \{\neg A\}$ mittels einer Herleitung $F \cup \mathrm{KNF}(\neg A) \vdash_{\mathrm{res}} \emptyset$ nachweisen können (s.u.).

Satz

Der $\mathcal{R}$-Kalkül ist korrekt aber nicht vollständig, d.h.,

$$F \vdash_{\text{res}} K \quad \textit{impliziert} \quad F \models_{\text{res}} K \qquad \textit{aber nicht notwendig umgekehrt}$$

Zudem kann $F \models A \in KNF$ gelten, wobei A mindetens zwei Klauseln hat.

Beweis.

Das Resolutionslemma impliziert Korrektheit. Umgekehrt verträgt sich Subsumption mit $\models_{\text{res}}$ aber nicht mit $\vdash_{\text{res}}$: Zwar gilt

$$\{\{p, q\}, \{p\, \neg q\}\} \models_{\text{res}} \{p\} \quad \text{und} \quad \{\{p, q\}, \{p\, \neg q\}\} \vdash_{\text{res}} \{p\}$$

Aber $\{p\} \subseteq \{p, r\}$ impliziert nur $\{\{p, q\}, \{p\, \neg q\}\} \models_{\text{res}} \{p, r\}$. Des Weiteren gilt $\{p \vee q, p \vee \neg q, \neg p \vee q\} \models p \wedge q$, was keine Klausel ist. □

Satz (Widerlegungsvollständig- und Korrektheit, John Alan Robinson)

$F \in KNF$ ist genau dann unerfüllbar, wenn $F \vdash_{\text{res}} \emptyset$.

Beweis.

Aus $F \vdash_{\text{res}} \emptyset$ folgt $F \vdash_{\text{res}} \ell$ und $F \vdash_{\text{res}} \neg\ell$ für ein Literal ℓ, nach obigem Satz also auch $F \models \ell$ und $F \models \neg\ell$. Damit ist die Menge F unerfüllbar.

Sei F unerfüllbar. Betrachte die Anzahl k verschiedener Literale in F.

$k = 0$: Dann gilt $F = \emptyset$.

Annahme: Die Behauptung stimmt für alle $i < k$.

$k > 0$: OBdA enthalte F keine tautologischen oder subsumierten Klauseln. In F tritt ein Atom p positiv und negativ auf, sonst wäre F erfüllbar.

F_+ enthalte alle F-Klauseln, in denen p auftritt;

F_- enthalte alle F-Klauseln, in denen $\neg p$ auftritt;

F_0 enthalte alle F-Klauseln, in denen weder p noch $\neg p$ auftritt.

Wegen $F = F_+ \cup F_- \cup F_0$ folgt aus der Unerfüllbarkeit einer oder der Vereinigung zweier dieser Klauselmengen nach Annahme die Behauptung.

Beweis, Fortsetzung.

Sind $F_+ \cup F_0$, $F_- \cup F_0$ und $F_+ \cup F_-$ erfüllbar, so zeigen wir:

$$F_+ \cup F_0 \vdash_{\text{res}} p \quad \text{und} \quad F_- \cup F_0 \vdash_{\text{res}} \neg p$$

Zwecks Elimination von p aus F_+ setze $F_\oplus := \{ K - \{p\} : K \in F_+ \}$.

Wäre $F_\oplus \cup F_0$ erfüllbar, dann gäbe es $\varphi \in \big(F_\oplus \cup F_0\big)^\triangleright$ mit $\varphi(p) = 0$, also würde $F_+ \cup F_0$ wie auch $F_- \cup F_0$ und somit F von φ erfüllt, $\lightning$.

Also gilt $F_\oplus \cup F_0 \vdash_{\text{res}} \emptyset$, etwa mittels einer $\mathcal{R}$-Herleitung $\langle J_i : i < n \rangle$ minimaler Länge. Da F_0 nach Voraussetzung erfüllbar ist, muss für jede Resolvente J_k von Klauseln J_i und J_j mit $i, j < k$, einer dieser Vorgänger zu $F_\oplus$ gehören; andernfalls wäre J_s überflüssig. Damit liefert

$$J'_i := \begin{cases} J_i \cup \{p\} & \text{falls } J_r \in F_\oplus \\ J_r & \text{sonst} \end{cases}$$

eine $\mathcal{R}$-Herleitung von p aus $F_+ \cup F_0$. Analog gilt $F_- \cup F_0 \vdash_{\text{res}} \neg p$. Beides zusammen impliziert $F \vdash_{\text{res}} \emptyset$. □

Resolutions-Heutistiken

Definition

Für $A \in \mathrm{KNF}$ bezeichnet $\boldsymbol{Res}(A)$ die Menge aller Klauseln, die in endlich vielen Schritten als Resolventen aus den Klauseln von A konstruierbar sind.

Ein naiver Algorithmus basiert auf der Tatsache, dass $A \in \mathrm{KNF}$ genau dann erfüllbar ist, wenn $\emptyset \in \boldsymbol{Res}(A)$ gilt. Dies wird nicht empfohlen.

Stattdessen kann man sich auf starke Herleitungen beschränken, in denen

- keine Klausel mehrfach auftritt;
- keine tautologischen Klauseln auftreten;
- einmal subsumierte Klauseln nicht weiter verwendet werden.

Dies erlaubt es, aus $\boldsymbol{Res}(A)$ bestimmte Klauseln zu entfernen und mit einer potentiell kleineren Menge $\widetilde{\boldsymbol{Res}}(A)$ zu arbeiten.

Zur Handrechnung ziehen wir ein systematisches graphisches Verfahren vor.

Beispiel (Die finanziellen Probleme der Familie Z, vergl. ▸ Folie 49)

Zunächst wandeln wir die Prämissen und die negierte(!) Schlußfolgerung in KNF um:

$$
\begin{array}{lll}
B_0 = \neg p \to \neg(q \wedge r \wedge s) & = p \vee \neg \vee \neg r \vee \neg s & = \{p, \neg q, \neg r, \neg s\} \\
B_1 = q & & = \{q\} \\
B_2 = r \vee s & & = \{r, s\} \\
B_3 = t \to \neg r & = \neg t \vee \neg r & = \{\neg r, \neg t \\
B_4 = \neg t \to s & = t \vee s & = \{s, t\} \\
\neg A = \neg(\neg p \to (\neg r \wedge s)) & = \neg p \wedge (r \vee \neg s) & = \{\neg p\}, \{r, \neg s\}
\end{array}
$$

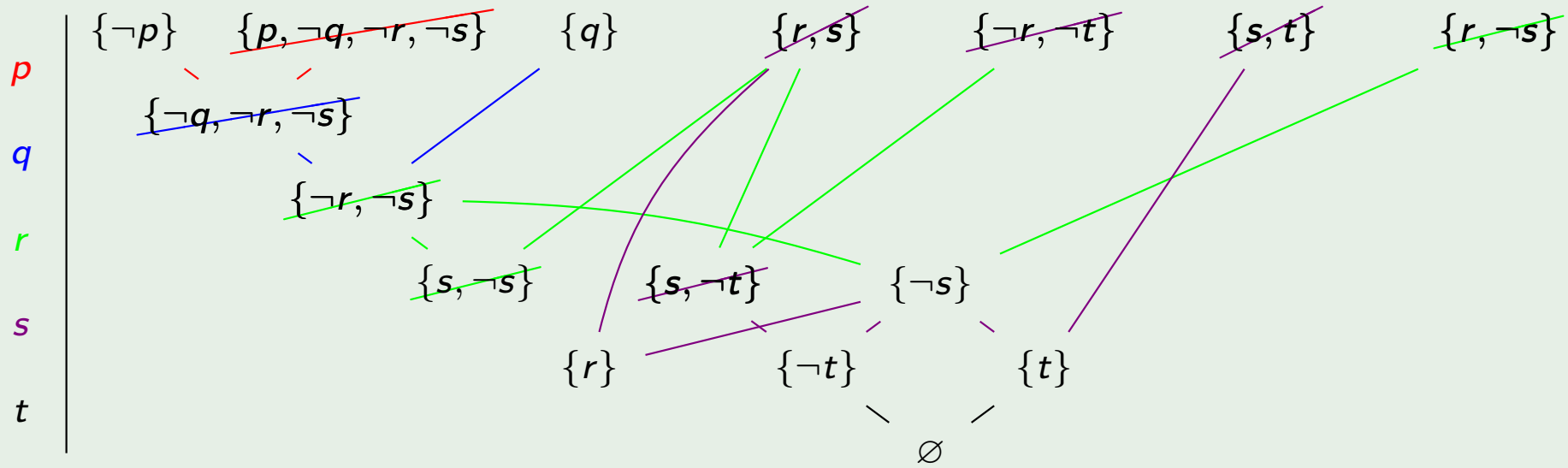
Die resultierende Klauselmenge

$$G = \{\{\neg p\}, \{p, \neg q, \neg r, \neg s\}, \{q\}, \{r, s\}, \{\neg r, \neg t\}, \{s, t\}, \{r, \neg s\}\}$$

kann nun in einem einfachen graphischen Verfahren auf (Un-)Erfüllbarkeit untersucht werden:

Beispiel

- Ordne die auftretenden Klauseln horizontal an;
- Ordne die auftretenden Atome (beliebig) vertikal an;
- konstruiere schichtweise alle Resolventen bzgl. des aktuellen Atoms; streiche tautologische oder subsumierte Formeln sofort;



Das Streichen der verbliebenen nichtleeren Klauseln kann unterbleiben.

$\mathcal{R}$-Herleitung: $\{\neg p\}, \{p, \neg q, \neg r, \neg s\}, \{\neg q, \neg r, \neg s\}, \{q\}, \{\neg r, \neg s\}, \{r, \neg s\}, \{r, s\}, \{\neg r, \neg t\}, \{s, t\}, \{s, \neg t\}, \{\neg s\}, \{t\}, \{\neg t\}, \emptyset$.

Zur Übung möge man eine andere Reihenfolge der Atome ausprobieren!

Bemerkungen

- Das Berechnen des Resolutionsabschlusses $\boldsymbol{Res}(A)$ (außer tautologischen Klauseln) ist unter dem Namen Stufenstrategie bekannt. Sofern $\emptyset$ nicht auftritt kann man sämtliche erfüllenden Belegungen bestimmen. (Für Handrechnung zu fehleranfällig.)
- Bei der sog. Stützmengenrestriktion werden Unit-Klauseln bevorzugt zur Resolventenbildung herangezogen.
- Bei P- bzw. N-Resolution sollte eine der beteiligten Klauseln nur positive (negative) Literale enthalten.
- Zumindest für Formeln in 2-KNF und für Horn-Formeln (mit höchstens einem positiven Literal pro Klausel) ist die Resolventen-Methode effizient. I.A. ist das leider nicht der Fall:

Resolution ist i.A. nicht effizient

Satz

Es gibt eine Folge $\emptyset \neq A_k \in KNF$, $k \in \mathbb{N}$, *mit folgenden Eigenschaften:*

- *in* A_k *kommen höchstens die Atome* p_i, $i < 2k+1$ *vor;*
- A_k *hat* $k+1$ *Klauseln;*
- *keine der* $2^{k+1}-1$ *Klauseln in* $\mathbf{Res}(A_k)$ *subsumiert eine andere.*

Beweis.

$k = 0$: $A_0 = p_0$ hat $2^1 - 1 = 1$ Resolventen.

Annahme: $A_k = \bigwedge_{i<k+1} K_{k,i}$ habe die gewünschten Eigenschaften.

$k+1$: $A_{k+1} := \mathrm{KNF}(A_k \vee p_{2k+1}) \wedge (\neg p_{2k+1} \vee p_{2k+2})$ hat zwei Atome und eine Klausel mehr als A_k, und $\mathbf{Res}(A_{k+1})$ hat alle um p_{2k+1} bzw. p_{2k+2} vergrößerten Resolventen von A_k, sowie $\{\neg p_{2k+1}, p_{2k+2}\}$. Dabei entstehen keine neuen Mengeninklusionen zwischen Klauseln. □

Die Tseitin-Transformation

Liegen die Prämissen eines endlichen Problems $\Gamma \models A$ in KNF, könnte die Resolventen-Methode die Unerfüllbarkeit von $\Gamma \cup \{\mathrm{KNF}(\neg A)\}$ nachweisen. Da aber die Berechnung von $\mathrm{KNF}(\neg A)$ teuer sein kann, möchte man stattdessen eine KNF-Formel verwenden, die nur ▸ erfüllbarkeitsäquivalent zu $\neg A$ ist, aber dafür schneller und einfacher zu bestimmen.

Die Distributivgesetze wiederholen ganze Teilformeln. Um das zu vermeiden führte G. S. Tseitin[12] neue Atome als Abkürzungen für Teilformeln ein. Die entsprechenden Äquivalenzen lassen sich leicht in KNF umformen und sind mittels Konjunktion zur Ausgangsformel hinzuzufügen. Das liefert eine erfüllbarkeitsäquivalente Formel in KNF mit mehr Atomen als vorher, deren Größe aber linear von der ursprünglichen Anzahl der Literale abhängt.

Für $\mathcal{J}_2$-Formeln ohne doppelte und ohne führende Negation genügt es, nur echte Teilformeln zu behandeln, die Kon- oder Disjunktionen sind.

[12] On the complexity of derivation in propositional calculus, Leningrad Seminar on Mathematic Logic, September 1966

Definition (Tseitin-Transformation für Kon- und Disjunktion)

A	$p \leftrightarrow A$	KNF$(p \leftrightarrow A)$ in Mengendarstellung
$B \wedge C$	$(B \wedge C \rightarrow p) \wedge (p \rightarrow B \wedge C)$	$\{\neg B, \neg C, p\}\{B, \neg p\}, \{C, \neg p\}$
$B \vee C$	$(p \rightarrow B \wedge C) \wedge (B \vee C \rightarrow p)$	$\{B, C, \neg p\}, \{\neg B, p\}, \{\neg C, p\}$

Beispiel

$A = (p \wedge \neg q) \vee \neg(r \wedge s)$ hat zwei echte binäre Teilformeln:

$$t_0 \leftrightarrow p \wedge \neg q \qquad \text{und} \qquad t_1 \leftrightarrow r \wedge s$$

Das liefert

$$\begin{aligned} \boldsymbol{Tst}(A) := & \{t_0, \neg t_1\} \\ & \{\neg p, q, t_0\}\{p, \neg t_0\}\{\neg q, \neg t_0\} \\ & \{\neg r, \neg s, t_1\}\{r, \neg t_1\}\{s, \neg t_1\} \in \text{KNF} \end{aligned}$$

Algorithmus

Gegeben: $A \in \mathcal{F}[\mathcal{A}]$, **Ausgabe**: $\boldsymbol{Tst}(A) \in \mathrm{KNF}$

1. Eliminiere die Junktoren $\bot$, $\top$, $\rightarrow$ und $\leftrightarrow$.
2. Eliminiere doppelte und führende Negationen.
3. Von unten nach oben im Syntaxbaum: ersetze jede echte binäre Teilformel B durch ein frisches p und füge $\mathrm{KNF}(p \leftrightarrow B)$ mittels Konjunktion zur aktuellen Formel hinzu.

Bemerkung

Drei oder mehr aufeinanderfolgende Kon- oder Disjunktionen können ebenso gehandhabt werden wie binäre: nur die KNF wächst um weitere Klauseln:

$$B \wedge C \wedge D \longmapsto \{\neg B, \neg C, \neg D, p\}, \{B, \neg p\}, \{C, \neg p\}, \{D, \neg p\}$$
$$B \vee C \vee D \longmapsto \{B, C, D, \neg p\}, \{\neg B, p\}, \{\neg C, p\}, \{\neg D, p\}$$

Satz

Für jede Formel A über $\mathcal{J}_2$ ohne doppelte oder führende Negation sind A und $\mathbf{Tse}(A)$ erfüllbarkeitsäquivalent.

Beweis.

Falls $\varphi \in \{A\}^{\triangleright}$ setzen wir für jede verwendete Instanz $p \leftrightarrow B$

$$\psi(p) := \hat{\varphi}(B)$$

und übernehmen die übrigen Werte von φ. Dann gilt nach Konstruktion $\hat{\psi}(\mathbf{Tst}(A)) = 1$.

Umgekehrt erfüllt $\psi \in \{\mathbf{Tst}(A)\}^{\triangleright}$ alle Klauseln von $\mathbf{Tst}(A)$. Aufgrund der erfüllten Äquivalenzklauseln rekonstruiert die Rücksubstitution der echten binären Teilformeln von A für die frischen Atome in der ersten Klausel die Ausgangsformel A, also folgt $\hat{\psi}(A) = 1$. □

Teil 2
Prädikatenlogik

(1. Stufe mit Gleichheit)

Inhaltsverzeichnis, Teil 2

Kapitel 8
Motivation und Überblick

Defizite der Aussagenlogik

- Die Aussagenlogik (AL) ist überall anwendbar, wo potentiell wahre oder falsche Aussagen formuliert und miteinander kombiniert werden können, unabhängig vom Kontext.
- Sie ist weniger dafür geeignet zwischen verschiedenen Anwendungsbereichen zu unterscheiden, genauer: Aussagen über einzelne Elemente eines spezifischen (nichtleeren) Datenbereiches zu formulieren, und ihr Verhältnis zueinander, z.B. für natürlichen Zahlen, gerichtete Graphen oder Datenbanken. Ihre recht einfache mathematische Struktur ist Konsequenz dieser beschränkten Ausdrucksfähigkeit.

Man braucht somit eine Möglichkeit, die atomaren Formeln an die Gegebenheiten der Datenbereiche von Interesse anzupassen. Das, und noch etwas mehr, leistet die Prädikatenlogik 1. Stufe (PL_1).

Mögliche Anwendungen in der Informatik

Viele Anwendungen bedienen sich der PL_1, z.B.

- Lösung von Anfragen auf Datenmengen in der KI oder in Informationssystemen;
- Formulierung von Integritätsbedingungen auf Daten; etwa Schleifeninvarianten eines Programms, Constraints auf XML-Dateien oder Datenbankeinträgen;
- Lösung von Constraint-Systemen beim Testen oder Planen.
- Logisches Programmieren, z.B. in PROLOG.

Syntax, Übersicht

Frege[13] definierte 1879 die Syntax der PL_1 in seiner „Begriffsschrift – Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens“.

In moderner Terminologie:

- Eine Signatur $\boldsymbol{Fun} + \boldsymbol{Pred} \xrightarrow{\mathcal{S}} \mathbb{N}$ weist Operatoren und Prädikaten (formalen Funktions- bzw. Relationssymbolen) eine Stelligkeit zu;
- eine (abzählbare) Menge $\mathcal{V}$ von Variablen liefert in einem 1. Schritt die $\boldsymbol{Fun}$-Algebra $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$, einen abstrakten Datenbereich aus Termen in den Operatoren (für den der Rekursionssatz gilt!);
- Prädikate und formale Gleichungen mit Termen als Argumenten bilden nun die dem Datenbereich angepaßte atomare Formelmenge $\mathcal{A}(\mathcal{S})$;
- die bekannten Junktoren der Aussagenlogik und neue unäre Junktoren $\forall x$ („für alle x“) und $\exists x$ („es gibt ein x“) für jede Variable x, liefern die Menge $\boldsymbol{FO}(\mathcal{S})$ aller Formeln; $\forall$ und $\exists$ heißen Quantoren.

[13] Friedrich Ludwig Gottlob Frege (1848–1925)

Semantik, Übersicht

Tarskis[14] Semantik von 1934 in moderner Terminologie:

- $\mathcal{S}$-Strukturen $\mathcal{M} = \langle D, I \rangle$ interpretieren $\mathcal{S}$ in einer Trägermenge D; speziell ist $I(R) \subseteq D^{\mathcal{S}(R)}$ für $R \in \boldsymbol{Pred}$ eine $\mathcal{S}(R)$-stellige Relation;
- Jede Belegung $\mathcal{V} \xrightarrow{\sigma} D$ der Variablen liefert einen eindeutigen $\boldsymbol{Fun}$-Homomorphismus $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}) \xrightarrow{\check{\sigma}} \langle D, I_{\boldsymbol{Fun}} \rangle$;
- Alle atomaren Formeln (in den Symbolen von $\boldsymbol{Pred}$ und der formalen Gleichheit $\doteq$) liefern gemäß $\check{\sigma}$ und $I_{\boldsymbol{Pred}}$ überprüfbare Aussagen bzgl. $\mathcal{M}$, also einen Wahrheitswert; analog zu $\mathcal{A} \xrightarrow{\varphi} \mathbb{B}$ der AL.
- Solch eine Belegung $\mathcal{A}(\mathcal{S}) \xrightarrow{\overline{\sigma}} \mathbb{B}$ setzt sich eindeutig zu einer Bewertung $\hat{\sigma}$ aller Formeln $F \in \boldsymbol{FO}(\mathcal{S})$ fort, analog zu $\mathcal{F}[\mathcal{A}] \xrightarrow{\hat{\varphi}} \mathbb{B}$.
- Mit einem fiesen Trick lassen sich die Semantiken für Terme und Formeln in eine ähnliche Form bringen.

[14] Alfred Tarski bzw. ursprünglich Alfred Tajtelbaum, 1901 – 1983

Beschreibung mathematischer Beziehungen

Beispiel

Syntax: Konstanten 1, 2, 3; 2-stellige Funktionssymbole $+$, $/$; 2-stelliges Prädikat $<$:

$$\text{Signatur} \quad \mathcal{S} = \langle \{1_{/0}, 2_{/0}, 3_{/0}, +_{/2}, /_{/2}\}, \{<_{/2}\} \rangle$$

$$\text{Terme (Infix)} \quad 1 \quad , \quad 1 + (2/3) \quad , \quad (x+3)/2 \quad , \quad \cdots$$

Logik: Junktoren $\rightarrow$, $\wedge$, Quantoren $\forall$, $\exists$;

$$\text{Formeln} \quad x < 3 \quad , \quad \forall x \, \forall y \; (x < y \rightarrow \exists z \, (x < z \wedge z < y))$$

Semantik: Datenbereich $D = \mathbb{Q}$, Konstante 1 bis Prädikat $<$ mit der üblichen Bedeutung.

Achtung: Andere Interpretationen der Signatur $\mathcal{S}$ sind möglich und dürfen nicht ignoriert werden!

Beschreibung von Beziehungen zwischen Daten

Beispiel

Syntax: ein Operator *WeiteDerReise*(-) $_{/1}$,
drei Prädikate *istHund*(-) $_{/1}$, *istFisch*(-) $_{/1}$, $<_{/2}$

Logik: Junktoren $\rightarrow$, $\wedge$, Quantor $\forall$;

typische Formel:

$$\forall x \, \forall y \big(\mathit{istHund}(x) \wedge \mathit{istFisch}(y) \rightarrow \mathit{WeiteDerReise}(x) < \mathit{WeiteDerReise}(y) \big)$$

Semantik: Datenbereich $D = \{\text{Lassie}, \text{Nemo}\} \cup \mathbb{N} \cup \{\perp\!\!\!\perp\}$

Die Standard-Interpretation der Funktion *WeiteDerReise*(-) liefert die Weite der Reise, die das Tier im Argument zurückgelegt hat, oder $\perp\!\!\!\perp$ sonst.

istHund und *istFisch* möge die charakteristische Funktion von $\{\text{Lassie}\}$ bzw. $\{\text{Nemo}\}$ sein.

Kapitel 9
Syntax der Prädikatenlogik

Syntax der Prädikatenlogik, Details

Auf zwei entscheidbaren abzählbaren Mengen ***Fun*** und ***Pred***, die zum Alphabet der AL disjunkt sind, betrachten wir eine Signatur

$$\boldsymbol{Fun} + \boldsymbol{Pred} \xrightarrow{\mathcal{S}} \mathbb{N}$$

Interpretiere die Elemente aus ***Fun*** als Operatoren oder Funktionssymbole, und die Elemente aus ***Pred*** als Prädikate oder Relationssymbole.

Üblicherweise verwenden wir Kleinbuchstaben f, g, $h\ldots$ für erstere, und Großbuchstaben R, S, $T\ldots$ für letztere.

$f_{/k}$ und $R_{/k}$ ist Kurzschreibweise für $\mathcal{S}(f) = k$ bzw. $\mathcal{S}(R) = k$.

Achtung: Ein weiteres binäres Prädikat $\doteq_{/2} \notin \boldsymbol{Pred}$ spielt später in der Semantik eine Sonderrolle als formale Gleichheit.

0-stellige Operatoren/Prädikate heißen Konstanten/Propositionen.

Weiter sei $\mathcal{V}$ eine (abzählbare entscheidbare) Menge von Variablen.

Aufbau der Formeln in drei Schritten

- Zuerst bilden wir für $\boldsymbol{Fun}$ die Termalgebra $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$;

$$t ::= v \mid f(t_0, \ldots, t_{\mathcal{S}(f)-1}) \quad \text{mit} \quad v \in \mathcal{V} \quad \text{und} \quad f \in \boldsymbol{Fun}$$

- Danach bilden wir $\boldsymbol{AT} := \boldsymbol{Atm}(\boldsymbol{Pred} + \{\doteq\}, \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}))$, die Menge der atomaren Formeln (mit Gleichheit) aus den Termen:

$$A ::= t_0 \doteq t_1 \mid R(t_0, \ldots, t_{\mathcal{S}(R)-1}) \quad \text{mit} \quad R \in \boldsymbol{Pred}$$

  Damit ist $\boldsymbol{AT}$ eine disjunkte Vereinigung $\sum_{R \in \boldsymbol{Pred}+\{\doteq\}} \boldsymbol{AT}_R$.

- Und schließlich bilden wir mit Hilfe der um $\forall x_{/1}$ und $\exists x_{/1}$, $x \in \mathcal{V}$, erweiterten Junktormenge $\mathcal{J}_{\mathcal{Q}}$ die Termalgebra $\boldsymbol{FO}(\mathcal{S})$ der prädikatenlogischen Formeln erster Stufe (first order formulae):

$$F ::= A \mid \top \mid \bot \mid \neg F \mid (F_0 \star F_1) \mid (\forall x\, F) \mid (\exists x\, F)$$

  mit $A \in \boldsymbol{AT}$, $\star \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$ und $x \in \mathcal{V}$.

Gebundene und freie Variablen

Als einstellige Junktoren binden $\forall x$ und $\exists x$ stärker als alle 2-stelligen.

Definition

Die Menge $\boldsymbol{V}(t)$ der in einem Term auftretenden Variablen ist

$$\boldsymbol{V}(x) = \{x\} \quad \text{und} \quad \boldsymbol{V}(f(t_0, \ldots, t_{\mathcal{S}(f)-1}) = \bigcup \{\, \boldsymbol{V}(t_i) \,:\, i < \mathcal{S}(f) \,\}$$

Analog sind die in einer (atomaren) Formel auftretenden Variablen definiert.

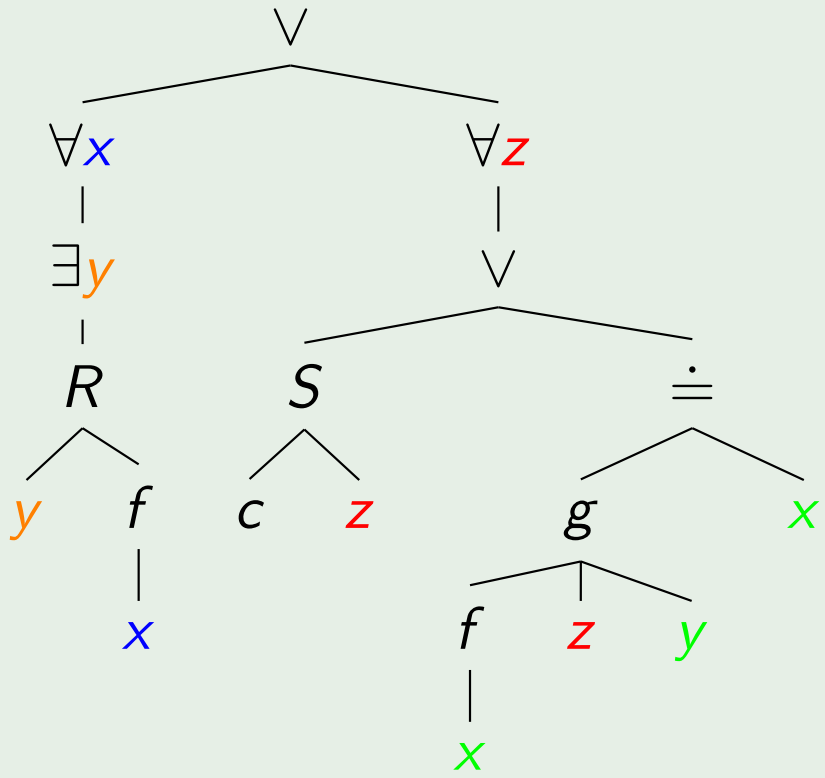
In $(\mathcal{Q}x\, F)$ ist F der Geltungsbereich für $\mathcal{Q}x$; jedes Auftreten von x in einem solchen Geltungsbereich heißt gebunden, außerhalb jedes solchen Geltungsbereichs frei. Die Menge der gebunden/frei auftretenden Variablen in $H \in \boldsymbol{FO}(\mathcal{S})$ wird mit $\boldsymbol{GV}(H)$ bzw. $\boldsymbol{FV}(H)$ bezeichnet; sie brauchen nicht disjunkt zu sein!

Formeln ohne frei vorkommende Variablen heißen abgeschlossen.

Formeln der PL_1 als 2-stufige Syntaxbäume

In der Baumdarstellung dienen die Prädikate sowie $\doteq$ als Schnittstellen zwischen den Bäumen der Terme und dem logischen Syntax-Baum.

Beispiel ($(\forall x\, \exists y\, R(y, f(x))) \vee \forall z\, (S(c, z) \vee g(f(x), z, y) \doteq x)$)



(frei auftretende Variablen)

Anmerkungen

Der Begriff der Entscheidbarkeit einer Menge wird erst in TheoInf2 offiziell eingeführt. Informell soll er hier bedeuten, dass wir ihre Elemente effizient, d.h., mit geringem Rechenaufwand oder sogar unmittelbar erkennen können.

Lemma

- *Die Entscheidbarkeit der Signatur $\mathcal{S}$ und der Variablenmenge $\mathcal{V}$ vererbt sich auf die Mengen der Terme und der (atomaren) Formeln.*
- *Zusammengesetzte Terme und Formeln lassen sich eindeutig zerlegen.*
- *Gebundene und freie Vorkommen von Variablen lassen sich effizient bestimmen.*

Kapitel 10
Elementare Semantik der Prädikatenlogik

$\mathcal{S}$-Strukturen

Das Ziel ist oft die Beschreibung von Beziehungen zwischen Elementen eines nichtleeren konkreten strukturierten Datenbereichs D.

Dafür war eine passende Signatur $\mathcal{S} = \textbf{\textit{Fun}} + \textbf{\textit{Pred}}$ gewählt worden, evtl. mit suggestiven Namen für die Funktionen- bzw. Relationssymbole.

Eine Semantik für $\textbf{\textit{FO}}(\mathcal{S})$ muss berücksichtigen, dass andere Interpretationen von $\mathcal{S}$ auf D wie auch andere Datenbereiche D' existieren können.

Definition

Unter einer $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I \rangle$ versteht man eine Menge $D \neq \emptyset$, den Datenbereich, zusammen mit einer Interpretation der Symbole in $\mathcal{S}$

$$D^{\mathcal{S}(f)} \xrightarrow{I(f)} D \text{ für } f \in \textbf{\textit{Fun}} \quad \text{und} \quad D^{\mathcal{S}(R)} \xrightarrow{I(R)} \mathbb{B} \text{ für } R \in \textbf{\textit{Pred}}$$

$I(\dot{=})$ ist immer die charakteristische Funktion der Gleichheit $=$ auf D.
Oft schreibt man $f^{\mathcal{M}}$ und $R^{\mathcal{M}}$ statt $I(f)$ bzw. $I(R)$.

Belegungen, Semantik der Terme für festes $\mathcal{M} = \langle D, I \rangle$

In der AL konnte die Menge $\mathcal{A}$ der atomaren Formeln direkt mit Wahrheitswerten belegt werden. In der PL_1 gelingt das nur indirekt. Dennoch wollen wir versuchen möglichst viel von der AL zu übernehmen.

Definition

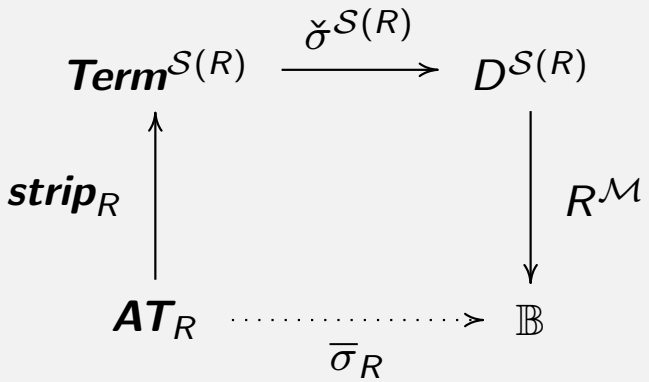
Eine Belegung der Variablen ist eine Abbildung $\mathcal{V} \xrightarrow{\sigma} D$ bzw. $\sigma \in D^{\mathcal{V}}$; gelegentliche Schreibweise $\sigma \in \mathcal{M}^{\mathcal{V}}$ (zwecks Erinnerung an I).

Aufgrund des ▸ Rekursionssatzes lässt sich solch eine Belegung σ eindeutig zu einem $\boldsymbol{Fun}$-Homomorphismus $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}) \xrightarrow{\check{\sigma}} \langle D, I_{\boldsymbol{Fun}} \rangle$ fortsetzen, der Semantik der Terme bzgl. $\sigma \in \mathcal{M}^{\mathcal{V}}$.

Das liefert auf kanonische Weise die Semantik der atomaren Formeln bzgl. σ als Abbildung $\boldsymbol{AT} = \boldsymbol{Atm}(\boldsymbol{Pred} + \{\doteq\}, \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})) \xrightarrow{\overline{\sigma}} \mathbb{B}$, die auf verschiedene Weise beschrieben werden kann:

Semantik atomarer Formeln für festes $\mathcal{M}$

Für jedes Prädikat $R \in \boldsymbol{Pred} + \{\dot{=}\}$ ist die Menge $\boldsymbol{AT}_R$ der atomaren Formeln, in denen R auftritt, isomorph zu $\boldsymbol{Term}^{\mathcal{S}(R)}$: man entfernt einfach das Symbol R, bzw. fügt es wieder hinzu. Nun lässt sich $\boldsymbol{AT} = \sum_{R \in \boldsymbol{Pred}+\{\dot{=}\}} \boldsymbol{AT}_R \xrightarrow{\overline{\sigma}} \mathbb{B}$ mittels Fallunterscheidung definieren:



Oder mittels Teilmengen $R^{\mathcal{M}}$ anstelle von charakteristischen Funktionen:

$$\overline{\sigma}(t_0 \dot{=} t_1) = 1 \text{ gdw } \check{\sigma}(t_0) = \check{\sigma}(t_1)$$

$$\overline{\sigma}(R(t_0, \ldots, t_{\mathcal{S}(R)-1})) = 1 \text{ gdw } \langle \check{\sigma}(t_0), \ldots, \check{\sigma}(t_{\mathcal{S}(R)-1}) \rangle \in R^{\mathcal{M}} \subseteq D^{\mathcal{S}(R)}$$

Semantik von Formeln für festes $\mathcal{M}$

Schließlich können wir den ▸ Rekursionssatz ein weiteres Mal anwenden und $\boldsymbol{AT} \xrightarrow{\overline{\sigma}} \mathbb{B}$ eindeutig zu einem Homomorphismus $\boldsymbol{FO}(\mathcal{S}) \xrightarrow{\hat{\sigma}} \mathbb{B}$ bzgl. der Signatur $\mathcal{J}_{\mathcal{Q}}$ fortsetzen, indem wir die Junktoren in $\mathcal{J}$ analog zur AL ▸ Folie 20 behandeln. Die neuen Junktoren erfordern einen Zwischenschritt:

Definition

Jedes Paar $\langle x, d\rangle \in \mathcal{V} \times D$ liefert eine Modifikation $D^{\mathcal{V}} \xrightarrow{\{x/d\}} D^{\mathcal{V}}$:

$$\sigma \longmapsto \sigma\{x/d\}^{a} \quad , \quad y \longmapsto \begin{cases} d & \text{falls } y = x \\ \sigma(y) & \text{sonst} \end{cases}$$

Damit lässt sich die Semantik quantifizierter Formeln formulieren:

$$\hat{\sigma}(\forall x\, A) := \inf\{\, \widehat{\sigma\{x/d\}}(A) \;:\; d \in D \,\}$$

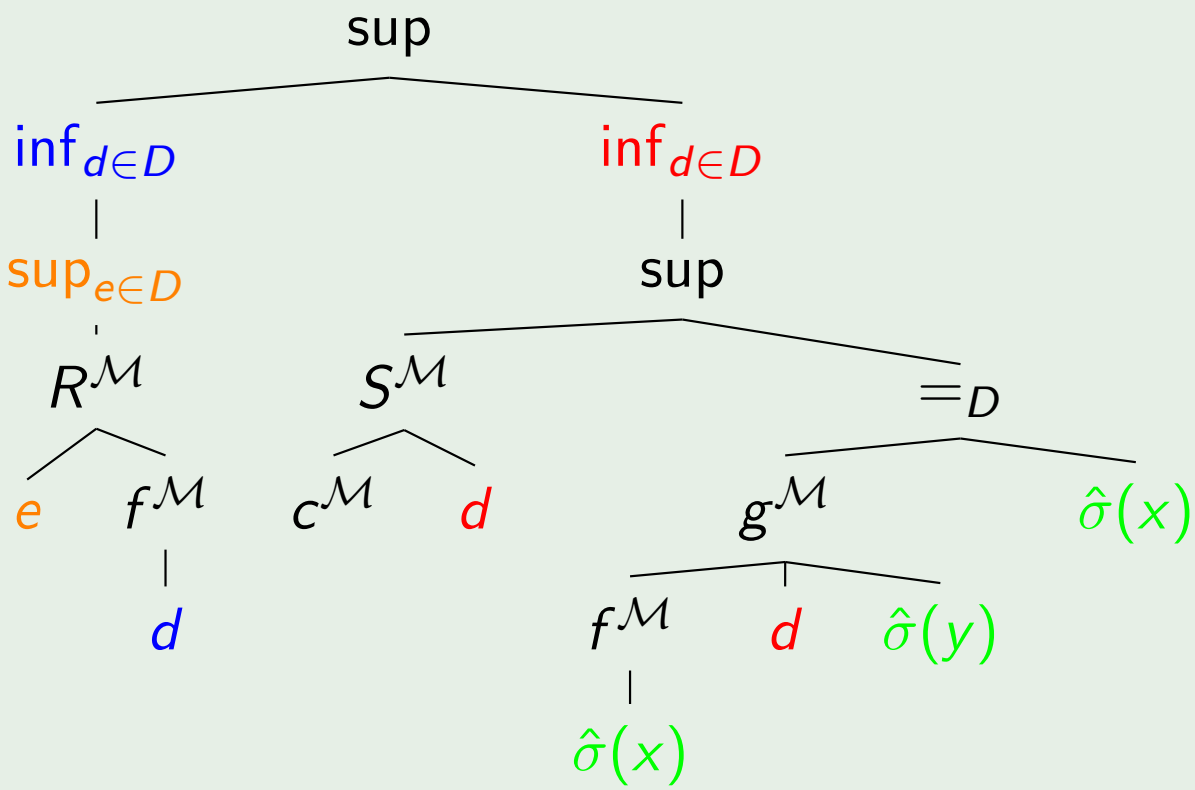
$$\hat{\sigma}(\exists x\, A) := \sup\{\, \widehat{\sigma\{x/d\}}(A) \;:\; d \in D \,\}$$

[a] Beachte die Reihenfolge! Die Funktion seht hier rechts.

Auswertung von Formeln in Baumform

Wir illustrieren die Auswertung einer Formel anhand des Beispiels auf ▸ Folie 191 für eine $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I \rangle$ und eine Belegung $\sigma \in D^{\mathcal{V}}$:

Beispiel ($\hat{\sigma}\big(\forall x\, \exists y\, R(y, f(x))\big) \vee \forall z\, \big(S(c, z) \vee g(f(x), z, y) \doteq x\big)$)



Beispiel (Modellierung eines konkreten Problems)

Eine Frau, die ein Porträt betrachtet, sagt: „Geschwister habe ich keine. Und die Mutter dieser Frau ist meiner Mutter Tochter.“ Wer ist da abgebildet?

Wir führen Konstanten $a_{/0}$ und $b_{/0}$ für die abgebildete Frau bzw. die Beobachterin ein. Weiter benötigen wir ein zweistelliges Geschwister-Prädikat $G_{/2}$, um die erste Aussage formulieren zu können:

$$\forall y.\, \neg G(b, y) \tag{1}$$

Analog könnte man ein Tochter-Prädikat T, oder auch ein Mutter-Prädikat M einführen. Es gibt aber eine bessere Lösung: eine einstellige Mutter-Operation $m_{/1}$. Das erlaubt die Formalisierung der zweiten Aussage mittels einer formalen Gleichung:

$$m(m(a) \doteq m(b) \tag{2}$$

Das liefert die Signatur $\mathcal{S} = \boldsymbol{Fun} + \boldsymbol{Pred} = \{a_{/0}, b_{/0}, m_{/1}\} + \{G_{/2}\}$.

Beispiel (Fortsetzung)

Die intendierte $\mathcal{S}$-Struktur $\mathcal{M}$ besteht aus der Menge D aller Menschen, mit zwei ausgewählten Frauen $a^{\mathcal{M}}$ und $b^{\mathcal{M}}$, der Mutter-Funktion $D \xrightarrow{m^{\mathcal{M}}} D$ und der Geschwister-Relation $G^{\mathcal{M}} \subseteq D \times D$.

Allgemein bekannte Tatsache über $\mathcal{M}$: Menschen ohne Geschwister sind die einzigen Kinder ihrer Mütter:

$$\forall y.\ \neg G(z, y) \to \forall u.\ \big(m(u) \doteq m(z) \to u \doteq z\big) \tag{3}$$

Was folgt aus den Prämissen (1), (2) und (3)? Modus Ponens mit (1) und der Instantiierung von (3) mit $z = b$ ergibt

$$\forall u.\ \big(m(u) \doteq m(b) \to u \doteq b\big) \tag{4}$$

was für $u = m(a)$ mit Modus Ponens aufgrund von (2) die Antwort liefert:

$$m(a) \doteq b$$

Beispiel (Graphentheorie)

Die Theorie gerichteter Graphen benötigt nur ein zweistelliges Prädikat $R_{/2}$ und keine Operatoren. In diesem Fall gilt

- $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}) = \boldsymbol{Term}(\emptyset, \mathcal{V}) = \mathcal{V}$, also immer $\check{\sigma} = \sigma \in D^{\mathcal{V}}$;
- $\boldsymbol{AT} = \{ R(x, y) \,:\, x, y \in \mathcal{V} \}$;
- $\mathcal{M}[\![R(x, y]\!](\sigma) = \overline{\sigma}\big(R(x, y)\big) \in \mathbb{B}$ besagt, ob es eine gerichtete Kante von $\sigma(x)$ nach $\sigma(y)$ gibt, oder nicht.

Schleifen (engl. "loops"), also Kanten von einem Knoten zu sich selbst, sind a priori erlaubt. Diese kann man mit $\forall x. \neg R(x, x)$ verbieten.

$$\forall x \forall y. \big(R(x, y) \land R(y, x) \to x \doteq y\big)$$

charakterisiert die gerichteten Graphen ohne 2-Zyklen. Ungerichtete Graphen erhält man, wenn R symmetrisch ist: zwei entgegengesetzte gerichtete Kanten $x \rightleftharpoons y$ entsprechen einer ungerichteten $x \text{ —— } y$.

$$\forall x \forall y. \big(R(x, y) \to R(y, x)\big)$$

Andere Sicht auf die Semantik der Terme/Formeln, $\mathcal{M}$ fest

Wir haben die Semantik der Terme/Formeln eingeführt als durch Belegungen $\sigma \in \mathcal{M}^{\mathcal{V}}$ parametrisierte Familien von Funktionen

$$D^{\mathcal{V}} \xrightarrow{(\breve{\_})} D^{\boldsymbol{Term}(\boldsymbol{Fun},\mathcal{V})} \quad , \quad \sigma \longmapsto \breve{\sigma}$$
$$D^{\mathcal{V}} \xrightarrow{(\hat{\_})} \mathbb{B}^{\boldsymbol{FO}(\mathcal{S})} \quad , \quad \sigma \longmapsto \hat{\sigma}$$

Uncurrying[15] liefert 2-stellige (Auswertungs-)Funktionen

$$\boldsymbol{Term}(\boldsymbol{Fun},\mathcal{V}) \times D^{\mathcal{V}} \xrightarrow{\mathcal{M}[\![-]\!](-)} D, \quad , \quad \langle t, \sigma\rangle \longmapsto \breve{\sigma}(t)$$
$$\boldsymbol{FO}(\mathcal{S}) \times D^{\mathcal{V}} \xrightarrow{\mathcal{M}[\![-]\!](-)} \mathbb{B}, \quad , \quad \langle A, \sigma\rangle \longmapsto \hat{\sigma}(A)$$

Die zweite entspricht der Auswertung $\mathcal{F}[\mathcal{A}] \times \mathbb{B}^{\mathcal{A}} \xrightarrow{\boldsymbol{E}} \mathbb{B}$ der AL ▸ Folie 19.

Die Wahl desselben Namens für beide Funktionen spiegelt deren strukturelle Ähnlichkeit wider; die sog. Semantik-Klammern $[\![\cdot]\!]$ sind kontext-sensitiv! Sie entstammen der denotationellen Semantik (Scott, Strachey 1971), und sind neben der formalen Semantik auch in der formalen Linguistik verbreitet.

[15] nach Haskell Brooks Curry (1900-1982)

Fieser Trick

Aufgrund der Symmetrie des cartesischen Produkts erhalten wir mittels currying weitere durch Terme/Formeln parametrisierte Familien

$$\begin{aligned}
\textbf{\textit{Term}}(\textbf{\textit{Fun}},\mathcal{V}) &\xrightarrow{\mathcal{M}[\![\cdot]\!]} D^{(D^{\mathcal{V}})} \;,\; t \longmapsto \mathcal{M}[\![t]\!] \quad,\; \sigma \longmapsto \mathcal{M}[\![t]\!](\sigma) := \check{\sigma}(t)\\
\textbf{\textit{FO}}(\mathcal{S}) &\xrightarrow{\mathcal{M}[\![\cdot]\!]} \mathbb{B}^{(D^{\mathcal{V}})} \;,\; A \longmapsto \mathcal{M}[\![A]\!] \quad,\; \sigma \longmapsto \mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A)
\end{aligned}$$

die ebenfalls Semantik der Terme/Formeln genannt werden!

[Der notationelle Aufwand erscheint disproportional hoch, und die Analogie zur AL wie wir sie eingeführt haben geht hier leider verloren. Dort mit der dualen Erfüllungsrelation $\boldsymbol{E}^{\mathrm{op}} \subseteq \mathbb{B}^{\mathcal{A}} \times \mathcal{F}[\mathcal{A}]$ zu arbeiten hätte zu einer durch Formeln parametrisierten Familie führen können, etwa

$$\mathcal{F}(\mathcal{A}) \xrightarrow{[\![\cdot]\!]} \mathbb{B}^{(\mathbb{B}^{\mathcal{A}})} \;,\; A \longmapsto [\![A]\!] \;,\; \varphi \longmapsto [\![A]\!](\varphi) = \hat{\varphi}(A)$$

Ob die AL damit leichter verständlich gewesen wäre, sei dahingestellt.]

Man muss obige in der Literatur verbreitete Notation aber beherrschen! Ggf. geben wir die einfachere Variante in orange an und nutzen sie in Beweisen.

$\mathcal{M}$ als Parameter

Bisher haben wir mit einer festen $\mathcal{S}$-Struktur $\mathcal{M}$ gearbeitet. Das erlaubt uns Begriffe weitgehend parallel zur AL auch für die PL_1 einzuführen.

Lässt man $\mathcal{M}$ zu variieren, erhält man die Familie der durch die möglichen $\mathcal{S}$-Strukturen $\mathcal{M}$ indizierten Auswertungsfunktionen

$$\boldsymbol{FO}(\mathcal{S}) \times D^{\mathcal{V}} \xrightarrow{\mathcal{M}[\![-]\!](-)} \mathbb{B} \quad \text{bzw.} \quad \boldsymbol{FO}(\mathcal{S}) \xrightarrow{\boldsymbol{E}_{\mathcal{M}}} D^{\mathcal{V}}$$

die zugehörigen Erfüllungsrelationen $\boldsymbol{E}_{\mathcal{M}}$, also die Urbilder von $1 \in \mathbb{B}$.

Die zugehörigen ▸ Hüllenoperatoren auf der Formelmenge $\boldsymbol{FO}(\mathcal{S})$ bezeichnen wir mit $(\ )^{\triangleright\triangleleft}_{\mathcal{M}} := \boldsymbol{E}^{\triangleleft}_{\mathcal{M}} \circ \boldsymbol{E}^{\triangleright}_{\mathcal{M}}$, und ihren Durchschnitt mit $(\ )^{\triangleright\triangleleft}$. Auch damit kann man weitgehend analog zur AL verfahren, die dortigen Beweise lassen sich oft übernehmen.

Allerdings gibt es hier keine Zerlegung der Form $(\ )^{\triangleright\triangleleft} := \boldsymbol{E}^{\triangleleft} \circ \boldsymbol{E}^{\triangleright}$ mehr! Man muss für allgemeines $\mathcal{M}$ argumentieren.

Logische Folgerung, Erfüllbar- und Allgemeingültigkeit

Definition (vergl. ▸ Folie 47)

Betrachte $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$ (aus sog. Prämissen) und $A, B \in \boldsymbol{FO}(\mathcal{S})$.

- Falls $A \in \Gamma^{\rhd\lhd}$ ($A \in \Gamma^{\rhd\lhd}_{\mathcal{M}}$), d.h., jeder Erfüller der Prämissen (in $\mathcal{M}$) erfüllt (in $\mathcal{M}$) auch A, schreibt man $\Gamma \models A$ ($\Gamma \models_{\mathcal{M}} A$) und sagt „A folgt (in $\mathcal{M}$) logisch aus Γ“.
- Ist A geschlossen (die Belegung der Variablen also irrelevant), so nennt man $\mathcal{M}$ ein Modell für A, sofern $\models_{\mathcal{M}} A$ gilt.
- A heißt allgemeingültig oder Tautologie, falls $\models A$ gilt.
- $A \in \boldsymbol{FO}(\mathcal{S})$ ($\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$) heißt erfüllbar, wenn eine $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I \rangle$ existiert mit $\boldsymbol{E}^{\rhd}_{\mathcal{M}}(\{A\}) \neq \emptyset$ (bzw. $\boldsymbol{E}^{\rhd}_{\mathcal{M}}(\Gamma) \neq \emptyset$).

Wieder ist A genau dann allgemeingültig, wenn $\neg A$ nicht erfüllbar ist.

Wir vermeiden Schreibweisen wie $\mathcal{M}, \sigma \models A$ oder $\mathcal{M} \models A$, falls A geschlossen, um $\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A) = 1$ auszudrücken.

Universeller und existentieller Abschluss

Im Folgenden werden wir oft an geschlossenen Formen (d.h., ohne freie Variablen) interessiert sein. Für jede Formel gibt es zwei sinnvolle Möglichkeiten, alle freien Variablen zu binden:

Definition

Für $A \in \boldsymbol{FO}(\mathcal{S})$ mit $\boldsymbol{FV}(A) \subseteq \{x_i : i < n\}$ heißt

$$\forall x_0 \ldots \forall x_{n-1}\, A \quad \text{bzw.} \quad \exists x_0 \ldots \exists x_{n-1}\, A$$

universeller bzw. existentieller Abschluss von A.

Lemma (HA!)

Für $A \in \boldsymbol{FO}(\mathcal{S})$ mit $\boldsymbol{FV}(A) \subseteq \{x_i : i < n\}$ ist die

- *Allgemeingültigkeit äq. zur Allgemeingültigkeit eines universellen Abschlusses;*
- *Erfüllbarkeit äq. zur Erfüllbarkeit eines existentiellen Abschlusses.* □

Charakterisierung der Unerfüllbarkeit von Formelmengen

Das folgende Ergebnis lässt sich nahezu analog zum entsprechenden Ergebnis der AL beweisen, lediglich für (2), (3) $\Rightarrow$ (0) muss man auf ein allgemeines $\mathcal{M}$ zurückgreifen.

Lemma

Folgende Bedingungen sind für $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$ *äquivalent:*

(0) Γ *ist unerfüllbar, bzw.* $\Gamma^{\triangleright}_{\mathcal{M}} = \emptyset$ *für alle* $\mathcal{S}$*-Strukturen* $\mathcal{M}$*.*

(1) $\Gamma \models A$ *für alle Formeln* A*, bzw.* $\Gamma^{\bowtie} = \boldsymbol{FO}(\mathcal{S})$*.*

(2) $\Gamma \models \bot$ *bzw.* $\bot \in \Gamma^{\bowtie}$*.*

(3) $\Gamma \models B$ *und* $\Gamma \models \neg B$ *bzw.* $\{B, \neg B\} \subseteq \Gamma^{\bowtie}$ *für ein* $B \in \boldsymbol{FO}(\mathcal{S})$*.* $\square$

Das semantische Deduktionstheorem (sDT) der PL_1

Lemma (sDT; Prämissen lassen sich zwischen $\models$ und $\to$ verschieben)

Für $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S}) \ni A, B$ *gilt:* $\Gamma \cup \{A\} \models B$ *gdw.* $\Gamma \models A \to B$.

Beweis.

($\Rightarrow$): Die Belegung $\sigma \in D^{\mathcal{V}}$ für eine $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I\rangle$ erfülle Γ.
$\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A) = 1$ impliziert nach Annahme $\mathcal{M}[\![B]\!](\sigma) = \hat{\sigma}(B) = 1$;
andernfalls gilt $\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A) = 0 \leq \mathcal{M}[\![B]\!](\sigma) = \hat{\sigma}(B)$.
In beiden Fällen ergibt sich $\mathcal{M}[\![A \to B]\!](\sigma) = \hat{\sigma}(A \to B) = 1$.

($\Leftarrow$): Für $\mathcal{M} = \langle D, I\rangle$ erfüllt $\sigma \in \boldsymbol{E}^{\triangleright}_{\mathcal{M}}(\Gamma \cup \{A\}) = \boldsymbol{E}^{\triangleright}_{\mathcal{M}}(\Gamma) \cap \boldsymbol{E}^{\triangleright}_{\mathcal{M}}(\{A\})$
sowohl Γ als auch A, nach Annahme also auch $A \to B$, und wegen
$\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A) \leq \mathcal{M}[\![B]\!](\sigma) = \hat{\sigma}(B)$ ebenfalls B. □

Korollar

$\{A, A \to B\} \models B$ bzw. $\Gamma \cup \{A\}$ unerfüllbar gdw. $\Gamma \models \neg A$. □

Modus Ponens, Modus Tollens und Kontraposition

Die aus den deduktiven Systemen der AL bekannten Typen von Schlussfolgerungen und Ähnliche gelten natürlich auch in der Semantik sowohl der AL als auch der PL_1:

Satz

Betrachte $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S}) \ni A, B$.

- $\Gamma \models B$ *und* $\Gamma \models B \to A$ *impliziert* $\Gamma \models A$ *(Modus Ponens).*
- $\Gamma \models B \to A$ *und* $\Gamma \models \neg A$ *impliziert* $\Gamma \models \neg B$ *(Modus Tollens).*
- $\Gamma \models B \to A$ *und* $\Gamma \models A$ *impliziert i.A. nicht* $\Gamma \models B$ *(Modus Bogus).*
- $\Gamma \cup \{B\} \models \neg A$ *gdw.* $\Gamma \cup \{A\} \models \neg B$ *(Kontraposition).*

Beweis.

HA. □

Die kanonische Halbordnung und Äquivalenz auf $\mathbf{FO}(\mathcal{S})$

Definition

Für eine $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I \rangle$ und jede Belegung $\sigma \in D^{\mathcal{V}}$ induziert $\mathbf{FO}(\mathcal{S}) \xrightarrow{\hat{\sigma}} \mathbb{B}$ eine Quasiordnung $\sqsubseteq_{\mathcal{M},\sigma}$ auf $\mathbf{FO}(\mathcal{S})$; $\sqsubseteq_{\mathcal{M}}$ bezeichnet analog zur AL den Durchschnitt über alle $\sigma \in \mathcal{M}^{\mathcal{V}}$, und $\sqsubseteq$ den Durchschnitt aller $\sqsubseteq_{\mathcal{M}}$, $\mathcal{M}$ eine $\mathcal{S}$-Struktur.

Wie gewohnt bezeichnen wir die resultierenden ÄR'n mit $\models\!\!\models_{\mathcal{M}}$ bzw. $\models\!\!\models$.

Lemma

Für $A, B \in \mathbf{FO}(\mathcal{S})$ und eine $\mathcal{S}$-Struktur $\mathcal{M}$ sind äquivalent:

1. $B \sqsubseteq_{\mathcal{M}} A$
2. $\{B\} \models_{\mathcal{M}} A$
3. $\mathcal{M}[\![B \to A]\!](\sigma) = 1$ *bzw.* $\hat{\sigma}(B \to A) = 1$ *für alle* $\sigma \in D^{\mathcal{V}}$

Weiter gilt natürlich $B \sqsubseteq A$ gdw. $\{B\} \models A$. □

Freie Variablen – das Koinzidenzlemma

In der AL konnte die Bewertung $\hat{\varphi}(B)$ einer Formel B nur von den Werten $\varphi(p)$ derjenigen Atome p abhängen, die in B vorkommen.

Entsprechend kann in der PL_1 die Semantik $\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A)$ nur von den Werten der in A auftretenden atomaren Formeln abhängen, und somit nur von den σ-Werten der in diesen vorkommenden Variablen.

Gebundene Auftreten von Variablen nehmen dabei keinen Einfluss; daher können gebundene Variablen auch umbenannt werden (vergl. ▸ Folie 216):

Lemma (Koinzidenzlemma)

Ist A eine PL_1-Formel über der Signatur $\mathcal{S}$ und $\mathcal{M} = \langle D, I\rangle$ eine $\mathcal{S}$-Struktur, so gilt für alle Belegungen $\sigma, \tau \in \mathcal{M}^{\mathcal{V}}$ mit $\sigma(x) = \tau(x)$ für alle $x \in \boldsymbol{FV}(A)$:

$$\mathcal{M}[\![A]\!](\sigma) = \mathcal{M}[\![A]\!](\tau) \qquad \textit{bzw.} \qquad \hat{\sigma}(A) = \hat{\tau}(A)$$ □

Rechenregeln für Quantoren

Neben den Rechenregeln auf ▸ Folie 65 und ▸ Folie 66 gelten weitere Regeln für die mittels Quantoren gebildeten neuen unären Junktoren $\forall x$ und $\exists x$:

Lemma (weitere logische Äquivalenzen in $\boldsymbol{FO}(\mathcal{S})$)

$$\neg\forall x\, A \models\!\dashv \exists x\, \neg A \qquad \neg\exists x\, A \models\!\dashv \forall x\, \neg A \tag{1}$$

$$\forall x\, A \wedge \forall x\, B \models\!\dashv \forall x\,(A \wedge B) \qquad \exists x\, A \vee \exists x\, B \models\!\dashv \exists x\,(A \vee B) \tag{2}$$

$$\forall x\, \forall y\, A \models\!\dashv \forall y\, \forall x\, A \qquad \exists x\, \exists y\, A \models\!\dashv \exists y\, \exists x\, A \tag{3}$$

(1) verallgemeinert De Morgan. *Sofern* $x \notin \boldsymbol{FV}(A)$ *gilt weiterhin*

$$A \star \mathcal{Q}x\, B \models\!\dashv \mathcal{Q}x\,(A \star B) \quad \text{für } \mathcal{Q} \text{ Quantor und } \star \in \{\wedge, \vee, \to\} \tag{4}$$

$$\mathcal{Q}x\, B \to A \models\!\dashv \bar{\mathcal{Q}}x\,(B \to A) \quad \text{für } \mathcal{Q} \text{ und } \bar{\mathcal{Q}} \text{ komplementär} \tag{5}$$

Aber i.A.

$$\forall x\, A \vee \forall x\, B \not\models\!\dashv \forall x\,(A \vee B) \qquad \exists x\, A \wedge \exists x\, B \not\models\!\dashv \exists x\,(A \wedge B) \qquad \forall x\, \exists y\, A \not\models\!\dashv \exists y\, \forall x\, A$$

Beispiel

Betrachte die Signatur $\mathcal{S} = \{0_{/0}, 1_{/0}, +_{/2}\} + \{\leq_{/2}\}$ und die $\mathcal{S}$-Struktur $\mathbb{N}$ mit der üblichen Interpretation der Operatoren und des Prädikats.

$$(\mathrm{a}) \quad \forall x.\, \exists y.\, (x + 1 \leq y) \quad \text{bzw.} \quad (\mathrm{b}) \quad \exists y.\, \forall x.\, (x + 1 \leq y)$$

bedeutet, dass es (a) zu jeder Zahl eine echt größere gibt (korrekt), bzw. dass es (b) eine größte Zahl gibt (falsch).

1-stellige Prädikate, die gerade bzw. ungerade Zahlen charakterisieren,

$$G(x) := \exists y.\, (y + y \doteq x) \quad \text{sowie} \quad U(x) := \exists y.\, (y + y \doteq x + 1)$$

erlauben es ausdrücken, dass (c) alle Zahlen gerade oder ungerade sind (korrekt), und dass (d) alle Zahlen gerade sind, oder alle Zahlen ungerade sind (falsch), formal

$$(\mathrm{c}) \quad \forall x.\, \big(G(x) \vee U(x)\big) \quad \text{bzw.} \quad (\mathrm{d}) \quad \forall x.\, G(x) \vee \forall x.\, U(x)$$

Substitutionen via Rekursionssatz und Modifikationen

Folie 197 definierte die Modifikation einer Belegung $\sigma \in \mathcal{M}^{\mathcal{V}}$ durch $\langle x, d\rangle \in \mathcal{V} \times D$. Betrachte speziell den Fall $D = \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$:

Definition

Für das Bild $\mathcal{V} \xrightarrow{\vartheta} \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$ der Inklusion $\mathcal{V} \xrightarrow{\iota} \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$ unter endlich vielen Modifikationen

$$(\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}))^{\mathcal{V}} \xrightarrow{\{x_i/t_i\}} (\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}))^{\mathcal{V}} \qquad i \in k$$

bezeichnet man die eindeutigen Fortsetzungen

- $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}) \xrightarrow{\check{\vartheta}} \boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$
- $\boldsymbol{AT} \xrightarrow{\overline{\vartheta}} \boldsymbol{AT}$
- $\boldsymbol{FO}(\mathcal{S}) \xrightarrow{\hat{\vartheta}} \boldsymbol{FO}(\mathcal{S})$

als Substitution auf Termen, bzw. (atomaren) Formeln. Insbesondere ist ϑ eine Belegung, die an nur endlich vielen Stellen von der Identität abweicht.

Anwendung von Substitutionen

Wendet man Substitutionen von rechts an, braucht man $\check{\vartheta}$, $\overline{\vartheta}$ und $\hat{\vartheta}$ nicht länger zu unterscheiden. **Achtung:** Sie binden stärker als Junktoren!

Bei der Anwendung auf quantifizierte Formeln dürfen aber keine neuen Bindungen erzeugt werden. Ggf. sind gebundene Variable umzubenennen:

Definition (Vereinheitlichte Substitutionen)

$$
\begin{aligned}
x\vartheta &:= \check{\vartheta}(x) \\
f(t_0, \ldots, t_{\mathcal{S}(f)-1})\vartheta &:= f(t_0\vartheta, \ldots, t_{\mathcal{S}(f)-1}\vartheta) \\
(t_0 \doteq t_1)\vartheta &:= t_0\vartheta \doteq t_1\vartheta \\
R(t_0, \ldots, t_{\mathcal{S}(R)-1})\vartheta &:= R(t_0\vartheta, \ldots, t_{\mathcal{S}(R)-1}\vartheta) \\
(\neg A)\vartheta &:= \neg(A\vartheta) \\
(A \star B)\vartheta &:= A\vartheta \star B\vartheta \\
(\mathcal{Q}x\, A)\,\vartheta &:= \mathcal{Q}y\,(A\{x/y\}\vartheta) \quad \text{mit } y = y\vartheta \notin \boldsymbol{V}(A)
\end{aligned}
$$

($\{x/y\}\vartheta$ ist natürlich die Komposition von Modifikationen, die erst $\{x/y\}$ und dann ϑ anwendet.)

Substitutionslemma und gebundene Umbenennung

Der Zusammenhang zwischen Substitutionen (in Termen und (atomaren) Formeln) und der Modifikation von Belegungen nimmt nun folgende eindrucksvolle Form an, was sich mittels Induktion über den Aufbau von Termen bzw. Formeln zeigen lässt (gute Übung für die Notation!):

Lemma (Substitutionslemma, für Terme und Formeln)

$\mathcal{M}[\![@\{x/t\}]\!](\vartheta) = \mathcal{M}[\![@]\!](\vartheta\{x/\mathcal{M}[\![t]\!](\vartheta)\})$ $\quad \tilde{\vartheta}(@\{x/t\}) = \vartheta\{x/\breve{\sigma}(t)\}(@)$ □

Korollar.

1. Ist $A \in \boldsymbol{FO}(\mathcal{S})$ allgemeingültig, dann auch $A\{x/t\}$.
2. Die Formel $\forall x\ A \to A\{x/t\}$ ist allgemeingültig. □

Lemma (Gebundene Umbenennung erhält Äquivalenz)

$\mathcal{Q}x\,A \rightleftharpoons \mathcal{Q}y\,A\{x/y\}$ *sofern* $y \notin \boldsymbol{FV}(A)$. □

PL_1-spezifische Folgerungen

Satz

1. *Falls* $x \notin \boldsymbol{FV}(\Gamma)$ *für* $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$, *dann* $\Gamma \models A$ *gdw.* $\Gamma \models \forall x\, A$ *(Generalisierung);*
   – *insbesondere* $\{A\} \models \forall x\, A$, *bzw.* $\models A \to \forall x\, A$, *sofern* $x \notin \boldsymbol{FV}(A)$.
2. *Entsteht* A' *aus* A *durch erlaubte (beachte Quantoren!) Ersetzung einiger (nicht notwendig aller) Vorkommen von* x *durch* y, *dann* $\models \forall x \forall y \left(x \doteq y \to (A \leftrightarrow A')\right)$ *(Variante der Kongruenz)*

Beweis

1. Aus $\Gamma \models \forall x\, A$ folgt trivialerweise $\Gamma \models A$. Umgekehrt betrachte eine $\mathcal{S}$-Struktur $\mathcal{M} = \langle D, I\rangle$ und $\sigma \in D^{\mathcal{V}}$, so dass Γ, und somit A, davon erfüllt wird. Nach Voraussetzung gilt dasselbe dann auch für jedes $\sigma\{x/d\}$, $d \in D$, und somit $\mathcal{M}[\![\forall x\, A]\!](\sigma) = \hat{\sigma}(\forall x\, A) = 1$.
2. HA! □

Beispiele

Beispiel (Falls $x \in \boldsymbol{FV}(A)$ braucht $\{A\} \models \forall x\, A$ nicht zu gelten.)

Für $\mathcal{S} = \{R_{/1}\}$ und $A = R(x)$ betrachte $\mathcal{M} = \langle \{0,1\}, R^{\mathcal{M}} = \{0\} \rangle$.
Die konstante 0-wertige Belegung $\sigma \in \{0,1\}^{\mathcal{V}}$ liefert

$$\mathcal{M}[\![A]\!](\sigma) = \hat{\sigma}(A) = 1 \quad \text{aber} \quad \mathcal{M}[\![\forall x\, A]\!](\sigma) = \hat{\sigma}(\forall X\, A) = 0$$

da auch $\mathcal{M}[\![A]\!](\sigma\{x/1\}) = \hat{\sigma}\{x/1\}(A) = 0$ im Infimum auftritt.

Beispiel ($\{\forall x\, A\} \models A$)

Wende DT mit $t = x$ auf $\forall x\, A \to A\{x/t\}$ an, was nach ▸ Folie 216 allgemeingültig ist.

Beispiel ($\models \exists x\ (R(x) \to \forall x\, R(x))$)

Die Formel wird in $\mathcal{M}$ mit 1 bewertet, wenn $R^{\mathcal{M}}(d) = 0$ für ein, oder wenn $R^{\mathcal{M}}(d) = 1$ für jedes $d \in D$ gilt, also immer

Lemma

$\forall x(A \to B)\} \models \forall x\, A \to \forall x\, B$

Beweis.

$$
\begin{array}{lrl}
 & \{\forall x\, A, \forall x\,(A \to B)\} \models \forall x\, B & \text{wegen DT zu zeigen} \\
\text{gdw.} & \{(\forall x\, A) \wedge \forall x\,(A \to B)\} \models \forall x\, B & \Gamma \text{ endlich} \\
\text{gdw.} & \{\forall x\,\big(A \wedge (A \to B)\big)\} \models \forall x\, B & \text{nach Rechenregel (2)} \\
\text{gdw.} & \{\forall x\,\big(A \wedge (\neg A \vee B)\big)\} \models \forall x\, B & \text{gemäß AL} \\
\text{gdw.} & \{\forall x\,\big((A \wedge \neg A) \vee (A \wedge B)\big)\} \models \forall x\, B & \text{Distributivität} \\
\text{gdw.} & \{\forall x\,\big(\bot \vee (A \wedge B)\big)\} \models \forall x\, B & \text{gemäß AL} \\
\text{gdw.} & \{\forall x\,(A \wedge B)\} \models \forall x\, B & \bot \text{ neutral bzgl. } \vee \\
\text{gdw.} & \{(\forall x\, A) \wedge \forall x\, B\} \models \forall x\, B & \text{nach Rechenregel (2)} \\
\text{gdw.} & \{\forall x\, A, \forall x\, B\} \models \forall x\, B & \text{trivial und korrekt}
\end{array}
$$

Lemma

$\models \exists x \forall y\, A \to \forall y \exists x\, A$

Beweis.

	$\{\exists x \forall y\, A\} \models \forall y \exists x\, A$		wegen DT zu zeigen
gdw.	$\{\exists x \forall y\, A\} \models \exists x\, A$		Generalisierung
gdw.	$\{\neg \exists x\, A\} \models \neg \exists x \forall y\, A$		Kontraposition
gdw.	$\{\forall x\, \neg A\} \models \forall x\, \neg \forall y\, A$		de Morgan
gdw.	$\{\forall x\, \neg A\} \models \neg \forall y\, A$		Generalisierung
gdw.	$\{\forall x\, \neg A, \forall y\, A\} \models \bot$		Korollar zu DT
gdw.	$\{\exists y \forall x\, \neg A, \exists x \forall y\, A\} \models \bot$		existentieller Abschluß
gdw.	$\exists y \forall x\, \neg A \wedge \exists x \forall y\, A$ nicht erfüllbar		offensichtlich wahr

□

Beispiel ($\models \forall x\, \forall y\, (x \doteq y \to (f(x,y) = g(x) \leftrightarrow f(y,y) = g(x))$)

Dies gilt immer, wenn $\mathcal{S}$ Funktionssymbole $f_{/2}$ und $g_{/1}$ enthält.

Kapitel 11
Skolem, Herbrand, Gödel

Bereinigte Formeln

Die Syntax der PL enthält einige Fallstricke, die man explizit umgehen muss.

Definition

$A \in \boldsymbol{FO}(\mathcal{S})$ heißt bereinigt, sofern

- keine Variable frei und gebunden auftritt, d.h., $\boldsymbol{FV}(A) \cap \boldsymbol{GV}(A) = \emptyset$;
- jede Variable höchstens einmal gebunden wird.

Durch iterierte gebundene Umbenennung lässt sich jede Formel bereinigen:

Lemma

Zu jeder Formel $A \in \boldsymbol{FO}(\mathcal{S})$ *existiert eine bereinigte Formel* $B \in \boldsymbol{FO}(\mathcal{S})$ *mit* $A \models\!\!\models B$. □

Wie ▸ Rechenregel (2) für Quantoren zeigt, braucht die Konjunktion/Disjunktion bereinigter Formeln nicht mehr bereinigt zu sein. Ungeschicktes Bereinigen kann Formeln aber auch unnötig aufblähen (s.u.).

Normalformen

Die ▸ Motivation, Formeln der AL in bestimmte „Normalformen“ zu überführen, greift auch hier. Ziel sind einfachere Beweise und effizientere Algorithmen. Dabei werden wir überwiegend an geschlossenen Formeln interessiert sein.

Auf ▸ Folie 206 haben wir bereits gesehen, wie Formeln durch äußere Quantoren sinnvoll in geschlossenen Formeln überführt werden können.

Wir betrachten nun zwei Typen von PL-Normalformen:

- Pränex-Normalform: alle Quantoren außen; diese erweisen sich als logisch äquivalent zur Ausgangsformel;
- Skolem-Normalform: PNF ohne $\exists$; diese sind i.A. nur „erfüllbarkeitsäquivalent“ zur Ausgangsformel, wobei dieser Begriff ähnlich wie in der AL noch genau gefasst werden muss.

Pränex- und Skolem[16] Normalform

Definition

Ist B quantorenfrei, so nennt man eine Formel $A \in \boldsymbol{FO}(\mathcal{S})$ der Form

$$\mathcal{Q}_0 y_0 \, \mathcal{Q}_1 y_1 \ldots \mathcal{Q}_{n-1} y_{n-1} B$$

mit $\mathcal{Q}_i \in \{\forall, \exists\}$, $i < n$, eine

- Pränex-Normalform oder kurz PNF; im bereinigten Fall spricht man von einer bPNF;
- Skolem-Normalform oder kurz SNF, falls es sich um eine bPNF ohne Existenzquantoren handelt, d.h., $\mathcal{Q}_i = \forall$ für $i < n$

Mit den ▸ logischen Äquivalenzen für quantifizierte Formeln sieht man leicht (HA):

Satz

Jede Formel $A \in \boldsymbol{FO}(\mathcal{S})$ besitzt eine äquivalente Formel B in bPNF. □

[16] Thoralf Albert Skolem (1887–1963)

Vorüberlegung zur Skolemisierung

Ziel: Die Existenzquantoren aus einer zur geschlossenen Formel $A \in \boldsymbol{FO}(\mathcal{S})$ äquivalenten bPNF

$$B = \mathcal{Q}_0 y_0 . \mathcal{Q}_1 y_1 . \ldots . \mathcal{Q}_{n-1} y_{n-1} . C \qquad \text{mit } C \text{ quantorenfrei}$$

so entfernen (am zweckmäßigsten, aber nicht zwingend, von außen nach innen), dass die resultierende Formel genau dann erfüllbar ist, wenn das für B gilt. Wie in der AL erhält man tatsächlich eine strengere Bedingung.

Der erste Existenzquantor möge in Position k von links auftreten:

$$B = \forall y_0 \ldots \forall y_{k-1} . \exists y_k . F \quad \text{mit} \quad F = \mathcal{Q}_{k+1} y_{k+1} \ldots \mathcal{Q}_{n-1} y_{n-1} . C$$

In jedem Modell $\mathcal{M} = \langle D, I \rangle$ für B gilt nun für alle $\boldsymbol{u} \in D^k$ und $\sigma \in D^{\mathcal{V}}$

$$\widehat{\sigma\{\boldsymbol{y}/\boldsymbol{u}\}}\,(\exists y_k . F) = \sup\big\{\widehat{\sigma\{\boldsymbol{y}/\boldsymbol{u}\}\{y_k/d\}}(F) \,:\, d \in D\big\} = 1$$

Unabhängig von σ wird $\boldsymbol{y} \in \mathcal{V}^k$ auf $\boldsymbol{u}$ abgebildet.

Zusammenhang mit (AC)

Also definiert $\sigma\{\boldsymbol{y}/\widehat{\boldsymbol{u}}\}\{y_k/d\}(F) = 1$ eine totale Relation $D^k \xrightarrow{R_B} D$.

- Dem Skolemisierungs-Trick (s.u.) liegt folgende Annahme zugrunde:

  Aus jeder totalen Relation $X \xrightarrow{R} Y$ zwischen Mengen X und Y lässt sich eine Funktion $X \xrightarrow{r} Y$ extrahieren.

  Man beachte, dass kein Algorithmus dafür angegeben wird!
- Obige Annahme ist im Wesentlichen das sog. Auswahlaxiom (AC) der Mengenlehre, das unabhängig von den übrigen Zermelo-Fraenkel Axiomen ist (vergl. Anhang C), d.h., es gibt Modelle der übrigen Axiome, die (AC) erfüllen, und solche, die das nicht tun.
- So harmlos und einleuchtend wie (AC) aussieht, hat es doch einige verblüffende und beunruhigende Konsequenzen, wie z.B. das ▸ Banach-Tarski Paradox. Insofern ist es immer ratsam zu wissen, welche Beweise oder Definitionen (AC) voraussetzen.

Skolems Trick:

formal einen funktionalen Zusammenhang zwischen $\boldsymbol{y} \in \mathcal{V}^k$, und y_k schaffen mittels eines frischen Funktionssymbols $f_{/k} \notin \boldsymbol{Fun}$: statt

$$B = \forall y_0 \ldots \forall y_{k-1}. \exists y_k. F$$

betrachte nun in $\boldsymbol{FO}\big((\boldsymbol{Fun} + \{f_{/k}\}) + \boldsymbol{Pred}\big)$ die Formel

$$B' = \forall y_0 \ldots \forall y_{k-1}. F\{y_k / f(y_0, \ldots, y_{k-1})\}$$

- Sofern (AC) gilt, lässt sich jedes Modell $\mathcal{M}$ für B zu einem Modell $\mathcal{M}'$ für B' erweitern, indem man aus der oben beschriebenen totalen Relation $D^k \xrightarrow{R} D$ eine Interpretation von $f_{/k}$ extrahiert.
- Umgekehrt wird aus jedem Modell $\mathcal{M}'$ für B' durch Entfernen von $f^{\mathcal{M}}$ aus I' ein Modell für B, denn $D^k \xrightarrow{f^{\mathcal{M}}} D$ ist als totale Relation in $D^k \xrightarrow{R_B} D$ enthalten, so dass die relevanten Suprema den Wert 1 annehmen.

Anmerkungen

- Jede ***FO***-Formel A hat eine minimale endliche Signatur $\mathcal{S}_A$, in der sie formuliert werden kann. (Mittels Signaturhomomorphismen ließe sich das Problem change-of-signature sauber beschreiben; wir wollen das hier aber nicht vertiefen.)
- Weder Reihenfolge noch Zahl der Quantoren einer (b)PNF B von $A \in \boldsymbol{FO}(\mathcal{S})$ müssen eindeutig bestimmt sein (siehe ▸ Rechenregeln):
  - Anstelle von (2) liefert Bereinigen und zweimalige Anwendung von (4) eine äquivalente Formel mit zwei Quantoren $\Rightarrow$ ungeschickt!
  - Falls in (4) die Formel A die Form $\mathcal{Q}'y.\, C$ mit $y \notin \boldsymbol{FV}(B)$ hat, dann können $\mathcal{Q}x$ und $\mathcal{Q}'y$ in beliebiger Reihenfolge nach außen gezogen werden, selbst wenn es sich um verschiedene Quantoren handelt. (Tipp: erst $\exists$ nach außen ziehen, dann $\forall$.)
- Folglich braucht die Signatur für Skolemisierungen von $A \in \boldsymbol{FO}(\mathcal{S})$ nicht eindeutig bestimmt zu sein. Hat man die Wahl, sind Skolemsymbole möglichst geringer Stelligkeit vorzuziehen (daher obiger Tipp).

Beispiel (leicht psychedelisch)

Für eine Signatur $\mathcal{S} = \{R_{/1}, S_{/2}\}$ ohne Operatoren betrachte

$$A = \forall x.\exists y.\big(S(x,y) \vee R(y)\big) \quad \text{und} \quad B = \exists y.\forall x.\big(R(y) \to S(y,x)\big)$$

Um $A \wedge B$ zu bereinigen benennt man in B die gebundenen Variablen um.

$$\begin{aligned}
A \wedge B &\bowtie \forall x.\exists y.\big(S(x,y) \vee R(y)\big) \wedge \exists u.\forall v.\big(R(u) \to S(u,v)\big) \\
&\bowtie \forall x.\exists y.\Big(\big(S(x,y) \vee R(y)\big) \wedge \exists u.\forall v.\big(R(u) \to S(u,v)\big)\Big) \\
&\bowtie \forall x.\exists y.\exists u.\forall v.\Big(\big(S(x,y) \vee R(y)\big) \wedge \big(R(u) \to S(u,v)\big)\Big) \quad (0) \\
\text{(wg. Symmetrie)} \quad &\bowtie \exists u.\forall v.\forall x.\exists y.\Big(\big(S(x,y) \vee R(y)\big) \wedge \big(R(u) \to S(u,v)\big)\Big) \quad (1) \\
&\bowtie \forall x.\Big(\exists y.\big(S(x,y) \vee R(y)\big) \wedge \exists u.\forall v.\big(R(u) \to S(u,v)\big)\Big) \\
&\bowtie \forall x.\exists u.\forall v.\exists y.\Big(\big(S(x,y) \vee R(y)\big) \wedge \big(R(u) \to S(u,v)\big)\Big) \quad (2) \\
\text{(wg. Symmetrie)} \quad &\bowtie \exists u.\forall x.\exists y.\forall v.\Big(\big(S(x,y) \vee R(y)\big) \wedge \big(R(u) \to S(u,v)\big)\Big) \quad (3)
\end{aligned}$$

Die markierten Formeln (0)–(3) liegen in bPNF vor.

Beispiel (Fortsetzung)

Elimination der Existenzquantoren durch Hinzufügen geeigneter Funktionssymbole zur ***Fun***-Komponente der Signatur liefert Skolem-Formeln unterschiedlicher Länge:

$$S_0 = \forall x.\forall v.\Big(\big(S(x, f(x)) \vee R(f(x))\big) \wedge \big(R(g(x)) \to S(g(x), v)\big)\Big) \;;\; f_{/1}\,,\, g_{/1}$$

$$S_1 = \forall v.\forall x.\Big(\big(S(x, h(v, x)) \vee R(h(v, x))\big) \wedge \big(R(c) \to S(c, v)\big)\Big) \;;\; h_{/2}\,,\, c_0$$

$$S_2 = \forall x.\forall v.\Big(\big(S(x, k(x, v)) \vee R(k(x, v))\big) \wedge \big(R(g(x)) \to S(g(x), v)\big)\Big) \;;\; k_{/2}\,,\, g_{/1}$$

$$S_3 = \forall x.\forall v.\Big(\big(S(x, f(x)) \vee R(f(x))\big) \wedge \big(R(c) \to S(c, v)\big)\Big) \;;\; f_{/1}\,,\, c_{/0}$$

Dies zeigt, dass Skolemisierung je nach Vorgehensweise unterschiedliche Ergebnisse liefern kann, speziell wenn verschiedene Leute dieselbe Formel bearbeiten. A priori ist nicht klar, ob es eine Strategie gibt, die immer zum kürzesten Ergebnis führt.

Man beachte weiterhin, dass jede der Formeln $A \wedge B$ und S_i, $i < 3$ natürlich erfüllbar ist, wenn eine der anderen diese Eigenschaft hat.

Erfüllbarkeitsäquivalenz in der PL_1

Bisher haben wir in der PL_1 den Begriff der Erfüllbarkeitsäquivalenz in Anführungszeichen gesetzt, bis klar war, wofür wir ihn wirklich brauchen.

- Das Hauptaugenmerk wird auf geschlossenen Formeln liegen, deren Semantik unabhängig von konkreten Belegungen $\sigma \in D^{\mathcal{V}}$ ist.
- Daher lässt sich der aussagenlogische Begriff nicht einfach übertragen.
- Erfüllbarkeitsäquivalente Formeln können unterschiedliche Minimalsignaturen haben, wie etwa eine geschlossene Formel und ihre nicht notwendig eindeutig bestimmten Skolemisierungen.
- Ersetzt man in einer Formelmenge Γ alle Formeln durch erfüllbarkeitsäquivalente Varianten soll die (Nicht-) Erfüllbarkeit von Γ invariant bleiben!

Wir führen die Erfüllbarkeitsäquivalenz nun mit Hilfe einer einfach beschreibbaren Quasiordnung zwischen geschlossenen Formeln ein.

Definition

Geschlossene Formeln A, B, deren Minimalsignaturen $\boldsymbol{Fun}_A \subseteq \boldsymbol{Fun}_B$ und $\boldsymbol{Pred}_A = \boldsymbol{Pred}_B$ erfüllen, genügen der Relation $A \sqsubseteq_e B$, sofern

- jedes $\mathcal{S}_A$-Modell für A zu einem $\mathcal{S}_B$-Modell für B erweitert werden kann;
- jedes $\mathcal{S}_B$-Modell für B eingeschränkt auf $\boldsymbol{Fun}_A$ bereits ein $\mathcal{S}_A$-Modell für A ist.

Erfüllbarkeitsäquivalenz $\models\!\!\!\models_e$ ist nun die kleinste ÄR, die $\sqsubseteq_e \cup (\sqsubseteq_e)^{\mathrm{op}}$ enthält, d.h., die transitive Hülle von $\sqsubseteq_e \cup (\sqsubseteq_e)^{\mathrm{op}}$.

Es ist unmittelbar klar, dass $\sqsubseteq_e$ reflexiv und transitiv und damit eine Quasiordnung ist. Weiter liegt jede geschlossene Formel bzgl. $\sqsubseteq_e$ automatisch unter jeder ihrer Skolemisierungen.

Skolemisierung

Algorithmus (informell)

Gegeben $A \in \boldsymbol{FO}(\mathcal{S})$.

- Bestimme eine zu A äquivalente bPNF B; dabei
  - bevorzuge ▸ Regel (2) gegenüber Bereinigung und Regel (4);
  - bei Anwendung von Regel (4) auf zwei Formeln mit verschiedenen ersten Quantoren, ziehe Existenz-Quantoren vor All-Quantoren nach außen, um die Stelligkeit der später benötigten Skolem-Symbole zu minimieren.
- Eliminiere die Existenz-Quantoren der resultierenden bPNF (etwa von links nach rechts) mit Hilfe passender Skolem-Symbole, die zur Signatur hinzuzufügen sind.

Satz (HA)

Jede bPNF-Formel $B \in \boldsymbol{FO}(\mathcal{S})$ ist zu ihrer Skolemisierung in $\boldsymbol{FO}(\mathcal{S} + \boldsymbol{Sko})$ erfüllbarkeitsäquivalent. □

Das Allgemeingültigkeitsproblem $\mathrm{AGP}(\mathrm{PL})$

Gegeben:	$A \in \boldsymbol{FO}(\mathcal{S})$
Frage:	ist A allgemeingültig?
oder äquivalent:	ist $\neg A$ unerfüllbar?

Es wird sich zeigen, dass tatsächlich ein Algorithmus existiert, der in endlich vielen Schritten terminiert, wenn $\neg A$ unerfüllbar ist, aber andernfalls nicht notwendig terminiert. In der Tat existiert kein Algorithmus, der in jedem Fall terminiert. Daher ist $\mathrm{AGP}(\mathrm{PL})$, genau wie das Unerfüllbarkeitsproblem für Formelmengen der AL, nur semi-entscheidbar (TheoInf 2).

Das offensichtliche Problem bei der Suche nach einem Algorithmus ist die Vielzahl der möglichen $\mathcal{S}$-Strukturen $\mathcal{M} = \langle D, I \rangle$:

- die Mächtigkeit des Datenbereichs D ist unbeschränkt;
- wir haben keine Information über I.

Vorarbeiten zur Herbrand[17] Theorie: Konstanten

Herbrands Kernaussage: ein kanonischer Datenbereich genügt, sofern

- die Signatur $\mathcal{S} = \boldsymbol{Fun} + \boldsymbol{Pre}$ mindestens eine Konstante enthält,
- und das Symbol $\doteq$ in A nicht vorkommt, geschrieben $A \in \boldsymbol{FO}^{\neq}(\mathcal{S})$.

Beide Bedingungen lassen sich immer erzwingen (Signaturen sind variabel):

Da die Datenbereiche unserer $\mathcal{S}$-Strukturen nicht leer sein dürfen, können wir einer Signatur $\mathcal{S} = \boldsymbol{Fun} + \boldsymbol{Pre}$ ohne Konstanten in $\boldsymbol{Fun}$ problemlos eine Konstante c hinzufügen. Offenbar gilt $\boldsymbol{FO}(\mathcal{S}) \subseteq \boldsymbol{FO}(\mathcal{S} + \{c_{/0}\})$

Andererseits gibt es zu jedem $B \in \boldsymbol{FO}(\mathcal{S} + \{c_{/0}\})$ eine kanonische erfüllbarkeitsäquivalente Formel in $\boldsymbol{FO}(\mathcal{S})$, nämlich

$$\exists x\, B\{c/x\} := \exists x\, B\{x/c\}^{-1} \in \boldsymbol{FO}(\mathcal{S}) \quad \text{mit} \quad x \notin \boldsymbol{FV}(B)$$

[17] Jacques Herbrand (1908–1931)

Vorarbeiten zur Herband Theorie: Eliminierung von $=$

Ebenso können wir ***Pred*** um ein binäres Prädikat $\mathbb{E}$ erweitern, so dass jede Formel $A \in \boldsymbol{FO}(\mathcal{S})$ erfüllbarkeitssäquivalent zu einer Formel $B \in \boldsymbol{FO}^{\neq}(\mathcal{S} + \{\mathbb{E}\})$ ist. $\mathcal{S}_A$ sei die endliche(!) Minimalsignatur für A.

- OBdA möge A in bPNF vorliegen. Ersetzt man jedes Auftreten von $\doteq$ in A durch $\mathbb{E}$, entsteht eine Formel A', unter Missbrauch der Notation $A' := A\{\doteq/\mathbb{E}\}$. Weiter gilt $\mathcal{S}_{A'} = \boldsymbol{Fun}_A + (\boldsymbol{Pred}_A + \{\mathbb{E}\})$.
- Jedes Modell $\mathcal{M}$ für A wird zu einem Modell von A', wenn man $\mathbb{E}^{\mathcal{M}}$ als Gleichheit interpretiert. A'-Modelle, in denen das nicht der Fall ist, liefern keine Information über die Existenz von A-Modellen.
- Bilde nun die Konjunktion $A^{\neq}$ von A' mit endlich vielen Formeln, die jede Interpretation von $\mathbb{E}$ zwingend zu einer Kongruenzrelation bzgl. $\boldsymbol{Fun}_{A'} = \boldsymbol{Fun}_A$ machen. Dann kann jedes Modell für $A^{\neq}$ nach dieser Relation faktorisiert werden (bilde Äquivalenzklassen), wodurch $\mathbb{E}$ als Gleichheit interpretiert wird und ein Modell für A entsteht (HA).

Der Satz von Herbrand

OBdA möge nun ***Fun*** eine Konstante und A kein $\doteq$ enthalten.

Definition

- Die ***Fun***-Algebra $D_{\mathcal{H}} := \mathbf{Term}(\mathbf{Fun}, \emptyset)$ besteht aus Grundtermen.
- Jede $\mathcal{S}$-Struktur der Form $\mathcal{H} = \langle D_{\mathcal{H}}, I \rangle$ heißt Herbrand-Struktur.
- Falls $\models_{\mathcal{H}} A$ nennen wir $\mathcal{H}$ ein Herbrand-Modell von A.

Die Interpretation von ***Pred*** in einer Herbrand-Struktur ist frei wählbar.

Im mathematischen Sinn ist $D_{\mathcal{H}}$ die freie ***Fun***-Algebra über $\emptyset$; sie ist genau dann nicht leer, wenn mindestens eine Konstante in ***Fun*** existiert.

Satz (Herbrand)

Für jede Menge $\Gamma \subseteq \mathbf{FO}^{\neq}(\mathcal{S})$ *geschlossener Formeln in SNF gilt:*

Γ *ist erfüllbar gdw.* Γ *hat ein Herbrand-Modell*

Beweis.

Wenn Γ ein Herbrand-Modell hat, ist Γ natürlich erfüllbar.

Ist $\mathcal{M} = \langle D, I \rangle$ ein Modell für Γ, so ist $\langle D, I_{\boldsymbol{Fun}} \rangle$ eine ***Fun***-Algebra. Die leere Abbildung $\emptyset \xrightarrow{z} D$ lässt sich aufgrund des ▸ Rekursionssatzes eindeutig zu einem ***Fun***-Homomorphismus $D_{\mathcal{H}} \xrightarrow{\tilde{z}} D$ fortsetzen.

Die $\tilde{z}$-Urbilder der Prädikate $R^{\mathcal{M}}$, $R \in \boldsymbol{Pred}$, machen $D_{\mathcal{H}}$ zu einer $\mathcal{S}$-Struktur $\mathcal{M}_{\mathcal{H}}$ und $\tilde{z}$ zu einem starken ***Pred***-Homomorphismus, d.h.,

$$R^{\mathcal{M}_{\mathcal{H}}}(u_0, \ldots, u_{\mathcal{S}(R)-1}) = R^{\mathcal{M}}(\tilde{z}(u_0), \ldots, \tilde{z}(u_{\mathcal{S}(R)-1})) \qquad (\star)$$

für alle $\boldsymbol{u} \in D_{\mathcal{H}}^{\mathcal{S}(R)-1}$.

Achtung: Dieselbe Konstruktion angewendet auf die Gleichheit über D liefert nicht notwendig die Gleichheit über $D_{\mathcal{H}}$, sondern i.A. nur eine Kongruenzrelation. Aus diesem Grund durfte in Γ kein $\doteq$ vorkommen.

Die Quantifikation in $A \in \Gamma$ erstrecke sich oBdA über $\boldsymbol{x} = \langle x_i : i < n \rangle$; der quantorenfreie Teil sei B.

Beweis, Fortsetzung.

Als geschlossene Formel hat $A \in \Gamma$ eine von Belegungen unabhängige Semantik. Für beliebiges $\tau \in (D_{\mathcal{H}})^{\mathcal{V}}$ setze $\sigma := \tilde{z} \circ \tau \in D^{\mathcal{V}}$. Wegen $(\star)$ gilt aufgrund des Substitutionslemmas und $\models_{\mathcal{M}} A$ für jedes $\boldsymbol{u} \in (D_{\mathcal{H}})^n$:

$$\begin{aligned}\mathcal{M}_{\mathcal{H}}[\![B\{\boldsymbol{x}/\boldsymbol{u}\}]\!](\tau) &= \mathcal{M}[\![B\{\boldsymbol{x}/\tilde{z}(\boldsymbol{u})\}]\!](\sigma) \\ &= \mathcal{M}[\![B]\!](\sigma\{\boldsymbol{x}/\mathcal{M}[\![\tilde{z}(\boldsymbol{u})]\!](\sigma)\}) = 1\end{aligned}$$

bzw. in weniger bombastischer Notation

$$\hat{\tau}(B\{\boldsymbol{x}/\boldsymbol{u}\}) = \hat{\sigma}(B\{\boldsymbol{x}/\tilde{z}(\boldsymbol{u})\}) = \sigma\{\boldsymbol{x}/\check{\sigma}(\tilde{z}(\boldsymbol{u}))\}(B) = 1$$

Für das Infimum über alle $\boldsymbol{u} \in (D_{\mathcal{H}})^n$ folgt daher

$$\mathcal{M}_{\mathcal{H}}[\![A]\!](\tau) = \mathcal{M}[\![A]\!](\tau \circ \tilde{z}) = \mathcal{M}[\![A]\!] = 1 \quad \text{bzw.} \quad \hat{\tau}(A) = \widehat{(\tau \circ \tilde{z})}(A) = 1$$

Damit ist $\mathcal{M}_{\mathcal{H}}$ ein Herbrand-Modell für A, also auch für Γ. □

Satz von Löwenheim[18]-Skolem

Satz (Löwenheim 1915; Skolem 1922)

Jede erfüllbare Menge Γ von Formeln hat ein abzählbares Modell.

Beweis.

Γ ist genau dann erfüllbar, wenn das für die Menge Γ' der (minimalen) existenziellen Abschlüsse gilt. Die Skolemisierung letzterer liefert eine Menge Γ''; diese ist ebenfalls erfüllbarkeitsäquivalent zu Γ.

Das Symbol $\doteq$ vererbt sich von Γ auf Γ''. Daher braucht $\mathcal{M}_{\mathcal{H}}$ aus dem obigen Beweis kein Modell von Γ'' zu sein. Aber wir können $\mathcal{M}_{\mathcal{H}}$ nach dem $\bar{z}$-Urbild der Gleichheit auf D (und somit einer Kongruenzrelation auf $D_{\mathcal{H}}$) faktorisieren und erhalten so ein Modell für Γ'' und folglich auch für Γ. Aus der Abzählbarkeit von $D_{\mathcal{H}}$ folgt auch die Abzählbarkeit der Quotientenmenge. □

[18] Leopold Löwenheim (1878–1957)

Nichtstandard-Modelle

Betrachte die Signatur der Arithmetik: $\mathcal{S}_{\text{arith}} = \{0_{/0}, 1_{/0}, +_{/2}, \cdot_{/2}; <_{/2}\}$.
Intendierte Struktur sind die natürlichen Zahlen $\mathbb{N}$ mit den kanonischen Operationen/Relationen; dafür ist $\mathcal{S}_{\text{Arith}}$ vermutlich zu grob.

Ziel: zeigen, dass es „seltsame“ $\mathcal{S}_{\text{Arith}}$-Strukturen $\mathcal{M}^* = \langle D^*, I^* \rangle$ gibt, die dieselben geschlossenen Formeln erfüllen wie $\mathbb{N}$ mit der kanonischen Interpretation, aber nicht zu $\mathbb{N}$ isomorph sind.

Solche Modelle heißen Nichtstandard-Modelle.

Einsicht: Wichtige Eigenschaften z.B. der natürlichen Zahlen lassen sich mit geschlossenen Formeln nicht in ***FO*** erfassen.

Anwendungen:

- ▸ Nichtstandard Analysis (Abraham Robinson, 1960);
- Computer-Algebra;
- Verifikation hybrider Systeme (Zug-und Flugzeugcontroller)

Satz

Es gibt Nicht-Standard-Modelle der Arithmetik.

Beweis.

Setze $A_n := \underbrace{1 + \cdots + 1}_{n \text{ mal}} < x$ mit freiem x und

$$\Gamma = \{ A \in \boldsymbol{FO}(\mathcal{S}_{\text{arith}}) : A \text{ geschlossen, und } \models_{\mathbb{N}} A \} \cup \{ A_i : i \in \mathbb{N} \}$$

Offenbar ist jede endliche Teilmenge von Γ erfüllbar, nämlich im Standard-Modell $\mathbb{N}$: da nur endlich viele Formeln der Form A_i vorkommen, findet man immer ein $\sigma \in \mathbb{N}^{\mathcal{V}}$, so dass $\sigma(x)$ hinreichend groß ist.

Jedes Modell $\mathcal{M}^*$ von Γ muss hingegen

- alle geschlossenen in $\mathbb{N}$ gültigen $\mathcal{S}_{\text{arith}}$-Formeln erfüllen,
- ein bzgl. $I^*(<)$ größtes Element haben, was $\mathcal{M}^* \not\cong \mathbb{N}$ impliziert

Daher lassen sich $\mathbb{N}$ und $\mathcal{M}^*$ nicht durch geschlossene $\mathcal{S}_{\text{arith}}$-Formeln unterscheiden, sie sind elementar äquivalent. □

Skolems Paradox

Die übliche Mengenlehre (Anhang C) ist ein Modell für die dortigen Axiom-Schemata. Nach dem Satz von Löwenheim-Skolem muss es also auch ein abzählbares Modell $\mathcal{M}_{\mathrm{LS}}$ von ZFC geben.

Gemäß Anhang B ist im Standardmodell der Mengenlehre die Potenzmenge von $\mathbb{N}$ überabzählbar. In $\mathcal{M}_{\mathrm{LS}}$ spielen Objekte $\mathcal{N}$ und $\mathcal{P}$ die Rolle der „Menge der natürlichen Zahlen“ bzw. von deren „Potenzmenge“. Dort sollte $\mathcal{P}$ „überabzählbar“ sein.

Warum ist das kein Widerspruch?

Die „Überabzählbarkeit“ von P in $\mathcal{M}_{\mathrm{LS}}$ bedeutet, dass in $\mathcal{M}_{\mathrm{LS}}$ keine „Abbildung“ von $\mathcal{P}$ nach $\mathcal{N}$ „injektiv“ ist. Dagegen ist die Abzählbarkeit von $\mathcal{M}_{\mathrm{LS}}$ eine Aussage im Standard-Modell der Mengenlehre und kann insofern gar nicht mit der „Überabzählbarkeit“ in $\mathcal{M}_{\mathrm{LS}}$ in Beziehung gesetzt werden. Ein Problem entsteht erst dann, wenn man beide Modelle unzulässig vermischt (hier: die Anführungszeichen weglässt).

Herbrand Expansion, Vorüberlegungen

Betrachte eine geschlossene bereinigte Formel

$$A = \forall x_0 \forall x_1 \ \ldots \ \forall x_{n-1}.\, B \in \boldsymbol{FO}^{\neq}(\mathcal{S}) \quad , \ B \text{ quantorenfrei}$$

Deren Auswertung in einer $\mathcal{S}_A = \mathcal{S}_B$-Struktur $\mathcal{M}_{\mathcal{H}}$ mit Trägermenge $D_{\mathcal{H}}$ ist unabhängig von der Belegung $\sigma \in (D_{\mathcal{H}})^{\mathcal{V}}$. Nach Konstruktion ersetzen sowohl $\check{\sigma}$ als auch $\widetilde{\sigma\{\boldsymbol{x}/\boldsymbol{t}\}}$ alle freien Variablen in $\boldsymbol{Fun}$-Termen über $\mathcal{V}$ durch bestimmte variablenfreien Terme. Da $\check{\sigma}$ nicht auf den Termen in $\boldsymbol{t}$ wirkt, stimmt $\widetilde{\sigma\{\boldsymbol{x}/\boldsymbol{t}\}}$ mit $\check{\sigma} \circ \{\boldsymbol{x}/\boldsymbol{t}\}$ überein. Folglich gilt für Formeln

$$\hat{\sigma}(A) = \inf\{\, \widehat{\sigma\{\boldsymbol{x}/\boldsymbol{t}\}}(B) : \boldsymbol{t} \in (D_{\mathcal{H}})^n \,\} = \inf\{\, \hat{\sigma}(B\{\boldsymbol{x}/\boldsymbol{t}\}) : \boldsymbol{t} \in (D_{\mathcal{H}})^n \,\} \ (*)$$

Die Werte $\hat{\sigma}(B\{\boldsymbol{x}/\boldsymbol{t}\})$ hängen wie in der AL von den Werten der nach Voraussetzung $\doteq$-freien atomaren Formeln $\boldsymbol{AT}^{\neq}_{\mathcal{H}} := \boldsymbol{Atm}(\boldsymbol{Pred}, \mathcal{D}_{\mathcal{H}})$ ab. Aber das sind genau die Interpretationen der Prädikate aus $\boldsymbol{Pred}_B$ in $D_{\mathcal{H}}$!

Satz von Gödel[19]-Herbrand-Skolem

Definition

Die Herbrand-Expansion der obigen Formel A besteht aus allen (parallelen) Substitution von Grundtermen (= Grundsubstitutionen) für Variablen von B

$$E(A) := \left\{ B\{\boldsymbol{x}/\boldsymbol{t}\} \;:\; \boldsymbol{t} \in (D_{\mathcal{H}})^n \right\} \subseteq \mathcal{F}[\boldsymbol{AT}^{\neq}_{\mathcal{H}}] \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S}_B)$$

Diese sind sämtlich quantorenfrei und geschlossen und folglich nach aussagenlogischem Muster aus $\doteq$-freien atomaren Formeln aufgebaut.

Für $\Gamma \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S})$ ist $E(\Gamma)$ die Vereinigung der Mengen $E(A)$, $A \in \Gamma$.

Satz

Für eine Menge $\Gamma \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S})$ geschlossener Formeln in SNF gilt

Γ *ist erfüllbar gdw.* $E(\Gamma)$ *ist aussagenlogisch erfüllbar*

[19] Kurt Gödel (1906–1978)

Beweis.

Aus einer $E(\Gamma)$ erfüllenden Belegung φ der atomaren Formeln in $\boldsymbol{AT}^{\neq}_{\mathcal{H}}$ erhalten wir nach obiger Überlegung eine Interpretation I_φ der Prädikate in $D_{\mathcal{H}}$. In der Herbrand- Struktur $\mathcal{M}_\varphi = \langle D_{\mathcal{H}}, I_\varphi \rangle$ gilt nun $\mathcal{M}_\varphi[\![A]\!] = 1$, da rechts in $(*)$ die Werte $\hat{\varphi}(B\{\boldsymbol{x}/\boldsymbol{t}\}) = 1$ auftreten.

Umgekehrt bestimmt jedes Modell $\mathcal{M}$ für Γ durch die Interpretation der Prädikate eine Belegung $\varphi_{\mathcal{M}}$ der atomaren Formeln. Aus $\mathcal{M}[\![A]\!] = 1$ folgt dann gemäß $(*)$ auch $\widehat{\varphi_{\mathcal{M}}}(B\{\boldsymbol{x}/\boldsymbol{t}\}) = 1$ für jedes $\boldsymbol{t} \in (D_{\mathcal{H}})^n$. □

In Verbindung mit dem Kompaktheitssatz der Aussagenlogik ergibt sich

Corollar

Eine Menge $\Gamma \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S})$ geschlossener Formeln in SNF ist unerfüllbar, genau dann wenn eine endliche Teilmenge von $E(\Gamma)$ unerfüllbar ist. □

Beispiel

Für die Signatur $\mathcal{S} = \{c_{/0}, f_{/1}, g_{/2}\} + \{P_{/1}\}$ betrachten wir einige Elemente der Herbrand-Expansion für $A = \forall x.\big(P(x) \vee \neg P(f(x))\big)$, sortiert nach der Länge (ohne Klammern):

Länge 7:	$P(c) \vee \neg P(f(c))$
Länge 9:	$P(f(c)) \vee \neg P(f(f(c)))$
Länge 11:	$P(f(f(c))) \vee \neg P(f(f(f(c))))$, $P(g(c,c)) \vee \neg P(f(g(c,c)))$
Länge 13:	$P(f(f(f(c)))) \vee \neg P(f(f(f(f(c)))))$,
	$P(g(f(c),c)) \vee \neg P(f(g(f(c),c)))$,
	$P(g(c,f(c))) \vee \neg P(f(g(c,f(c))))$

Hier liefern die Substitutionen Klauseln in den positiven/negativen atomaren Formeln ohne Variablen, mit dem Potential zur Resolventenbildung.

Falls der quantorenfreie Teil von einer Formel mehr als eine freie Variable enthält, vervielfachen sich die Möglichkeiten zur Substitution.

Gilmores Algorithmus (ineffizient)

Algorithmus

Eingabe: $\Gamma \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S})$ bestehe aus geschlossen Formeln in SNF, und A_i, $i \in \mathbb{N}$, sei eine Aufzählung von $E(\Gamma)$.
Algorithmus: Solange $G_n := \bigwedge_{i<n} A_i$ erfüllbar ist, bilde $G_{n+1} = G_n \wedge A_n$.

Folgende Aussagen sind für Γ äquivalent:

- Der Algorithmus terminiert;
- für ein minimales $k > 0$ gilt $\models \neg G_k$;
- für ein minimales $k > 0$ ist G_k nicht erfüllbar;
- für ein minimales $k > 0$ ist $\{ A_i : i < k \} \subseteq E(\Gamma)$ nicht erfüllbar;
- $E(\Gamma)$ ist nicht erfüllbar;
- Γ ist nicht erfüllbar.

Damit ist das Allgemeingültigkeitsproblem der PL semi-entscheidbar.

Der Kompaktheitssatz der PL

Satz

Eine Formelmenge $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$ ist genau dann erfüllbar, wenn sie endlich erfüllbar ist, d.h., wenn jede endliche Teilmenge $\Gamma_0 \subseteq \Gamma$ erfüllbar ist.

Beweis.

Die Notwendigkeit ist klar.

Für die Hinlänglichkeit betrachte erst den Fall geschlossener Formeln in SNF ohne $\doteq$. Hier ist Γ genau dann erfüllbar, wenn das für $E(\Gamma) := \bigcup\{ E(A) : A \in \Gamma \}$ gilt, was nach dem KPS(AL) zur endlichen Erfüllbarkeit von $E(\Gamma)$ äquivalent ist.

Ist Γ endlich erfüllbar, finden wir zu jeder endlichen Teilmenge $\Delta_0 \subseteq E(\Gamma)$ eine endliche und somit erfüllbare Teilmenge $\Gamma_0 \subseteq \Gamma$ mit $\Delta_0 \subseteq E(\Gamma_0)$. Damit ist Δ_0 erfüllbar, und folglich ist $E(\Gamma)$ endlich erfüllbar, woraus nach obiger Überlegung die Erfüllbarkeit von Γ folgt.

Beweis, Fortsetzung.

Nun betrachte eine beliebige Menge $\Gamma \subseteq \boldsymbol{FO}(\mathcal{S})$.

1. Bilde den existentiellen Abschluß aller Formeln in Γ.
2. Tritt $\doteq$ in der resultierenden Menge Γ' auf, so ersetze es durch ein frisches Prädikat $E_{/2}$ und erweitere Γ' für jedes $A \in \Gamma'$ um Formeln, die jeweils sicherstellen, dass E in jeder Struktur als Kongruenz bzgl. der Operatoren und Prädikate in $\mathcal{S}_A$ interpretiert werden muss (vergl. HA).
3. Wandle die Elemente der resultierenden Menge $\Gamma'' \subseteq \boldsymbol{FO}^{\neq}(\mathcal{S} + \{E_{/2}\})$ SNF um. Ggf. füge eine Konstante $c_{/0}$ zur Signatur $\mathcal{S}$ hinzu.
4. Behandle die resultierende Menge $\Gamma''' \subseteq \boldsymbol{FO}(\mathcal{S} + \{E_{/2}\} + \boldsymbol{Sko} + \{c\})$ wie oben. □

Kapitel 12
Algorithmen

Tableaus in der PL_1

Um die Tableau-Methode in die PL übertragen, klassifiziert man in $\boldsymbol{FO}^{\neq}(\mathcal{S})$ quantifizierte Formeln und ihre Negationen, vergl. ▸ Folie 139 ▸ Folie 140:

Definition

- „Literale“: atomare Formeln und ihre Negationen;
- doppelte Negationen – überflüssig, sofort eliminieren;
- α: $A \wedge B$, $\neg(A \vee B)$, $\neg(A \to B)$, $(\neg\neg A)$;
- β: $A \vee B$, $A \to B$, $\neg(A \wedge B)$;
- γ: $\forall x\, A$, $\neg\exists x\, A$;
- δ: $\exists x\, A$, $\neg\forall x\, A$.

Neue Regeln mit $t \in D_{\mathcal{H}}$ bzw. frischer Konstante c für den aktuellen Ast:

$$\gamma: \quad \frac{\forall x\, A}{A\{x/t\}} \quad , \quad \frac{\neg\exists x\, A}{\neg A\{x/t\}} \qquad \delta: \quad \frac{\exists x\, A}{A\{x/c\}} \quad , \quad \frac{\neg\forall x\, A}{\neg A\{x/c\}}$$

Die „Nenner“ bezeichnen wir als γ- bzw. δ-Teilchen.

Während die Tableau-Methode der AL die Teile-Hülle einer Formelmenge iterativ bestimmte und deren Elemente so in einer Baumstruktur arrangierte, dass mögliche erfüllende Belegungen bzw. deren Nichtexistenz ablesbar wurden, berücksichtigt die PL_1-Variante zudem die Skolemisierung (via δ-Formeln) und die Konstruktion der Herbrand-Expansion (via γ-Formeln).

- δ-Formeln haben Vorrang; das entspricht der Eliminierung von Existenz-Quantoren zugunsten neuer Konstanten (Skolemisierung). Diese vergrößern sukzessive die Signatur $\mathcal{S}$, also auch $D_{\mathcal{H}}$, stehen daher anschließend für Substitutionen in γ-Formeln zur Verfügung!
- γ-Formeln können wiederholt und beliebig oft zu Substitutionen herangezogen werden; dies entspricht der Herbrand-Expansion.
- Für explizite Verzweigungen sind nur die β-Teilchen zuständig.

Die Begriffsbildungen der AL (vergl. ▸ Folie 147) sind nur geringfügig anzupassen. Wie zuvor beschränken wir uns auf die Menge $\boldsymbol{FO}^{\neq\neg\neg}(\mathcal{S})$ der Formeln, die nicht mit einer doppelten Negation beginnen.

Definition

Ein Tableau der PL ist Abbildung $\mathbb{B}^* \xrightarrow{\tau} \boldsymbol{P}(\boldsymbol{FO}^{\neq\neg\neg}(\mathcal{S}))$, so dass

- das Urbild $\mathcal{B}(\tau)$ der nichtleeren Formelmengen ist Präfix-abgeschlossen und erfüllt $w0 \in \mathcal{B}(\tau)$ gdw. $w1 \in \mathcal{B}(\tau)$ für alle $w \in \mathbb{B}^*$.
- $\Gamma \subseteq \tau(\varepsilon) \subseteq \Gamma^{<\alpha,\gamma,\delta>}$.
- für jedes $\varepsilon \neq w \in \mathcal{B}(\tau)$ existiert $B \in \tau(w)$ mit $\tau(w) \subseteq \{B\}^{<\alpha,\gamma,\delta>}$ und B ist β-Teilchen einer Formel aus einem echten Vorgängerknoten.

Ein Ast Θ von $\mathcal{B}(\tau)$ (maximale lineare Präfix-geordnete Teilmenge) heißt

- vollständig, falls $\bigcup \tau[\Theta]$
  - unter α- und γ-Teilchen abgeschlossen ist;
  - mit jeder β-/δ-Formel mindestens ein β-/δ-Teilchen enthält.
- abgeschlossen, falls $\bigcup \tau[\Theta]$ eine Formel und ihre Negation enthält; andernfalls heißt Θ offen.

Ein Tableau heißt abgeschlossen, falls jeder Ast abgeschlossen ist.

Übertragung der Tableau-Sätze der AL

Lemma

Jedes Tableau τ kann vervollständigt werden. □

Lemma (Hintikka)

Für vollständige Tableau-Astmengen gilt: „erfüllbar“=„offen“. □

Satz

Γ ist unerfüllbar gdw. Γ hat ein abgeschlossenes Tableau. □

Satz

Für $A \in \boldsymbol{FO}^{\neq\neg\neg}(\mathcal{S})$ und $\Gamma \subseteq \boldsymbol{FO}^{\neq\neg\neg}(\mathcal{S})$ gilt:

1. $\models A$ *gdw.* $\neg A$ *hat ein abgeschlossenes Tableau.*
2. $\Gamma \models A$ *gdw.* $\Gamma \cup \{\neg A\}$ *hat ein abgeschlossenes Tableau.* □

Die Tableau-Methode in der Praxis

Beispiel (vergl. Aufgabe 5, Blatt 0)

Betrachte folgende Behauptung:

Jede transitive, symmetrische, totale Relation ist reflexiv.

Um sie zu überprüfen, wählen wir eine Signatur, die nur aus einem binären Prädikat $R_{/2}$ besteht und formalisieren die relevanten Eigenschaften:

(0) Transitivität: $\forall x \forall y \forall z.\, (R(x,y) \wedge R(y,z) \rightarrow R(x,z))$

(1) Symmetrie: $\forall x \forall y.\, (R(x,y) \rightarrow R(y,x))$

(2) Totalität: $\forall x \exists y.\, R(x,y)$

(3) Reflexivität: $\forall x.\, R(x,x)$

Zum Nachweis von $\{(0),(1),(2)\} \models (3)$ untersuchen wir mit Hilfe der Tableau-Methode, ob $\Gamma = \{(0),(1),(2),\neg(3)\}$ nicht erfüllbar ist, d.h., ein abgeschlossenes Tableau hat.

Beispiel (Fortsetzung)

*	$\forall x \forall y \forall z.\, (R(x,y) \wedge R(y,z) \rightarrow R(x,z))$	5
*	$\forall x \forall y.\, (R(x,y) \rightarrow R(y,x))$	3
*	$\forall x \exists y.\, R(x,y)$	1
*	$\neg \forall x.\, R(x,x)$	0
0δ	$\neg R(a,a)$	
1γ	$\exists y.\, R(a,y)$	2
2δ	$R(a,b)$	
3γ	$R(a,b) \rightarrow R(b,a)$	4

4a $\neg R(a,b)$ — $\lightning$, 2δ, 4a

4b $R(b,a)$

5γ $R(a,b) \wedge R(b,a) \rightarrow R(a,a)$ 6

6a $\neg(R(a,b) \wedge R(b,a))$ 7

6b $R(a,a)$ — $\lightning$, 0δ, 6b

7a $\neg R(a,b)$ — $\lightning$, 2δ, 7a

7b $\neg R(b,a)$ — $\lightning$, 4b, 7b

Heuristik zur Konstruktion endlicher Modelle

Im Gegensatz zur AL ist braucht die Tableau-Konstruktion für endliche Formelmengen Γ nicht zu terminieren, da die Herbrand-Expansion $D_{\mathcal{H}}$ je nach Beschaffenheit von ***Fun*** unendlich sein kann.

Das Gegenstück zu einer erfüllenden Belegung in der AL wäre bei der Tableau- Methode der PL_1 ein endliches Modell; aufgrund der abgeschlossenen Formeln spielen die Belegungen der Variablen hier keine Rolle. Um ein solches zu finden, ändert man die Strategie ab, bei δ-Formeln für den aktuellen Ast nur frische Konstanten einzuführen:

- ▷ verwende zunächst im Ast schon vorhandene Konstanten;
- ▷ erst wenn so Widersprüche entstehen, wähle frische Konstanten.

Substituiere darüberhinaus zunächst nur Konstanten in γ- Formeln. Sobald die γ-Formeln nichts Neues liefern, bilden die bisher entlang des Zweigs eingeführten Konstanten selber den Datenbereich eines Modells für Γ.

Beispiel (Endliche Modelle für $\{\exists x.\, \neg R(x,x), \forall x \exists y.\, R(x,y)\}$)

Gibt es auf endlichen Mengen totale Relationen, die nicht reflexiv sind?

$*$	$\exists x.\, \neg R(x,x))$	0
$*$	$\forall x \exists y.\, R(x,y)$	1
0δ	$\neg R(a,a)$	
1γ	$\exists y.\, R(a,y)$	2
2δ	$R(a,b)$	
1γ	$\exists y.\, R(b,y)$	3
3δ	$R(b,a)$	

Dies liefert ein Modell mit $D = \{a,b\}$ und $R_0 = \{\langle a,b\rangle, \langle b,a\rangle\}$. Erzeugt man mit Hilfe der γ-Formel eine weitere Instanz von $\exists y.\, R(b,y)$, kann man diesmal b für y substituieren, was ein weiteres Model auf D mit $R_1 = \{\langle a,b\rangle, \langle b,a\rangle, \langle b,b\rangle\}$ liefert.

Achtung: a und b sind zunächst Konstanten in der gemäß Skolem erweiterten Signatur. Indem wir sie hier außerdem zur Konstruktion der Modelle verwenden, interpretieren wir die Konstanten implizit.

Resolution in der PL

Sind alle quantorenfreien Teile B der geschlossenen Skolem-Normalformen $A \in \Gamma$ Konjunktionen prädikatenlogischer Klauseln, gehören also zu KNF, so gilt dies auch für die Elemente der Herbrand-Expansion, und wir können die Resolutions-Methode der AL anwenden.

Beispiel

$\mathcal{S} = \{a_{/0}, f_{/1}; R_{/1}\}$ und $A = \forall x\big(R(x) \wedge \neg R(f(x))\big)$ liefert

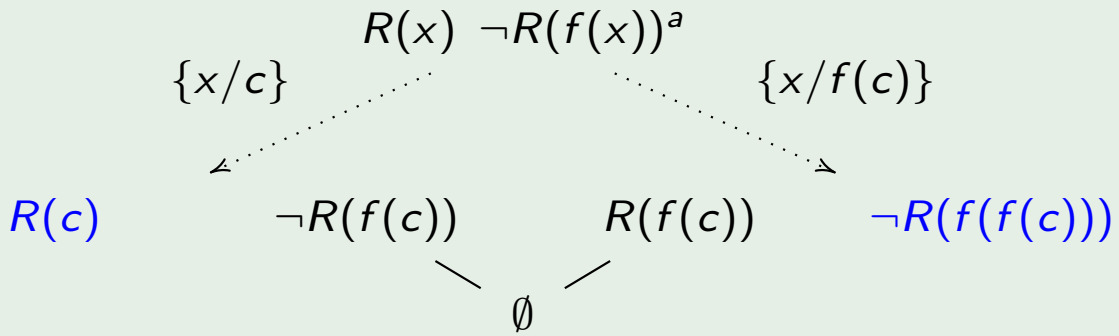
$$E(A) = \{R(a) \wedge \neg R(f(a)), R(f(a)) \wedge \neg R(f(f(a))), \dots\}$$

Die ersten beiden Elemente liefern bereits zwei Unit-Klauseln mit leerer Resolvente, $\neg R(f(a))$ und $R(f(a))$. Damit ist A nicht erfüllbar.

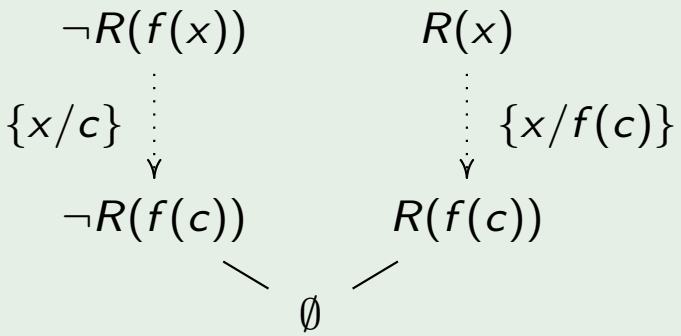
Idee zur Vermeidung nicht-zielführender Klauseln: $E(A)$ „verfeinern“, indem man B als Klauselmenge auffasst und die Substitution lokal auf einzelne „vielversprechende“ Klauseln beschränkt. Das führt zur sog. Unifikation.

Beispiel (Forsetzung)

Die globalen Substitutionen $\{x/c\}$ und $\{x/f(c)\}$ liefern folgende Resolventenberechnung mit zwei überflüssigen Resolventen



Mit Hilfe lokaler Substitutionen in den B-Klauseln erhält man stattdessen



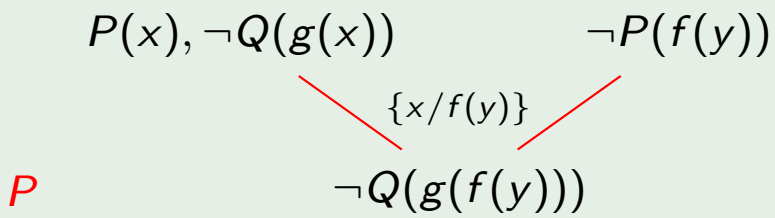
[a] wir lassen die Mengenklammern um die Klauseln weg, da sie mit denen der Substitutionen kollidieren.

Problem: die algorithmische Suche nach Grundinstanzen der B-Klauseln, die möglichst direkt zur Herleitung der leeren Klausel führen:

- systematisches Probieren der Grundsubstitutionen ist zu aufwändig;
- vorausschauende Auswahl von B-Klauseln und Grundsubstitutionen für effiziente Resolventenbildung ist sehr schwierig.

Neue Strategie: Suche nach beliebigen Term-Substitutionen (nicht nur Grundterme!) in B-Klauseln, die den nächsten Resolutionsschritt erlauben indem sie Literale aus verschiedenen Klauseln komplementär machen.

Beispiel



Achtung: auf beiden Seiten muss dieselbe Substitution wirken!

Unifikation

Definition

Betrachte eine endliche Menge $\mathcal{L}$ von Literalen der Prädikatenlogik, d.h., positiven oder negativen atomaren Formeln, und eine Substitution $\hat{\vartheta}$.

- $\hat{\vartheta}$ heißt Unifikator von $\mathcal{L}$, wenn alle $\hat{\vartheta}$-Bilder der Elemente von $\mathcal{L}$ übereinstimmen;
- $\hat{\vartheta}$ heißt allgemeinster Unifikator von $\mathcal{L}$, wenn jeder Unifikator $\hat{\rho}$ für $\mathcal{L}$ durch $\hat{\vartheta}$ faktorisiert, d.h., $\hat{\rho} = \hat{\vartheta}; \hat{\omega}$ für eine Substitution $\hat{\omega}$ gilt.
- $\mathcal{L}$ heißt unifizierbar, wenn ein Unifikator für $\mathcal{L}$ existiert.

Satz (John Alan Robinson)

Jede unifizierbare Literal-Menge hat einen allgemeinsten Unifikator. □

Existiert ein allgemeinster Unifikator für $\mathcal{L}$, kann er algorithmisch bestimmt werden.

Der Unifikationsalgorithmus

Algorithmus

Eingabe: $\mathcal{L} = \{ L_i : i < n \}$
Initialisierung: $\vartheta := \iota$, d.h., $\hat{\vartheta} = \mathit{id}$

- Solange ϑ nicht $\mathcal{L}$ unifiziert, durchsuche parallel alle L_i von links nach rechts und bestimme die erste Position, in der Unterschiede auftreten; wähle zwei dieser Literale aus.
- Falls keins dieser Literale an besagter Position eine Variable enthält, ist $\mathcal{L}$ nicht unifizierbar.
- Andernfalls steht in einem der gewählten Literale in besagter Position die Variable x, während im anderen Literal ein Term t beginnt.
- Falls $x \in \mathit{FV}(t)$, ist $\mathcal{L}$ nicht unifizierbar.
- Andernfalls aktualisiere ϑ zu $\vartheta\{x/t\}$

Beispiel

Klausel	Klausel	Klausel	Modifikation
$Q(f(a,x),u)$	$Q(f(y,g(u)),h(w))$	$Q(v,h(b))$	$\{v/f(a,x)\}$
$Q(f(a,x),u)$	$Q(f(y,g(u)),h(w))$	$Q(f(a,x),h(b))$	$\{y/a\}$
$Q(f(a,x),u)$	$Q(f(a,g(u)),h(w))$	$Q(f(a,x),h(b))$	$\{x/g(u)\}$
$Q(f(a,g(u)),u)$	$Q(f(a,g(u)),h(w))$	$Q(f(a,g(u)),h(b))$	$\{u/h(w)\}$
$Q(f(a,g(h(w))),h(w))$	$Q(f(a,g(h(w))),h(w))$	$Q(f(a,g(h(w))),h(b))$	$\{w/b\}$
$Q(f(a,g(h(w))),h(b))$	$Q(f(a,g(h(w))),h(b))$	$Q(f(a,g(h(w))),h(b))$	

Der most general unifyer (MGU) ist somit die von der Belegung

$$\vartheta = \iota\{v/f(a,x)\}\{y/a\}\{x/g(u)\}\{u/h(w)\}\{w/b\}$$

induzierte Substitution, wobei $\mathcal{V} \xrightarrow{\iota} \boldsymbol{Term}(\boldsymbol{Fun},\mathcal{V})$ die Inklusion ist.

Beispiel

Klausel	Klausel	Modifikation
$P(x,f(y))$	$P(f(a),y)$	$\{x/f(a)\}$
$P(f(a),f(y))$	$P(f(a),y)$	~~$\{y/f(y)\}$~~

Verallgemeinerte prädikatenlogische Resolvente

Definition

Betrachte PL_1 Klauseln K und K' mit disjunkten Variablenmengen (ggf. durch Umbenennen erzwingen). Falls Literale $L_i \in K$, $i < m$, und $L'_j \in K'$, $j < n$, existieren, so dass

$$\{\neg L_i : i < m\} \cup \{L'_j : j < n\}$$

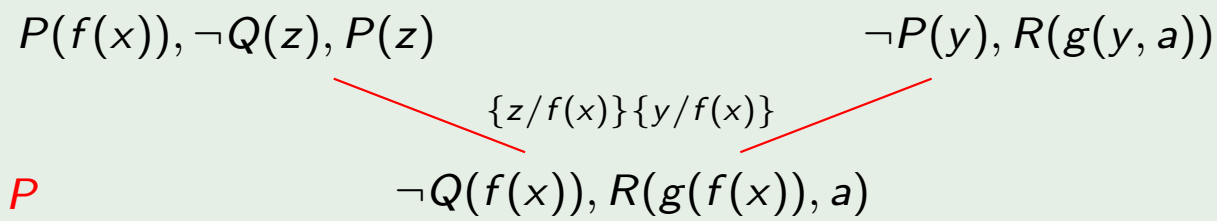
unifizierbar ist mit allgemeinstem Unifikator $\hat{\vartheta}$, dann heißt

$$R := \left(K - \{L_i : i < m\} \cup K' - \{L'_j : j < n\}\right)\hat{\vartheta}$$

prädikatenlogische Resolvente von K und K'.

Die Resolutionsregel der AL ergibt sich für $m = n = 1$ und $\vartheta = \iota$, also $\hat{\vartheta} = \{\} = \mathbf{id}$.

Beispiel



Hier gilt $m = 2$ und $n = 1$.

- Die Handrechnung lässt sich ähnlich gestalten wie in der AL.
- Da zu resolvierende Literale dasselbe Prädikat aufweisen müssen, kann man eine Liste der zu bearbeitenden Prädikate vorgeben.
- Sind nicht alle Literale einer Klausel mit demselben Prädikat unifizierbar, muss dieses Prädikat mehrfach in der obigen Listen auftreten.
- Dürfen Klauseln von der weiteren Betrachtung ausgeschlossen werden, wenn sie nach einer echten Unifikation tautologisch geworden sind?
- Wie ist mit Klauseln zu verfahren, die nach einer echten Unifikation von der resultierenden Resolvente subsumiert werden?

Beispiel

Weisen Sie die Allgemeingültigkeit nach:

$$A = \forall y.\, Q(y) \vee \neg\forall x.\, \Big[\big(Q(x) \vee R(x)\big) \wedge \exists z.\, \Big(\neg P(z) \wedge \big(P(z) \vee \neg R(x)\big)\Big)\Big]$$

Negation und Anwendung der De Morgan'schen Regel liefert:

$$\exists y.\, Q\neg(y) \wedge \forall x.\, \Big[\big(Q(x) \vee R(x)\big) \wedge \exists z.\, \Big(\neg P(z) \wedge \big(P(z) \vee \neg R(x)\big)\Big)\Big]$$

Daraus ergibt sich die PNF:

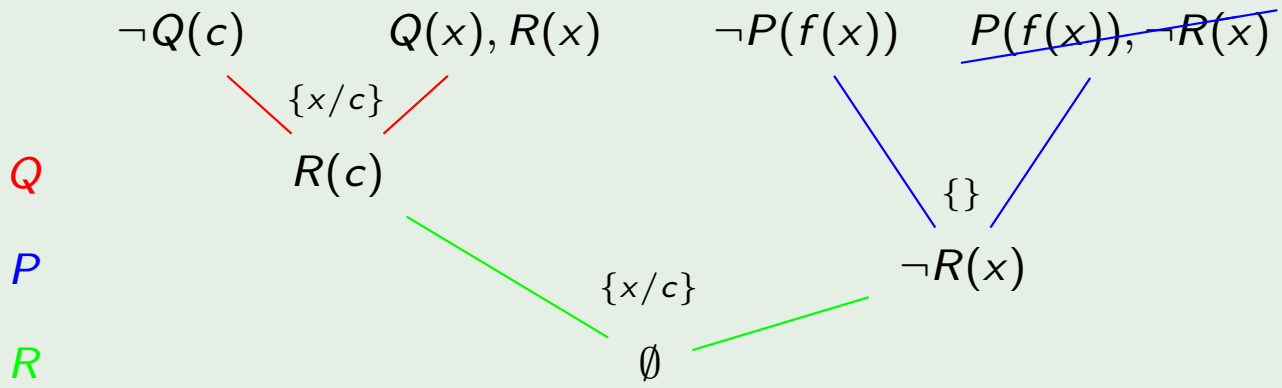
$$\exists y\, \forall x\, \exists z.\, \big[\neg Q(y) \wedge \big(Q(x) \vee R(x)\big) \wedge \neg P(z) \wedge \big(P(z) \vee \neg R(x)\big)\big]$$

Skolemisierung liefert schließlich:

$$\forall x\, \big[Q\neg(c) \wedge \big(Q(x) \vee R(x)\big) \wedge \neg P(f(x)) \wedge \big(P(f(x)) \vee \neg R(x)\big)\big]$$

Der quantorenfreie Teil liegt nun in KNF vor, was Resolution ermöglicht.

Beispiel (Fortsetzung)



Damit ist die ursprünglichen Formel A nicht erfüllbar.

Beispiel

Weisen Sie die Unerfüllbarkeit nach von

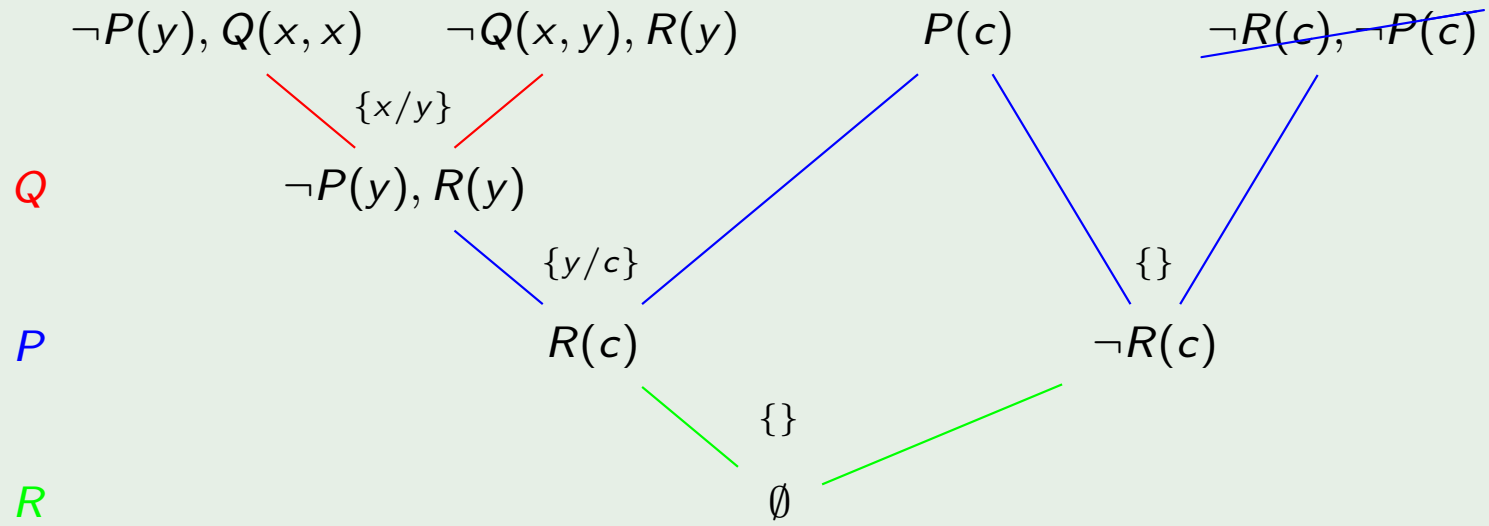
$$\exists z \forall x \forall y.\ \big[\big(P(y) \to Q(x,x)\big) \wedge \big(Q(x,y) \to R(y)\big) \wedge P(z) \wedge \big(\neg R(z) \vee \neg P(z)\big)\big]$$

Wir beginnen mit der Skolemisierung:

$$\forall x \forall y.\ \big[\big(P(y) \to Q(x,x)\big) \wedge \big(Q(x,y) \to R(y)\big) \wedge P(c) \wedge \big(\neg R(c) \vee \neg P(c)\big)\big]$$

Elimination von $\to$ liefert einen quantorenfreien Teil in KNF, für Resolution:

Beispiel (Fortsetzung)



Damit ist die ursprünglichen Formel A nicht erfüllbar.

Im ersten Schritt wäre auch die Substitution $\{y/x\}$ möglich aber weniger effizient gewesen, da y häufige auftritt als x.

Versuchen Sie zur Übung, andere Reihenfolgen der Prädikate zu wählen!

Satz (Widerlegungsvollständig- und Korrektheit, John Alan Robinson)

$A = \forall x_0 \ldots \forall x_{n-1}. B$ in SNF mit B quantorenfrei in KNF ist genau dann unerfüllbar, wenn $B \vdash_{\text{res}} \emptyset$.

Achtung: Leider braucht das oben beschriebene Verfahren bei der Suche nach der Resolvente $\emptyset$ nicht zu terminieren. Daher ist das Problem der Unerfüllbarkeit hier nicht entscheidbar.

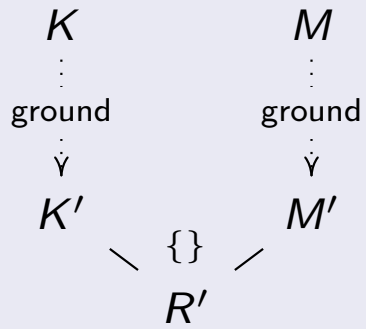
Der Beweis zu Robinsons Satz reduziert die obige prädikatenlogische Resolventenbildung, bei der nicht-triviale Unifikationen auftreten können, auf die im Wesentlichen aussagenlogische Resolventenbildung, bei der nur Grundsubstitutionen zur Anwendung kommen.

Wesentliches Hilfsmittel ist das folgende „Lifting Lemma“:

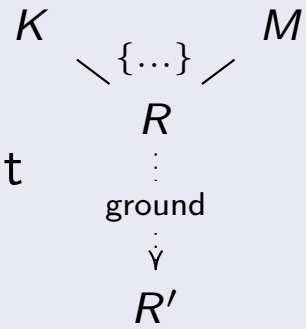
Lifting Lemma

Sind K und M prädikatenlogische Klauseln mit Grundinstanzen K' und M', die eine aussagenlogische Resolvente R' haben, dann existiert eine prädikatenlogische Resolvente R von K und M, so dass R' eine Grundinstanz von R ist, in Diagrammform

Falls



dann existiert R mit



Die umgekehrte Richtung ist klar: ergänze den Unifikator für die Resolventenbildung auf der rechten Seite um Modifikationen, die alle in den weggresolvierten Literalen auftretenden Variablen auf die Konstante c abbilden. Komponiert mit der Grundsubstitution auf der rechten Seite liefert das wieder eine Grundsubstitution, die links verwendet werden kann.

Anhänge

Anhang A
Signaturen und Term-Algebren

Viele mathematische Gebiete befassen sich mit Mengen, die eine zusätzliche Struktur tragen, und struktur-erhaltenden Abbildungen zwischen diesen.

Häufig beinhalten solche Strukturen auf einer Grundmenge X Operationen der Form $X^n \xrightarrow{f} X$ der Stelligkeit $n \in \mathbb{N}$.

Beispiel

- Die Addition auf den natürlichen (ganzen, rationalen, reellen...) Zahlen;
- Die Multiplikation auf den natürlichen (ganzen, rationalen, reellen...) Zahlen;
- Die Konkatenation von Wörtern in Mengen der Form A^*.
- Ausgezeichnete Elemente wie z.B. die 0 oder die 1 in den natürlichen (ganzen, rationalen, reellen...) Zahlen, i in den komplexen Zahlen, oder das leere Wort $\varepsilon \in A^*$.
- Unäre Operationen wie die Multiplikation mit -1, oder der Übergang von einer komplexen Zahl $c = a + bi$ zur konjugiert komplexen Zahl $\bar{c} = a - bi$, oder die Spiegelung eines Worts aus A^*.

Signaturen und deren Algebren

Bestehen die relevanten Strukturen ausschließlich aus Operationen, die ggf. noch gewissen Gleichungen (s.u.) genügen müssen, spricht man gemeinhin von algebraischen Strukturen.

Zwecks Abstraktion unterscheidet man in der sog. universellen Algebra Strukturen zunächst gemäß ihres Typs, d.h., wieviele Operationen welcher Stelligkeit auftreten. Dieser lässt sich in Form einer Abbildung $\boldsymbol{Fun} \xrightarrow{\mathcal{S}} \mathbb{N}$ beschreiben, genannt Signatur. Die Elemente von $\boldsymbol{Fun}$ heißen Operatoren. Für konkrete Anwendungen sind diese zu interpretieren:

Definition

Eine Algebra $\langle X, I \rangle$ des Typs ($\boldsymbol{Fun} \xrightarrow{\mathcal{S}} \mathbb{N}$), kurz $\mathcal{S}$-Algebra, besteht aus

- einer Menge X;
- einer Operation $X^{\mathcal{S}(g)} \xrightarrow{I(g)} X$ für jeden Operator $g \in \boldsymbol{Fun}$.

Homomorphismen sind genauso wichtig wie Algebren

Definition

Sind $\langle X, I\rangle$ und $\langle Y, J\rangle$ Algebren des Typs ($\boldsymbol{Fun} \xrightarrow{\mathcal{S}} \mathbb{N}$), so heißt eine Abbildung $X \xrightarrow{f} Y$ Homomorphismus, sofern für jedes $g \in \boldsymbol{Fun}$ gilt

$$\begin{array}{ccc} X^{\mathcal{S}(g)} & \xrightarrow{f^{\mathcal{S}(g)}} & Y^{\mathcal{S}(g)} \\ \Big\downarrow{\scriptstyle I(g)} & & \Big\downarrow{\scriptstyle J(g)} \\ X & \xrightarrow[f]{} & Y \end{array}$$

wobei $f^{\mathcal{S}(g)}$ komponentenweise wie f operiert. Explizit heißt das

$$f(I(g)(\boldsymbol{x})) = J(g)\langle\, f(x_i) \,:\, i < \mathcal{S}(g)\,\rangle$$

für alle $\boldsymbol{x} = \langle\, x_i \,:\, i < \mathcal{S}(g)\,\rangle \in X^{\mathcal{S}(g)}$.

Termalgebren

Definition

Für eine Signatur $\boldsymbol{Fun} \xrightarrow{\mathcal{S}} \mathbb{N}$ und eine Menge $\mathcal{V}$ ist $\boldsymbol{Term}(\mathcal{S}, \mathcal{V})$, die Menge der syntaktischen $\mathcal{S}$-Terme über $\mathcal{V}$, die kleinste Teilmenge von $(\boldsymbol{Fun} + \mathcal{V})^*$ mit folgenden Abschlusseigenschaften:

- $\mathcal{V} \subseteq \boldsymbol{Term}(\mathcal{S}, \mathcal{V})$;
- für jedes $g \in \boldsymbol{Fun}$ und $\boldsymbol{t} = \langle\, t_i \,:\, i < \boldsymbol{ar}\, g \,\rangle \in \boldsymbol{Term}(\mathcal{S}, \mathcal{V})^{\mathcal{S}(g)}$ gilt

$$g\,\boldsymbol{t} = g\; t_0 \;\ldots\; t_{\mathcal{S}(g)-1} \in \boldsymbol{Term}(\mathcal{S}, \mathcal{V})$$

In $\boldsymbol{Term}(\mathcal{S}, \mathcal{V})$ haben die Operatoren aus $\mathcal{S}$ eine kanonische Interpretation

$$\Phi(g)\boldsymbol{t} := g\,\boldsymbol{t}$$

$\langle \boldsymbol{Term}(\mathcal{S}, \mathcal{V}), \Phi \rangle$ heißt freie $\mathcal{S}$-Algebra über $\mathcal{V}$.

Vereinfachend schreiben wir meist g statt $\Phi(g)$.

Strukturelle Induktion

Obige induktiven Definition von $\boldsymbol{Term}(\mathcal{S},\mathcal{V})$ impliziert ein Beweisprinzip:

Satz

E sei eine Eigenschaft, die $\mathcal{S}$-Terme haben können oder nicht. Sofern

- *jedes $v \in \mathcal{V}$ Eigenschaft E hat, und*
- *für jedes $g \in \boldsymbol{Fun}$ folgt aus der Eigenschaft E für alle Komponenten von $\boldsymbol{t} \in \boldsymbol{Term}(\mathcal{S},\mathcal{V})^{\mathcal{S}(g)}$ die Eigenschaft E für $g\boldsymbol{t}$,*

haben alle $\mathcal{S}$-Term über $\mathcal{V}$ Eigenschaft E. □

Folgerung. (HA)

Jeder $\mathcal{S}$-Term über $\mathcal{V}$ kann auf genau eine Weise aus einfacheren derartigen Termen konstruiert werden. Insbesondere ist kein echtes Präfix und kein echtes Postfix eines $\mathcal{S}$-Terms über $\mathcal{V}$ ein derartiger Term.

Klammern oder keine Klammern, das ist hier die Frage

Obige sog. polnische Notation ▸ polnische Notation $g\boldsymbol{t}$ für einen mit Hilfe von $g \in \boldsymbol{Fun}$ konstruierten Term enthält keine Klammern (und somit keine Redundanz), denn die Elemente der disjunkten (!) Vereinigung $\boldsymbol{Fun} + \mathcal{V}$ sind eindeutig unterscheidbar, auch wenn sie zu Wörtern konkateniert wurden.

- Zwecks besserer Lesbarkeit von Termen erlauben wir aber, statt $g\boldsymbol{t}$ auch $g(\boldsymbol{t})$ oder $g(t_0, \ldots, t_{\mathcal{S}(g)-1})$ zu schreiben. Klammern und Kommata sind aber nicht Teil des Terms. Sie helfen dabei, die $\mathcal{S}(g)$ rechts auf g folgenden Terme t_i schneller zu identifizieren.
- Bei der sog. ▸ umgekehrt polnische Notation $\boldsymbol{t}g = t_0 \ldots t_{\mathcal{S}(g)-1}\, g$ für Terme sind beim Lesen von g (von links nach rechts) alle $\mathcal{S}(g)$ relevanten Terme t_i bereits bekannt, sie liegen z.B. auf einen Stack (vergl. frühere HP Taschenrechner), dessen erste $\mathcal{S}(g)$ Level beim Auswerten von g verschwinden (ENTER-Taste!). Andere Beispiele: PostScript, Forth.
- Die Infix-Schreibweise für binäre Operatoren in $\mathcal{S}$, etwa in der AL, braucht zwingend Klammern: z.B. ist $A \to B \to C$ uneindeutig!

Der Rekursionssatz

Diverse Konstruktionen in der AL wie auch der PL_1 basieren auf der Anwendung des folgenden Rekursionssatzes. Statt dort das Rad immer wieder neu zu erfinden, formulieren wir ihn hier in allgemeiner Form:

Satz

Ist $\langle X, I\rangle$ *eine Algebra des Typs* $\mathbf{Fun} \xrightarrow{\mathcal{S}} \mathbb{N}$, *dann lässt sich jede Belegung der Variablen* $\mathcal{V} \xrightarrow{\varphi} X$ *eindeutig zu einem* $\mathcal{S}$*-Homomorphismus* $\langle \mathbf{Term}(\mathcal{S}, \mathcal{V}), \Phi\rangle \xrightarrow{\bar{\varphi}} \langle X, I\rangle$ *fortsetzen.*

Beweis.

Existenz: Definiere $\bar{\varphi}$ induktiv:

- $\bar{\varphi}(v) := \varphi(v)$ für $v \in \mathcal{V}$;
- $\bar{\varphi}\big(g(\boldsymbol{t})\big) := I(g)\langle \bar{\varphi}(t_0), \ldots, \bar{\varphi}(t_{\mathcal{S}(g)-1})\rangle$ für $g \in \mathcal{S}$ und $\boldsymbol{t} = \langle\, t_i \,:\, i < \mathcal{S}(g)\,\rangle \in \mathbf{Term}(\mathcal{S}, \mathcal{V})^{\mathcal{S}(g)}$.

Nach Konstruktion ist dies ein $\mathcal{S}$-Homomorphismus, der φ fortsetzt.

Fortsetzung.

Eindeutigkeit: Ist $\langle \boldsymbol{Term}(\mathcal{S},\mathcal{V}),\Phi\rangle \xrightarrow{\psi} \langle X, I\rangle$ ein $\mathcal{S}$-Homomorphismus, der φ fortsetzt, so wenden wir strukturelle Induktion an:
Anfang: für jedes $v \in \mathcal{V}$ gilt

$$\bar{\varphi}(v) = \varphi(v) = \psi(v)$$

Annahme: Für $g \in S$ und $\boldsymbol{t} = \langle\, t_i \,:\, i < \mathcal{S}(g)\,\rangle \in \boldsymbol{Term}(\mathcal{S},\mathcal{V})^{\mathcal{S}(g)}$ gelte

$$\bar{\varphi}(t_i) = \psi(t_i) \quad , \quad i < \mathcal{S}(g)$$

Schluss:

$$\begin{aligned}\bar{\varphi}\big(g(\boldsymbol{t})\big) &= I(g)\langle \bar{\varphi}(t_0),\ldots,\bar{\varphi}(t_{\mathcal{S}(g)-1})\rangle \\ &= I(g)\langle \psi(t_0),\ldots,\psi(t_{\mathcal{S}(g)-1})\rangle = \psi\big(g(t_0,\ldots,t_{\mathcal{S}(g)-1})\big)\end{aligned}$$

Somit stimmt $\bar{\varphi}$ auf g-Termen mit ψ überein, also überall. back □

Anhang B
Abzählbarkeit

Abzählbarkeit

$\mathbb{N}$ bezeichnet die unendliche Menge $\{0, 1, 2, \dots\}$ der natürlichen Zahlen.

Definition

Eine Menge B heißt abzählbar, wenn es eine injektive Abbildung $B \xrightarrow{f} \mathbb{N}$ gibt. Andernfalls heißt sie überabzählbar.

Insbesondere ist jede endliche Menge abzählbar. Um diese auszuschließen, spricht man von abzählbar unendlichen Mengen.

Satz.

Folgende Bedingungen für B sind äquivalent:

(a) B ist abzählbar.

(b) $B = \emptyset$ oder es gibt eine surjektive Abbildung $\mathbb{N} \xrightarrow{g} B$.

(c) Es gibt eine surjektive partielle Abbildung $\mathbb{N} \xrightarrow{h} B$. □

Solch ein g oder h heißt dann Aufzählung von B.

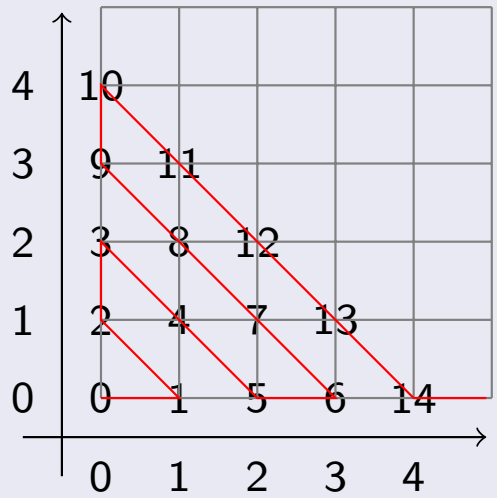
Satz

Teilmengen, endliche cartesische Produkte und abzählbare Vereinigungen abzählbarer Mengen sind wieder abzählbar. Dagegen sind Potenzmengen abzählbar unendlicher Mengen überabzählbar.

Beweis.

Teilmengen: Komponiere die injektive Inklusion $C \xrightarrow{\iota} B$ mit $B \xrightarrow{f} \mathbb{N}$.

Cartesisches Produkt: Es genügt, $\mathbb{N} \times \mathbb{N}$ als abzählbar nachzuweisen. Eine Injektion ist z.B. gegeben durch



Beweis (Fortsetzung).

Abzählbare Vereinigungen: Es genügt, die disjunkte Vereinigung von $\mathbb{N}$ Kopien von $\mathbb{N}$ als abzählbar nachzuweisen (warum?). Aber eine derartige Vereinigung ist isomorph zu $\mathbb{N} \times \mathbb{N}$.

Unendliche Potenzmengen: Es genügt zu zeigen, dass $\boldsymbol{P}(\mathbb{N})$ überabzählbar ist. Wir verfahren indirekt: ist $\boldsymbol{P}(\mathbb{N}) \xrightarrow{g} \mathbb{N}$ injektiv, so betrachten wir die Menge

$$K := \{\, g(B) \,:\, B \subseteq \mathbb{N} \wedge g(B) \notin B \,\} \subseteq \mathbb{N}$$

Wegen der Injektivität von g ist K das einzige g-Urbild von $g(K)$. Gilt $g(K) \in K$, muss das aufgrund von $g(K) \notin K$ der Fall sein. Umgekehrt kann $g(K) \notin K$ nur aufgrund von $g(K) \in K$ gelten, Widerspruch. (Dieses Argument funktioniert für jede unendliche Menge X und zeigt, dass keine injektive Abbildung von $P(X)$ nach X existiert.) □

Lemma

Ist X abzählbar, so auch die Menge aller endlichen Wörter (= Tupel) über X, d.h., die disjunkte Vereinigung

$$X^* := \sum_{i \in \mathbb{N}} X^i = X^0 + X^1 + X^2 + \cdots$$

Beweis.

Es handelt sich um eine disjunkte Vereinigung endlicher cartesischer Produkte abzählbarer Mengen, und diese ist nach obigem Satz abzählbar. □

Corollar.

Die Menge $\mathcal{F}[\mathcal{A}]$ aller aussagenlogischen Formeln ist abzählbar.

Beweis.

Es handelt sich um eine Teilmenge von $(\mathcal{A} + \mathcal{J})^*$. □

Lemma

Für jede abzählbare Signatur $\mathcal{S}$ ist die Menge $\boldsymbol{FO}(\mathcal{S})$ abzählbar.

Beweis.

Da $\mathcal{V}$ abzählbar ist, gilt das auch für $(\mathcal{S} + \mathcal{V})^*$ und somit für die Teilmenge $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V})$. Die atomaren Formeln sind nun Wörter über den Alphabet $\boldsymbol{Term}(\boldsymbol{Fun}, \mathcal{V}) + \boldsymbol{Pred} + \{(,)\}$, bilden also auch eine abzählbare Menge. Nun greift im Wesentlichen dasselbe Argument, wie für die Aussagenlogik: mittels abzahlbar vieler Junktoren lassen sich nur abzählbar viele endliche Formeln erzeugen. □

Anhang C
Zermelo-Fraenkel Mengenlehre und AC

ZFC

Die übliche Zermelo[20]-Fraenkel[21] Mengenlehre lässt sich z.B. mit der Signatur $\mathcal{S}_{\mathrm{set}} = \{\emptyset_{/0}; \in_{/2}\}$ formulieren und basiert auf 8 Axiom-Schemata:

- Extensionalität: zwei Mengen stimmen genau dann überein, wenn sie dieselben Elemente haben:

$$\forall x \, \forall y \, \big(\forall z \, (z \in x \leftrightarrow z \in y) \rightarrow x \doteq y\big)$$

- Leere Menge: die Menge $\emptyset$ hat keine Elemente:

$$\forall x \, \neg(x \in \emptyset)$$

- Paarmengen: Aus je zwei Mengen lässt sich eine neue Menge mit genau diesen Elementen bilden (ungeordnete Paare):

$$\forall x \, \forall y \, \exists z \, \forall w \, \big(w \in z \leftrightarrow (w \doteq x \vee w \doteq y)\big)$$

  Schreibweise: $\{x, y\}$ für z.

[20] Ernst Friedrich Ferdinand Zermelo (1871–1953)
[21] Abraham Halevi Fraenkel (1891–1965)

- Vereinigung: je zwei Mengen können vereinigt werden:

$$\forall x\, \exists y\, \forall z\, \big(z \in y \leftrightarrow \exists w\, (z \in w \wedge w \in x)\big)$$

  Schreibweise: $\bigcup x$ für y, oder $u \cup v$ falls $x = \{u, v\}$.

- Unendlichkeit: es gib eine unendliche Menge:

$$\exists x\, \big(\emptyset \in x \wedge \forall y\, (y \in x \to y \cup \{y\} \in x)\big)$$

- Potenzmengen: zu jeder Menge kann man die Menge ihrer „Teilmengen“ bilden:

$$\forall x\, \exists y\, \forall z\, \big(z \in y \leftrightarrow \forall w\, (w \in z \to w \in x)\big)$$

  Schreibweise: $\boldsymbol{P}(x)$ für y.

- Regularität oder Fundierung: jede nichtleere Menge enthält ein zu ihr disjunktes Element:

$$\forall x \left(\neg(x \doteq \emptyset) \to \exists y\, \big(y \in x \wedge \neg\exists z\, (z \in x \wedge z \in y)\big)\right)$$

- Aussonderung: zu jedem in ZF ausdrückbaren Eigenschaft existiert zu jeder Menge die Teilmenge derjenigen Elemente mit dieser Eigenschaft: Falls $\boldsymbol{FV}(A) \subseteq \{x, z\} \cup \{w_i : i < n\}$ (Metasprache!), dann

$$\forall x \, \forall w_0 \cdots \forall w_{n-1} \, \exists y \, \forall z \, (z \in y \leftrightarrow z \in x \wedge A)$$

  Schreibweise: $\{ z : z \in x \wedge A \}$ oder $\{ z \in x : A \}$ für y.

- Auswahl (Choice), wird von Praktikern oft hinzugenommen: aus jedem Element u einer Menge x paarweise disjunkter nichtleerer Mengen kann man jeweils genau ein ($\exists!$) Element auswählen:

$$\forall x \Big((\emptyset \notin x) \wedge \forall y \, \forall z \, \big(y \in x \wedge z \in x \rightarrow x \doteq y \vee \neg \exists w \, (w \in x \wedge w \in y) \\ \rightarrow \exists w \, \forall u \, (u \in x \rightarrow \exists! v \, (v \in u \wedge v \in w))) \Big)$$

Ein großer Teil der üblichen Mathematik lässt sich in ZFC formulieren, etwa die Konzepte einer natürlichen Zahl, der Menge $\mathbb{N}$, einer Funktion, der Injektivität, der Surjektivität, der Abzählbarkeit, etc.