Grundlagen der Logischen Programmierung

Udo Hebisch

WS 2011/12

Dieses Skript enthält nur den “roten Faden” der Vorlesung. Wesentliche Inhalte werden ausschließlich in der Vorlesung vermittelt. Daher ist dieses Skript nicht zum Selbststudium gedacht, sondern nur als “Erinnerungsstütze”.

Inhaltsverzeichnis

1 Syntax der Prädikatenlogik

Zur Wiederholung der Begriffe der Aussagenlogik vergleiche man den ersten Abschnitt des Skriptes zur Diskreten Mathematik

www.mathe.tu-freiberg.de/~hebisch/skripte/diskmath/diskmath.pdf

1.1 Terme und Formeln

Definition 1.1 Eine *Signatur* $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{P})$ besteht aus drei paarweise disjunkten Mengen von *Symbolen*:

- einer Menge $\mathcal{S}$ von *Sorten*symbolen,
- einer Menge $\mathcal{F}$ von *Funktions*- oder *Operationssymbolen*,
- einer Menge $\mathcal{P}$ von *Prädikaten*symbolen,

wobei jedes $f \in \mathcal{F}$ und $P \in \mathcal{P}$ eine *Stelligkeit* $n \in \mathbb{N}_0$ und eine *Sortigkeit* besitzt, geschrieben $f(s_1, \ldots, s_n) \to s$ bzw. $P(s_1, \ldots, s_n)$. Dabei heißen $s_1, \ldots, s_n$ die *Argumentsorten* und s die *Zielsorte*.

Für $n = 0$ heißt f *Konstanten*symbol und P *Aussagen*symbol oder *Aussagenvariable*.

Beispiel 1.2 a) Zur Axiomatisierung der Gruppentheorie.

b) Zur Axiomatisierung der Vektorrechnung.

Im folgenden sei stets $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{P})$ eine beliebige, aber feste Signatur und $V = (V_s)_{s \in \mathcal{S}}$ eine zu $\mathcal{S}, \mathcal{F}$ und $\mathcal{P}$ disjunkte Mengenfamilie von *Variablen* zu den Sorten $s \in \mathcal{S}$. Im Unterschied zu den Aussagenvariablen werden sie auch *Objektvariablen* genannt.

Definition 1.3 Die Menge $T_\Sigma(V)$ der (Σ-) *Terme mit Variablen aus* V ist induktiv definiert:

1. Jede Konstante $c \to s$ aus $\mathcal{F}$ ist ein Term zur Sorte s.

2. Jede Variable $v \in V_s$ ist ein Term zur Sorte s.
3. Ist $f(s_1, \ldots, s_n) \to s$ ein Funktionssymbol aus $\mathcal{F}$ und sind $t_1, \ldots, t_n$ Terme zu den Sorten $s_1, \ldots, s_n$, so ist $f(t_1, \ldots, t_n)$ ein Term zur Sorte s.

$T_\Sigma(V)$ besteht genau aus den gemäß 1. - 3. gebildeten Termen.

Bemerkung 1.4 Der induktive Aufbau von $T_\Sigma(V)$ gestattet es, Funktionen auf $T_\Sigma(V)$ rekursiv zu definieren, z. B. die Funktion $\mathrm{var} : T_\Sigma(V) \to \mathfrak{P}(V)$ gemäß

1. $\mathrm{var}(c \to s) := \emptyset$.
2. $\mathrm{var}(v) := \{v\}$ für $v \in V$.
3. $\mathrm{var}(f(t_1, \ldots, t_n)) := \mathrm{var}(t_1) \cup \ldots \cup \mathrm{var}(t_n)$.

Ein Term $t \in T_\Sigma(V)$ mit $\mathrm{var}(t) = \emptyset$ heißt *variablenfrei* oder *Grundterm*.

Im folgenden existiere zu jeder Sorte $s \in \mathcal{S}$ mindestens ein Grundterm, also wenigstens eine Konstante $c \to s$. (Sonst ist die Sorte bei späteren Betrachtungen überflüssig, vgl. auch Bemerkung 1.11.)

Beispiel 1.5 Terme zu Beispiel 1.2 a).

Definition 1.6 a) Ist $P(s_1, \ldots, s_n)$ ein Prädikatensymbol und sind $t_1, \ldots, t_n$ aus $T_\Sigma(V)$ Terme zu den Sorten $s_1, \ldots, s_n$, so ist $P(t_1, \ldots, t_n)$ eine *atomare Formel* oder ein *Atom*.

b) Die Menge der *quantorenfreien* (Σ-)*Formeln* (mit Variablen aus V) ist induktiv definiert:

1. Jede atomare Formel ist eine quantorenfreie Formel.
2. Die *Wahrheitswerte* $W, F \notin \mathcal{S} \cup \mathcal{F} \cup \mathcal{P} \cup V$ sind quantorenfreie Formeln.
3. Sind A und B quantorenfreie Formeln, so auch

$$(\neg A),\ (A \wedge B),\ (A \vee B),\ (A \to B),\ (A \leftrightarrow B).$$

Gemäß 1. - 3. entstehen sämtliche quantorenfreien Formeln. Sie werden auch *Formeln der offenen Prädikatenlogik* genannt. (Auch die *Junktoren* $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ sollen in keiner der Mengen $\mathcal{S}, \mathcal{F}, \mathcal{P}, V$ vorkommen.)

c) Ist A ein Atom, dann heißt A *positives* und $\neg A$ *negatives Literal*.

d) Die in den quantorenfreien Formeln enthaltenen Variablen heißen *freie* oder *ungebundene* Variablen. Auch für quantorenfreie Formeln G kann wie in Bemerkung 1.4 rekursiv die Menge $\mathrm{var}(G)$ der in G enthaltenen Variablen definiert werden. Im Fall $\mathrm{var}(G) = \emptyset$ heißt G eine *Grundformel*.

Beispiel 1.7 Gruppenaxiome.

Definition 1.8 a) Die Menge der (*prädikatenlogischen*) (Σ-)*Formeln* erhält man induktiv durch

1. Jede quantorenfreie Formel ist eine Formel.
2. Ist A eine Formel, in der die Variable $x \in V$ frei vorkommt, dann ist $\forall x A$ eine Formel, der *Allabschluß von* A bezüglich x. Ebenso ist $\exists x A$ eine Formel, der *Existenzabschluß von* A bezüglich x. In beiden Fällen heißt die zuvor frei vorkommende Variable x nun durch den jeweiligen *Quantor* ($\forall$ bzw. $\exists$) *gebunden*. Sind alle Variablen einer Formel gebunden, so heißt die Formel *geschlossen*.
3. Sind A und B Formeln, so auch
$$(\neg A),\ (A \wedge B),\ (A \vee B),\ (A \rightarrow B),\ (A \leftrightarrow B).$$
Freie Variablen bleiben hierbei frei.

Genau die gemäß 1. - 3. gebildeten Formeln sind die *Formeln der Prädikatenlogik erster Stufe*, vgl. auch Bemerkung 1.12.

b) Der Abschluß einer Formel A bezüglich aller in ihr freien Variablen nur mit Allquantoren bzw. nur mit Existenzquantoren heißt ihr *Allabschluß* bzw. *Existenzabschluß*.

Bemerkung 1.9 Um Klammern zu sparen, werden Prioritäten für Quantoren und Junktoren eingeführt:

$$\forall, \exists \text{ vor } \neg \text{ vor } \wedge \text{ vor } \vee \text{ vor } \rightarrow \text{ vor } \leftrightarrow .$$

Bei gleicher Priorität wird wie üblich eine Klammerung von links nach rechts angenommen.

Beispiel 1.10 Allabschluß und Existenzabschluß.

Bemerkung 1.11 Die Sorten $\mathcal{S}$ können eine unstrukturierte Menge bilden (*flache Sortenstruktur*) oder durch eine partielle Ordnung $\sqsubseteq$ geordnet sein (*hierarchische Sortenstruktur*). Die zulässige Struktur dieser *Untersortenbeziehung* $\sqsubseteq$, z. B. linear, baumförmig, verbandsartig, beeinflußt die möglichen Variablenbelegungen und daher die Konstruktion von Kalkülen der jeweiligen Logik. Ein Beispiel ist etwa $\mathcal{S} = \{Integer \sqsubseteq Real \sqsubseteq Complex\}$.

Mehrsortige Logiken sind nicht ausdrucksstärker als solche ohne Sorten: Für jede Sorte $s \in \mathcal{S}$ wird ein neues einstelliges Prädikat $S \notin \mathcal{P}$ eingeführt und die Untersortenbeziehung $S_1 \sqsubset S_2$ durch die Implikation $\forall x(S_1(x) \rightarrow S_2(x))$ ausgedrückt. Für sortierte Variable $x \in V_s$ wird $\forall xG$ dann als $\forall x(S(x) \rightarrow G)$ notiert und $f(s_1, \ldots, s_n) \rightarrow s$ als $\forall x_1 \ldots \forall x_n(S_1(x_1) \wedge \ldots \wedge S_n(x_n) \rightarrow S(f(x_1, \ldots, x_n)))$.

Bemerkung 1.12 In der Prädikatenlogik erster Stufe ist die Quantifizierung (Definition 1.8 a)) nur über Objektvariable erlaubt, nicht über Funktions- oder Prädikatenvariable, also nicht:

$$\forall P \text{ symmetrisch}(P) \leftrightarrow (\forall x \forall y P(x, y) \leftrightarrow P(y, x)).$$

In der Prädikatenlogik zweiter Stufe ist die Quantifizierung dann auch über Mengen von Objekten erlaubt. Identifiziert man nun ein Prädikat P mit der Menge der Objekte, auf die es zutrifft, so ist also auch die Quantifizierung über Funktions- oder Prädikatenvariable erlaubt, also auch:

$$\forall M \forall N \ (M \subseteq N \leftrightarrow \forall x(x \in M \rightarrow x \in N)).$$

Bekanntlich ist die Struktur der reellen Zahlen erst durch das Supremumsaxiom bis auf Isomorphie eindeutig charakterisiert. Dabei wird über alle beschränkten Teil*mengen* quantifiziert, also nicht mehr in der Prädikatenlogik erster Stufe.

Ebenso erfordert das Induktionsaxiom für die natürlichen Zahlen

$$\forall P \ (P(0) \wedge (\forall x(P(x) \rightarrow P(x+1)) \rightarrow \forall y P(y))$$

die Prädikatenlogik zweiter Stufe.

Quantifiziert man auch über Mengen von Mengen von Objekten, dann ist man in der Prädikatenlogik dritter Stufe usw.

1.2 Substitutionen

Definition 1.13 a) Eine Abbildung $\sigma : V \to T_\Sigma(V)$ gemäß $x \mapsto x\sigma$ heißt *Substitution*, wenn x und $x\sigma$ für alle $x \in V$ dieselbe Sorte haben und $\mathrm{dom}(\sigma) := \{x \in V \mid x\sigma \neq x\}$ endlich ist. Für $\mathrm{dom}(\sigma) = \{x_1, \ldots, x_n\}$ und $x_i\sigma = t_i, i = 1, \ldots, n$ schreibt man auch $\sigma = [x_1|t_1, \ldots, x_n|t_n]$ oder

$$\sigma = \begin{pmatrix} x_1 & \ldots & x_n \\ t_1 & \ldots & t_n \end{pmatrix}.$$

Die identische Substitution wird mit $\varepsilon = [\ ]$ bezeichnet. Ist σ eine Permutaion von V, so heißt σ eine *Umbenennung*.

Unter $\mathrm{var}(\sigma) := \bigcup_{x \in \mathrm{dom}(\sigma)} \mathrm{var}(x\sigma)$ wird die Menge der Variablen im Bildbereich einer Substitution verstanden. Gilt $\mathrm{var}(\sigma) = \emptyset$, so heißt σ *Grundsubstitution*.

b) Substitutionen σ werden wie folgt auf Terme, quantorenfreie Formeln und Mengen quantorenfreier Formeln zu σ^* fortgesetzt:

1. $x\sigma^* := x\sigma$ für alle $x \in V$.
2. $(f(t_1, \ldots, t_n))\sigma^* := f(t_1\sigma^*, \ldots, t_n\sigma^*)$ für alle anderen Terme
   $(P(t_1, \ldots, t_n))\sigma^* := P(t_1\sigma^*, \ldots, t_n\sigma^*)$ für alle Atome.
3. $W\sigma^* := W$, $F\sigma^* := F$,
   $(\neg A)\sigma^* := \neg(A\sigma^*)$,
   $(A \wedge B)\sigma^* := (A\sigma^* \wedge B\sigma^*)$, $(A \vee B)\sigma^* := (A\sigma^* \vee B\sigma^*)$,
   $(A \to B)\sigma^* := (A\sigma^* \to B\sigma^*)$, $(A \leftrightarrow B)\sigma^* := (A\sigma^* \leftrightarrow B\sigma^*)$.
4. $M\sigma^* := \{m\sigma^* \mid m \in M\}$ für Mengen M quantorenfreier Formeln.

Man schreibt dann wieder einfach σ für σ^*.

c) Die *Komposition* $\sigma\tau$ zweier Substitutionen σ und τ ist definiert gemäß $t(\sigma\tau) := (t\sigma)\tau$ für alle Terme t und $A(\sigma\tau) := (A\sigma)\tau$ für alle quantorenfreien Formeln A. Man nennt dann $\sigma\tau$ eine *Instanz* von σ.

d) Ist σ eine Substitution und t ein Term, so heißt $t\sigma$ eine *Instanz* von t. Ist $t\sigma$ variablenfrei, so heißt $t\sigma$ eine *Grundinstanz* von t. Mit $\mathrm{grund}(t)$ werde die Menge aller Grundinstanzen eines Terms t bezeichnet. Für quantorenfreie Formeln und Mengen derartiger Formeln werden diese Bezeichnungen analog verwendet.

Beispiel 1.14 Substitutionen, Umbenennungen, Grundinstanz.

Lemma 1.15 *Für alle Substitutionen σ, τ, ϱ gelten:*

a) $\sigma\varepsilon = \varepsilon\sigma = \sigma$,

b) $(\sigma\tau)\varrho = \sigma(\tau\varrho)$,

c) $\sigma\tau = \sigma \Longrightarrow x\tau = x$ *für alle* $x \in var(\sigma)$.

1.3 Ableitungen

Definition 1.16 a) Eine (n*-stellige*) *Ableitungsregel* (besser: ein Regelschema) mit den *Prämissen* $G_1, \ldots, G_n$ und der *Konklusion* G ist ein syntaktisches Schema der Gestalt

$$\begin{array}{c} G_1 \\ \vdots \\ G_n \\ \hline G \end{array} \qquad \text{bzw.} \qquad \frac{G_1 \ldots G_n}{G}$$

mit prädikatenlogischen Formeln $G, G_1, \ldots, G_n$. Ein *Regelsystem* $\mathcal{R}$ ist eine Menge von Ableitungsregeln.

b) Eine Ableitungsregel auf gegebene Formeln *anwenden* heißt, eine Instanz des Regelschemas mit diesen Formeln zur Deckung zu bringen und daraus die Konklusion als entsprechende Instanz zu erhalten. Dabei soll stets entscheidbar sein, ob das Regelschema auf diese Formeln anwendbar ist, und die Konklusion soll daraus berechenbar sein. Eine solche Regelanwendung heißt ein *Ableitungsschritt*.

Beispiel 1.17 Schnittregel.

Beispiel 1.18 Substitutionsregel.

Definition 1.19 Es sei $\mathcal{R}$ ein Regelsystem, X eine Formelmenge und G eine Formel. Eine *Ableitung von* G *mit* $\mathcal{R}$ *aus* X ist induktiv definiert:

1. Für jedes $G \in X$ ist (G) eine Ableitung von G mit $\mathcal{R}$ aus X.
2. Ist $(C_1, \ldots, C_k)$ eine Ableitung von C_k mit $\mathcal{R}$ aus X und $R \in \mathcal{R}$ eine n-stellige Ableitungsregel, so daß durch Anwendung von R auf die Formeln $G_1, \ldots, G_n \in X \cup \{C_1, \ldots, C_k\}$ die Konklusion G entsteht, dann ist $(C_1, \ldots, C_k, G)$ eine Ableitung von G mit $\mathcal{R}$ aus X.

Gibt es eine Ableitung von G mit $\mathcal{R}$ aus X, so heißt *G mit $\mathcal{R}$ aus X ableitbar*, in Zeichen: $X \vdash_{\mathcal{R}} G$ (bei zwei Regeln auch: $X \vdash_{R_1+R_2} G$).

Bemerkung 1.20 Jede Ableitung ist endlich. Bei Aussagen über die Ableitbarkeit kann man sich daher auf endliche Formelmengen beschränken und Beweise mit Induktion über den Aufbau der Ableitung führen.

1.4 Normalformen

Definition 1.21 a) Eine Formel der Gestalt

$$Q_1 \dots Q_n \; M$$

ist in *pränexer Normalform*, wenn M (die *Matrix*) eine quantorenfreie Formel ist und die Q_i jeweils aus einem Quantor gefolgt von einer Variablen bestehen.

b) Eine Formel in pränexer Normalform wird in *Skolem-Normalform* überführt oder *skolemisiert* (Toralf Skolem, 1887 - 1963), indem man **von außen nach innen** alle Quantoren eliminiert gemäß

1. Allquantoren $\forall x$ fallen für jede Variable $x \in V$ einfach weg, die gebundene Variable x wird wieder frei,
2. anstelle der Variablen y des wegfallenden Existenzquantors $\exists y$ wird ein neuer Term substituiert, der aus einem Funktionssymbol $f \notin \mathcal{F}$ besteht (Änderung der Signatur!), welches mit allen bisher "befreiten" Variablen parametrisiert wird.

Beispiel 1.22 Skolemisierung.

Definition 1.23 a) Eine quantorenfreie Formel der Gestalt $K_1 \vee \dots \vee K_n$ ist in *disjunktiver Normalform*, wenn die K_i Konjunktionen von Literalen sind.

b) Eine quantorenfreie Formel der Gestalt $D_1 \wedge \dots \wedge D_n$ ist in *konjunktiver Normalform*, wenn die D_i Disjunktionen von Literalen sind.

c) Eine quantorenfreie Formel ist in *Negations-Normalform*, wenn sie nur mittels der Junktoren $\neg, \wedge, \vee$ aufgebaut ist, wobei $\neg$ nur direkt an Atomen steht.

Definition 1.24 Sind $P_i,\ i = 1, \ldots, n$ und $Q_j,\ j = 1, \ldots, m$ Atome, so heißt

$$P_1 \wedge \ldots \wedge P_n \rightarrow Q_1 \vee \ldots \vee Q_m$$

Gentzenformel oder *Klausel* (Gerhard Gentzen, 1909 - 1945). Äquivalent hierzu ist

$$\neg P_1 \vee \ldots \vee \neg P_n \vee Q_1 \vee \ldots \vee Q_m$$

oder in mengentheoretischer Schreibweise kurz

$$\{\neg P_1, \ldots, \neg P_n, Q_1, \ldots, Q_m\}.$$

Spezialfälle sind

$n = 0$ (*positive Gentzenformel*): $W \rightarrow Q_1 \vee \ldots \vee Q_m$,

$m = 0$ (*negative Gentzenformel*): $P_1 \wedge \ldots \wedge P_n \rightarrow F$,

$n = 0, m = 1$ (*positive Einheitsklausel*, Prolog: Fakt): $W \rightarrow Q$,

$n = 1, m = 0$ (*negative Einheitsklausel*): $P \rightarrow F$,

$n = m = 0$ (*Widerspruch*): $W \rightarrow F$, abgekürzt: $\Box$

Gentzenformeln mit $m \leq 1$ heißen *Hornformeln*. Spezialfälle sind:

$m = 1$ (*nichtnegative Hornformel*, Prolog: Regel): $P_1 \wedge \ldots \wedge P_n \rightarrow Q$,

$m = 0$ (*negative Hornformel*, Prolog: Anfrage): $P_1 \wedge \ldots \wedge P_n \rightarrow F$.

Bemerkung 1.25 Die **S**chnittregel (GS) wird also auf **G**entzenformeln angewandt und enthält folgende Spezialfälle:

Positiver Schnitt: $\dfrac{W \rightarrow B \vee P \quad P \wedge C \rightarrow D}{C \rightarrow B \vee D}$

Negativer Schnitt: $\dfrac{A \rightarrow B \vee P \quad P \wedge C \rightarrow F}{A \wedge C \rightarrow B}$

Positiver Einheitsschnitt: $\dfrac{W \rightarrow P \quad P \wedge C \rightarrow D}{C \rightarrow D}$

Negativer Einheitsschnitt: $\dfrac{A \rightarrow B \vee P \quad P \rightarrow F}{A \rightarrow B}$

Prolog-Interpretationsregel: $\dfrac{A \rightarrow P \quad P \wedge C \rightarrow F}{A \wedge C \rightarrow F}$

in Prolog-Notation:

$$\begin{array}{l} P :- A_1, \ldots, A_n. \\ ?- P, C_1, \ldots, C_m. \\ \hline ?- A_1, \ldots, A_n, C_1, \ldots, C_m. \end{array}$$

2 Semantik der Prädikatenlogik

2.1 Interpretationen und Modelle

Definition 2.1 a) Ein Tripel $\mathbb{M} = (\mathbb{D}, \mathbb{F}, \mathbb{P})$ heißt *Struktur zur Signatur* $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{P})$, wenn gilt

1. $\mathbb{D} = (\mathbb{D}_s)_{s \in \mathcal{S}}$ ist eine Familie nichtleerer (*Daten-* oder *Individuen-*)*Mengen* $\mathbb{D}_s =: I(s)$.
2. $\mathbb{F}$ ist eine Menge von Funktionen, wobei zu $f(s_1, \ldots, s_n) \to s$ aus $\mathcal{F}$ stets eine Funktion $I(f) := g : \mathbb{D}_{s_1} \times \ldots \times \mathbb{D}_{s_n} \to \mathbb{D}_s$ existiert. Speziell existiert zu jeder Konstanten $c \to s$ ein Element $I(c) := d \in \mathbb{D}_s$.
3. $\mathbb{P}$ ist eine Menge von *Relationen* (oder *Prädikaten*), wobei zu $P(s_1, \ldots, s_n)$ aus $\mathcal{P}$ stets eine Relation $I(P) := p \subseteq \mathbb{D}_{s_1} \times \ldots \times \mathbb{D}_{s_n}$ existiert. Ein Prädikat p kann auch als *Wahrheitswertfunktion* $p : \mathbb{D}_{s_1} \times \ldots \times \mathbb{D}_{s_n} \to \{W, F\}$ aufgefaßt werden. Im Fall $n = 0$, also für Aussagenvariablen P, ist daher $I(P) = p \in \{W, F\}$ wie in der Aussagenlogik einfach ein Wahrheitswert.

Die Familie der Abbildungen $I : \mathcal{S} \to \mathbb{D}$, $I : \mathcal{F} \to \mathbb{F}$, $I : \mathcal{P} \to \mathbb{P}$ wird auch *Interpretation von* Σ (in $\mathbb{M}$) genannt.

b) Eine *Variablenzuweisung* ist eine Abbildung $\omega : V \to \mathbb{D}$ gemäß $x \mapsto \omega(x) \in \mathbb{D}_s$ für alle Variablen $x \in V_s$.

Beispiel 2.2 Gruppe $\mathbb{Z}_4$.

Definition 2.3 Mit den Bezeichnungen aus Definition 2.1 wird der *Wert eines Terms* $\mathrm{wert}_I^\omega(t) \in \mathbb{D}$ bzw. der *Wert einer Formel* $\mathrm{wert}_I^\omega(A) \in \{W, F\}$ *bezüglich* I *und* ω wie folgt definiert:

1. $\mathrm{wert}_I^\omega(c) := I(c)$ für alle Konstanten c aus $\mathcal{F}$ (unabhängig von ω!),
2. $\mathrm{wert}_I^\omega(x) := \omega(x)$ für alle Variablen $x \in V$,
3. $\mathrm{wert}_I^\omega(f(t_1, \ldots, t_n)) := I(f)(\mathrm{wert}_I^\omega(t_1), \ldots, \mathrm{wert}_I^\omega(t_n))$ für alle $f \in \mathcal{F}$ und dazu passende Terme t_i,
4. $\mathrm{wert}_I^\omega(W) := W$, $\mathrm{wert}_I^\omega(F) := F$ unabhängig von I und ω,

5. $\text{wert}_I^\omega(P) := I(P)$ für alle Aussagensymbole $P \in \mathcal{P}$,

6. $\text{wert}_I^\omega(P(t_1, \dots, t_n)) := I(P)(\text{wert}_I^\omega(t_1), \dots, \text{wert}_I^\omega(t_n))$ für alle $P \in \mathcal{P}$ und dazu passende Terme t_i,

7. $\text{wert}_I^\omega(\neg A) := W$ genau dann, wenn $\text{wert}_I^\omega(A) := F$,

8. $\text{wert}_I^\omega(A \wedge B) := W$ genau dann, wenn $\text{wert}_I^\omega(A) = W$ **und** $\text{wert}_I^\omega(B) = W$ (hieraus folgt die Assoziativität, Kommutativität und Idempotenz des Junktors $\wedge$),

9. $\text{wert}_I^\omega(A \vee B) := W$ genau dann, wenn $\text{wert}_I^\omega(A) = W$ **oder** $\text{wert}_I^\omega(B) = W$ (hieraus folgt die Assoziativität, Kommutativität und Idempotenz des Junktors $\vee$),

10. $\text{wert}_I^\omega(A \to B) := W$ genau dann, wenn $\text{wert}_I^\omega(A) = F$ oder $\text{wert}_I^\omega(B) = W$ (dies ist die übliche Interpretation des Junktors $\to$),

11. $\text{wert}_I^\omega(A \leftrightarrow B) := W$ genau dann, wenn $\text{wert}_I^\omega(A) = \text{wert}_I^\omega(B)$,

12. $\text{wert}_I^\omega(\forall x A) := W$ genau dann, wenn $\text{wert}_I^{\omega'}(A) = W$ für alle Variablenzuweisungen ω', die sich höchstens in x von ω unterscheiden,

13. $\text{wert}_I^\omega(\exists x A) := W$ genau dann, wenn eine Variablenzuweisung ω' mit $\text{wert}_I^{\omega'}(A) = W$ existiert, die sich höchstens in x von ω unterscheidet.

Nach 12. und 13. ist für eine geschlossene Formel A der Wert unabhängig von der Variablenzuweisung. Man definiert daher in diesem Fall *den Wert von A in der Struktur* $\mathbb{M}$: $\text{wert}_\mathbb{M}(A) := \text{wert}_I(A) := \text{wert}_I^\omega(A)$ für eine beliebige Variablenzuweisung ω.

Der Wert einer Formel mit freien Variablen ist dann der Wert ihres Allabschlusses.

Definition 2.4 a) Eine Formel G (über einer Signatur Σ) *gilt in einer Struktur* $\mathbb{M}$ (zur Signatur Σ) genau dann, wenn der Allabschluß von G in $\mathbb{M}$ den Wert W besitzt (kurz: *wahr ist*). Man nennt dann $\mathbb{M}$ ein *Modell* von G. Die Struktur $\mathbb{M}$ ist Modell einer Formelmenge X, wenn $\mathbb{M}$ Modell aller Formeln $G \in X$ ist.

b) Eine Formelmenge X heißt *erfüllbar*, wenn sie ein Modell besitzt, andernfalls *unerfüllbar* oder *widersprüchlich*.

c) Zwei Formelmengen X und Y heißen *erfüllbarkeitsgleich*, wenn entweder beide erfüllbar oder beide unerfüllbar sind.

d) Eine Formel heißt *allgemeingültig* oder eine *Tautologie*, wenn jede Struktur ein Modell von ihr ist.

e) Eine Formel G *folgt* (*logisch*) aus einer Formelmenge X, in Zeichen $X \models G$, genau dann, wenn jedes Modell von X auch Modell von G ist. Analog definiert man $X \models Y$ für Formelmengen Y.

Folgen zwei Formelmengen wechselseitig auseinander, so heißen sie *äquivalent*.

Unmittelbar aus Definition 2.4 b) und e) folgt

Satz 2.5 *Für eine Formelmenge X und eine geschlossene Formel G gilt:*

Es ist $X \models G$ genau dann, wenn $X \cup \{\neg G\}$ widersprüchlich ist.

Beispiel 2.6 Modelle gruppentheoretischer Formeln.

Es gelten folgende *Normalformen-Sätze*:

Satz 2.7 *Jede prädikatenlogische Formel G läßt sich in eine zu ihr äquivalente Formel G' in pränexer Normalform überführen.*

Satz 2.8 *Jede Formel G' in pränexer Normalform läßt sich in eine dazu erfüllbarkeitsgleiche Formel G'' in Skolem-Normalform überführen.*

Satz 2.9 *Jede quantorenfreie Formel G'' läßt sich in eine dazu äquivalente endliche Menge $X = \{G_i\}$ von Gentzenformeln G_i überführen.*

Anstelle von Beweisen:

Beispiel 2.10 Umformung einer Formel in eine Menge von Gentzenformeln.

2.2 Herbrand-Modelle

Jaques Herbrand, 1908 - 1931.

Definition 2.11 a) Eine *Herbrand-Struktur zur Signatur* $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{P})$ ist eine Struktur $\mathbb{H} = (\mathbb{U}, \mathbb{F}_T, \mathbb{P})$ mit

1. dem *Herbrand-Universum* $\mathbb{U} = (\mathbb{U}_s)_{s\in\mathcal{S}}$ und $\mathbb{U}_s := \{t \mid t$ ist Grundterm zur Sorte $s\} \neq \emptyset$ (vgl. Verabredung über Grundterme) (eindeutig bestimmt!),

2. dem *Funktionenbereich* $\mathbb{F}_T$ gemäß: Für $f(s_1,\ldots,s_n) \to s$ aus $\mathcal{F}$ ist $I(f) := f_T : \mathbb{U}_{s_1}\times\ldots\times\mathbb{U}_{s_n} \to \mathbb{U}_s$ aus $\mathbb{F}_T$ definiert durch $f_T(t_1,\ldots,t_n) := f(t_1,\ldots,t_n)$ aus $\mathbb{U}_s$ (eindeutig bestimmt!).

3. dem *Prädikatenbereich* $\mathbb{P}$ gemäß: Für $P(s_1,\ldots,s_n)$ aus $\mathcal{P}$ ist $I(P) := p : \mathbb{U}_{s_1}\times\ldots\times\mathbb{U}_{s_n} \to \{W,F\}$ aus $\mathbb{P}$ *beliebig* festgelegt.

b) Die *Herbrand-Basis* $\mathbb{B}$ *zur Signatur* Σ ist die Menge aller Grundatome

$$\mathbb{B} := \{P(t_1,\ldots,t_n) \mid P(s_1,\ldots,s_n) \text{ aus } \mathcal{P}, t_{s_i} \in \mathbb{U}_{s_i}\}.$$

Beispiel 2.12 Herbrand-Universum und Herbrand-Basis.

Lemma 2.13 *Jede Festlegung der Prädikate einer Herbrand-Struktur* $\mathbb{H}$ *(und damit einer Interpretation* I *von* Σ *in* $\mathbb{H}$*) läßt sich eindeutig beschreiben durch eine Teilmenge* $\mathbb{A}$ *der Herbrand-Basis* $\mathbb{B}$ *gemäß*

$$P(t_1,\ldots,t_n) \in \mathbb{A} \text{ genau dann, wenn } wert_{\mathbb{H}}(P(t_1,\ldots,t_n)) = W.$$

Definition 2.14 Ein *Herbrand-Modell* $\mathbb{M}$ einer Formelmenge X ist eine Herbrand-Struktur $\mathbb{H} = \mathbb{M}$, die Modell von X ist.

Satz 2.15 *Eine Menge von quantorenfreien Formeln ist genau dann erfüllbar, wenn sie ein Herbrand-Modell besitzt.*

Beispiel 2.16 Herbrand-Modelle.

Im Gegensatz zur Situation in diesem Beispiel gilt

Satz 2.17 *Ist* X *eine Menge von Hornformeln und* $\mathbb{M}_i$ *ein Herbrand-Modell von* X *für* $i \in I \neq \emptyset$, *dann ist* $\mathbb{M} := \bigcap_{i\in I}\mathbb{M}_i$ *ebenfalls ein Herbrand-Modell von* X. *Speziell ist* $\mathbb{M}_X := \bigcap\{\mathbb{M} \mid \mathbb{M}$ *ist ein Herbrand-Modell von* $X\}$ *das* kleinste Herbrand-Modell von X *für jede erfüllbare Menge* X *von Hornformeln.*

Zum Beweis dieses Satzes und zur Berechnung von $\mathbb{M}_X$ vgl. J. W. Lloyd, Foundations of Logic Programming, Springer 1987.

2.3 Korrektheit und Vollständigkeit

Definition 2.18 Ein Regelsystem $\mathcal{R}$ heißt *korrekt*, falls gilt:

$$X \mapsto_{\mathcal{R}} G \text{ impliziert } X \models G,$$

es heißt *vollständig*, falls gilt:

$$X \models G \text{ impliziert } X \mapsto_{\mathcal{R}} G,$$

und es heißt *widerlegungsvollständig*, falls gilt:

$$\text{“ } X \text{ widersprüchlich” impliziert} X \mapsto_{\mathcal{R}} \square$$

für alle Formelmengen X und alle Formeln G (einer bestimmten Formelklasse).

Bemerkung 2.19 Nach Satz 2.5 gilt für beliebige Formelmengen X und **geschlossene** Formeln G:

$$X \models G \text{ ist äquivalent zu } X \cup \{\neg G\} \text{ ist widersprüchlich.}$$

Da der Wahrheitswert einer (quantorenfreien) Gentzenformel gleich dem Wahrheitswert ihres Allabschlusses ist, kann man sich bei der Betrachtung von Gentzenformeln auf widerlegungsvollständige (statt vollständige) Regelsysteme beschränken.

Bezeichnet (GS) die Schnittregel für Gentzenformeln und $(Subst)$ die Substitution für Gentzenformeln, dann gilt:

Satz 2.20 *Das Regelsystem $\mathcal{R} = \{(GS), (Subst)\}$ ist korrekt für Gentzenformeln.*

Bemerkung 2.21 Natürlich sind auch alle in Bemerkung 1.25 angegebenen Spezialfälle der Schnittregel korrekt, insbesondere also Prolog für Hornformeln.

Wesentlich aufwendiger ist der Nachweis der Widerlegungsvollständigkeit (zum Beweis vgl. 1.49 bis 1.55 in Hofbauer/Kutsche, Grundlagen des maschinellen Beweisens, Vieweg 1989):

Satz 2.22 *Das Regelsystem $\mathcal{R} = \{(GS), (Subst)\}$ ist widerlegungsvollständig für Gentzenformeln.*

Bemerkung 2.23 Eine Analyse der in Hofbauer/Kutsche angegebenen Beweise zeigt insbesondere die Widerlegungsvollständigkeit der Prolog-Interpretationsregel für Hornformeln.

Für automatische Beweisverfahren entscheidend ist der folgende *Satz von Herbrand.*

Satz 2.24 *Eine Menge X von Gentzenformeln ist genau dann widersprüchlich, wenn eine endliche Menge von Grundinstanzen von X existiert, die widersprüchlich ist.*

3 Resolutionsregeln

3.1 Unifikation

Für eine feste Signatur $\Sigma = (\mathcal{S}, \mathcal{F}, \mathcal{P})$ bezeichne *Subst* die Menge aller Substitutionen $\sigma : V \to T_\Sigma(V)$ gemäß Definition 1.13. Nach Lemma 1.15 handelt es sich bezüglich der Komposition um eine Halbgruppe mit dem Einselement $\varepsilon = [\ ]$. Durch

$$\sigma \leq \tau :\Longleftrightarrow \text{ es gibt ein } \varrho \in \mathit{Subst} \text{ mit } \sigma\varrho = \tau$$

für alle $\sigma, \tau \in \mathit{Subst}$ (gelesen: σ ist allgemeiner als τ) wird eine Relation $\leq$ auf *Subst* definiert, die ersichtlich reflexiv ($\varrho = \varepsilon$!) und transitiv ist, also $(\mathit{Subst}, \leq)$ zu einer *quasigeordneten* Menge macht. Diese Quasiordnung ist i. a. nicht antisymmetrisch, also keine partielle Ordnung: Für $x \neq y$ aus V_s und $\sigma = [x|y] \neq \tau = [y|x]$ hat man nämlich $\sigma\tau = \tau$ und $\tau\sigma = \sigma$, also $\sigma \leq \tau$ und $\tau \leq \sigma$.

Wie für jede Quasiordnung $\leq$ wird durch

$$\sigma \sim \tau :\Longleftrightarrow \sigma \leq \tau \text{ und } \tau \leq \sigma$$

aber eine Äquivalenzrelation $\sim$ auf *Subst* definiert und die Äquivalenzklassen dann durch

$$[\sigma]_\sim \leq [\tau]_\sim :\Longleftrightarrow \sigma \leq \tau$$

partiell geordnet. Ersichtlich ist $[\varepsilon]_\sim$ kleinstes Element von $(\mathit{Subst}/\sim, \leq)$.

Lemma 3.1 *Für $\sigma, \tau \in \mathit{Subst}$ gilt $\sigma \sim \tau$ genau dann, wenn es eine Umbenennung $\pi \in \mathit{Subst}$, also eine Permutation von V, mit $\sigma\pi = \tau$ gibt. Insbesondere ist $[\varepsilon]_\sim = \{\pi \in \mathit{Subst} \mid \pi \text{ ist Umbenennung}\}$.*

Bemerkung 3.2 Beim Beweis von "$\Rightarrow$" ist wesentlich, daß $\sigma : V \to T_\Sigma(V)$ nur endlich viele $x \in V$ verändert, vgl. Definition 1.13. Sonst hätte man etwa das Gegenbeispiel $\sigma = [x_0|x_1, \ldots, x_k|x_{k+1}, \ldots]$ und $\tau = \varepsilon$ mit $\sigma \leq \tau$ und $\tau \leq \sigma$, aber $\sigma\pi \neq \tau (\Longleftrightarrow \sigma \neq \pi^{-1})$ für jede Permutation π von V.

Definition 3.3 Zwei Terme $t, t' \in T_\Sigma(V)$ bzw. eine Termmenge $M \subseteq T_\Sigma(V)$ heißen *unifizierbar*, falls eine Substitution σ mit $t\sigma = t'\sigma$ bzw. mit $t_1\sigma = t_2\sigma$ für alle $t_1, t_2 \in M$ (also $|M\sigma| = 1$) existiert. In diesem Fall heißt σ ein *Unifikator* von t und t' bzw. von M. Ein Unifikator heißt *allgemeinster Unifikator* (von t und t' bzw. von M), wenn $\sigma \leq \tau$ für alle Unifikatoren τ von t und t' bzw. von M gilt.

Bemerkung 3.4 Nach Lemma 3.1 sind allgemeinste Unifikatoren also bis auf Umbenennungen eindeutig bestimmt.

Sind t und t' bzw. M unifizierbar, dann dient das folgende, 1982 publizierte Verfahren zur Berechnung des allgemeinsten Unifikators und damit zum Nachweis seiner Existenz.

Definition 3.5 Die *Martelli-Montanari-Regeln* zur Berechnung allgemeinster Unifikatoren sind

$(U1)$ Dekompositionsregel

$$\frac{E \vee \{f(t_1, \dots, t_n) \equiv f(t'_1, \dots, t'_n)\}}{E \cup \{t_1 \equiv t'_1, \dots, t_n \equiv t'_n\}}$$

$(U2)$ Variablen-Elimination

$$\frac{E \vee \{x \equiv t\}}{E\tau}$$

x kommt in t nicht vor, wohl aber in E, und es ist $\tau = [x|t]$

$(U3)$ Umordnungsregel

$$\frac{E \vee \{t \equiv x\}}{E \cup \{x \equiv t\}}$$

t ist keine Variable

$(U4)$ Elimination trivialer Gleichungen

$$\frac{E \vee \{x \equiv x\}}{E}$$

Dabei ist E eine Menge gerichteter Gleichungen, $\vee$ die disjunkte Vereinigung, t, t_i, t'_i, sind Terme, f ein Funktionssymbol und x eine Variable.

Beispiel 3.6 Berechnung eines allgemeinsten Unifikators.

Bemerkung 3.7 Der Test in $(U2)$, ob x in t vorkommt, heißt auch "occur-check".

Satz 3.8 *Es sei E eine Menge gerichteter Gleichungen und E' aus E durch Anwendung einer der Regeln $(U1) - (U4)$ entstanden. Dann gelten:*

(i) E ist genau dann mit dem Unifikator σ unifizierbar, wenn E' mit σ unifizierbar ist.

(ii) σ ist allgemeinster Unifikator von E genau dann, wenn σ allgemeinster Unifikator von E' ist.

Definition 3.9 Unifikationsalgorithmus Es seien t, t' Terme bzw. M eine endliche Menge von Termen. Der Unifikationsalgorithmus läuft dann folgendermaßen ab:

1. Bilde die *Start-Gleichungsmenge* $E_0 = \{t \equiv t'\}$ bzw. $E_0 = \{t \equiv t' \mid t, t' \in M, t \neq t'\}$.
2. Solange dies möglich ist, wende eine der Regeln $(U1) - (U4)$ an.
3. Ist keine Regel mehr anwendbar, so ist die *Ergebnis-Gleichungsmenge* E^* erreicht.

Satz 3.10 *a) Der Unifikationsalgorithmus terminiert für jede Start-Gleichungsmenge E_0.*

b) Gilt $E^* = \{x_1 \equiv t_1, \ldots, x_k \equiv t_k\}, k > 0$ mit paarweise verschiedenen Variablen x_i, die in keinem der Terme t_j mehr vorkommen, so sind t und t' bzw. M unifizierbar mit dem allgemeinsten Unifikator $\sigma = [x_1|t_1, \ldots, x_k|t_k]$. In diesem Fall heißt E^* vollständig gelöst. *Andernfalls sind t und t' bzw. M nicht unifizierbar.*

Bemerkung 3.11 a) Der Unifikationsalgorithmus benötigt im schlimmsten Fall exponentiellen Speicher- und Zeitaufwand.

b) Es gibt Unifikationsalgorithmen mit quadratischem, ja sogar fast linearem Aufwand.

3.2 Resolution

Definition 3.12 Es seien P, Q Atome, A und C Konjunktionen sowie B und D Disjunktionen von Atomen. Weiterhin sei π eine Umbenennung der Variablen von $Q \wedge C \to D$ derart, daß $A \to B \vee P$ und $(Q \wedge C \to D)\pi$ keine gemeinsamen

Variablen enthalten, und σ ein allgemeinster Unifikator von P und $Q\pi$. Dann heißt die Ableitungsregel

$$(Res)\frac{A \to B \vee P \quad Q \wedge C \to D}{(A \wedge C\pi \to B \vee D\pi)\sigma}$$

(*pure*) *Resolution.*

Beispiel 3.13 Beispiel zur Resolution.

Bemerkung 3.14 Die pure Resolution ist noch nicht widerlegungsvollständig, denn für Variablen x, y, u, v ist $X = \{W \to P(x) \vee P(y), P(u) \wedge P(v) \to F\}$ widersprüchlich. Aus der ersten Formel folgt nämlich ($x = y = u = v$!) zunächst $P(x)$ und mit der zweiten Formel dann $W \to F$. Mit (Res) bleiben aber bei jedem Ableitungsschritt zwei Literale erhalten, der Widerspruch wird also nicht gefunden. Dies liegt daran, daß in den Prämissen von (Res) nicht nur P und $Q\pi$ unifizierbar sind, sondern noch mehrere Atome. Hier sind dies $P(x), P(y), P(u)$ und $P(v)$.

Definition 3.15 Es seien $P_1, \ldots, P_n$ für $n \geq 1$ Atome, A eine Konjunktion und B eine Disjunktion von Atomen, sowie σ ein allgemeinster Unifikator von $\{P_1, \ldots, P_n\}$. Dann heißt die Ableitungsregel

$$(Fak) \quad \frac{A \to B \vee P_1 \vee \ldots P_n}{(A \to B \vee P_1)\sigma} \text{ bzw. } \frac{P_1 \wedge \ldots \wedge P_n \wedge A \to B}{(P_1 \wedge A \to B)\sigma}$$

Faktorisierungsregel. Die Konklusionen heißen *Faktoren* der jeweiligen Prämissen.

Bemerkung 3.16 a) In Klauselschreibweise

$$\frac{\{\neg A_1, \ldots, \neg A_m, B_1, \ldots, B_\ell, P_1, \ldots, P_n\}}{\{\neg A_1\sigma, \ldots, \neg A_m\sigma, B_1\sigma, \ldots, B_\ell\sigma, P_1\sigma\}}$$

bzw.

$$\frac{\{\neg A_1, \ldots, \neg A_m, B_1, \ldots, B_\ell, \neg P_1, \ldots, \neg P_n\}}{\{\neg A_1\sigma, \ldots, \neg A_m\sigma, B_1\sigma, \ldots, B_\ell\sigma, \neg P_1\sigma\}}$$

sieht man, daß es sich tatsächlich nur um eine Regel handelt.

b) Mit Resolution und Faktorisierung findet man auch den Widerspruch in

$$X = \{W \to P(x) \vee P(y), P(u) \wedge P(v) \to F\},$$

denn durch Faktorisierung erhält man etwa $W \to P(x)$ mit $\sigma_1 = [y|x]$ und $P(u) \to F$ mit $\sigma_2 = [v|u]$, woraus sich mit (Res) der Widerspruch $W \to F$ ergibt.

1965 führte J. A. Robinson die folgende Kombination von (Res) und (Fak) ein:

Definition 3.17 Es seien $P_1, \ldots, P_n, Q_1, \ldots, Q_m$ Atome, A und C Konjunktionen sowie B und D Disjunktionen von Atomen. Weiterhin sei π eine Umbenennung der Variablen von $Q_1 \wedge \ldots \wedge Q_m \wedge C \rightarrow D$ derart, daß $A \rightarrow B \vee P_1 \vee \ldots \vee P_n$ und $(Q_1 \wedge \ldots \wedge Q_m \wedge C \rightarrow D)\pi$ keine gemeinsamen Variablen enthalten, und σ ein allgemeinster Unifikator von $\{P_1, \ldots, P_n, Q_1, \ldots, Q_m\}$. Dann heißt die Ableitungsregel

$$(Rob) \quad \frac{A \rightarrow B \vee P_1 \vee \ldots \vee P_n \quad Q_1 \wedge \ldots \wedge Q_m \wedge C \rightarrow D}{(A \wedge C\pi \rightarrow B \vee D\pi)\sigma}$$

volle oder *klassische Resolution.*

Lemma 3.18 Lifting-Lemma *Es seien G_1' und G_2' Gentzenformeln mit Instanzen G_1 bzw. G_2, so daß der Schnitt (GS) $\frac{G_1 G_2}{G}$ möglich ist. Dann existiert eine Gentzenformel G^*, die durch einen Resolutionsschritt aus Faktoren von G_1' und G_2' entsteht, so daß G Instanz von G^* ist.*

Beispiel 3.19 Beispiel zum Lifting-Lemma.

Satz 3.20 *Die beiden Regelsysteme $(Res + Fak)$ bzw. (Rob) sind korrekt und widerlegungsvollständig für Gentzenformeln.*

Bemerkung 3.21 Das Lifting-Lemma zeigt lokal (für einen einzelnen Schnitt), daß Ableitungen mittels Schnitten und *beliebigen* Substitutionen in solche mit Schnitten und *allgemeinsten Unifikatoren* “geliftet” werden können.

4 Suchstrategien

Die Widerlegungsvollständigkeit eines Regelsystems $\mathcal{R}$ besagt für eine widersprüchliche Formelmenge X nur die Existenz einer Ableitung $X \vdash_{\mathcal{R}} \Box$. Für die Effektivität eines (maschinellen) Beweises ist aber die Frage, welche Regel aus $\mathcal{R}$ mit welchen Prämissen aus den bisher abgeleiteten Formeln für den jeweils nächsten Ableitungsschritt anzuwenden ist, von entscheidender Bedeutung. Hierfür sind mittlerweile sehr viele Strategien untersucht worden, von denen nun einige vorgestellt werden.

4.1 Breitensuche (Level-Saturation)

Hierbei werden systematisch alle Formeln erzeugt, die sich in einem Schritt, dann in zwei Schritten usw. aus X mittels $\mathcal{R}$ herleiten lassen. Dazu sei noch definiert

$\mathcal{R}(X) := \{C \mid C$ entsteht durch einmalige Anwendung einer Regel aus $\mathcal{R}$ auf Formeln aus $X\}$.

Definition 4.1 Die *Breitensuche* bildet für eine gegebenen Formelmenge X und ein gegebenes Regelsystem $\mathcal{R}$ sukzessive die Formelmenge S_n des n-ten Levels für $n \in \mathbb{N}$:

$$(BS) \qquad \begin{aligned} S_0 &:= X, \\ S_{n+1} &:= S_n \cup \mathcal{R}(S_n). \end{aligned}$$

Bemerkung 4.2 Bei der Breitensuche werden also alle möglichen Ableitungen parallel erzeugt. Das bedeuttet, daß $|S_n|$ sehr stark anwächst ("kombinatorische Explosion"). Für ein widerlegungsvollständiges Regelsystem $\mathcal{R}$ existiert aber genau für widersprüchliche Formelmengen X ein $n \in \mathbb{N}$ mit $\Box \in S_n$. Bei der konkreten Berechnung von S_{n+1} wird man natürlich nur Regeln anwenden, die mindestens eine Formel aus $S_n \setminus S_{n-1}$ als Prämisse verwenden. Man wird daher (BS) verfeinern, etwa für $\mathcal{R} = \{(Res), (Fak)\}$:

Für Klasuelmengen Y und Z bezeichne

$Res(Y, Z)$ die Menge aller in einem Schritt zwischen Klauseln aus Y und Klauseln aus Z bildbaren Resolventen

$Fak(Y)$ die Menge aller in einem Schritt aus Y bildbaren Faktoren.

Dann sei

$$(BS) \qquad \begin{aligned} S_0 &:= X, \\ S_1 &:= X \cup Res(X, X) \cup Fak(X), \\ S_{n+1} &:= S_n \cup Res(S_n \setminus S_{n-1}, S_n) \cup Fak(S_n \setminus S_{n-1}). \end{aligned}$$

4.2 Präferenz-Strategien

Hierbei benutzt man ein korrektes und widerlegungsvollständiges Regelsystem $\mathcal{R}$ und ein korrektes, nicht widerlegungsvollständiges Regelsystem $\mathcal{T}$, dessen Regeln aber "effektiver" anwendbar sind (Anwendbarkeit mit weniger Aufwand zu entscheiden, Konklusionen schneller zu berechnen, Konklusionen sind "einfachere" Formeln, etc.). Man zieht dann die Regeln aus $\mathcal{T}$ denen aus $\mathcal{R}$ vor, ohne die Widerlegungsvollständigkeit zu verlieren.

Definition 4.3 Eine *Präferenz-Strategie* bildet für eine gegebene Formelmenge X und zwei Regelsysteme $\mathcal{R}$ und $\mathcal{T}$ in Abhängigkeit von $k > 0$ sukzessive die Formelmengen S_n:

$$(PS) \qquad \begin{array}{l} S_0 := X, \\ S_{n+1} := S_n \cup \mathcal{R}(S_n), \text{ falls } n+1 \equiv 0 \; mod \; k \\ S_{n+1} := S_n \cup \mathcal{T}(S_n), \text{ sonst.} \end{array}$$

Bemerkung 4.4 Der Parameter k steuert in einfacher Weise, wie stark $\mathcal{T}$-Schritte gegenüber $\mathcal{R}$-Schritten bevorzugt werden. Die Strategie ist aber *fair*, denn jeder $\mathcal{R}$-Schritt, der bei (BS) (k=1) ohne Benutzung von $\mathcal{T}$ durchgeführt worden wäre, kommt auch bei (PS) vor, wenn auch verzögert. Auch für diese Strategie sind Verfeinerungen möglich, aber die Fairness ist jeweils zu beweisen. Dann existiert bei widerlegungsvollständigem $\mathcal{R}$ unabhängig von $\mathcal{T}$ für jede widerspruchsvolle Formelmenge X ein $n \in \mathbb{N}$ mit $\Box \in S_n$.

Beispiel 4.5 Gentzenschnitt und Einheitsschnitte.

Für aussagenlogische Gentzenformeln sei $\mathcal{R} = \{(GS)\}$, also ein korrektes und widerlegungsvollständiges Regelsystem, und $\mathcal{T} = \{(T_1), (T_2)\}$, wobei (T_1) den positiven und (T_2) den negativen Einheitsschnitt bezeichne, also ein korrektes, aber nicht widerlegungsvollständiges Regelsystem. Betrachte die Strategien (BS) und (PS) bei $k = 3$ für die Formelmenge

$$\begin{array}{rcl} X & = & \{(1)P, \quad (2)P \to Q, \quad (3)P \wedge R \to S, \quad (4)P \wedge S \to R, \\ & & (5)P \wedge Q \to R \vee S, \quad (6)P \wedge Q \wedge R \wedge S \to F\}. \end{array}$$

Dabei ist (5) keine Hornformel und (6) das zu beweisende "negierte Theorem" $P \wedge Q \wedge R \wedge S$.

Level 1

(PS): Wegen $k = 3$ wird nur $\mathcal{T}$ verwendet und erzeugt die folgenden Formeln

(7) Q (aus (1) und (2) mit (T_1)),

(8) $R \to S$ (aus (1) und (3) mit (T_1)),

(9) $S \to R$ (aus (1) und (4) mit (T_1)),

(10) $Q \to R \vee S$ (aus (1) und (5) mit (T_1)),

(11) $Q \wedge R \wedge S \to F$ (aus (1) und (6) mit (T_1)).

(T_2) ist noch nicht anwendbar!

(BS): Mit (GS) werden ebenfalls (7) - (11) erzeugt und zusätzlich noch

(12) $P \to R \vee S$ (aus (2) und (5)),

(13) $P \wedge R \wedge S \to F$ (aus (2) und (6)),

(14) $P \wedge R \to R$ (aus (3) und (4), eine Tautologie!),

(15) $P \wedge Q \wedge R \to F$ (aus (3) und (6)),

(16) $P \wedge S \to S$ (aus (4) und (3) [Reihenfolge der Prämissen!] und aus (5) und (3) [Mehrfachableitung!], Tautologie!),

(17) $P \wedge Q \wedge S \to F$ (aus (4) und (6)),

(18) $P \wedge Q \wedge S \to S$ (aus (5) und (6), Tautologie),

(19) $P \wedge Q \wedge R \to R$ (aus (5) und (6), mehrere Konklusionen!, Tautologie),

(20) $P \wedge Q \to S$ (aus (5) und (3)),

(21) $P \wedge Q \to R$ (aus (5) und (4)).

Beachtung der Reihenfolge der Prämissen und Möglichkeit mehrfacher Konklusionen bedeuten einen erhöhten Suchaufwand bei den Ableitungen, Mehrfachableitungen bedeuten einen Suchaufwand beim Abspeichern der abgeleiteten Formeln!

Level 2

(PS): Mit $\mathcal{T}$ werden außer den Formel (7) - (11) aus Level 1 noch erzeugt

(12) $P \to R \vee S$ (aus (7) und (5)),

(13) $P \wedge R \wedge S \to F$ (aus (7) und (6)),

(22) $W \to R \vee S$ (aus (7) und (10)),

(23) $R \wedge S \to F$ (aus (7) und (11)).

(BS)): (GS) erzeugt jetzt außer (22) und (23) noch

(24) $P \to R$ (aus (4) und (12)),

(25) $P \to S$ (aus (8) und (12)),

(26) $R \to R$ (aus (8) und (9), Tautologie!),

(27) $S \to S$ (aus (9) und (8), Tautologie!)

(28) $P \wedge R \to F$ (aus (8) und (13)),

(29) $Q \wedge R \to F$ (aus (8) und (11)),

(30) $Q \to S$ (aus (10) und (8)),

(31) $A \to R$ (aus (10) und (9)),

(32) $Q \wedge S \to F$ (aus (9) und (11)),

(33) $Q \wedge S \to S$ (aus (1) und (18) oder aus (10) und (11)), Tautologie!),

(34) $Q \wedge R \to R$ (aus (1) und (19) oder aus (10) und (11)), Tautologie!)

(35) $P \wedge S \to F$ (aus (9) und (13));

(36)* $P \wedge Q \to F$ (aus (17) und (20)),

(37)* $P \wedge Q \wedge R \to S$ (aus (8) und (18)),

(38)* $P \wedge Q \wedge S \to R$ (aus (9) und (19)).

Level 3

(PS): Jetzt erzeugt (PS) alle bisher noch nicht erzeugten Formeln, die (BS) im Level 1 erzeugt hat, und weitere. Insbesondere sind dies alle Formeln, die (BS) im Level 2 erzeugt hat, außer den mit * markierten.

Unter anderem werden auch die Formeln $W \to R$ und $R \to F$ bzw. $W \to S$ und $S \to F$ generiert, aus denen sich dann im nächsten Level der gesuchte Widerspruch ableiten läßt.

(BS): Natürlich erzeugt auch (BS) jetzt die zur Herleitung des Widerspruchs benötigten Formeln. Aber auch (BS) gelingt im Level 3 noch nicht die Herleitung des Widerspruchs.

4.3 Vereinfachungsregeln

Statt eine widersprüchliche Formelmenge X mit einem widerlegungsvollständigen Regelsystem $\mathcal{R}$ zu widerlegen, wird eine echte Teilmenge X' von X widerlegt, die zu X äquivalent oder zumindest erfüllbarkeitsgleich ist. Die im folgenden beschriebenen Vereinfachungsregeln sind unproblematisch, wenn man sie nur **vor** den Ableitungen mittels $\mathcal{R}$ anwendet, jedoch nicht mehr, wenn dies **während** der Ableitungen geschieht. Jede Formelmenge X ist nämlich äquivalent zu $X' \subset X$, wenn in $X \setminus X'$ nur Formeln liegen, die aus X' logisch folgen. Für ein korrektes Regelsystem $\mathcal{R}$ folgt aber jede aus X abgeleitete Formel stets logisch aus X. Jede neu abgeleitete Formel würde bei diesem Vorgehen der Vereinfachung also sofort wieder gestrichen. Ein Widerspruch würde daher nur gefunden, wenn er sich durch einen einzigen Schritt aus X ergäbe.

Definition 4.6 Die *Breitensuche mit Vereinfachungsregeln* bildet für eine gegebene Formelmenge X, ein System $\mathcal{R}$ von Ableitungsregeln und ein System $\mathcal{S}impl$ von *Vereinfachungsregeln* sukzessive die Formelmengen S_n:

$$(BSimpl) \qquad \begin{array}{l} S_0 := \mathcal{S}impl(X), \\ S_{n+1} := \mathcal{S}impl(S_n \cup \mathcal{R}(S_n)). \end{array}$$

Dabei bedeutet $\mathcal{S}impl(Y)$ eine minimale Formelmenge, die aus Y entsteht, wenn in irgend einer Weise die Regeln aus $\mathcal{S}impl$ angewendet werden.

Bemerkung 4.7 a) Je nach Wahl von $\mathcal{S}impl$ ist $\mathcal{S}impl(Y)$ nicht eindeutig bestimmt, $(BSimpl)$ also nicht deterministisch!

b) Ein Regelsystem $\mathcal{R}$ heißt *kompatibel* mit der Vereinfachungsregel $\mathcal{S}impl$, wenn genau für widersprüchliche Formelmengen X ein $n \in \mathbb{N}$ mit $\Box \in S_n$ existiert.

Definition 4.8 Die Vereinfachungsregel

$$(ET) \qquad \frac{X \cup \{C\}}{X},$$

wobei X eine Klauselmenge und die Klausel C eine Tautologie ist, heißt *Elimination von Tautologien*.

Bemerkung 4.9 a) Offensichtlich sind Prämisse und Konklusion in (ET) äquivalent, also erst recht erfüllbarkeitsgleich.

b) Die Anwendbarkeit von (ET) ist leicht entscheidbar, denn eine Klausel C ist genau dann eine Tautologie, wenn es ein Literal L mit $\{L, \neg L\} \subseteq C$ gibt.

Die nächsten beiden Vereinfachungsregeln werden hier nur für aussagenlogische Gentzenformeln formuliert. Verallgemeinerungen auf prädikatenlogische Formeln sind aber möglich.

Für eine Klauselmenge X und ein Literal L sei

$$\Delta(X, L) := \{C \setminus \{\neg L\} \mid C \in X \text{ und } L \notin C\},$$

also alle Klauseln aus X, die L nicht enthalten, aber aus diesen wird $\neg L$ entfernt. In den Klauseln aus $\Delta(X, L)$ kommt also weder L noch $\neg L$ vor.

Definition 4.10 Die Vereinfachungsregel

$$(PL) \qquad \frac{X}{\Delta(X, L)},$$

wobei X eine Menge variablenfreier Klauseln und L ein variablenfreies Literal ist, so daß $\neg L$ in keiner Klausel aus X vorkommt, heißt *Pure-Literal-Regel.*

Folgerung 4.11 In (PL) sind X und $\Delta(X, L)$ erfüllbarkeitsgleich.

Definition 4.12 Die Vereinfachungsregel

$$(UC) \qquad \frac{X \cup \{\{L\}\}}{\Delta(X, L)},$$

wobei X eine Menge variablenfreier Klauseln und L ein variablenfreies Literal ist, heißt *Einheitsklausel-Regel.*

Folgerung 4.13 In (UC) sind Prämisse und Konklusion erfüllbarkeitsgleich.

Bemerkung 4.14 Bei einer "prädikatenlogischen" Version von (PL) muß man voraussetzen, daß in keiner Klausel aus X ein Literal L' vorkommt, welches mit $\neg L$ unifizierbar ist.

Bei einer "prädikatenlogischen" Version von (UC) muß man in der Konklusion zu $\Delta(X, L)$ noch alle Resolventen hinzunehmen, die sich aus X mit L bilden lassen.

Dann lassen sich wiederum die obigen Folgerungen beweisen.

Definition 4.15 Eine Klausel C *subsumiert* eine Klausel D [*stark*] (in Zeichen $C \leq_{sub} D$ bzw. $C \leq_{ssub} D$), wenn eine Substitution $\sigma \in \mathit{Subst}$ existiert, so daß $C\sigma \subseteq D$ [und $|C| \leq |D|$] gilt. Eine Klauselmenge X *subsumiert* eine Klauselmenge Y [*stark*] (in Zeichen: $X \leq_{sub} Y$ bzw. $X \leq_{ssub} Y$), wenn zu jeder Klausel $D \in Y$ eine Klausel $C \in X$ mit $C \leq_{sub} D$ [$C \leq_{ssub} D$] existiert.

Bemerkung 4.16 a) Die Relationen $\leq_{sub}$ und $\leq_{ssub}$ sind Quasiordnungen (also reflexiv und transitiv) auf der Menge $\mathcal{C}$ aller Klauseln bzw. auf $\mathfrak{P}(\mathcal{C})$.

b) Aus $C \leq_{sub} D$ folgt $C \leq_{ssub} D$ und daraus $C \models D$.

c) Für aussagenlogische Formeln fallen diese drei Relationen zusammen, in der Prädikatenlogik gilt dies jedoch nicht.

d) Die Subsumtion ist mit der Faktorisierung nicht kompatibel, denn jeder Faktor $D = C\sigma$ einer Klausel C wird von C trivialerweise subsumiert. Zur Widerlegung notwendige Faktoren werden bei Vereinfachung durch eine allgemeine “Subsumtionsregel” aber sofort gestrichen, der Widerspruch kann also nicht mehr gefunden werden.

e) Die Relation $\leq_{sub}$ (und daher auch $\leq_{ssub}$) ist für Klauseln entscheidbar. Ein Entscheidungsalgorithmus kann mittels linearer Resolution konstruiert werden.

Definition 4.17 Die Vereinfachungsregel

$$(ES) \qquad \frac{X \cup \{C, D\}}{X \cup \{C\}},$$

wobei X eine Menge von Klauseln und C, D Klauseln mit $C \leq_{ssub} D$ sind, heißt *Elimination subsumierter Klauseln.*

Folgerung 4.18 *In (ES) sind Prämisse und Konklusion äquivalent.*

Satz 4.19 *Die Regelsysteme $\{(Res), (Fak)\}$ bzw. $\{(Rob)\}$ sind kompatibel mit $\mathcal{S}impl = \{(ET)\}$, $\mathcal{S}impl = \{(ES)\}$ und $\mathcal{S}impl = \{(ET), (ES)\}$*

Den Beweis führt man mit folgenden Hilfsaussagen.

Lemma 4.20 *a) Seien X, Y Klauselmengen mit $X \leq_{sub} Y$ und D eine Klausel mit $Y \vdash_{(Res)+(Fak)} D$. Dann existiert eine Klausel C mit $C \leq_{sub} D$ und $X \vdash_{(Res)+(Fak)} C$, d. h. die Ableitung $\vdash_{(Res)+(Fak)}$ ist monoton bezüglich $\leq_{sub}$.*

b) Zu jeder $(Res{+}Fak)$-Ableitung einer Klausel D aus einer Formelmenge X gibt es eine $(Res + Fak)$-Ableitung einer Klausel C aus X, in der keine Tautologie vorkommt, so daß $C \leq_{sub} D$ gilt.

Folgerung 4.21 *a) Aus* $X \leq_{sub} Y$ *und* $Y \vdash_{(Res)+(Fak)} \Box$ *folgt* $X \vdash_{(Res)+(Fak)} \Box$

b) Zu jeder $(Res + Fak)$*-Widerlegung von* X *existiert eine Widerlegung von* X*, die keine Tautologien enthält.*

4.4 Strukturorientierte Strategien

Bei der Widerlegung einer Klauselmenge mittels Resolution sollen als Prämissen eines Resolutions-Schrittes nur solche Klauseln zugelassen werden,

i) deren äußere Gestalt bestimmten Kriterien genügt (lokale Struktur) oder

ii) deren Herkunft beim Ableitungsvorgang bestimmten Kriterien genügt (globale Struktur).

Durch geschickte Wahl dieser Kriterien kann man bestimmte einfache Strukturen des "Ableitungsbaumes" erzwingen.

Beispiel 4.22 Ableitung mit Center-Klauseln.

Definition 4.23 a) Bei der *Unit-Resolution* ist ein Schnitt gemäß (GS) genau dann erlaubt, wenn mindestens eine der Prämissen eine Einheitsklausel ist. Solche Schnitte werden mit $(Unit)$ bezeichnet.

b) Bei der *Input-Resolution* ist ein Schnitt gemäß (GS) genau dann erlaubt, wenn mindestens eine der Prämissen aus der ursprünglich zu widerlegenden Formelmenge X stammt. Solche Schnitte werden mit $(Input)$ bezeichnet.

Satz 4.24 *Die Regelsysteme* $\{(Unit), (Fak)\}$ *und* $\{(Input), (Fak)\}$ *sind korrekt. Sie sind beide widerlegungsvollständig für Hornformeln, jedoch nicht allgemein für Gentzenformeln.*

Es gilt sogar

Satz 4.25 *Für variablenfreie Klauselmengen* X *hat man*

$$X \vdash_{(Unit)+(Fak)} \Box \text{ genau dann, wenn } X \vdash_{(Input)+(Fak)} \Box.$$

Bemerkung 4.26 Für Hornformeln ist sogar schon $(Unit)$ widerlegungsvollständig.

Die folgende Strategie der *linearen Resolution* ist diejenige, die auch für Prolog Anwendung findet.

Definition 4.27 Es sei X eine Klauselmenge und $C_1, C_2, \dots, C_k$ eine Folge von Klauseln, so daß für $1 < i \leq k$ jeweils C_i durch Schnitt zwischen C_{i-1} und einer Klausel aus $X \cup \{C_1, \dots, C_{i-1}\}$ entstanden ist. Dann ist C_k durch *lineare Resolution* aus X entstanden. Die C_j für $1 \leq j \leq k$ heißen *Center-Klauseln*, C_1 *Start-Center-Klausel* dieser Ableitung. Für $C_k = \Box$ spricht man von einer *linearen Widerlegung* von X.

Bemerkung 4.28 a) Jede Input-Ableitung ist eine lineare Ableitung. Lineare Resolution verallgemeinert also (vgl. Satz 4.32) Input-Resolution zu einem widerlegungsvollständigen Verfahren.

b) Auch der Schnitt von C_{i-1} mit sich selbst ist erlaubt.

c) Die neben den Center-Klauseln an der Ableitung beteiligten Klauseln werden auch *Side-Klauseln* genannt.

Beispiel 4.29 Lineare Widerlegung.

Lemma 4.30 *Seien C_0, B_1, B_2 variablenfreie Klauseln aus X, B Resolvente von B_1 und B_2, C Resolvente von B und C_0. Das Literal L_C von B, über das C_0 mit B geschnitten wird, stamme (o. B. d. A.) aus B_1, C_1 sei Resolvente von C_0 und B_1. Dann läßt sich diese Ableitung in eine lineare Ableitung mit Start-Center-Klausel C_0 und einer Klausel $G \subseteq C$ transformieren, so daß einer der beiden folgenden Fälle eintritt:*

1. G ist Resolvente von C_1 und B_2.

2. $G = C$ ist Resolvente von C_0 und C_2, wobei C_2 Resolvente von C_1 und B_2 ist.

Bemerkung 4.31 Man kann zeigen, daß sogar jede Formel aus jeder minimal widersprüchlichen Teilemenge X' von X zur Start-Center-Klausel einer linearen Widerlegung gemacht werden kann.

Dabei heißt eine widersprüchliche Menge X' von Klauseln *minimal widersprüchlich*, wenn jede echte Teilmenge von X' erfüllbar ist.

Mit Induktion über die Länge der Ableitung erhält man weiter:

Satz 4.32 *Jede widersprüchliche Klauselmenge X ist linear widerlegbar.*

Man kann auch mehrere der vorgestellten Strategien kombinieren:

Beispiel 4.33 Kombination der Präferenz-Strategie mit (ET) und (ES).

Man darf jedoch nicht bedenkenlos beliebige Strategien miteinander kombinieren!

Index