Partielles Gleichheits-Schließen (Paramodulation)

Vollständiger, korrekter Hintergrundbeweiser
für die Gleichheitstheorie:

$$\frac{\begin{array}{c} t \approx s \\ \phi[t'] \end{array}}{(\phi[s])\sigma} \qquad \frac{\begin{array}{c} s \approx t \\ \phi[t'] \end{array}}{(\phi[s])\sigma}$$

$$\frac{\neg(t \approx t')}{*}$$

wobei jeweils σ ein mgu von t und t' ist

Beispiel.

Tableaubeweis für die Unerfüllbarkeit der aus (1)–(5) bestehenden Formelmenge.

(1) $(\forall x)(g(x) \approx f(x) \vee \neg(x \approx a))$

(2) $(\forall x)(g(f(x)) \approx x)$

(3) $b \approx c$

(4) $p(g(g(a)), b)$

(5) $\neg p(a, c)$

(6) $g(f(x_1)) \approx x_1$

(7) $g(x_2) \approx f(x_2) \vee \neg(x_2 \approx a)$

(8) $g(x_2) \approx f(x_2)$ (9) $\neg(x_2 \approx a)$

(10) $\boxed{p(g(f(a)), b)}$

(11) $\boxed{p(a, b)}$

(12) $\boxed{p(a, c)}$

Bei Ableitung von (10) aus (4) mit (8):
Anwendung von $\sigma_1 = \{x_2 \leftarrow a\}$

Bei Ableitung von (11) aus (10) mit (6):
Anwendung von $\sigma_2 = \{x_1 \leftarrow a\}$

Partielles Gleichheits-Schließen (RUE-Resolution)

$$\frac{\begin{array}{c} p(t_1, \ldots, t_k) \\ \neg p(s_1, \ldots, s_k) \end{array}}{\neg(t_1 \approx s_1) \mid \cdots \mid \neg(t_k \approx s_k)}$$

$$\frac{\neg(f(t_1, \ldots, t_k) \approx f(s_1, \ldots, s_k))}{\neg(t_1 \approx s_1) \mid \cdots \mid \neg(t_k \approx s_k)}$$

$$\frac{t \approx s}{s \approx t}$$

$$\frac{\neg(t \approx t')}{*}$$

wobei σ ein mgu von t und t' ist

Totales Gleichheits-Schließen

Ein vollständiger, totaler Hintergrundbeweiser für Gleichheit muß die folgenden Refuter liefern.

($E(\Phi)$ bezeichne die Menge der (nicht-quantifizierten) Gleichungen $l \approx r$ in einem Key Φ)

Eine Substitution σ ist ein Refuter für Φ, wenn

$$(E(\Phi))\sigma \models^{\circ}_{\mathcal{E}} \langle t_1, \ldots, t_n \rangle\sigma \approx \langle s_1, \ldots, s_n \rangle\sigma$$

für ein Paar $p(t_1, \ldots, t_n)$ und $\neg p(s_1, \ldots, s_n)$ in Φ

– oder –

$$(E(\Phi)\sigma) \models^{\circ}_{\mathcal{E}} t\sigma \approx s\sigma$$

für $\neg(t \approx s)$ in Φ

Zu lösendes Problem: **starre E-Unifikation**

(Simultane) Starre E-Unifikation

Definition. Ein starres E-Unifikationsproblem $\langle E, s, t\rangle$ besteht aus

– einer Menge E von (quantorenfreien) Gleichungen

– Termen s und t

Eine Substitution σ ist eine Lösung von $\langle E, s, t\rangle$, wenn

$$E\sigma \models^{\circ}_{\mathcal{E}} s\sigma \approx t\sigma$$

■

Zum Test, ob alle Äste eines Tableaus zugleich geschlossen werden können:

Definition. Ein **simultanes** starres E-Unifikationsproblem ist eine endliche Menge starrer E-Unifikationsprobleme.

Eine Substitution ist Lösung eines simultanen Problems, wenn sie Lösung jedes seiner Teilprobleme ist.
■

Beispiel.

B_1: linker Ast des Beispieltableaus

B_2: rechter Ast des Beispieltableaus

$E(B_2) = \{\, b \approx c,\ g(f(x_1)) \approx x_1 \,\}$

$E(B_1) = E(B_2) \cup \{\, g(x_2) \approx f(x_2) \,\}$

Aus B_2 Ast zu extrahierende starre E-Unifikationsprobleme:

P_1: $\langle E(\Phi_2),\ \langle g(g(a)), b\rangle,\ \langle a, c\rangle\rangle$

P_2: $\langle E(\Phi_2),\ x_2,\ a\rangle$

Aus B_1 Ast zu extrahierende starre E-Unifikationsprobleme:

P_3: $\langle E(\Phi_1),\ \langle g(g(a)), b\rangle,\ \langle a, c\rangle\rangle$

$\{x_1 \leftarrow a,\ x_2 \leftarrow a\}$

ist eine simultane Lösung von P_2 und P_3 ■

Entscheidbarkeit starrer E-Unifikation

Universelle E-Unifikation ist	unentscheidbar
Starre E-Unifikation ist	entscheidbar (!)
	(NP-vollständig)
Simultane starre E-Unifikation ist	unentscheidbar (!)

Es ist möglich, eine minimale Lösung zu einem starren E-Unifikationsproblem effizient zu berechnen.

Es ist *nicht* möglich eine Lösung für eine simultanes Problem durch Kombination der minimalen Lösungen seiner Teilprobleme zu berechnen.

Ein vollständiger Tableaukalkül entsteht,
wenn man *alle* Lösungen der aus den Ästen extrahierten Probleme aufzählt und (fair) verwendet (nicht nur die minimalen).

Offenes Forschungsproblem:
Kann ein vollständiger Tableaukalkül konstruiert werden, der nur minimale (oder jedenfalls nur endlich viele) Lösungen der extrahierten Probleme verwendet?

Ein solcher Kalkül wäre besonders effizient.

Vergleich von Kalkülen und Beweisprozeduren

Vergleichsmöglichkeiten

- Länge des kürzesten Beweises bezüglich eines realistischen Begriffs von elementaren Beweisschritten (AL + PL).
- Größe des Suchraums (nur AL Horn).
- Berechnung des Erwartungswertes der durchschnittlichen Beweislänge für zufallsgenerierte Formelklassen in Abhängigkeit verschiedener Parameter wie Anzahl Klauseln/Größe der Signatur (nur AL).
- Qualitative Aussagen (wie Beweiskonfluenz)

  Theorie

  Praxis
- Laufzeitvergleiche implementierter Prozeduren für Benchmark-Suiten.
- Fallstudien in Anwendungen.

Relative Beweiskomplexität von Kalkülen

Die Länge des kürzesten Beweises entspricht der *indeterministischen* Stärke des Kalküls.

Maßfunktion $|\cdots|$

$|\cdots|$ liefert # Knoten von Klauseltableaus

$|\cdots|$ liefert # Literalauftauchen in Klauselmengen

Bemerkung:

Die Funktion $|\cdots|$ muß „realistisch" sein. Sie sollte die tatsächliche Kosten für Konstruktion eines Beweises in elementaren Maschinenschritten modulo eines polynomiellen Faktors erkennen lassen.

Problematisch für $|\cdots|$ ist beispielsweise:
Nach polynomiell vielen Beweisschritten Terme exponentieller Länge

Verursacht z.B. durch unbeschränkte Schnittregel:

$$\frac{}{p \mid \neg p} \qquad p \in \mathrm{At}_\Sigma$$

$|\cdots|$ (wie oben) für Klauseltableau realistisch.

$|\cdots|$ (wie oben) für PL Resolution nicht realistisch

Definition. (Nur AL)
Eine Beweisprozedur P ist **super** gdw es ein Polynom p gibt, so daß für jede Klauselmenge S die Länge des kürzesten P-Beweises durch $p(|S|)$ beschränkt ist. ■

Definition. (PL und AL)
Eine Beweisprozedur P kann eine Prozedur Q **p-simulieren** gdw es ein Polynom p gibt, so daß für alle Klauselmengen S gilt:

wenn es einen Q-Beweis für S der Länge n gibt, dann gibt es einen P-Beweis für S höchstens der Länge $p(n)$.

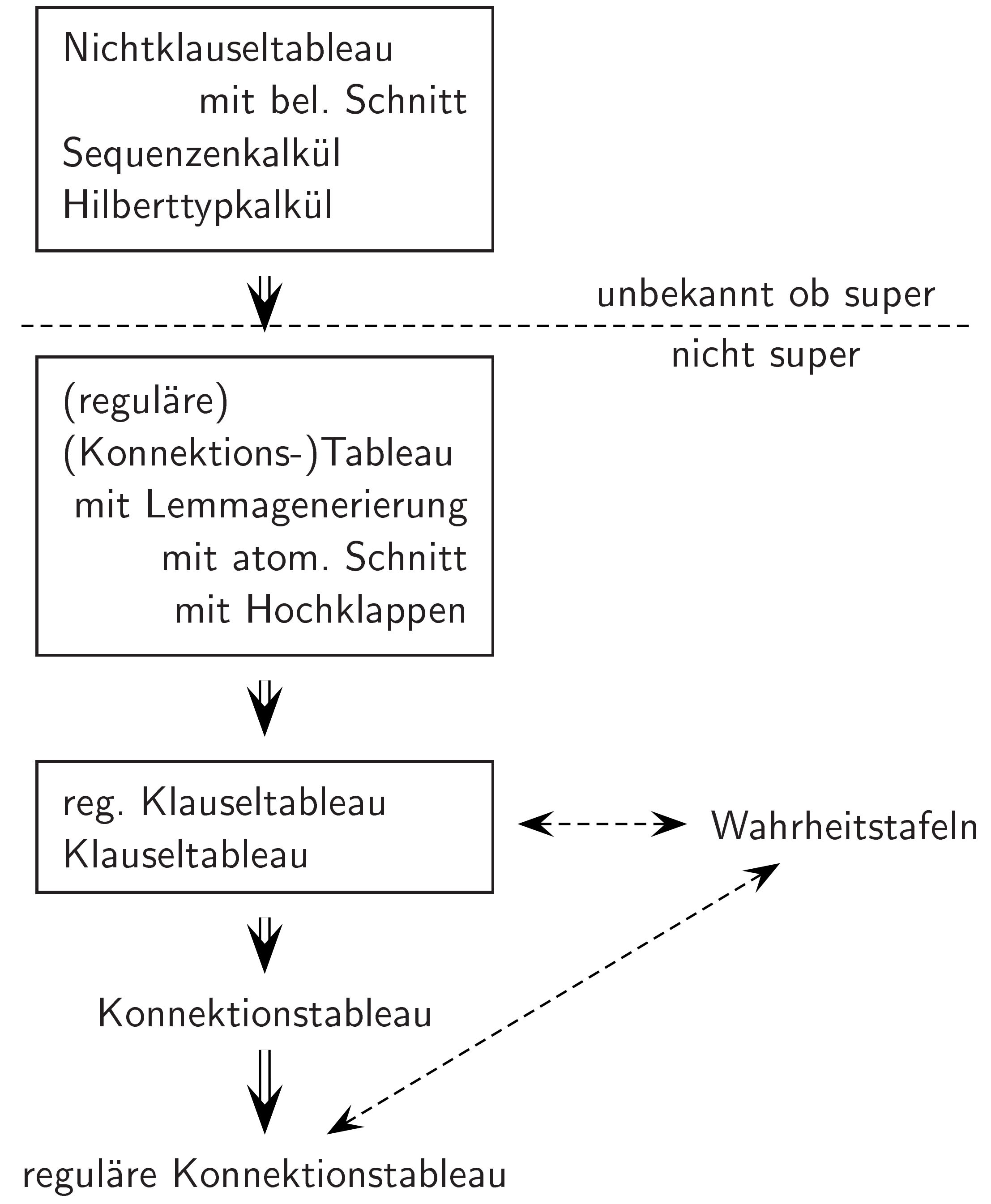
P und Q heißen **p-äquivalent** gdw sie sich gegenseitig p-simulieren können. ■

Proposition. Sei P nicht super und P könne Q p-simulieren. Dann ist Q nicht super.

Die Relation „x kann y p-simulieren“ ist transitiv. ■

Die relative Beweiskomplexität ist nur wenig mit Praxistauglichkeit eines Kalküls korreliert:

- Kalkül $\neq$ Suchprozedur
- Je stärker ein Kalkül, desto größer sein Suchraum.
- Der Suchraum p-äquivalenter Kalküle kann sich deutlich unterscheiden.
- Wichtige Eigenschaften (Zielgerichtetheit, ...) werden nicht erfaßt.



$\Uparrow$ kann p-simul.en, nicht umgekehrt ⇡ kann nicht p-simul.en

p-äquivalent